frontiers | Research Topics

Image-based digital tools for diagnosis and surgical treatment: applications, challenges and prospects

Edited by
Laura Cercenelli, Adrian Elmi-Terander and Thomas Maal

Coordinated by
Virginia Mamone and Francesca Manni

Published in
Frontiers in Bioengineering and Biotechnology

FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-6453-0
DOI 10.3389/978-2-8325-6453-0

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Image-based digital tools for diagnosis and surgical treatment: applications, challenges and prospects

Topic editors
Laura Cercenelli — University of Bologna, Italy
Adrian Elmi-Terander — Stockholm Spine Center, Sweden
Thomas Maal — Radboud University Medical Centre, Netherlands

Topic coordinators
Virginia Mamone — Cisanello Hospital, Italy
Francesca Manni — Philips, Netherlands

Citation
Cercenelli, L., Elmi-Terander, A., Maal, T., Mamone, V., Manni, F., eds. (2025). *Image-based digital tools for diagnosis and surgical treatment: applications, challenges and prospects*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-6453-0

Table of contents

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Editorial
PUBLISHED 29 May 2025
DOI 10.3389/fbioe.2025.1597238

Check for updates

OPEN ACCESS

EDITED BY
Dan Du,
Washington State University, United States

REVIEWED BY
Shichao Ding,
University of California, San Diego,
United States

*CORRESPONDENCE
Laura Cercenelli,
laura.cercenelli@unibo.it

RECEIVED 20 March 2025
ACCEPTED 19 May 2025
PUBLISHED 29 May 2025

CITATION
Cercenelli L, Elmi-Terander A, Maal T, Mamone V and Manni F (2025) Editorial: Image-based digital tools for diagnosis and surgical treatment: applications, challenges, and prospects .
Front. Bioeng. Biotechnol. 13:1597238.
doi: 10.3389/fbioe.2025.1597238

COPYRIGHT
© 2025 Cercenelli, Elmi-Terander, Maal, Mamone and Manni. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Image-based digital tools for diagnosis and surgical treatment: applications, challenges, and prospects

Laura Cercenelli[1]*, Adrian Elmi-Terander[2], Thomas Maal[3], Virginia Mamone[4] and Francesca Manni[5]

[1]eDIMES Lab - Laboratory of Bioengineering, Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy, [2]Department of Clinical Neuroscience, Karolinska Institutet (KI), Stockholm, Sweden, [3]3D lab, Radboud University Nijmegen Medical Centre, Nijmegen, Netherlands, [4]Department of Computer Science, The EndoCAS Center for Computer-Assisted Surgery, University of Pisa, Pisa, Italy, [5]Department of Electrical Engineering, Eindhoven University of Technology (TU/e), Eindhoven, Netherlands

KEYWORDS

3D modeling and printing, augmented (virtual) reality, radiomics, image-based diagnosis, surgical planning and simulation

Editorial on the Research Topic
Image-based digital tools for diagnosis and surgical treatment: applications, challenges, and prospects authors

Introduction

The advent of image-based digital tools has revolutionized the landscape of modern medicine, particularly in diagnosis and surgical treatment. Advances in deep learning, computer vision, 3D modeling, 3D printing and Augmented Reality (AR) have significantly enhanced the precision and efficacy of medical procedures, reducing human error and improving patient outcomes.

The articles collected within this Research Topic explore various facets of these emerging technologies, highlighting their applications, challenges, and prospects.

The collected studies can be grouped into the following focus areas of the Research Topic.

Innovations in medical image segmentation and radiomics

Deep learning-based segmentation has emerged as a cornerstone in medical image analysis. Also radiomics, i.e., the extraction of quantitative features from medical images, has gained traction as a tool for enhancing diagnostic accuracy.

The study by Abidin et al. presents a comprehensive review of brain tumor segmentation using multi-modal MRI and deep learning techniques. Their survey categorizes state-of-the-art models into CNN-based, transformer-based, and hybrid architectures, underscoring the strengths and weaknesses of each approach.

Mao et al. introduce a 2D medical image segmentation framework called Progressive Learning Network (PL-Net), which optimizes medical image segmentation by integrating coarse-to-fine semantic learning without increasing computational complexity.

Sun et al. present the innovative DA-TransUNet framework, which incorporates a Dual Attention Block (DA-Block) — combining position and channel attention—into a Transformer-enhanced U-Net architecture. This tailored design allows to improve segmentation accuracy particularly for high-detail medical images by refining feature extraction and filtering out irrelevant information.

The study by Lv et al. proposed an automated method for mandibular canal segmentation using a transformer-based neural network with cl-Dice loss and pixel-level feature fusion to improve accuracy. Their approach addresses challenges like sample imbalance and unclear boundaries through mandibular foramen localization, contrast enhancement, and pre-training with Deep Label Fusion on synthetic datasets. The method achieved state-of-the-art results, demonstrating high precision and robustness in 3D mandibular canal localization.

Jia et al. explore the application of radiomics in optimizing diagnosis and surgical planning for chronic osteomyelitis. Their findings indicate that an expanded region of interest (ROI) in MRI scans improves predictive performance, offering valuable insights for precision medicine approaches.

Finally, the study by Chen et al. reviews recent advances in retinal vessel quantification technology based on fundus imaging, highlighting its key role in detecting and monitoring ocular and systemic diseases. It focuses on how innovations in imaging and AI have enhanced diagnostic accuracy, offering clinicians and researchers an updated overview of its clinical applications.

Across these studies, a shared emphasis emerges on refining feature extraction and dealing with data complexity through attention mechanisms or radiomic modeling. Several approaches tackle the ongoing challenges of data imbalance and annotation scarcity, particularly in high-detail segmentation tasks. Additionally, transformer-based methods consistently appear as a unifying trend aimed at capturing long-range dependencies and improving contextual awareness.

Advances in 3D imaging, modeling and printing for surgical planning and regenerative medicine

The rapid evolution of 3D imaging, modeling and printing technologies is creating unparalleled opportunities to optimize surgical planning, improve patient outcomes, and driving progress in regenerative medicine.

In this context, Henckel et al. advanced the validation of low-dose 3D-CT imaging for acetabular implant orientation in orthopedic surgery. Their findings confirmed that 3D-CT is a highly accurate and precise modality for measuring cup inclination and version in total hip arthroplasty, outperforming conventional 2D CT methods. The results support 3D-CT as a standard for post-operative evaluation and surgical system optimization.

Similarly, Wu et al. developed novel 2D and 3D CT-based injury models for predicting the risk of femoral neck fracture in patients with femoral head fractures. Their quantitative parameters—such as percentage of maximum defect length and fracture area—showed high diagnostic accuracy and can serve as reliable predictors of fracture risk. This approach provides clinicians with valuable decision-making tools in trauma management and preoperative risk assessment.

In the domain of 3D printing, Capellini et al. presented a brief research report on the application of 3D printed models for the management of complex congenital heart disease. Their work underscores how physical models derived from high-resolution MRI and CT datasets allow for a detailed and tactile understanding of intricate pediatric cardiac anatomies. Among the techniques explored, selective laser sintering (SLS) proved to be the most cost-effective and time-efficient solution.

Pisani et al. explored bioartificial scaffolds produced using solvent casting, electrospinning, and 3D printing within the field of regenerative medicine. Their findings demonstrate that hybrid scaffolds—merging various fabrication methods—achieve excellent cell viability and mechanical properties similar to native soft tissues.

Moving to veterinary surgery, Chambers et al. evaluated a custom 3D-printed cutting guide for canine caudal maxillectomy, finding that it improved surgical accuracy for both experienced and novice surgeons. While it slightly increased procedure time, the enhanced precision, especially in achieving oncologic margins, suggests that such digital tools can assist less experienced surgeons and improve outcomes in veterinary oncology.

The studies in this area collectively highlight the growing reliance on patient-specific models, whether for orthopedic planning and post-operative evaluation, trauma prediction, or surgical simulation. A consistent theme is the shift toward integrating anatomical fidelity with fabrication feasibility—balancing precision and efficiency. Moreover, the translational relevance of these models is emphasized both in human and veterinary medicine, underscoring the broad applicability of 3D digital workflows.

Augmented reality in surgery and patient education

The integration of AR into surgery and patient education is another transformative development improving surgical accuracy and enriching communication between clinicians and patients.

Nasir et al. provide a systematic review of AR applications in orthopedic and maxillofacial oncological surgeries, detailing its potential to enhance precision through improved visualization. The study emphasizes the need for further clinical validation and the integration of external navigation systems to improve accuracy.

Similarly, Urlings et al. investigate the role of AR in patient education for intracranial aneurysms, highlighting how immersive visualizations can enhance patient understanding and shared decision-making.

Both studies highlight the visualization benefits of AR, while also pointing to common implementation hurdles—particularly the challenges of integrating AR with surgical navigation systems and the need for robust clinical validation. Whether applied in the operating room or during patient consultations, AR consistently emerges as a powerful tool for enhancing spatial understanding. However, technical limitations continue to hinder its broader adoption in clinical settings.

Conclusion and future directions

The contributions within this Research Topic underscore the transformative impact of digital imaging tools in diagnosis and surgical treatment. From AI-driven segmentation techniques to AR-enhanced surgical interventions, these advancements promise a future where precision medicine is more accessible, efficient, and patient-centered.

Across all sections, a shared trajectory emerges: the shift toward more personalized, data-driven, and visually guided approaches to healthcare. While each study presents specific technological innovations, they collectively reflect a broader trend of converging digital methodologies—such as the use of transformers in segmentation, patient-specific 3D models for planning, and immersive AR for both surgeons and patients. These tools not only enhance accuracy and confidence in clinical decision-making but also foster interdisciplinary collaboration among clinicians, engineers, and data scientists.

Despite these promising advancements, challenges remain on the path to real-world clinical translation. In AI-based segmentation tools, the major issues deal with computational efficiency, the need for large, annotated datasets, and model interpretability. In surgery, AR faces limitations in real-time tracking accuracy and workflow integration. For 3D modeling and printing, the translation of digital data into physical models requires standardized protocols and cost-effectiveness evaluations for broader clinical adoption.

Moreover, regulatory hurdles and data governance policies pose significant obstacles, especially for AI applications and patient-specific devices. Bridging the gap between innovation and implementation will require not only technical refinement but also rigorous validation, regulatory clarity, and integration into clinical training and reimbursement frameworks.

The insights from the included studies suggest that an integrated ecosystem—where data, tools, and expertise converge—can accelerate the transition from research innovation to bedside impact. Continued research and cross-disciplinary collaborations between computer scientists, engineers, radiologists, and surgeons will be essential for translating innovations from the lab into clinical practice and for achieving the full potential of image-based digital tools in delivering safer, more effective, and tailored patient care.

Author contributions

LC: Writing – original draft, Writing – review and editing. AE-T: Writing – review and editing. TM: Writing – review and editing. . VM: Writing – review and editing. FM: Writing – review and editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

TYPE Original Research
PUBLISHED 17 January 2023
DOI 10.3389/fbioe.2023.1115639

Check for updates

OPEN ACCESS

EDITED BY
Laura Cercenelli,
University of Bologna, Italy

REVIEWED BY
Michele Fiore,
IRCCS Azienda Ospedaliero-Universitaria di Bologna, Italy
Andrea Sambri,
IRCCS Azienda Ospedaliera Universitaria di Bologna, Italy
Francesca Manni,
Eindhoven University of Technology, Netherlands

*CORRESPONDENCE
Jiong Mei,
✉ meijiong@sjtu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Biomechanics, a section of the journal Frontiers in Bioengineering and Biotechnology

RECEIVED 06 December 2022
ACCEPTED 04 January 2023
PUBLISHED 17 January 2023

CITATION
Wu S, Wang W, Li R, Guo J, Miao Y, Li G and Mei J (2023), Fractured morphology of femoral head associated with subsequent femoral neck fracture: Injury analyses of 2D and 3D models of femoral head fractures with computed tomography. *Front. Bioeng. Biotechnol.* 11:1115639.
doi: 10.3389/fbioe.2023.1115639

COPYRIGHT
© 2023 Wu, Wang, Li, Guo, Miao, Li and Mei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Fractured morphology of femoral head associated with subsequent femoral neck fracture: Injury analyses of 2D and 3D models of femoral head fractures with computed tomography

Shenghui Wu[1†], Wei Wang[2†], Ruiyang Li[1], Jingyi Guo[3], Yu Miao[1], Guangyi Li[1] and Jiong Mei[1*]

[1]Department of Orthopedic Surgery, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China, [2]Department of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong, China, [3]Clinical Research Center, Shanghai Sixth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China

Background: The injury of femoral head varies among femoral head fractures (FHFs). In addition, the injury degree of the femoral head is a significant predictor of femoral neck fracture (FNF) incidence in patients with FHFs. However, the exact measurement methods have yet been clearly defined based on injury models of FHFs. This study aimed to design a new measurement for the injury degree of the femoral head on 2D and 3D models with computed tomography (CT) images and investigate its association with FHFs with FNF.

Methods: A consecutive series of 209 patients with FHFs was assessed regarding patient characteristics, CT images, and rate of FNF. New parameters for injury degree of femoral head, including percentage of maximum defect length (PMDL) in the 2D CT model and percentage of fracture area (PFA) in the 3D CT-reconstruction model, were respectively measured. Four 2D parameters included PMDLs in the coronal, cross-sectional and sagittal plane and average PMDL across all three planes. Reliability tests for all parameters were evaluated in 100 randomly selected patients. The PMDL with better reliability and areas under curves (AUCs) was finally defined as the 2D parameter. Factors associated with FNF were determined by binary logistic regression analysis. The sensitivity, specificity, likelihood ratios, and positive and negative predictive values for different cut-off values of the 2D and 3D parameters were employed to test the diagnostic accuracy for FNF prediction.

Results: Intra- and inter-class coefficients for all parameters were ≥0.887. AUCs of all parameters ranged from 0.719 to 0.929 ($p < 0.05$). The average PMDL across all three planes was defined as the 2D parameter. The results of logistic regression analysis showed that average PMDL across all three planes and PFA were the significant predictors of FNF ($p < 0.05$). The cutoff values of the average PMDL across all three planes and PFA were 91.65% and 29.68%. The sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, predictive positive value and negative predictive value of 2D (3D) parameters were 91.7% (83.3%), 93.4% (58.4%), 13.8 (2.0), 0.09 (0.29), 45.83% (10.87%), and 99.46% (98.29%).

Conclusion: The new measurement on 2D and 3D injury models with CT has been established to assess the fracture risk of femoral neck in patients with FHFs in the

10.3389/fbioe.2023.1115639

clinic practice. 2D and 3D parameters in FHFs were a feasible adjunctive diagnostic tool in identifying FNFs. In addition, this finding might also provide a theoretic basis for the investigation of the convenient digital-model in complex injury analysis.

KEYWORDS

femoral head fracture, femoral neck fracture, 2D, 3D, injury model, computed tomography

Background

Femoral head fractures (FHFs) are usually associated with high-energy trauma and posterior hip dislocation. The major choice of surgical FHF treatment was reduction and internal fixation, or fragment removal, while joint replacement surgeries were favored for some FHF patients due to the presence of the ipsilateral femoral neck fracture (FNF) (Menger et al., 2021). The ipsilateral FNF had an incidence of 8.6% and the worst prognosis among all FHFs (Menger et al., 2021 Giannoudis et al., 2009). There were many possible scenarios, simultaneously with FHF by trauma, during reduction, during an eventual FHF fixation, upon resumption of the weight bearing, *etc.*, for the occurrence of the ipsilateral FNF.

In the Pipkin classification, a single FHF was called Pipkin I or II, and associated fractures of the femoral neck and acetabulum were respectively defined as Pipkin III and Ⅳ (Pipkin, 1957). During treatment, the ipsilateral FNF is a severe intraoperative and postoperative complication of closed reduction and internal fixation of FHF with posterior hip dislocation. This may turn Pipkin I and II into a rare Pipkin III, with increased risk of femoral head avascular necrosis (AVN) (Park et al., 2016). Open reduction is the most commonly used method for iatrogenic FNF prevention in clinical practice. However, the ability of open reduction to reduce incidence of AVN is controversial. Although open reduction can lower the risk of intraoperative FNF caused by closed reduction, it is associated with complications of AVN, with a high risk of injury to vascular supply to the femoral head (Guo et al., 2010). Moreover, open reduction may fail to reverse AVN followed by joint replacement, when femoral neck refractures occur after FHF internal fixation without any injury. Hence, early recognition of the risk of FNF in FHFs is crucial, as the early identified characteristics of this injury type could help to make a more rational treatment strategy and improve the prognosis. However, there is a lack of prediction tools in clinical practice.

TABLE 1 Background characteristics of patients with femoral head fractures.

Characteristics	
Side of injury [no. (%)]	
Left	102 (48.8%)
Right	107 (51.2%)
Age (yr)	39.85 ± 16.41
Sex [no. (%)]	
Male	169 (80.9%)
Female	40 (19.1%)
Pipkin Classification [no. (%)]	
I	3 (1.4%)
II	60 (28.7%)
III	8 (3.8%)
IV	138 (66.0%)
2D Measurement-Percentage of maximum defect length (%)	
Coronal plane	76.41 ± 15.52
Cross-sectional plane	78.82 ± 16.45
Sagittal plane	78.57 ± 17.10
Average	77.93 ± 13.51
3D Measurement-Percentage of fracture area (%)	26.76 ± 12.95

Previous studies have reported an association between the injury degree of femoral head and the incidence of FNF in FHFs (Davis, 1950; Henle et al., 2017). FHF is generally characterized by development of a shearing force against the acetabular rim caused by injury to the hip joint (Davis, 1950; Thompson and Epstein, 1951). The contact area between the femoral head surface and acetabular wall surface determines whether FHFs occur with or without the additional osseous lesion on the femoral neck (Henle et al., 2017). When the contact area of the femoral head surface is larger, more of the axial compression force transmitted to the hip is distributed to the surrounding bone. This results in FNFs besides FHFs. Hence, the injury degree of femoral head is a significant predictor of FNF incidence in patients with FHFs (Davis, 1950; Thompson and Epstein, 1951; Henle et al., 2017).

To date, no study has described an accurate and reliable method for measuring the injury degree of femoral head for predicting the incidence of FNF in FHFs. In addition, the hip geometry for predicting diseases of proximal femur has been studied extensively, but the exact measurement regarding FHFs has yet been clearly defined (Kazemi et al., 2016; Wells et al., 2017; Hesper et al., 2018). Moreover, clinically suspected FHF is routinely assessed using computed tomography (CT) examinations over the last few decades. Comprehensive analysis of two-dimensional (2D) and three-dimensional (3D) human models based on CT revealed femoral morphology in hip diseases (Wells et al., 2017; Hesper et al., 2018). Therefore, it is now possible to design a digital tool for identifying the risk of FNF in FHFs based on CT imaging.

This study aimed to design a method for measuring the injury degree of femoral head based on 2D and 3D injury models with CT. The secondary aim was to investigate the association of the injury degree of femoral head with incidence of FNFs in FHFs with CT. We hypothesized that 2D and 3D CT-based parameters were reliable predictors of FNF in patients with FHFs.

Methods

Study design

This study was approved by the Institutional Review Board (IRB) of our institution [No.2022-KY-026(K)] and followed the Strengthening the Reporting of Observational Studies in

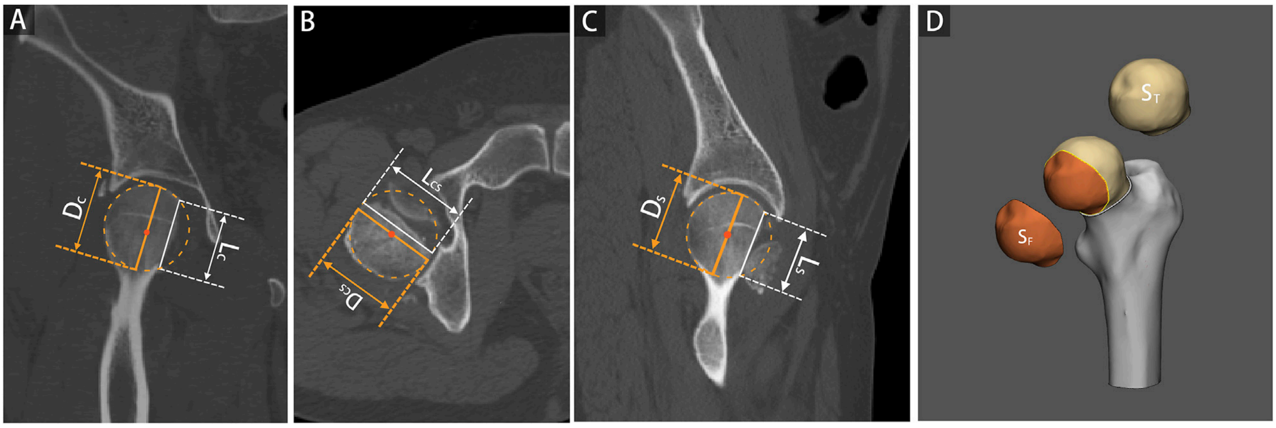


FIGURE 1
2D and 3D parameter measurements based on 2D and 3D injury models of femoral head fractures. The percentage of maximum defect length (PMDL) was measured as the ratio between the most significant residual defect length (L) and the optimum circle diameter for the femoral head (D), including coronal plane [Panel **(A)**, $PMDL_C = L_C/D_C$], cross-sectional [Panel **(B)**, $PMDL_{CS} = L_{CS}/D_{CS}$], and sagittal plane [Panel **(C)**, $PMDL_S = L_S/D_S$]. The percentage of fracture area (PFA) was determined as the ratio between the fracture area and the total area of the femoral head [Panel **(D)**, $PFA = S_F/S_T$].

Epidemiology (STROBE) reporting guidelines for cross-sectional studies (von Elm et al., 2007). Given the retrospective nature of this study, participant informed consent was waived by IRB.

Participants

Using a prospectively maintained orthopedic database at a large level-I trauma center, we retrospectively analyzed CT images of the hip joint in patients diagnosed with FHF between 2011 and 2022. Three investigators independently reviewed the imaging data of all FHFs to identify initially missed FHFs. A total of 228 FHFs in 228 patients were included. Patients with insufficient or poor quality axial CT images (i.e., images with severe artifacts) ($N = 12$), unclosed epiphyseal line of the femoral head ($N = 2$), pathological fracture ($N = 3$), or skeletal immaturity ($N = 2$) were also excluded. Finally, 209 hips in 209 patients were included, and there were 12 cases with FHF and FNF. Patient characteristics and features of FHFs are summarized in Table 1.

Injury model generation

Raw CT data of patients with femoral head fractures were obtained and imported into the Mimics software (Materialise, Leuven, Belgium) for image analysis. The 2D injury model of FHF was defined as an evident 2D image of maximum bone defect of femoral head in each CT plane. The 3D injury model of FHF was defined as the 3D-reconstruction model of the 3D-CT data using 3-matic software (Materialise, Belgium). Additional details of the reconstruction of 3D model analysis are given in the Appendix A.

Measurements

Each 2D injury model of FHF, including coronal plane, cross-sectional, and sagittal plane, were separately used to measure the percentage of maximum defect length (PMDL). The 2D parameter in each 2D injury model was defined as the ratio between the most significant residual defect length and the optimum circle diameter of the femoral head (Wells et al., 2017; Hesper et al., 2018), (Figures 1A–C). Subsequently, the PMDL of each plane and average PMDL across three planes were determined. The 3D parameter of the percentage of fracture area (PFA) in each 3D injury model was defined as the ratio between the fracture area and total area of the femoral head (Figures 1D). Additional details of the measurement process of the 3D parameter are also given in the supplementary appendix.

To test the repeatability and reliability of PMDL and PFA parameters, 100 randomly selected FNFs were independently measured by two investigators using 2D-CT and 3D-CT images for inter-observer analysis. One of the investigators conducted two additional measurements 1 month apart for intra-observer analysis.

Statistical analysis

Qualitative data were presented as a percentage whereas quantitative data were expressed as the mean with standard deviation (SD) using SPSS 24.0 software (IBM SPSS Inc., Armonk, New York). Reproducibility and agreement of parameters were tested using Bland-Altman (Bland and Altman, 1986) and intraclass correlation (ICC) (Rousson et al., 2002). The ICCs were interpreted according to a method by Landis and Koch (Landis and Koch, 1977). Factors significantly associated with FNF, including the injured side, age, gender, Pipkin classification, the 2D parameters, and the 3D parameter, were determined by binary logistic regression analysis. The diagnostic accuracy of significant 2D or 3D parameters was determined from the area under receiver operating characteristic curves or AUCs. Optimal cutoff value of each parameter was calculated using the receiver operating characteristic (ROC) curve. Patients were grouped according to the cutoff of the 2D parameter, with higher diagnostic accuracy, and the 3D parameter, respectively. Differences between groups were analyzed using the chi-square test or

TABLE 2 Inter-observer and intra-observer reliability.

Parameters		Inter-observer reliability		Intra-observer reliability	
		ICC	95% CI	ICC	95% CI
Two-dimensional Measurement[a]	Coronal plane	0.926	0.845–0.959	0.887	0.837–0.922
	Cross-sectional plane	0.926	0.873–0.954	0.986	0.979–0.990
	Sagittal plane	0.915	0.870–0.944	0.976	0.965–0.984
	Average	0.955	0.849–0.980	0.975	0.963–0.983
Three-dimensional Measurement[b]		0.987	0.981–0.992	0.994	0.991–0.996

ICC: intraclass correlation coefficient; CI: confidence interval.
[a]Two-dimensional measurement of the percentage of maximum defect length.
[b]Three-dimensional measurement of the percentage of fracture area.

Fisher exact probability method for dichotomized values and the Mann-Whitney test for continuous values. Differences were considered statistically different at $p < 0.05$. The sensitivity, specificity, likelihood ratios, and positive and negative predictive values for different parameter cut-off values were also calculated.

Power calculation was performed by using PASS 15 Power Analysis and Sample Size Software (2017). NCSS, LLC. Kaysville, Utah, United States, ncss.com/software/pass. The primary outcome measures were cutoff values and area under the curve (AUC) of ROC curves for the FNF occurrence prediction. The secondary outcomes included the reproducibility and agreement of parameters and subgroup analyses based on cutoff values.

Results

Baseline characteristics

After screening 228 patients (228 hips) in our hospital's orthopedic database, 19 patients were excluded. Among those excluded were eleven patients with insufficient or poor-quality CT images and two patients with an unclosed epiphyseal line of the femoral head. Five more patients were excluded for having a pathological fracture or skeletal immaturity. One patient with iatrogenic FNF was excluded for lacking CT images. Therefore, 209 hips in 209 patients comprising 102 left hip injuries (48.8%) and 107 right hip injuries (51.2%), were finally analyzed. Of these, 12 patients had FHF and FNF, including 7 Pipkin III, 2 Pipkin IV, 1 iatrogenic FNF during closed reduction, and 2 refractures of the femoral neck after internal fixation without trauma or fall.

Parameter reproducibility and agreement test

The results showed an almost perfect inter-and intra-observer reliability with ICC ≥0.887 (95% CI 0.837–0.922) for 2D parameters tested. Similarly, an almost perfect inter-and intra-observer reliability was found for the 3D parameter with ICC ≥0.987 (95% CI 0.981–0.992) (Table 2). The Bland–Altman analysis of 2D and 3D parameters measured by the two observers also showed high concordance (Figure 2).

Correlation analysis between parameters and FNF occurrence

The binary logistic regression analysis of sex, age, injury side, Pipkin classification, 2D parameters, and 3D parameter showed that all 2D parameters except the sagittal plane and 3D parameter were significant predictors of FNF occurrence ($p < 0.05$) (Tables 3, 4).

ROC analysis for FNF occurrence prediction

The AUC for the average PMDL three planes with a cutoff value of 91.65% was favorable and significantly better compared with that for 2D parameters ($p < 0.001$). The results of ROC analysis showed that the optimal cutoff value of the 3D parameter was 29.68% (Figure 3).

Post-hoc power analysis

The *post hoc* power calculation based on AUCs in the primary outcome data showed that the 2D parameter, average PMDL across three planes, and the 3D parameter could predict FNF occurrence. The study power of 2D parameter and 3D parameter was 100% and 74.24%.

Diagnostic accuracy analysis

Subgroups were analyzed separately based on the cutoff values of 2D (average PMDL across all three planes) and 3D (PFA) parameters. Figure 4 shows the distribution of single FHFs and the number of FHFs combined with FNFs in the indicated subgroups. In total, 11/24 of FHFs with average PMDL ≥91.65% across all three planes were FNFs whereas 1/185 of FHFs with an average PMDL <91.65% across all three planes were FNFs ($p < 0.001$). In addition, 10/92 of FHFs with PFA ≥29.68% were FNFs whereas 2/117 of FHFs with PFA <29.68% were FNFs ($p = 0.005$) (Table 5).

The sensitivity and specificity of the 2D parameter were 91.7% and 93.4%, and the 3D parameter were 83.3% and 58.4%. The positive likelihood ratio and negative likelihood ratio of 2D parameter for FNF prediction were 13.8 and 0.09, and 3D ones were 2.0 and 0.29. The predictive positive value and negative predictive value for 2D parameter were 45.83% and 99.46%, and the 3D parameter were 10.87% and 98.29%.

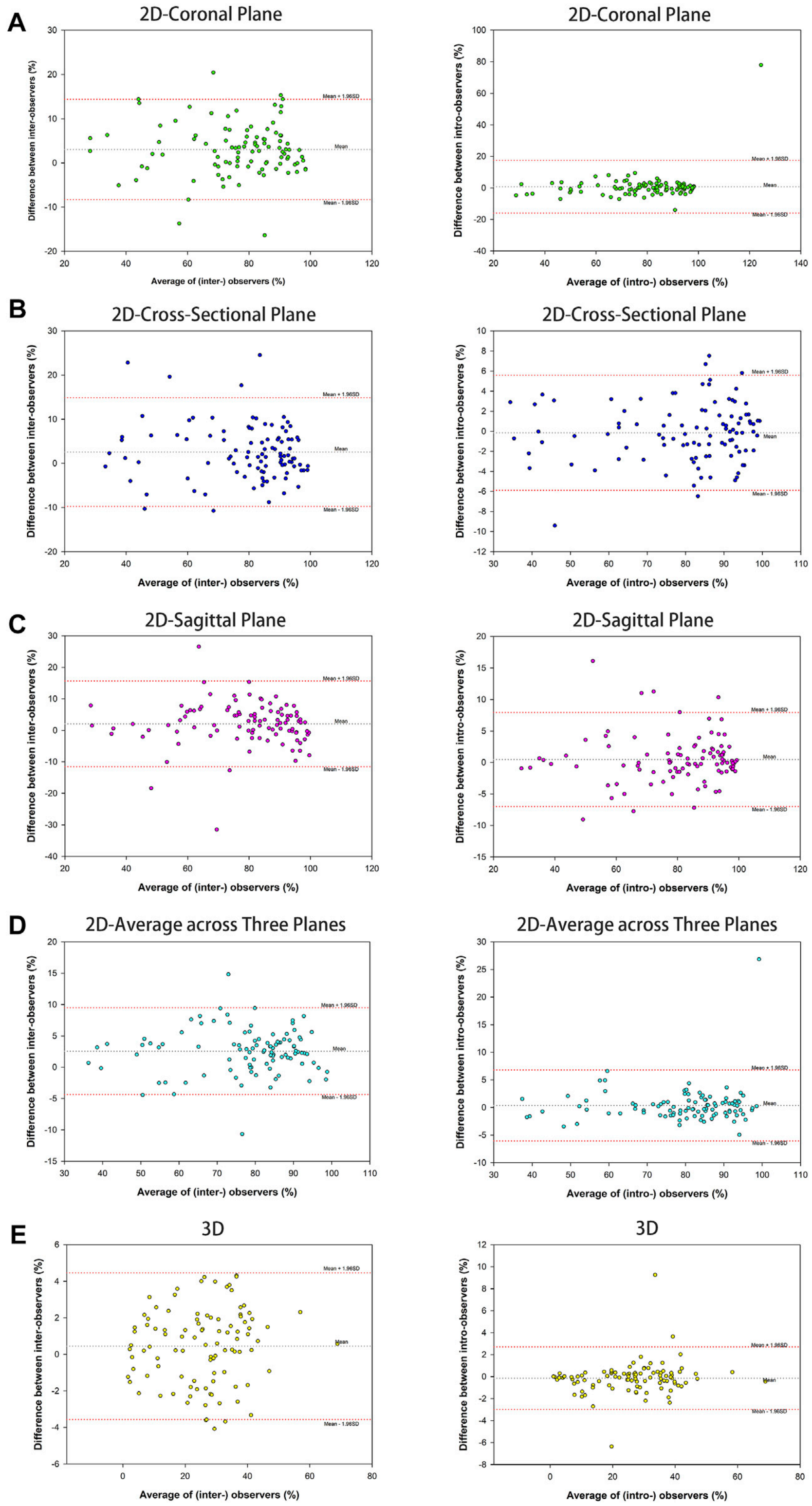


FIGURE 2
Bland-Altman plots of parameter measurement showed high assessment agreement of inter-observers and intra-observers, containing 2D parameter of the percentage of maximum defect length [coronal plane **(A)**, cross-sectional **(B)**, sagittal plane **(C)**, and average across three planes **(D)**] and 3D parameter of the percentage of fracture area **(E)**.

TABLE 3 Relationship between femoral neck fracture and patient characteristics for the 209 patients with femoral head fracture by univariate binary logistic regression analysis.

Variable	Femoral head fracture with femoral neck fracture	
	OR (95% CI)	*p*-value
Side of injury	0.665 (0.204–2.168)	0.499
Age	1.011 (0.977–1.046)	0.528
Sex	1.441 (0.372–5.586)	0.597
Pipkin classification		
I	Reference	
II		1.000
III		0.999
IV		0.999
2D Measurement-Percentage of maximum defect length		
Coronal plane	1.212 (1.089–1.348)	<0.001
Cross-sectional plane	1.363 (1.152–1.611)	<0.001
Sagittal plane	1.256 (1.098–1.437)	0.001
Average	1.470 (1.229–1.757)	<0.001
3D Measurement-Percentage of fracture area	1.080 (1.023–1.140)	0.005

TABLE 4 Relationship between femoral neck fracture and patient characteristics for the 209 patients with femoral head fracture by multivariate binary logistic regression analysis.

Variable	Femoral head fracture with femoral neck fracture	
	OR (95% CI)	*p*-value
2D measurement-Coronal plane		
Side of injury	0.318 (0.023–4.331)	0.390
Age	1.022 (0.955–1.094)	0.524
Sex	1.232 (0.070–21.788)	0.887
Pipkin classification		
I	Reference	
II		1.000
III		0.999
IV		0.999
Percentage of maximum defect length	1.257 (1.009–1.566)	0.042
2D Measurement-Cross-sectional plane		
Side of injury	1.080 (0.037–31.721)	0.964
Age	1.049 (0.939–1.173)	0.399
Sex	4.610 (0.085–249.166)	0.453
Pipkin classification		
I	Reference	
II		1.000
III		0.999
IV		1.000
Percentage of maximum defect length	1.997 (1.121–3.557)	0.019
2D Measurement-Sagittal plane		
Side of injury	0.784 (0.053–11.491)	0.859
Age	1.044 (0.965–1.130)	0.286
Sex	0.718 (0.029–17.977)	0.840

(Continued on following page)

TABLE 4 (*Continued*) Relationship between femoral neck fracture and patient characteristics for the 209 patients with femoral head fracture by multivariate binary logistic regression analysis.

Variable	Femoral head fracture with femoral neck fracture	
	OR (95% CI)	*p*-value
Pipkin classification		
I	Reference	
II		1.000
III		0.998
IV		0.999
Percentage of maximum defect length	1.308 (0.969–1.766)	0.080
2D Measurement-Average		
Side of injury	0.843 (0.031–22.949)	0.920
Age	1.018 (0.925–1.121)	0.710
Sex	1.650 (0.009–298.304)	0.850
Pipkin classification		
I	Reference	
II		1.000
III		0.998
IV		1.000
Percentage of maximum defect length	1.789 (1.102–2.903)	0.019
3D Measurement		
Side of injury	0.149 (0.008–2.651)	0.195
Age	1.008 (0.930–1.093)	0.842
Sex	0.613 (0.032–11.628)	0.744
Pipkin classification		
I	Reference	
II		1.000
III		0.999
IV		1.000
Percentage of fracture area	1.254 (1.012–1.554)	0.038

Discussion

The results of this study confirmed our hypothesis that the 2D and 3D CT-based parameters were reliable predictors of FNF in patients with FHFs. Furthermore, all parameters in 2D and 3D injury models were statistically significant when expressed as a continuous value. However, we presented 2D and 3D parameters for the injury degree of the femoral head as a dichotomized value. Therefore, knowledge of "50% FNF in patients with FHFs when average PMDL across all three planes exceeds 91.65% or PFA exceeds 29.68%" could be applied in clinical practice. To our knowledge, this is a single large consecutive case series of FHFs to measure 2D and 3D CT-based parameters and investigate their association with FNF in patients with FHFs. The newly exact measurement of injury degree of femoral head, based on 2D and 3D model analysis, has been established to assess the fracture risk of femoral neck, and it could be treated as a feasible adjunctive diagnostic tool in identifying FNFs in patients with FHFs. In this study, all parameters based on CT-based injury model was of great value in clinical application and research due to its convenience and favorable diagnostic performance.

In the past decade, a positive role of the position of the femoral head in relation to the acetabulum in iatrogenic FNF prevention was proposed by a small series of studies (Sy et al., 2005; Park et al., 2016). However, this conclusion remained controversial. Sy et al. (2005) reported that FNF might occur when femoral head defects were caught on acetabular rim during reduction movement. In addition, the femoral head remained attached above and posteriorly to the acetabulum and rotated less than 90° in four cases. In another retrospective study by Park et al. (2016), five of the nine patients experienced FNFs during attempted closed reduction. Fragments of the femoral head fracture were retained in the acetabulum in these series while the remaining component was posterior and superior relative to the acetabulum. In addition, the remaining component was engaged or locked against the sharp rear angle of the acetabulum. Therefore, the injury degree of femoral head is essentially a significant predictor of the incidence of FNF in FHFs during attempted reduction (Davis, 1950; Thompson and Epstein, 1951; Henle et al., 2017). However, because of poor methodological rigor inherent in qualitative reports, a relatively small sample size, and lack of

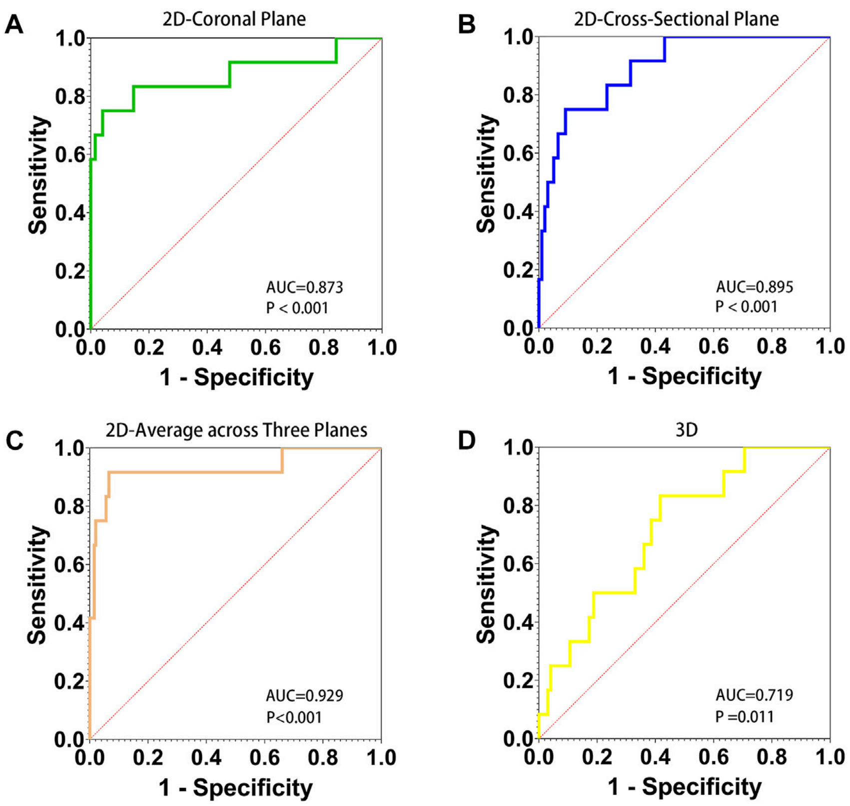


FIGURE 3
AUC, area under the ROC curve. Receiver operating characteristic (ROC) curve for 2D and 3D parameters to predict refracture of the femoral neck after femoral head fractures. The diagnostic accuracy of 2D parameter of the percentage of maximum defect length of coronal plane **(A)**, cross-sectional **(B)**, average across three planes **(C)**, and 3D parameter of the percentage of fracture area **(D)** was shown.

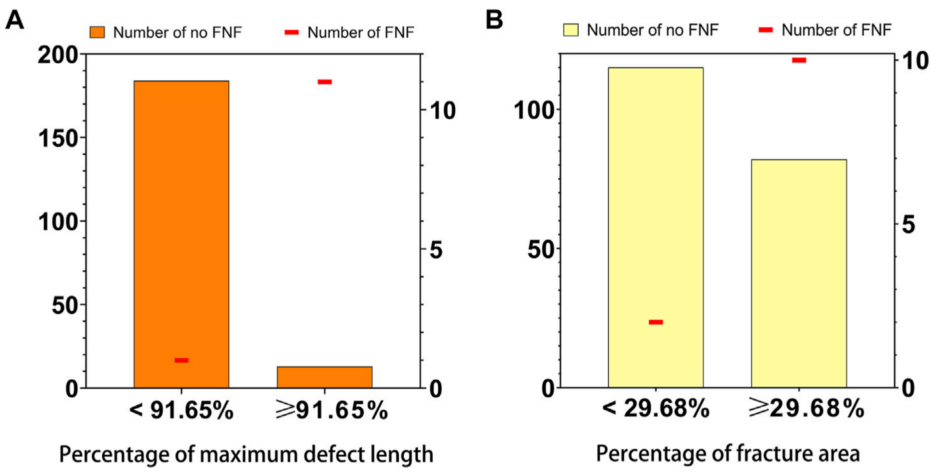


FIGURE 4
Distribution of 2D parameter **(A)**, the average percentage of maximum defect length across three planes, and 3D parameter **(B)**, the percentage of fracture area, and the number of the femoral neck fracture in the 209 patients with femoral head fractures. Numbers were indicated [femoral head fracture with (line in red) or without (bar in orange or yellow) femoral neck fracture] in the relevant group.

measurement criterion and comparison group, the level of evidence for clinical application of positioning of the femoral head in relation to the acetabulum is very low.

In the current study, 12 patients were found with FHF and FNF, including concomitant fracture (Pipkin III and IV), iatrogenic FNF during closed reduction (Figure 5), and femoral neck refracture after

TABLE 5 Data for 209 patients with femoral head fractures.

Variable	2D measurement of average across three planes			3D measurement		
	Percentage of maximum defect length	Percentage of maximum defect length	*p*-value	Percentage of fracture area	Percentage of fracture area	*p*-value
	<91.65%	≥91.65%		<29.68%	≥29.68%	
Number	185 (100)	24 (100)		117 (100)	92 (100)	
Side of injury			0.576[a]			0.100[a]
Left	89 (48.1)	13 (54.2)		63 (53.8)	39 (42.4)	
Right	96 (51.9)	11 (45.8)		54 (46.2)	53 (57.6)	
Age	38.0 (25.5–50.0)	37.5 (23.3–53.3)	0.873[b]	36.0 (24.5–50.0)	39.5 (26.0–54.0)	0.282[b]
Sex			0.581[c]			0.889[a]
Male	148 (80.0)	21 (87.5)		95 (81.2)	74 (80.4)	
Female	37 (20.0)	3 (12.5)		22 (18.8)	18 (19.6)	
Pipkin classification			<0.001[c]			0.006[c]
I	3 (1.6)	0 (0.0)		3 (2.6)	0 (0.0)	
II	53 (28.6)	7 (29.2)		26 (22.2)	34 (37.0)	
III	1 (0.5)	7 (29.2)		2 (1.7)	6 (6.5)	
IV	128 (69.2)	10 (41.7)		86 (73.5)	52 (56.5)	
Simplified Pipkin classification			0.912[a]			0.057[a]
I + II	56 (30.3)	7 (29.2)		29 (24.8)	34 (37.0)	
III + IV	129 (69.7)	17 (70.8)		88 (75.2)	58 (63.0)	
Femoral neck fracture	1 (0.5)	11 (45.8)	<0.001[c]	2 (1.7)	10 (10.9)	0.005[a]

[a]Number of patients (percentage) and *p*-values determined with the chi-square test.
[b]Median (interquartile range) and *p*-values derived with the Mann-Whitney test.
[c]Number of patients (percentage) and *p*-values determined with the Fisher exact test.

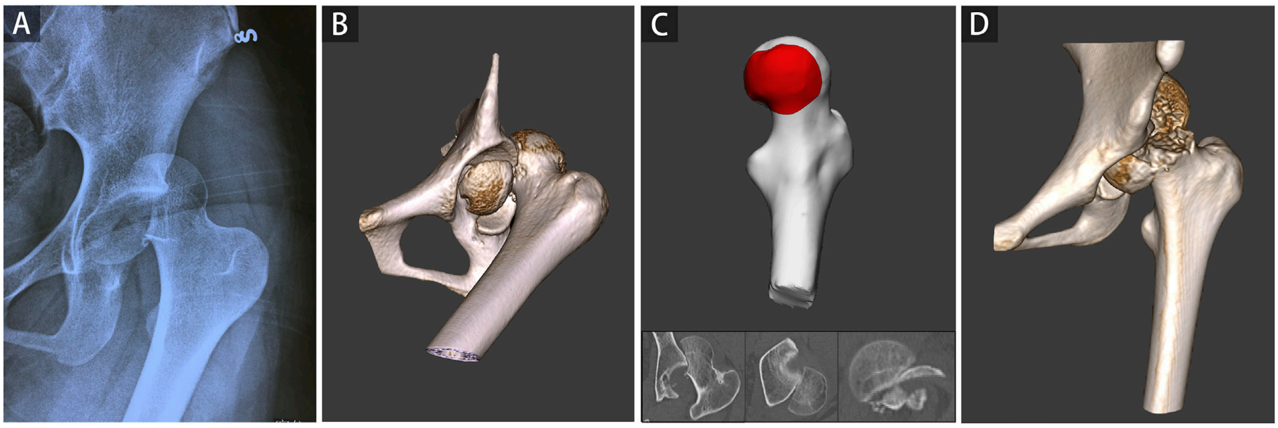


FIGURE 5
A representative case of iatrogenic femoral neck fracture during closed reduction was shown. A 16-year-old girl involved in a severe traffic accident was diagnosed with femoral head fracture based on X-ray **(A)** and CT **(B)**. The average percentage of maximum defect length across three planes of 92.80% in the 2D injury model and the percentage of fracture area of 38.33% in the 3D injury model were determined **(C)**. A refracture of the femoral neck occurred during closed reduction **(D)**.

FHF internal fixation without trauma. All three kinds of FHFs with FNF were included in this study to make 2D and 3D parameters more universally applicable. The present case illustrates patient's refracture of the femoral neck after FHF internal fixation without trauma or fall (Figure 6). No similar findings have been previously reported. Therefore, we speculated that a naturally favorable stress

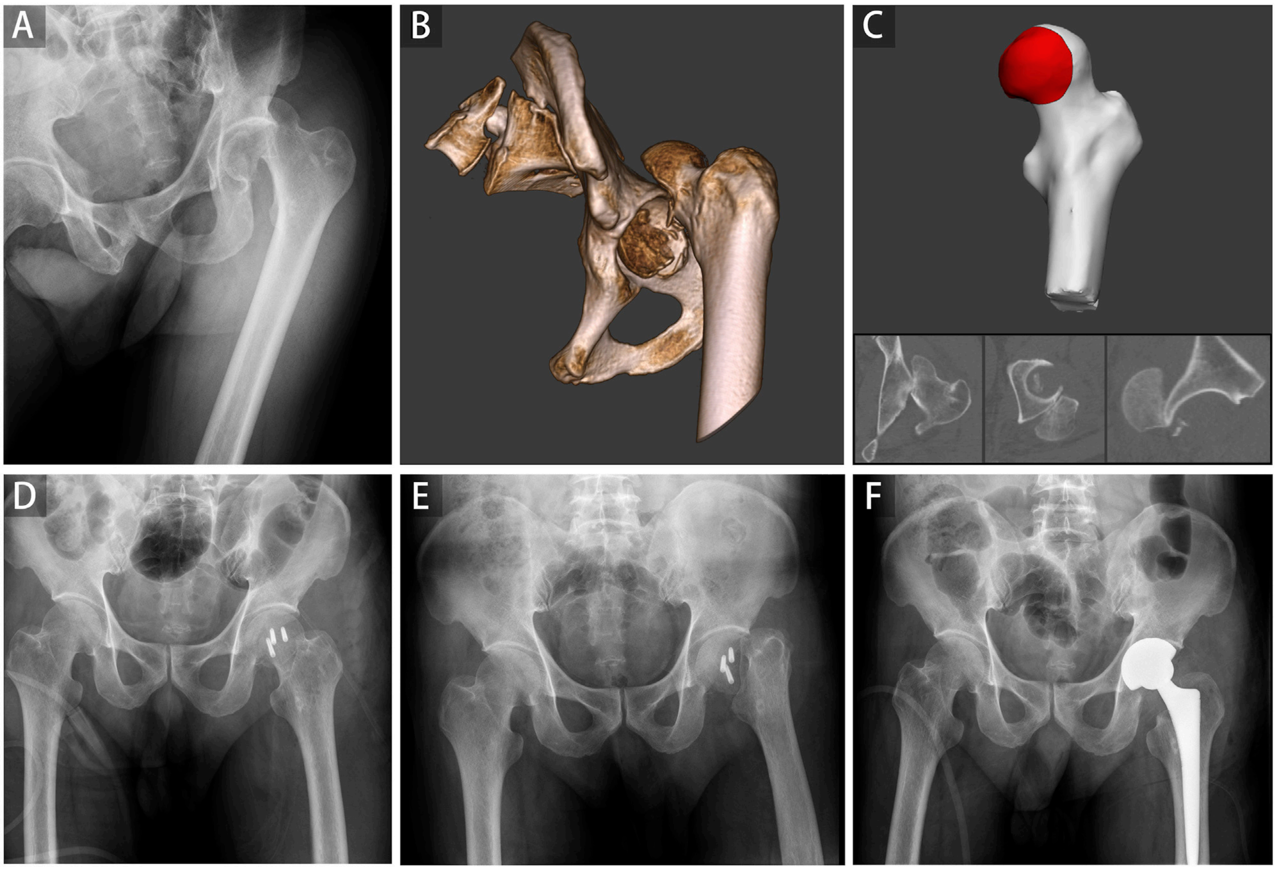


FIGURE 6
A representative case of the refracture of the femoral neck after internal fixation of the femoral head fracture was shown. A 62-year-old man involved in a severe traffic accident was diagnosed with femoral head fracture based on X-ray **(A)** and CT **(B)**. The average percentage of maximum defect length across three planes of 96.64% in the 2D injury model and the percentage of fracture area of 40.72% in the 3D injury model were determined **(C)**. Open reduction and internal fixation were conducted, and the postoperative radiograph showed good reduction status **(D)**. A refracture of the femoral neck occurred without trauma and fall 3 weeks postoperatively **(E)**. Finally, hip replacement surgery was performed **(F)**.

distribution mechanism could not be balanced by FHF internal fixation, which causes femoral neck refracture without trauma or fall. With approximately 70% of the articular surface of the femoral head engaging in load transfer, bone defects of the femoral head can cause significant load changes in the femoral head and neck after fracture (Greenwald and Haynes, 1972). Thus, ensuring anatomic congruity in the articular surface is important for effective management of FHFs (Beebe et al., 2016). However, due to the difficulty in restoring the natural anatomic hip structure, FHF combined with FNF is associated with a more dismal prognosis than a single FHF (Giannoudis et al., 2009). Hence, early identification of these patients could improve their prognosis by providing more aggressive treatment strategies. Our findings showed a high risk of FNF when the injury degree of the femoral head reached a critical value determined from 2D and 3D parameters.

This study also corroborates previous findings regarding the percentage of the femoral head fragments that can be removed in the treatment of FNFs (Epstein, 1974; Epstein et al., 1976). In one study, excellent clinical outcomes were reported for eight patients who had less than one-third of the femoral head removed (Epstein, 1974). In contrast, when more than one-third of the femoral head was removed, no complete functional recovery was due to too much stress in the hip joint (Epstein et al., 1976). Our results showed that patients with 3D parameter of PFA exceeding 30% were at high risk of developing FNF. Moreover, of the three 2D parameters (coronal plane, cross-sectional, and sagittal plane), the average PMDL across all three planes had the best diagnostic accuracy and measurement reliability. In addition, there was no standard position for patients to take the CT examination due to severe pain, and taking the average across three planes for each 2D parameter, combined with the 3D parameter based on 3D fracture reconstruction, could allowed for minimizing the impact on patients' position and selective bias of each CT plane. Therefore, an integrated assessment of 2D-CT images would help improve the predictive performance of the parameters.

One patient with iatrogenic FNF was excluded from this study for lacking CT images. FHFs were rare but serious injuries caused by high-energy trauma (Alonso et al., 2000). A CT scan of the hips was needed for better diagnosis and treatment of the fractures (De Mauro et al., 2021). The CT scan helped understand features of FHF, including the femoral head fracture pattern, the congruity of the hip joint, and the presence or absence of intra-articular loose fragments, which may not be accurately detected in X-ray image (Ross and Gardner, 2012). Moreover, it is inappropriate to define the standard measurement condition using X-rays to predict FNF in patients with FHFs, due to overlapping bones, exposure differences, and different body positions. The feasibility of the 2D and 3D parameters was assured by routine CT examinations of this special injury type. On the other hand, multidimensional and comprehensive assessments of CT images ensured better predictive efficacy.

Limitations

This study had some limitations. First, a retrospective design was adopted and the number of cases with a combination of FHF and FNF was small. However, the sample size of this type of fracture was larger compared with samples in previous studies, representing a strength of our study (Pipkin, 1957; Park et al., 2016; Scolaro et al., 2017; Keong et al., 2019). Moreover, a large sample of FHFs in consecutive series was taken, which helped minimize selection bias. In addition, the *post hoc* power calculation in this study demonstrated the 2D and 3D parameters had adequate power to predict FNF occurrence. Second, although the results demonstrated high repeatability and reliability of the two methods, no further comparison was made between 2D and 3D methods. However, using two distinct parameters, the 2D and 3D parameters with respective strengths and biases, could enable surgeons to lower estimation errors of the injury degree of the femoral head, thus further predicting the incidence of FNF in patients with FHFs with CT in clinical practice.

Conclusion

In summary, the new measurement for injury degree of femoral head, based on 2D and 3D injury models with CT, appeared to be reliable to assess the fracture risk of femoral neck in patients with FHFs in the clinic practice. All new parameters, including average percentage of maximum defect length across all three planes in 2D parameter and percentage of fracture area in 3D parameter, indicated strong emergence of femoral neck fracture in patients with femoral head fractures. Thus, these parameters could be a feasible adjunctive diagnostic tool in identifying FNFs. In addition, this finding might also provide a theoretic basis for the investigation of the convenient digital-model in complex injurie analysis.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Shanghai Sixth People's Hospital. All patients signed an informed consent form at hospital admission allowing the use of anonymized data for research purposes. So, the individual consent for this retrospective study was waived by the Ethics Committee of Shanghai Sixth People's Hospital.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by SHW, WW, RYL, JYG, YM, GYL, and JM. The first draft of the manuscript was written by SHW and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (Nos 81271991 and 12172224).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alonso, J. E., Volgas, D. A., Giordano, V., and Stannard, J. P. (2000). A review of the treatment of hip dislocations associated with acetabular fractures. *Clin. Orthop. Relat. Res.* 377, 32–43. doi:10.1097/00003086-200008000-00007

Beebe, M. J., Bauer, J. M., and Mir, H. R. (2016). Treatment of hip dislocations and associated injuries: Current state of care. *Orthop. Clin. N. Am.* 47 (3), 527–549. doi:10.1016/j.ocl.2016.02.002

Bland, J. M., and Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1 (8476), 307–310. doi:10.1016/s0140-6736(86)90837-8

Davis, J. B. (1950). Simultaneous femoral head fracture and traumatic hip dislocation. *Am. J. Surg.* 80 (7), 893–895. doi:10.1016/0002-9610(50)90473-9

De Mauro, D., Rovere, G., Smakaj, A., Marino, S., Ciolli, G., Perna, A., et al. (2021). Gibson approach and surgical hip dislocation according to Ganz in the treatment of femoral head fractures. *BMC Musculoskelet. Disord.* 22 (Suppl. 2), 961. doi:10.1186/s12891-021-04800-w

Epstein, H. C. (1974). Posterior fracture-dislocations of the hip; long-term follow-up. *J. Bone Jt. Surg. Am.* 56 (6), 1103–1127. doi:10.2106/00004623-197456060-00001

Epstein, H. C., Wiss, D. A., and Cozen, L. (1976). Posterior fracture dislocation of the hip with fractures of the femoral head. *Clin. Orthop. Relat. Res.* 201, 9–17. doi:10.1097/00003086-198512000-00002

Giannoudis, P. V., Kontakis, G., Christoforakis, Z., Akula, M., Tosounidis, T., and Koutras, C. (2009). Management, complications and clinical results of femoral head fractures. *Injury* 40 (12), 1245–1251. doi:10.1016/j.injury.2009.10.024

Greenwald, A. S., and Haynes, D. W. (1972). Weight-bearing areas in the human hip joint. *J. bone Jt. Surg. Br.* 54 (1), 157–163. doi:10.1302/0301-620x.54b1.157

Guo, J. J., Tang, N., Yang, H. L., Qin, L., and Leung, K. S. (2010). Impact of surgical approach on postoperative heterotopic ossification and avascular necrosis in femoral head fractures: A systematic review. *Int. Orthop.* 34 (3), 319–322. doi:10.1007/s00264-009-0849-3

Henle, P., Kloen, P., and Siebenrock, K. A. (2017). Femoral head injuries: Which treatment strategy can be recommended? *Injury* 38 (4), 478–488. doi:10.1016/j.injury.2007.01.023

Hesper, T., Bixby, S. D., Maranho, D. A., Miller, P., Kim, Y. J., and Novais, E. N. (2018). Morphologic features of the contralateral femur in patients with unilateral slipped capital femoral epiphysis resembles mild slip deformity: A matched cohort study. *Clin. Orthop. Relat. Res.* 476 (4), 890–899. doi:10.1007/s11999.0000000000000127

Kazemi, S. M., Qoreishy, M., Keipourfard, A., Sajjadi, M. M., and Shokraneh, S. (2016). Effects of hip geometry on fracture patterns of proximal femur. *archives bone Jt. Surg.* 4 (3), 248–252.

Keong, M. W., Razak, H., Koon, W. M., and Ping, C. C. (2019). Case report of a rare Pipkin type III femoral head fracture. *J. Orthop. case Rep.* 9 (5), 11–15. doi:10.13107/jocr.2019.v09.i05.1510

Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33 (1), 159–174. doi:10.2307/2529310

Menger, M. M., Braun, B. J., Herath, S. C., Kuper, M. A., Rollmann, M. F., and Histing, T. (2021). Fractures of the femoral head: A narrative review. *EFORT open Rev.* 6 (11), 1122–1131. doi:10.1302/2058-5241.6.210034

Park, K. H., Kim, J. W., Oh, C. W., and Kyung, H. S. (2016). A treatment strategy to avoid iatrogenic Pipkin type III femoral head fracture-dislocations. *Arch. Orthop. Trauma Surg.* 136 (8), 1107–1113. doi:10.1007/s00402-016-2481-1

Pipkin, G. (1957). Treatment of grade IV fracture-dislocation of the hip. *J. Bone Jt. Surg. Am.* 39-a (5), 1027–1197. doi:10.2106/00004623-195739050-00004

Ross, J. R., and Gardner, M. J. (2012). Femoral head fractures. *Curr. Rev. Musculoskelet. Med.* 5 (3), 199–205. doi:10.1007/s12178-012-9129-8

Rousson, V., Gasser, T., and Seifert, B. (2002). Assessing intrarater, interrater and test-retest reliability of continuous measurements. *Statistics Med.* 21 (22), 3431–3446. doi:10.1002/sim.1253

Scolaro, J. A., Marecek, G., Firoozabadi, R., Krieg, J. C., and Routt, M. L. C. (2017). Management and radiographic outcomes of femoral head fractures. *J. Orthop. Traumatol.* 18 (3), 235–241. doi:10.1007/s10195-017-0445-z

Sy, M. H., Kinkpé, C. V., Dakouré, P. W., Dieme, C., Sane, A., Ndiaye, A., et al. (2005). [Femoral neck fracture complicating orthopedic reposition of a dislocated hip: Four cases]. *Revue de Chir. Orthop. reparatrice de l'appareil moteur* 91 (2), 173–179. doi:10.1016/s0035-1040(05)84297-7

Thompson, V. P., and Epstein, H. C. (1951) Traumatic dislocation of the hip; a survey of two hundred and four cases covering a period of twenty-one years. *J. Bone Jt. Surg. Am.* 33-a(3):746–792. doi:10.2106/00004623-195133030-00023

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gotzsche, P. C., and Vandenbroucke, J. P. (2007). The strengthening the reporting of observational studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Epidemiology* 18 (6), 800–804. doi:10.1097/EDE.0b013e3181577654

Wells, J., Nepple, J. J., Crook, K., Ross, J. R., Bedi, A., Schoenecker, P., et al. (2017). Femoral morphology in the dysplastic hip: Three-dimensional characterizations with CT. *Clin. Orthop. Relat. Res.* 475 (4), 1045–1054. doi:10.1007/s11999-016-5119-2

10.3389/fbioe.2023.1115639

Appendix A:Details of 3D parameter measurement

Mimics software (version 20.0, materialise)

1. Reconstruct the 3D models of the proximal femur.
2. Export the 3-D model of the proximal femur of the injured (Figure A1, The proximal femur in yellow) to the 3-matic software.

3-matic (materialise)

1. Standardize the coronal, cross-sectional, and sagittal positions of the proximal femur of the injured to determine the femoral head part, according to the previous literature report [Reference: Wu S, Wang W, Zhang B, et al. A three-dimensional measurement based on CT for the posterior tilt with ideal inter-and intra-observer reliability in non-displaced femoral neck fractures. Comput Methods Biomech Biomed Engin. 2021; 24 (16): 1854–1861]. Then standardize the model size of the femoral head of the injured based on the scaling factor (Figure A1) to better match the femoral head of the template.

   1.1. Draw the caput sphere on the femoral head of the template and the injured, referred to the process in the above literature report, respectively. Then determine the radius of the sphere of the template and the injured, defined as $R_{template}$ and R_{case}, respectively.
   1.2. Calculate the scaling factor, defined as the ratio between $R_{template}$ and R_{case}, used to adjust the model size of the femoral head of the injured to match the femoral head of the template better.
2. Achieve the best possible alignment between the femoral head of the template and the injured, referred to the process in the above literature report (Panel C, N-points registration and Local registration). The greener the color is, the higher the degree of matching is (Figure A1, The degree of matching).
3. Generate the fracture region through Subtraction Boolean Operation between the femoral head of the template and the femoral head of the injured (Panel D, Subtraction boolean operation).
4. Extract the area of the femoral head of the template defined as S_{total}, then extract the area of the fractured region defined as $S_{fracture}$.
5. Calculate the percentage of fracture area, defined as the ratio between $S_{fracture}$ and S_{total} (Figure A1).

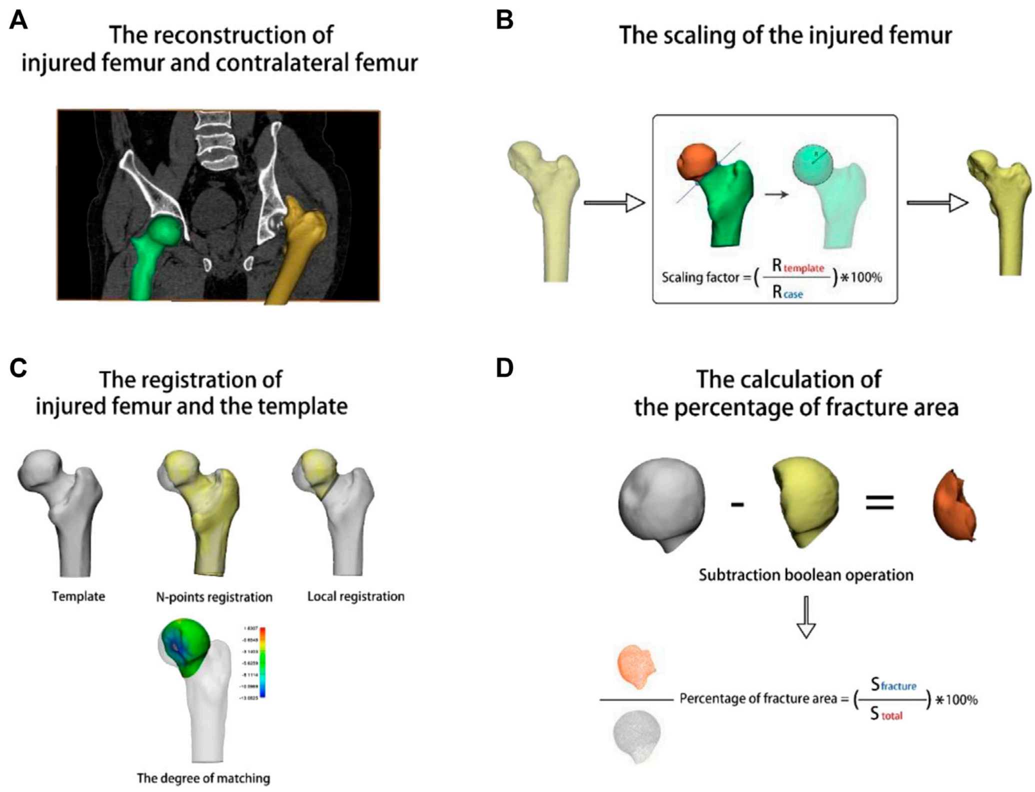


FIGURE A1
Flowchart of the 3D parameter measurement.

TYPE Original Research
PUBLISHED 28 April 2023
DOI 10.3389/fbioe.2023.1150061

Check for updates

OPEN ACCESS

EDITED BY
Laura Cercenelli,
University of Bologna, Bologna, Italy

REVIEWED BY
Jan Kubicek,
VSB-Technical University of Ostrava, Czechia
Nattapon Chantarapanich,
Kasetsart University, Thailand

*CORRESPONDENCE
Angelika Ramesh,
angelika.ramesh.18@ucl.ac.uk

†These authors have contributed equally to this work and share first authorship

RECEIVED 23 January 2023
ACCEPTED 17 April 2023
PUBLISHED 28 April 2023

CITATION
Henckel J, Ramesh A, Hothi H, Richards R, Di Laura A and Hart A (2023), The accuracy and precision of acetabular implant measurements from CT imaging. *Front. Bioeng. Biotechnol.* 11:1150061.
doi: 10.3389/fbioe.2023.1150061

COPYRIGHT
© 2023 Henckel, Ramesh, Hothi, Richards, Di Laura and Hart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The accuracy and precision of acetabular implant measurements from CT imaging

Johann Henckel[1†], Angelika Ramesh[1,2*†], Harry Hothi[1,2], Robin Richards[3], Anna Di Laura[1,2] and Alister Hart[3]

[1]Royal National Orthopaedic Hospital NHS Trust, Stanmore, United Kingdom, [2]Department of Mechanical Engineering, University College London, London, United Kingdom, [3]Institute of Orthopaedics and Musculoskeletal Science, University College London, London, United Kingdom

The placement of acetabular implant components determines the short- and long-term outcomes of total hip replacement (THR) and a number of tools have been developed to assist the surgeon in achieving cup orientation to match the surgical plan. However, the accuracy and precision of 3D-CT for the measurement of acetabular component position and orientation is yet to be established. To investigate this, we compared measurements of cobalt chrome acetabular components implanted into 2 different bony pelvic models between a coordinate measuring Faro arm and 3 different low dose CT images, including 3D-CT, 2D anterior pelvic plane (APP) referenced CT and 2D scanner referenced (SR) CT. Intra-observer differences were assessed using the Intraclass correlation coefficient (ICC). The effect of imaging the pelvis positioned in 3 different orientations within the CT scanner was also assessed. The measured parameters were the angles of inclination and version. 3D-CT measurements were found to closely match the "true values" of the component position measurements, compared with the 2D-CT methods. ICC analysis also showed good agreement between the coordinate measuring arm (CMA) and 3D-CT but poor agreement between the 2D SR method, in the results from two observers. When using the coordinate system of the CT scanner, the measurements consistently produced the greatest error; this method yielded values up to 34° different from the reference digitising arm. However, the difference between the true inclination and version angles and those measured from 3D APP CT was below half a degree in all cases. We concluded that low radiation dose 3D-CT is a validated reference standard for the measurement of acetabular cup orientation.

KEYWORDS

3D-CT, inclination, version, acetabular component position, coordinate measuring arm, cup orientation

1 Introduction

The last decade has witnessed a substantial increase in the total number of hip replacements performed annually worldwide, with concurrent advancements in the surgical approach and technological tools used in the field. The use of robotic technology, computer-based navigation systems, patient-specific instrumentation (PSI) and custom implants for the placement of acetabular components all promise improved accuracy and superior reproducibility (Henckel et al., 2018; Krämer et al., 2018; Fontalis et al., 2021). If these are to be implemented on a wider scale, it is crucial to evaluate their accuracy using metrics such as inclination and version, which are frequently used to quantify

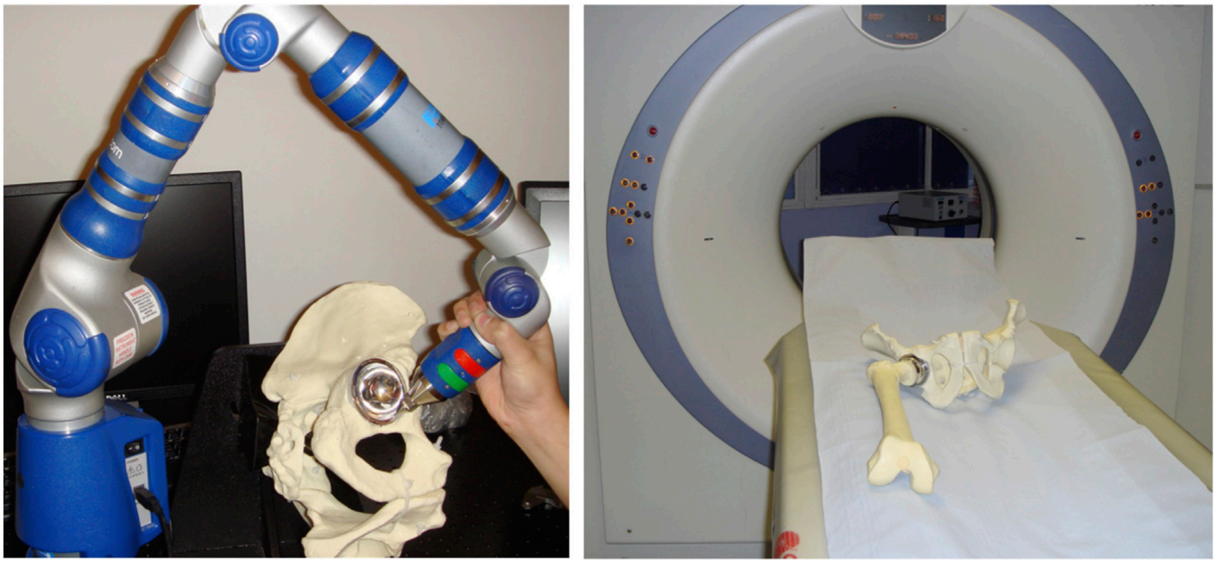


FIGURE 1
(Left) The orientation of the metal acetabular cup implanted in the pelvic Sawbone model being measured by the digitising arm. (Right) The model placed on the CT scanner table with the fiducials (hexagonal fully threaded titanium screws) placed on landmarks for robust landmarking.

acetabular cup position (Spencer-Gardner et al., 2016; Fontalis et al., 2021). Post-operative hip prosthesis measurements allow orthopaedic surgeons and companies to assess the achieved component position and orientation relative to the intended surgical plan. This is important to make robust assessment of new enabling technologies in surgery.

Measurement of hip implant positioning using 3D-CT has become the gold standard approach (Kaiser et al., 2021). Numerous studies have investigated the orientation of hip implants from post-operative CT scans and planar radiographs (Davda et al., 2015; Ma et al., 2022). However, the presence of metal artifacts and identification of anatomical landmarks on the reconstructed 3D pelvic model can introduce errors in the measurements of the implant position (Brownlie et al., 2020). These potential errors have not yet been quantified.

The novelty of this study includes the following.

- This is the first study quantifying the errors associated with CT measurements of acetabular cup orientation, from different coordinate systems, using artificial pelvic bone models. Accurate measurements can help evaluate the use of computer-aided systems which claim a 1–4° accuracy for cup alignment (Elson et al., 2015).
- Comparison of measurements taken from slices of the same CT scan before and after orienting the image using specialized software is a task which has not yet been undertaken.

The aim of our study was to quantify the accuracy and precision of 3D- and 2D-CT measurements of acetabular cup position for total hip arthroplasty (THA) outcome assessment. This will be achieved by evaluating and comparing the inclination and version of implanted acetabular cups using a) a coordinate measuring arm (CMA), b) 3D anterior pelvic (APP) referenced CT, c) 2D APP referenced CT and d) 2D scanner referenced (SR) CT. The effects of changes in pelvic orientation and inter- and intra-observer error on the measurements are also explored.

The four different methods adopted for measuring the cup inclination and version are outlined in this paper. This includes the identification and alignment to the APP. The results from the CT methods will then be compared to the reference standard in order to define the accuracy of each approach.

2 Materials and methods

Two pelvic artificial bone models (Sawbone Pelvis - Foam Cortical SKU:1301-1) were implanted with metal acetabular components (Figure 1). The pelvic height (HP) was 150 mm, the interspinous distance (ISD -distance between the two most prominent points of the anterior superior iliac spine (ASIS)) was 250 mm and the intercristal distance (ICD - distance between the most prominent points of the iliac crest) of the model was 270 mm (Musielak et al., 2019). The acetabular size was chosen by a senior orthopaedic surgeon to fit the Sawbone pelvic model used in this study.

Identification of both ASIS and pubic tubercles was made by placing the anterior face of the pelvis onto a flat surface covered with graphite to reveal the 3 prominences which define the APP. 5 mm fully threaded hexagonal titanium screws serving as "fiducial markers" were securely fixed into these 3 points, allowing for robust landmarking on the CT images without the associated imaging artefact, resulting from the use of metal implants. The two pelvic models were prepared using conventional tools to ream the acetabular socket to the size of the prosthetic cup. The cup was securely implanted in an orientation deemed suitable for the pelvis (approximately 45° of inclination and 20° of anteversion). A press-fit fixation was achieved and thoroughly assessed to ensure that there was no movement of the component relative to the bone.

2.1 Measurement of cup position using the coordinate measuring arm

The "post-op" pelvises were rigidly secured to a pelvic clamp. A digitising arm (reference standard, Gage Max FaroArm) (McPherson et al., 2005) was used to take a single point (defined in all 3 axes x, y, z) on each of the 3 fiducial markers and 20 on the cup rim (Figure 1). This process was repeated 12 times by each of the two observers, for both pelvises. While the 3 fiducial marker points defined the APP, the 20 points taken on the cup rim defined the plane of the acetabular component's cup face, known as the 'cup plane'. The angular relationship between these 2 planes was computed using a commercially available software. Inclination was defined as the angle between the cup plane and the transverse anatomical plane of the pelvis (orthogonal to the APP). Anteversion was defined as the angle between the cup plane and the sagittal anatomical plane of the pelvis (orthogonal to both the APP and transverse plane) (Murray, 1993; Zhang et al., 2014). The values calculated represent 'true values' for the 3D orientation of the cup in the pelvis using the APP as the frame of reference.

2.2 CT imaging of the implanted cups

The pelvic model with the implanted acetabular cup and its corresponding femoral component was scanned using a 16-slice CT scanner (Siemens SOMATOM Emotion eco 16-slice configuration) (Figure 1). The images were acquired axially using a spiral sequence and at 0.75 mm increments with a kV of 100, mAs of 100 and a pitch of 1. The pelvic construct was CT scanned in 3 supine orientations: A) parallel to the floor; B) 10° of lateral pelvic tilt; and C) 20° of anterior pelvic tilt. Following the scanning process the images were saved in Digital Imaging and Communications in Medicine (DICOM) format. These axial slices were then reconstructed in both the coronal and sagittal plane with respect to the pelvis' position in the scanner and with 0.75 mm spacing.

2.3 Measurement of cup position using 3 CT methods

Measurement method 1: 3D CT measurement of cup position referenced from the APP, referred to as "3D APP CT." The 2D unprocessed axial slices in DICOM format were reconstructed to produce a 3D model of the pelvis with the prosthetic cup. This 3D image was oriented in the APP but in this method the 3D reconstructed virtual model was directly used for the measurement. Markers were placed on the fiducials, along with 20 on the acetabular cup rim, which was clearly differentiated from the metallic femoral head. The compound angles between the APP and the plane of the face of the cup were computed to give the angles of inclination and version. The position measurements (angles of inclination and anteversion) of the component were converted from the anatomical definition to radiographic definition for the Faro arm and 3D-CT, using dedicated software. This ensured all measurements were according to the radiographic definition, for comparison (Murray, 1993).

Measurement method 2: 2D CT measurement of cup position using Robin 3D software with the dataset orientated to APP in the coronal plane. This is referred to as 2D APP referenced, "2D APP CT." The 6 datasets were exported to a picture archiving (PACS) workstation. The 2D unprocessed axial slices in DICOM format were reconstructed using Robin 3D to produce a 3D model oriented in the APP from which axial slices again oriented to the APP were viewed. 2D reconstructed slices in all three planes, parallel and perpendicular to the APP were then generated and snapshots taken in both the coronal and axial planes at the equatorial region of the components. In the coronal view 2 points were selected, one on the most superior and the other on the most inferior edges of the acetabular component margins. The angles between 2 lines, one joining the 2 spines (ASIS) and a second joining the 2 points on the cup, the "inclination angle" were measured. In the axial view and at the cup's equatorial region, 2 points were placed on the cup margin, one on the most anterior edge and the other on the most posterior edge. The angle between a line joining the 2 markers and a line at 90° to the horizontal represents the angle of anteversion. Measurements were made 12 times by 2 observers.

Measurement method 3: 2D CT, referenced off the position of the pelvis in the CT scanner, referred to as 2D scanner referenced (SR), "2D SR CT." Measurements of the inclination and version angles were taken directly from the coronal and axial slices, respectively. The method for measurement used in the second technique (2D APP CT) was repeated here, on these un-oriented slices.

Statistical methods: Bland-Altman plots were used to assess the level of agreement between the CT methods and the digitising CMA. Inter- and intra-observer differences were assessed using the Intra-class correlation coefficient (ICC) tool on SPSS and it considered the effect of different CT measurement methods.

3 Results

3.1 CT agreement with true CMA values

Table 1 summarizes the mean difference in inclination and version angles between the digitising arm and the 3 CT-based methods, for both pelvic models. These were positioned in a CT scanner for imaging in 3 different positions (A, B, C). All possible combinations of observers, pelvises, pelvic orientations, imaging methods and number of repeats produced 432 datapoints for the angles of inclination and version. This includes 2 observers, 2 pelvises, 3 orientations, 3 methods and 12 repeats. Table 1 reveals that 2D SR CT yields values up to 34° different to the true value and is thus a less accurate method as this is very poor agreement.

Supplementary Table S1 shows the average inclination and version measurements taken to define the orientation of the acetabular cup. The differences in the measurements between the digitising arm and the 3 CT methods is evident; the values obtained

TABLE 1 Absolute mean (+/-2SD) difference between 12-paired coordinate measuring arm and CT measurements from observer 1 for both pelvises. 3 CT methods were used (3D APP CT, 2D APP CT and 2D SR CT), each with the pelvis in 3 different orientations in the CT scanner. Measurements were made of the cup in terms of inclination and version angles.

			Inclination angle (º)	
Pelvis	Orientation	D Arm-3D APP CT	D-Arm-2D APP CT	D Arm-2D SR CT
1	A	0.06 ± 0.07	0.1 ± 0.32	2.9 ± 1.2
1	B	0.23 ± 0.07	0.36 ± 0.20	0.12 ± 1.09
1	C	0.22 ± 0.08	0.17 ± 0.31	6.04 ± 2.48
			Version angle (º)	
Pelvis	Orientation	D Arm-3D APP CT	D-Arm-2D APP CT	D Arm-2D SR CT
1	A	0.36 ± 0.14	3.38 ± 0.36	12.9 ± 1.9
1	B	0.23 ± 0.10	3.00+ /-0.30	9.60 ± 0.84
1	C	0.41 ± 0.11	2.77 ± 0.38	19.9 ± 1.08
			Inclination angle (º)	
Pelvis	Orientation	D Arm-3D APP CT	D-Arm-2D APP CT	D Arm-2D SR CT
2	A	0.22 ± 0.07	0.20 ± 0.16	4.23 ± 0.88
2	B	0.24 ± 0.07	0.36 ± 0.09	8.31 ± 1.47
2	C	0.31 ± 0.06	0.33 ± 0.16	2.06 ± 1.50
			Version angle (º)	
Pelvis	Orientation	D Arm-3D APP CT	D-Arm-2D APP CT	D Arm-2D SR CT
2	A	0.13 ± 0.08	14.7 ± 0.29	14.8 ± 0.64
2	B	0.28 ± 0.08	14.1 ± 0.37	18.6 ± 0.66
2	C	0.13 ± 0.07	14.1 ± 0.28	34.3 ± 0.71

from 3D APP CT closely follows the Faro arm, whereas 2D SR CT reveals the largest mismatch particularly for the version angles. This is true for both pelvises.

Bland-Altman and XY plots (Figures 2, 3) were used to assess the difference between the three CT methods and the CMA. Figure 2 compares the digitising arm values to the 3 different CT methods in terms of the measured cup inclination angle. The plots quantify the level of agreement/disagreement between the true position and the position measured using the 3 CT techniques. The solid blue line represents the mean difference, whereas the dashed lines represent the 95% limits of agreement (mean difference+/-1.96 SD). 3D APP CT and 2D APP CT showed good agreement with the Faro arm whereas the 2D SR CT measurements showed very poor agreement.

Figure 3 shows the same comparison of methods but for the cup version angle. This showed the same pattern of agreement/disagreement, however the spread of error recorded was greater than for inclination angle. Overall, there is high validity of the 3D APP CT method with decreasing validity from 3D APP CT to 2D APP CT to 2D SR CT. Once again, for version, the 3D APP CT and 2D APP CT measurements are close to the true (Faro arm) value, but less accurate compared with inclination. Figures 2, 3 are a true representation of the patterns seen for both pelvises in all three positions. The 2D SR CT measurements consistently produced the greatest error, disregarding the outlier shown for pelvis 1 in position B in Table 1. The XY plot (Figures 2, 3D) also shows that the discrepancy between the digitising arm value was lowest for 3D APP CT and highest for 2D SR CT.

Figure 4 shows the effect of the different pelvic orientations (A, B, C) on cup version. The method chosen for display was 2D SR CT for pelvis 2, as this showed the greatest level of disagreement between the orientations. For position C all measurements over 30° different to the digitising arm value. This result is noteworthy yet not surprising, particularly for radiographers, as it confirms that the position of the patient within the scanner will affect the measurements made. So, the 20° anterior pelvic tilt resulted in even greater disagreement between the 2D measurement and the true digitising arm position measurements.

3.2 ICC analysis

To measure inter-observer agreement/disagreement, intra-class correlation coefficients (ICC) of the measurements ($n = 216$) taken from both observers were calculated. ICC analysis between the observers showed excellent agreement (>0.9) for both inclination and version values. ICC analysis also showed good agreement for both observers between the CMA and 3D-CT but poor agreement between the CMA and 2D-CT methods. Irrespective of the observer, the ICC correlation between the Faro arm and 3D-CT was above 0.9 for inclination and version. The correlation between the Faro arm and 2D APP CT showed moderate agreement (0.5< ICC <0.75) for version but excellent agreement for inclination. This reaffirms that measurement of the version angle is a more difficult task to undertake, compared with inclination. Finally, the ICC value for Faro vs. 2D SR CT was below 0.5, denoting poor agreement.

4 Discussion

Our study is the first to quantify the potential errors associated with the use of 3D-CT for the measurement of cup angles of the acetabular component in metal-on-metal (MOM) hip replacement. Despite the frequent use of 3D-CT for the measurement of implant position, it is yet unknown how accurately this method can do so. There are several investigations in the literature which compare inclination and version measurements between different imaging methods, including 3D-CT and 2D radiographs (Davda et al., 2015), and also 3D hipEOS and 3D-CT (Anderson et al., 2022). Findings from these studies have proven 2D measurements to be less reliable, particularly when measuring cup version, while hipEOS can provide comparable angular measurements to 3D-CT. But these are clinical imaging studies of patients and only provide a comparison of different techniques, not the accuracy and precision of 3D-CT, which this lab-based study aimed to offer. Validation of 3D-CT is difficult to achieve directly from patients, given the invasive technique that would be required to take measurements using a sterilized digitising arm intra-operatively.

This study was designed to take account of the main variables that may affect the measurement of cup orientation. Firstly, the MOM relationship of the components blurring the boundaries between the cup and head components. Secondly, the position of the pelvis in the CT scanner and thirdly, the effect of the observer in terms of inter- and intra-observer error.

The composite sawbones used here provide a uniform test bed with physical properties similar to that of real bone (Heiner, 2008). These medical models are primarily used for the testing of prosthetic implant fixation and provide a reliable alternative to cadavers. The CT imaging protocol was developed for clinical use, the aim being to minimize radiation dose whilst maintaining adequate image quality, following the guidelines of The Ionising Radiation (Medical Exposure) Regulations 2000 (IR (ME)R) (Health Department, 2001). Many measurements (n = 432) were taken to analyse the effect of error from the method of measurement, the pelvic position, and the observer.

Metal artefact reduction strategies: The artefacts produced by imaging prosthetic implants are as a result of their constituent metals being high x-ray absorbers, reducing the amount of radiation energy transmitted through the implant to the detector, making valid measurement difficult (Berg et al., 2006). This obstruction of the x-ray beam results in the distortion of the images, leading to uncertainties in component position by the reading clinician.

The use of conventional imaging protocols produces streaking artefacts from the metal implant. To minimize the effect of these metal artefacts on the images we used both a specific image acquisition protocol and software solutions. Our imaging sequence made use of an extended Hounsfield scale (Hounsfield Units are the Supporting Material units for density as used in CT imaging) (Hounsfield, 1973). The software allowed visualization of the high-density profile of the metal component with a minimum of streaking artefacts: these images can be viewed at chosen Hounsfield thresholds. A Hounsfield threshold of 6,000 clearly visualizes cobalt-chromium. Whilst there is no single best strategy to remove these artefacts, we utilized software strategies that allowed the extension of the Hounsfield scale to define the component edges more sharply (Itokawa et al., 2008; Rinkel et al., 2008). Although some studies have shown that by manipulating the imaging acquisition parameters - mainly increasing the scanning kV - the volume of these artefacts can be reduced but not eliminated (Moon et al., 2008), we chose not to adopt this approach as the radiation dose to the patient is also increased as a result of this.

Component position measurement: The inclination angle was measured relative to a plane orthogonal to the APP-ordinarily a horizontal line when the patient is standing or when the observer is looking at an AP radiograph of the pelvis. The ‘radiographic inclination’ is easily approximated on a plain AP radiograph as the angle between the cup face and a horizontal line drawn between the ischial prominences or the lowest aspect of the teardrop. The anteversion angle is harder to describe and even harder to measure satisfactorily using plain radiographs. It is the angle between the cup face and a line in the mid coronal axis. CT on the other hand images a volume from which 2D slices can be viewed in almost any region and orientation. Others have shown the benefit of CT when compared to standard radiograph measurements of the acetabular cup (Snijders et al., 2019) but have not published the validation of their methods. Presumably they assumed that CT scanning is a validated imaging modality. However, the application of CT to MOM hip component measurement has not been compared against a reference standard. This issue is more relevant when new CT protocols and 3D measurement software are employed. Such techniques become more important when required to overcome the difficult situation presented by MOM hips, where the large diameter and high-density head obscures the imaging of the cup face.

The measurements made from our optimized 2D images revealed that the largest deviation from the true value (as determined by the CMA) occurred when the orientation of the pelvis was rotated with respect to the gantry of the scanner. This deviation was found to be greater when measuring the angle of anteversion. Conversely, both accuracy (deviation from the true value) and precision (repeatability) were improved when the pelvis was positioned in the supine position, parallel to the table. We recommend that attention is paid to the positioning of the patient’s pelvis in ensuring that the landmarks used to the derive APP are as parallel to the table as possible particularly when using the scanner as the frame of reference for measurement. Further improvements in accuracy and precision were achieved using 2D orientated CT (2D APP CT). The angles derived from the 3D APP referenced CT (3D APP CT) measurement method were least affected by the orientation of the pelvis.

We would like to emphasise that our protocol includes high-resolution 0.75 mm collimations to image both bone and prosthetic

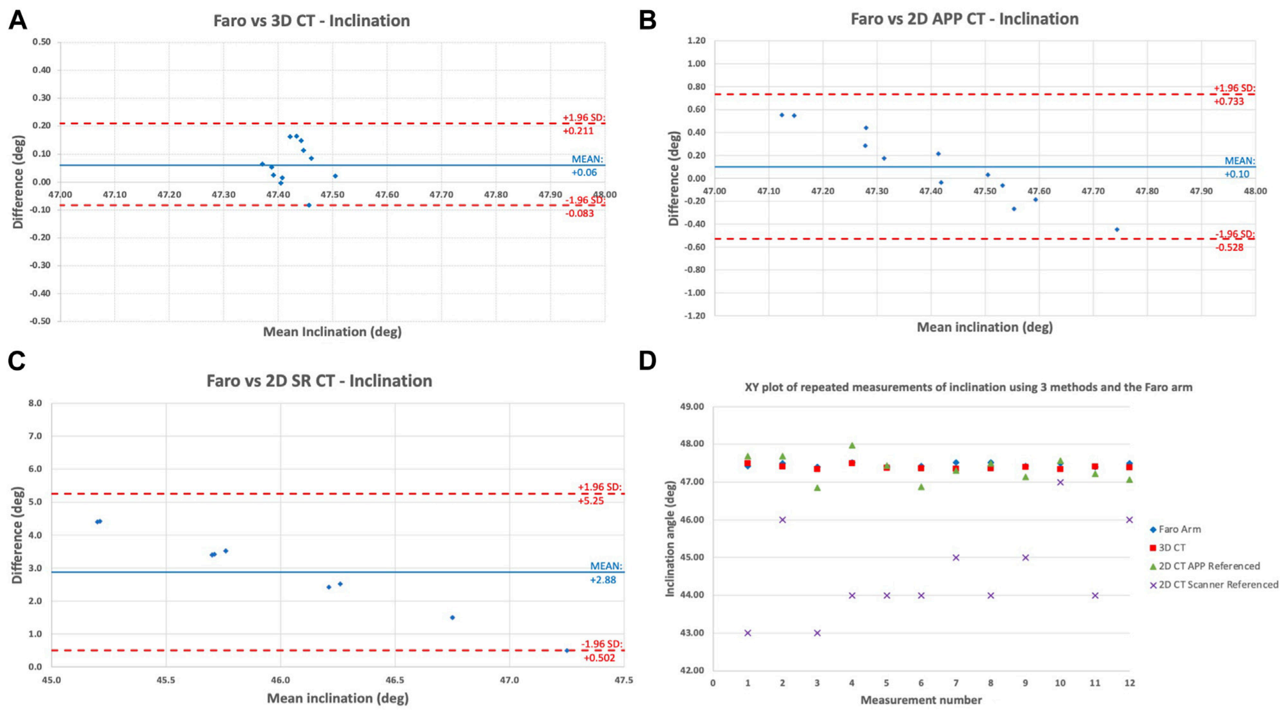


FIGURE 2
Bland-Altman plot showing the level of agreement/disagreement between the digitising arm inclination measurements and **(A)** 3D APP CT; **(B)** 2D APP CT; **(C)** 2D SR CT; **(D)** XY scatter plot of the 12 repeated measurements of inclination angles according to each of the 3 imaging methods and the Faro digitising arm for pelvis 1A by observer 1.

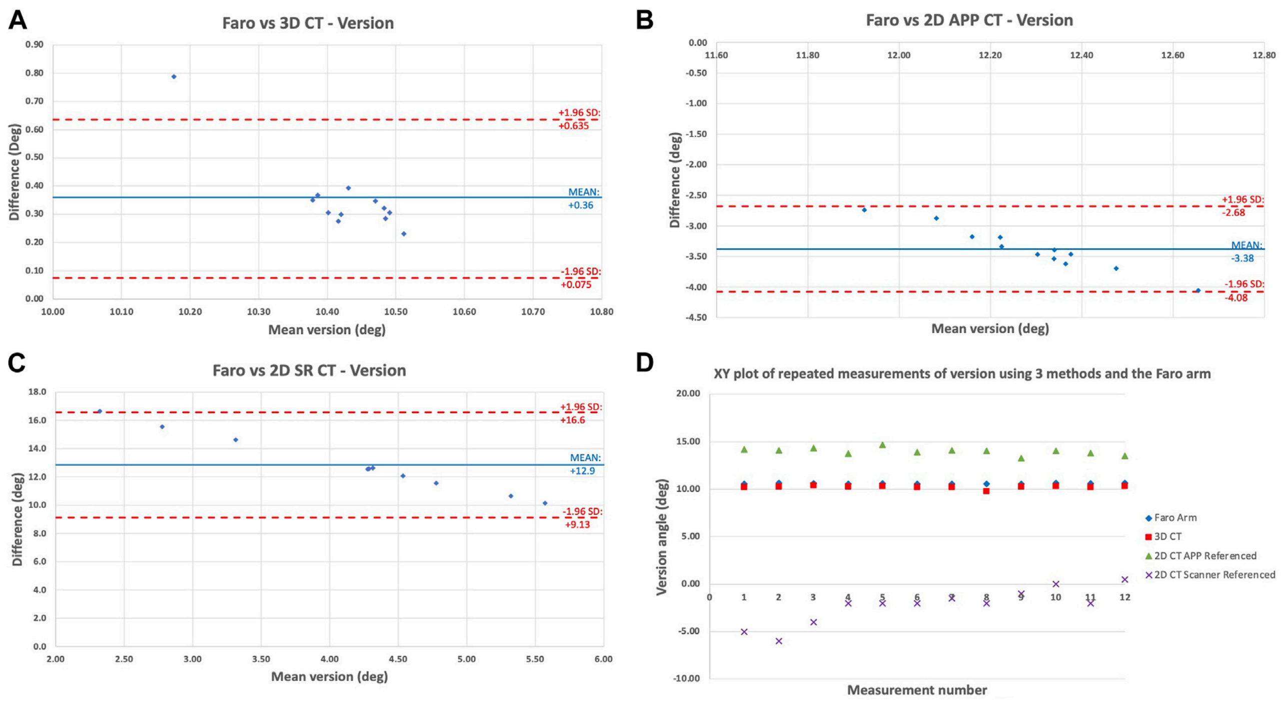


FIGURE 3
Bland-Altman plot showing the level of agreement/disagreement between the digitising arm version measurements and **(A)** 3D APP CT; **(B)** 2D APP CT; **(C)** 2D SR CT; **(D)** XY scatter plot of the 12 repeated measurements of version angles according to each of the 3 imaging methods and the Faro digitising arm for pelvis 1A by observer 1.

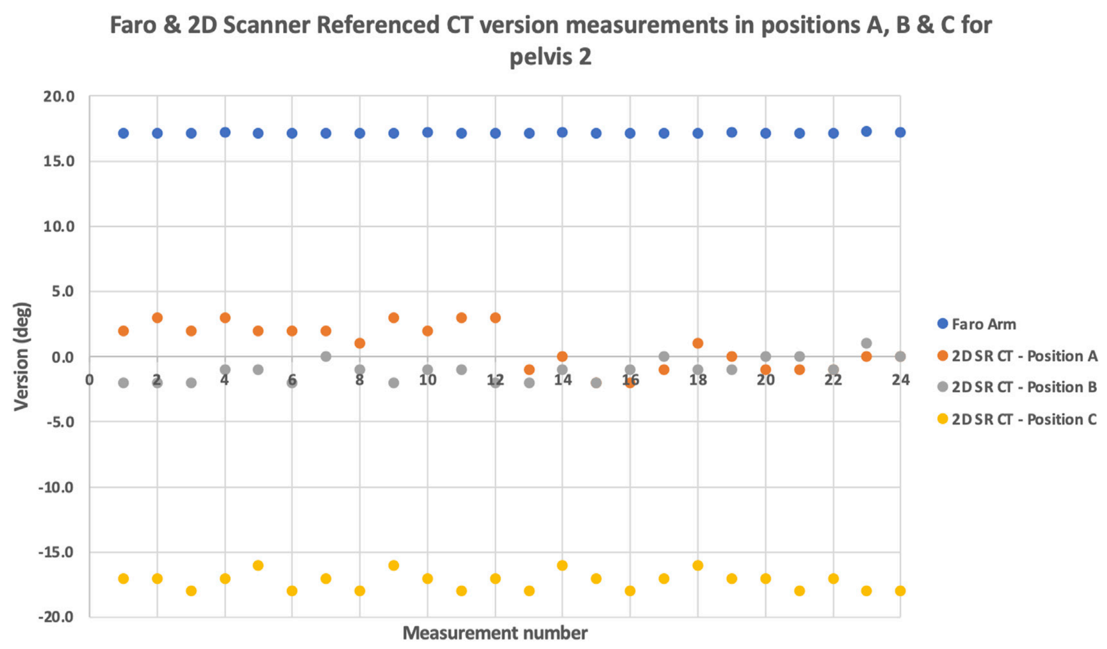


FIGURE 4
The effect of pelvic orientation: XY scatter plot of the repeated version angle measurements using the digitising arm and 2D SR CT for pelvis 2. The measurements were obtained at different pelvic orientations **(A,B,C)**.

components and enables 'separation' of the edges of two metal components, the cup and head held in close proximity. The narrower slices increase the resolution of the images and are particularly useful for large MoM hips.

In method 1 (3D APP CT) we compared like with like: the 'radiographical' definition of inclination and version angles were used to compare the values for both the CMA and the virtual 3D model. The very small errors were probably due to the presence of the low-level metal artefacts affecting the labelling of the 20 points on the cup margin. We used high resolution CT to minimize this source of error (Lou et al., 2007). The accuracy of the calibrated digitising arm, our 'reference standard', used here, far outstrips the current resolution of clinical CT. In method 2 (2D APP CT) we standardized the frame of reference to the APP. The error between this method and the digitising arm may have been due to the fact that only two points were used to label the cup margin and the issues inherent in using a 2D view (i.e., was the most representative 2D "snapshot" selected?). The error between method 3 (2D SR CT) and the digitising arm was most likely due to the difference in the frame of reference.

A number of studies have shown good inter- and intra-observer agreement on measurements of inclination made on plain radiographs, i.e., precision. However, accuracy (closeness to the true value) is poor when compared to APP referenced CT (Bayraktar et al., 2017). Measurements on these plain 2D images are referenced off the position of the patient in the radiograph (Nishino et al., 2013) and are further confounded by the diverging source of x-ray. The close agreement between the observers in this study shows that our 3D APP CT method is also very precise.

Davda et al. (2015) also measured cup version and inclination of MOM hips using 3D-CT and Ein Bild Roentgen Analyse (EBRA) software (for 2D radiographic analysis). Their results revealed underestimated values of cup version using the 2D method. This is also true for the results of the present study. Supplementary Table S1 shows that the 2D SR CT version measurements are consistently lower than the reference angles. However, Davda et al. (2015) used 3D-CT as their gold standard measurement, for comparison. Equally, Ma et al. (2022) compared the orientation measurements taken using low-dose bi-planar radiographs (EOS imaging) with 3D-CT, but their results endorsed the use of EOS imaging for post-THA component orientation measurements. This study is the first to compare 3D-CT with a laboratory digitising arm, which is a gold standard for measurements in engineering.

With ongoing advancements in hip replacement surgery, numerous tools, including navigation, robotic and augmented reality systems, and PSI are offered to the orthopaedic surgeon to assist them in achieving optimal implant positioning. These will determine the performance of the prosthesis. If these instrumentations and surgical techniques are to be adopted in the operating theatre, there needs to be evidence that the orientation and position of the implant achieved with their use matches the surgical plan. This study sought to quantify the errors in the placement of the acetabular component using different measurement techniques and coordinate systems of the CT data, relative to a reference standard digitising arm. We have demonstrated that 3D-CT measurements can be used to perform post-operative radiological assessment and implant surveillance.

We acknowledge that this study has its limitations. The measurements were limited to the acetabular component only and the position of the femoral component was not measured. Thus, the errors associated with CT measurements of the femoral stem may be quantified in future research. We are aware of the better image resolution and faster acquisition time associated with a 64-slice CT scanner (hence giving more accurate results), but we scanned the pelvis using a more widely available 16-slice CT scanner to represent a worst-case scenario for the measurements taken. A more modern 64-slice scanner may be used in future studies to give even more accurate angular measurements. This study measured the accuracy of the acetabular implant performed conventionally by the surgeon. By contrast, in computer-aided surgery, the size and implant position can be pre-planned (Inoue et al., 2019). However, this study did not include this aspect.

5 Conclusion

In conclusion, this study validates the use of 3D-CT for the measurement of acetabular component positioning post-operatively. Although this lab-based study involved an artificial pelvic bone model, the results may be extended in the context of clinical CT images of real patients. The differences found in the measurements from the variable CT methods, necessitates this paper. This is because the measurements taken from the same CT scan before and after specialized 3D-rendering software is used to manipulate the image (by orientating the APP to the coronal), are shown to differ. This emphasizes the need for additional software which is not available on the CT console or most PACS systems.

Here, the accurate measurement of inclination and anteversion in the context of an unknown pelvic orientation presents a fundamental 3-dimensional challenge. The challenge is increased in the context of large diameter MOM hips. In this study we have demonstrated and validated the use of 3D-CT to measure the anatomical angles of inclination and version (later converted into the radiographic definitions of the angles). The ability to standardize the measurements of cup orientation to the accepted frame of reference (APP) in three dimensions will allow surgeons and researchers to more accurately study and plan surgery. This will further the study of the relationship between component position and outcome, such as the investigation of painful and poorly functioning prosthetic hips.

Many studies have used 2D-CT measurement methods and this work has shown that 3D-CT methods offer superior accuracy and therefore we recommend the use of 3D-rendering software solutions. Our CT protocol has significant clinical improvement over others because it uses a low radiation dose and minimizes metal artefact. Although the study was performed on MOM prosthesis, we are currently using the method here developed on metal-on-polyethylene and ceramic-on-ceramic prostheses where the challenges in identifying the component boundaries are not as great. Our study is very pertinent because we are bettering our understanding in the large variability of post-operative component position and its effect on function and failure rates. With an ever-increasing demand for high performance hips from the young, more active and an ageing population, the orthopaedic surgeon of today needs additional tools in his armamentarium to further study the function and longevity (Deep et al., 2017) of hip replacement surgery.

Data availability statement

The datasets presented in this article are not readily available because none. Requests to access the datasets should be directed to Angelika.ramesh.18@ucl.ac.uk.

Author contributions

JH, AR, and AH- conceptualization; data curation; methodology; software; formal analysis; project administration; resources; writing–original draft; writing–review and editing. HH and AD—software; formal analysis; project administration; resources; writing–original draft; writing–review and editing. RR–software; data curation; methodology; writing–original draft.

Funding

This research was supported by The Arthroplasty for Arthritis Charity, the Maurice Hatter Foundation, the RNOH Charity, the Rosetrees Trust and the Stoneygate Trust and by researchers at the National Institute for Health Research, University College London Hospitals.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1150061/full#supplementary-material

References

Anderson, C. G., Brilliant, Z. R., Jang, S. J., Sokrab, R., Mayman, D. J., Vigdorchik, J. M., et al. (2022). Validating the use of 3D biplanar radiography versus CT when measuring femoral anteversion after total hip arthroplasty. *Bone. Jt. J.* 104-B, 1196–1201. doi:10.1302/0301-620x.104b11.bjj-2022-0194.r2

Bayraktar, V., Weber, M., von Kunow, F., Zeman, F., Craiovan, B., Renkawitz, T., et al. (2017). Accuracy of measuring acetabular cup position after total hip arthroplasty: Comparison between a radiographic planning software and three-dimensional computed tomography. *Int. Orthop.* 41, 731–738. doi:10.1007/s00264-016-3240-1

Berg, V. B., Malghem, J., Maldague, B., and Lecouvet, F. (2006). Multi-detector CT imaging in the postoperative orthopedic patient with metal hardware. *Eur. J. Radiol.* 60, 470–479. doi:10.1016/j.ejrad.2006.08.008

Brownlie, C. A., Evans, R., Morrison, D., Hayes, A., Song, S., and Kuster, M. S. (2020). Improved accuracy of CT based measurements for anterior prominence of acetabular prosthesis using a novel protocol based on Anatomical Landmarks. *Orthop. Traumatol. Surg. Res.* 106, 563–568. doi:10.1016/j.otsr.2019.10.019

Davda, K., Smyth, N., Cobb, J. P., and Hart, A. J. (2015). 2D measurements of cup orientation are less reliable than 3D measurements. *Acta Orthop.* 86, 485–490. doi:10.3109/17453674.2015.1017791

Deep, K., Shankar, S., and Mahendra, A. (2017). Computer assisted navigation in total knee and hip arthroplasty. *SICOT-J* 50, 50–56. doi:10.1051/sicotj/2017034

Elson, L., Dounchis, J., Illgen, R., Marchand, R. C., Padgett, D. E., Bragdon, C. R., et al. (2015). Precision of acetabular cup placement in robotic integrated total hip arthroplasty. *Hip Int.* 25, 531–536. doi:10.5301/hipint.5000289

Fontalis, A., Epinette, J-A., Thaler, M., Zagra, L., Khanduja, V., and Haddad, F. S. (2021). Advances and innovations in total hip arthroplasty. *SICOT-J* 7, 26–10. doi:10.1051/sicotj/2021025

Health Department (2001). *The ionising radiation (medical exposure) Regulations 2000: Programme of inspections (S.I. 2000/1059).* Scotland: Health Department.

Heiner, A. D. (2008). Structural properties of fourth-generation composite femurs and tibias. *J. Biomech.* 41, 3282–3284. doi:10.1016/j.jbiomech.2008.08.013

Henckel, J., Holme, T. J., Radford, W., Skinner, J. A., and Hart, A. J. (2018). 3D-printed patient-specific guides for hip arthroplasty. *J. Am. Acad. Orthop. Surg.* 26, 342–348. doi:10.5435/jaaos-d-16-00719

Hounsfield, G. N. (1973). Computerized transverse axial scanning (tomography): Part 1. Description of system. *Br. J. Radiol.* 46, 1016–1022. doi:10.1259/0007-1285-46-552-1016

Inoue, D., Kabata, T., Kimura, H., and Tsuchiya, H. (2019). A prospective clinical trial to assess the accuracy of an MRI-based patient-specific acetabular instrument guide in total hip arthroplasty. *Eur. J. Orthop. Surg. Traumatol.* 29, 65–71. doi:10.1007/s00590-018-2279-7

Itokawa, H., Hiraide, T., Moriya, M., Fujimoto, M., Nagashima, G., Suzuki, R., et al. (2008). The influence on the images of computed tomography caused by the use of artificial cranial reconstructive materials. *No Shinkei Geka* 36, 607–614.

Kaiser, M., Renkawitz, T., Benditz, A., König, M., Thieme, M., and Weber, M. (2021). Pelvic tilt impacts cup orientation on CT: How accurate is the gold standard? *Acta Radiol.* 63, 698–705. doi:10.1177/02841851211009466

Krämer, M., Kahrs, L. A., Friese, K-I., von Falck, C., and Hurschler, C. (2018). Inter- and intra-operator reliability in patient-specific template positioning for total hip arthroplasty. A cadaver study. *Int. J. Med. Robot.* 14, 18877–e1896. doi:10.1002/rcs.1887

Lou, L., Lagravere, M. O., Compton, S., Major, P. W., and Flores-Mir, C. (2007). Accuracy of measurements and reliability of landmark identification with computed tomography (CT) techniques in the maxillofacial area: A systematic review. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.* 104, 402–411. doi:10.1016/j.tripleo.2006.07.015

Ma, Z., Tang, H., Zhou, Y., Wang, S., Yang, D., and Guo, S. (2022). Assessing component orientation of total hip arthroplasty using the low-dose bi-planar radiographs. *BMC Musculoskelet. Disord.* 23, 886–889. doi:10.1186/s12891-022-05835-3

McPherson, A., Karrholm, J., Pinskerova, V., Sosna, A., and Martelli, S. (2005). Imaging knee position using MRI, RSA/CT and 3D digitisation. *J. Biomech.* 38, 263–268. doi:10.1016/j.jbiomech.2004.02.007

Moon, S. G., Hong, S. H., Choi, J. Y., Jun, W. S., Kang, H. G., Kim, H. S., et al. (2008). Metal artifact reduction by the alteration of technical factors in multidetector computed tomography: A 3-dimensional quantitative assessment. *J. Comput. Assist. Tomogr.* 32, 630–633. doi:10.1097/rct.0b013e3181568b27

Murray, D. W. (1993). The definition and measurement of acetabular orientation. *J. Bone Jt. Surg. Br.* 75, 228–232. doi:10.1302/0301-620x.75b2.8444942

Musielak, B., Kubicka, A. M., Rychlik, M., Czubak, J., Czwojdziński, A., Grzegorzewski, A., et al. (2019). Variation in pelvic shape and size in eastern European males: A computed tomography comparative study. *PeerJ* 7, 64333–e6524. doi:10.7717/peerj.6433

Nishino, H., Nakamura, S., Arai, N., and Matsushita, T. (2013). Accuracy and precision of version angle measurements of the acetabular component after total hip arthroplasty. *J. Arthroplasty* 9, 1644–1647. doi:10.1016/j.arth.2013.02.014

Rinkel, J., Dillon, W. P., Funk, T., Gould, R., and Prevrhal, S. (2008). Computed tomographic metal artifact reduction for the detection and quantitation of small features near large metallic implants: A comparison of published methods. *J. Comput. Assist. Tomogr.* 32, 621–629. doi:10.1097/rct.0b013e318149e215

Snijders, T. E., Schlösser, T. P., van Gaalen, S. M., Castelein, R. M., Weinans, H., and de Gast, A. (2019). Non-equivalent results from different anteversion measurements methods for the evaluation of the acetabular cup orientation in total hip arthroplasty. *Orthop. Surg.* 11, 241–247. doi:10.1111/os.12445

Spencer-Gardner, L., Pierrepont, J., Topham, M., Baré, J., McMahon, S., and Shimmin, A. J. (2016). Patient-specific instrumentation improves the accuracy of acetabular component placement in total hip arthroplasty. *Bone. Jt. J.* 98, 1342–1346. doi:10.1302/0301-620x.98b10.37808

Zhang, J., Wang, L., Mao, Y., Li, H., Ding, H., and Zhu, Z. (2014). The use of combined anteversion in total hip arthroplasty for patients with developmental dysplasia of the hip. *J. Arthroplasty* 29, 621–625. doi:10.1016/j.arth.2013.08.004

frontiers | Frontiers in Veterinary Science

TYPE Original Research
PUBLISHED 08 June 2023
DOI 10.3389/fvets.2023.1127025

Check for updates

OPEN ACCESS

EDITED BY
Laura Cercenelli,
University of Bologna, Italy

REVIEWED BY
Nadine Fiani,
Cornell University, United States
Joao Filipe Requicha,
University of Trás-os-Montes and Alto Douro, Portugal

*CORRESPONDENCE
Marine Traverson
matraver@ncsu.edu

RECEIVED 19 December 2022
ACCEPTED 17 May 2023
PUBLISHED 08 June 2023

CITATION
Chambers A, Traverson M, Neal S, Konala S and Harrysson O (2023) Performances of novel custom 3D-printed cutting guide in canine caudal maxillectomy: a cadaveric study. *Front. Vet. Sci.* 10:1127025.
doi: 10.3389/fvets.2023.1127025

COPYRIGHT
© 2023 Chambers, Traverson, Neal, Konala and Harrysson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Performances of novel custom 3D-printed cutting guide in canine caudal maxillectomy: a cadaveric study

Aidan Chambers[1], Marine Traverson[1,2]*, Shelby Neal[2,3], Satyanarayana Konala[2,4] and Ola Harrysson[2,4,5,6]

[1]Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC, United States, [2]Center for Additive Manufacturing and Logistics, College of Engineering, North Carolina State University, Raleigh, NC, United States, [3]Department of Chemical and Biomolecular Engineering, North Carolina State University, Raleigh, NC, United States, [4]Fitts Department of Industrial and Systems Engineering, North Carolina State University, Raleigh, NC, United States, [5]Department of Biomedical Engineering, College of Engineering, North Carolina State University, Raleigh, NC, United States, [6]Department of Material Science and Engineering, College of Engineering, North Carolina State University, Raleigh, NC, United States

Introduction: Caudal maxillectomies are challenging procedures for most veterinary surgeons. Custom guides may allow the procedure to become more accessible.

Methods: A cadaveric study was performed to evaluate the accuracy and efficiency of stereolithography guided (3D-printed) caudal maxillectomy. Mean absolute linear deviation from planned to performed cuts and mean procedure duration were compared pairwise between three study groups, with 10 canine cadaver head sides per group: 3D-printed guided caudal maxillectomy performed by an experienced surgeon (ESG) and a novice surgery resident (NSG), and freehand procedure performed by an experienced surgeon (ESF).

Results: Accuracy was systematically higher for ESG versus ESF, and statistically significant for 4 of 5 osteotomies ($p < 0.05$). There was no statistical difference in accuracy between ESG and NSG. The highest absolute mean linear deviation for ESG was <2 mm and >5 mm for ESF. Procedure duration was statistically significantly longer for ESG than ESF ($p < 0.001$), and for NSG than ESG ($p < 0.001$).

Discussion: Surgical accuracy of canine caudal maxillectomy was improved with the use of our novel custom cutting guide, despite a longer duration procedure. Improved accuracy obtained with the use of the custom cutting guide could prove beneficial in achieving complete oncologic margins. The time increase might be acceptable if hemorrhage can be adequately controlled *in vivo*. Further development in custom guides may improve the overall efficacy of the procedure.

KEYWORDS

3D-printing, maxillectomy, surgical guide, accuracy, oncologic margins

1. Introduction

Tumors of the oral cavity make up approximately 6% of all tumors in dogs, with malignant melanoma, squamous cell carcinoma, and fibrosarcoma being the most common neoplasms. Wide surgical excision with 1–2 cm gross surgical margins is recommended to obtain local tumor control in malignant cases and locally invasive benign cases such as acanthomatous ameloblastoma (1, 2). MacLellan et al. (2), however, reported that 31.6% of dogs undergoing any type of partial maxillectomy had incomplete histologic margins. Incomplete resection has been intimately associated with local tumor recurrence, and 65% of maxillary tumors removed with incomplete margins have been reported to recur locally versus 22% of tumors removed with complete margins (2–4). Anatomic location has also been proven to impact local tumor control. Sarowitz et al. (3) found that caudal maxillary tumors were associated with a 1.5 times hazard ratio for local tumor recurrence compared to all other maxillary tumors. This is potentially related to the difficulty of obtaining complete histologic margins in this challenging anatomic location (3–5). Indeed, caudal maxillary tumors are believed to be difficult to remove because of poorer surgical exposure compared to rostral tumors, more abundant vasculature caudally, and generally larger size masses that have escaped owners' notice until they start causing clinical signs (3–5). There is a need to increase technical ease and intraoperative accuracy for the excision of caudal maxillary tumors to optimize oncologic margins and improve long-term local tumor control.

Stereolithography (3D-printing) is becoming more popular in human and veterinary surgery because of the ability to make custom surgical guides, patient specific anatomical models, and individualized prostheses based on the diagnostic images of the patient (6–10). Stereolithography allows three dimensional models to be created in a variety of polymer materials using a laser light source that selectively cures and solidifies a liquid plastic layer-by-layer along the cross-section of the object (8, 10). Custom-made drilling and cutting guides have been demonstrated to aid in surgical accuracy, efficiency, and allow implants to be best fitted to the patient (9, 10). In the human surgical field, 3D-printed surgical guides have been largely used to improve accuracy in dental procedures with small margins of error and to maximize the chances of obtaining clean surgical margins in oncologic procedures (11–13). The use of a custom surgical guide has been shown to improve accuracy in 76%–92% of cases compared to traditional freehand procedures for osteotomies and complex reconstructions in human maxillofacial, dental, orthopedic, and oncologic surgery (7, 9). In addition to the maxilla being a complex surgical location, there is great variability in anatomy between breeds within the canine species (14, 15). Therefore, the use of 3D-printed custom-made surgical guides may also improve accuracy and local tumor control for canine caudal maxillectomy procedures.

Additionally, major surgical hemorrhage has been reported as the number one intraoperative complication in canine maxillectomy, and occurred in 83% of dogs undergoing a caudal maxillectomy *via* a combined dorsolateral and intraoral approach in a retrospective study by MacLellan et al. (2). As a consequence, patients undergoing a complete or caudal maxillectomy are reportedly 3 to 6.5 times more at risk of requiring a blood transfusion intraoperatively compared to other types of maxillectomy and/or mandibulectomy procedures (2, 16). This is thought to be related to the complexity of the locoregional vasculature (2, 16). Additionally, median duration of the surgical procedure is also significantly increased for a caudal maxillectomy compared to other types of maxillectomies, and certainly plays a role in the increased hemorrhagic risk (2). In addition to improvements in accuracy, the use of a 3D-printed cutting and drilling guide created by computer aided manufacturing (CAM) has been shown to decrease the amount of time spent in the operating room (7, 9). In a systematic review of 227 human surgical studies using 3D-printing technology, 28% of reports using 3D-printed custom-made surgical guides saw a reduction in operating room or treatment time with the use of CAM compared to conventional planning methods (7). The use of a 3D-printed custom-made surgical guide may decrease the duration of caudal maxillectomy procedures and allow for more rapid hemostasis to be obtained, lowering the morbidity associated with the surgery.

Finally, the use of a 3D-printed surgical guide has the potential to not only improve the efficiency and accuracy of caudal maxillectomy procedures performed by experimented surgeons, but also to make these complex procedures more accessible to novice surgeons or boarded surgeons with limited experience in caudal maxillectomy. In the human literature, there is conflicting information regarding the use of 3D-printed surgical guides and whether such guides facilitate the procedures in the hands of a novice surgeon in comparison to an experienced surgeon (8, 11–13). Because of the wide range of procedures that veterinary surgeons are trained to do, it is possible for a residency-trained surgeon to not have primary experience with a caudal maxillectomy by the time they complete their program. The use of a 3D-printed cutting guide, however, may allow a caudal maxillectomy to become more accessible to veterinary surgeons that are advanced in their surgical training but have not yet had the opportunity to perform such a procedure. Novice surgeons also have a tendency to perform procedures slower than experienced surgeons; that difference is even more pronounced with complex procedures, as demonstrated in human surgery (17, 18). For a complex procedure such as a caudal maxillectomy, the use of a 3D-printed custom-made cutting guide may allow a novice surgeon to perform the surgery in a duration comparable to an experienced surgeon.

The objectives of this study were to (1) design a 3D-printed custom-made caudal maxillectomy surgical guide, and to (2) evaluate the accuracy and efficiency of the surgical guide in cadaveric dogs. We hypothesized that the use of a 3D-printed custom-made surgical guide increases the accuracy and efficiency of the osteotomy compared to a standard freehand procedure (hypothesis 1), and that there would not be any difference in accuracy and efficiency between a novice and an experienced surgeon when using the cutting guide (hypothesis 2).

2. Materials and methods

2.1. Specimen randomization

Fifteen heads were obtained from fresh frozen canine cadavers euthanatized at local shelters for reasons unrelated to the study and thawed in preparation to the procedure. Both left and right sides of the heads were used as separate subjects. Head number and lateralization were randomized per testing group and order with 10 head sides per treatment group. The heads were inspected for any visual abnormality, and classified as doli-, brachy-, or mesocephalic. Their length was recorded from the most ventral aspect of the maxillary canine to the caudal aspect of the occipital bone.

2.2. Study groups

Two experimenters participated in this study: a board-certified surgeon and surgical oncologist (MT) with experience using 3D-printed guided implantology (considered the experienced surgeon) while the other experimenter (AC) was a second-year surgical resident at the time of the experimentation (considered the novice surgeon). To address the study objectives, three treatment groups were used that consisted of caudal maxillectomy procedures performed on canine cadaver heads (1) with the use of individualized 3D-printed surgical guide by the experienced surgeon (group ESG), (2) with the use of individualized 3D-printed surgical guide by the novice surgeon (group NSG), or (3) freehand by the experienced surgeon (group ESF). Both the experienced and novice surgeons were right hand dominant and assisted to the other's procedures (see Figure 1).

2.3. Maxillectomy planning

Computed tomography (CT) scans of all cadaver heads were performed *via* a 64-slice CT scanner (Siemens 64 slice; Siemens Medical Solutions, Malvern, Pennsylvania) and imported into a DICOM image processing software (Mimics Innovation Suite, Materialise, Leuven, Belgium). Following the procedure described by Lascelles et al. (19), five osteotomy cuts were defined to complete a caudal maxillectomy with ventral orbitectomy, (1) zygomatic cut, (2) rostrolateral cut, (3) dorsolateral cut, (4) palatine cut, and (5) orbital rim cut. The skulls were thresholded, segmented, and virtual osteotomy cuts were planned in 3-Matics software (Mimics Innovation Suite, Materialise, Leuven, Belgium) (see Figure 1).

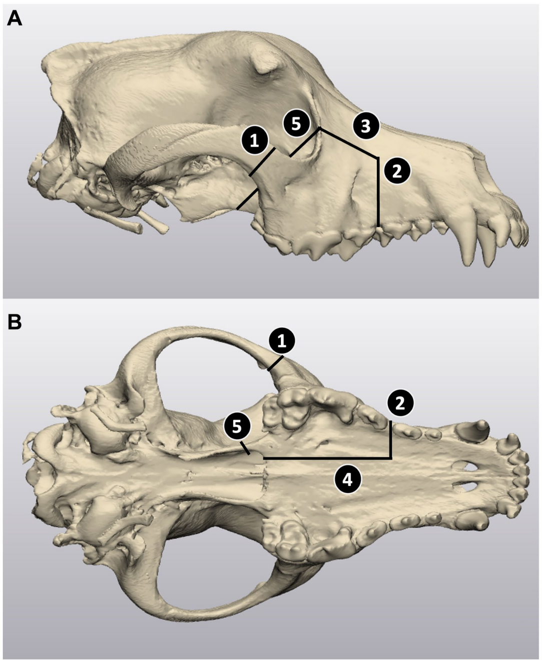


FIGURE 1
Virtual caudal maxillectomy planning in 3-Matic. **(A)** Lateral view of the skull; **(B)** ventral intraoral view. The cuts for the maxillectomies were planned based on the following anatomical landmarks: (1) zygomatic cut, located caudal to the zygomaticomaxillary suture, (2) rostrolateral cut, extending to one third of the width of the palate at the rostral aspect of the third maxillary premolar tooth, to 5mm rostral and dorsal to the infraorbital foramen, (3) dorsolateral cut, joining the dorsolateral extremity of cut 2 to the mid orbital rim of the lacrimal bone, (4) palatine cut, extending from the rostral intra-oral extremity of cut 2 to the caudal edge of the hard palate along the third of the width of the palate (5) orbit cut, angled from the extremity of cut 3 at the orbital rim to the medial aspect of the second maxillary molar tooth.

2.4. Custom guides manufacturing

Custom 3D-printed surgical guides were designed in 3-Matics software with the collaboration of an engineering undergraduate student (SN) at the Center for Additive Manufacturing and Logistics (CAMAL), for a total of 20 left and right sides of heads randomly allocated to treatment groups ESG and NSG. Each guide was made as two separate imbricating parts with a dorsal segment covering the orbit and dorsal maxilla and a ventral segment covering the caudal aspect of the dental quadrant and caudolateral palate as seen in Figure 2. Five drilling holes were added to secure the guide in place with 2.0 mm stainless steel screws (Arthrex Vet Systems, Naples, Florida, United States). The surgeon (MT) and surgery resident (AC) created the initial datum planes of the virtual cuts, and the cylinders at the location and orientation of the future screws as those require advanced knowledge of locoregional anatomy. The guide was then designed by the engineering student (SN) with strategical checkpoints along the design process at which times the CAD files were shared with the surgical team who reviewed the design before the engineering student could move on to the next step. The 3 checkpoints included a first review time before Boolean union of the parts, a second after trimming excess guide material, smoothening and planning the two-part sectioning of the guide, and a final checkpoint before printing. The cutting guides were printed in a stereolithography printer (Form 3; Formlabs, Sommerville, Massachusetts) with resin (Tough 1500; Formlabs, Sommerville, Massachusetts). Tough 1500 resin was chosen because it is considered biocompatible by the manufacturer as well as acceptable for steam autoclaving, gamma sterilization, and ethylene oxide sterilization.

2.5. Procedures

The head and procedure orders were randomized, and the experienced (MT) and novice (AC) surgeons performed the role of surgical assistant when not primarily involved in the procedure. Guidance for the guide placement or osteotomy was not provided to the novice from the experienced surgeon for the NSG treatment group, and *vice-versa*. All maxillectomy procedures were performed *via* a combined dorsolateral and intraoral approach as described by Lascelles et al. (19). This approach elevates the skin and nasolabialis muscle from the maxilla from the zygomatic arch to the rostrolateral aspect of the bone dorsal to the planned rostral margin; specificities of the guided procedure included exposure of the zygomatic arch and dorsolateral nose rostrally to the level of the canine, as well as the extension of the soft tissue dissection along the medial canthus of the eye to allow for guide placement. The guide was purposefully designed to allow placement without elevation of the gingiva or palatal tissue.

The freehand maxillectomy was performed as previously planned in the modeling software using the aforementioned anatomic landmarks (see Figure 1). The surgeon was allowed to visualize the model including the planned cuts in 3-Matics while drawing the planned resection on the cadaver head with a #15 scalpel blade to mark a thin cut line in the periosteum as one would visualize the CT images of the patients before marking the landmark of the cuts with an electrosurgical handpiece in a live procedure.

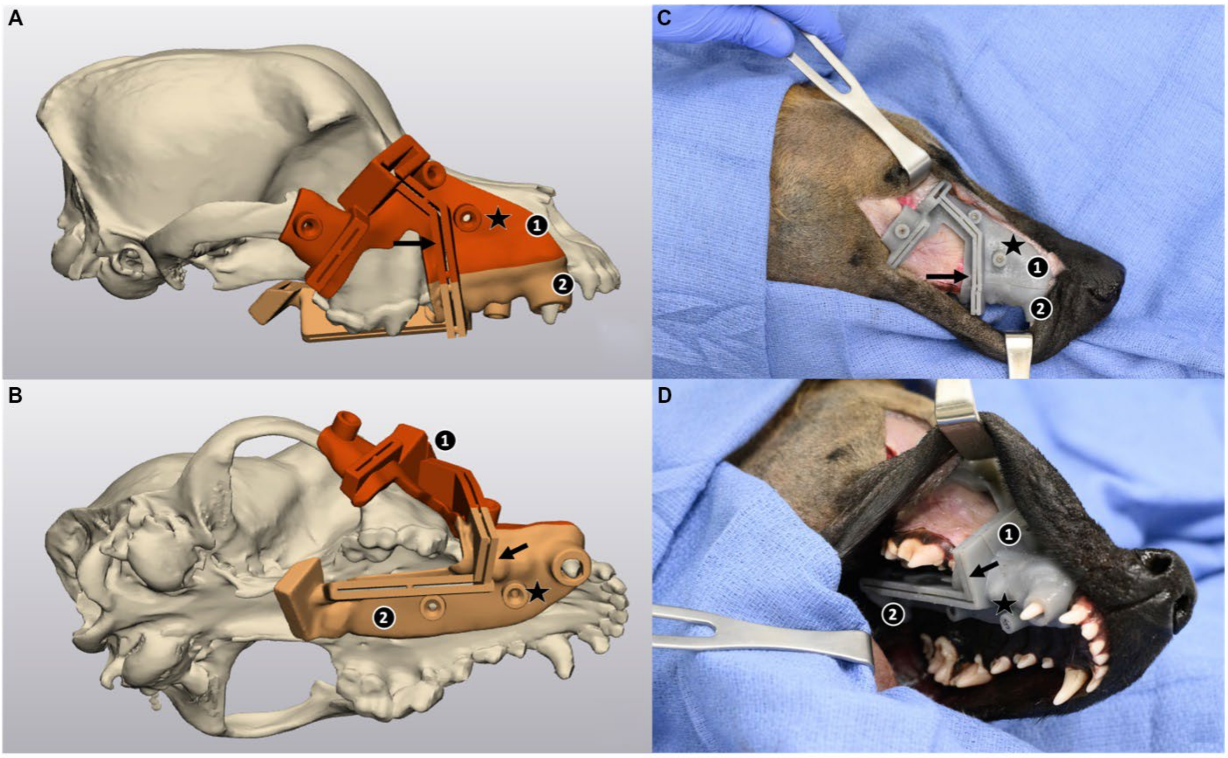


FIGURE 2
Three-dimensional model of a skull and computer-aided design (CAD) of a custom-made maxillectomy guide in lateral view **(A)** and ventrolateral view **(B)** in 3-Matic, and the corresponding views of the guide on a cadaver **(C,D)**. The guide is designed to recognize the contours of the skull, dental arch, gingiva, and palatine mucosa, and consists of 2 imbricating pieces (labeled 1 and 2), 5 drilling holes (star) to allow for the placement of cortical screws for stabilization, and 5 cutting grooves (arrow) guiding the 5 osteotomies. In **(C,D)** a combined dorsolateral and intraoral approach was performed to gain access to the maxilla. Screws have been placed to secure the guide in place.

For the guided osteotomy, the surgical guide was fitted to the head and secured in place with a total of five 2.0 mm cortical stainless-steel screws (Figure 2). The cuts were performed with a MicroAire oscillating saw (MicroAire, Charlottesville, Virginia) and a 0.6 mm Kerf blade for cuts 1 through 4; the orbital cut (5) was performed with a 15 mm × 2 mm osteotome and mallet. The guide was then removed after the screws had been retrieved to separate and extract the maxillectomy segment.

2.6. Quality assessment

For the ESG and NSG groups, the *ease of placement* for the guide was recorded as easy (able to fit the guide within 1 min without modification of the surgical approach), moderate (requiring modification of the surgical approach and/or able to fit the guide between 1 and 3 min), difficult (unable to fit or stabilize the guide without modifying the guide and/or able to fit the guide after 3 min).

Additionally, any instances of *guide failure*, including inability to fit the guide, cracking of the osteotome or saw groove, inability to secure the guide with screws, loose guide placement, or other failures were recorded.

For all three groups, the *quality of the cut* was recorded as high (smooth and straight cuts for all cuts and all cuts intersecting within 2 mm of their individual demarcation), moderate (one or two irregular cuts and/or cuts extending 2–5 mm beyond their intersection), or low (more than two irregular cuts and/or cuts extending >5 mm beyond their intersection).

2.7. Accuracy assessment

The linear deviation was calculated using the model-to-model distance extension in 3D Slicer from Kitware (Kitware Inc., Clifton Park, New York, United States). When the deviation was toward the excised segment, a positive value was assigned and when away from the excised segment, a negative value was assigned. Absolute values of these deviations were used to calculate the means.

Average linear deviation of the performed cuts from the planned cuts was used to measure accuracy between groups. The procedure used to obtain 3D models of the heads for planning the cuts was repeated with the skulls scanned post-maxillectomy (Figure 3). For guided subjects (ESG and NSG groups), an additional CT scan was performed with the guide secured in place prior to the osteotomy. This was included to investigate the origin of the deviation and establish if the variation observed between the planned and performed cuts was a result of an error at the time of guide placement (due to CAD/CAM and/or the individual placement considered *manufacturing/positioning error*) or a consequence of guide shifting during the osteotomy (considered *cutting error*).

To establish a comparison between *planned* (pre-op), *guided* (post-guide placement), and *performed* (post-maxillectomy) cuts in

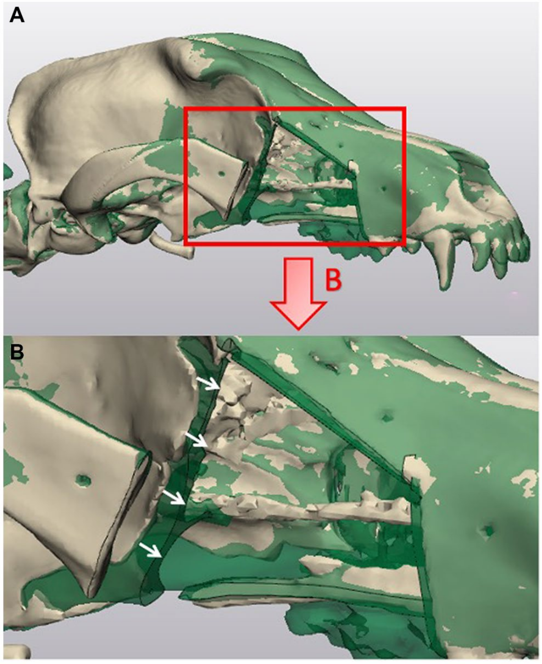


FIGURE 3
Representation of linear deviation evaluation between planned and performed maxillectomy. **(A)** Is a CAD superimposition of the post-maxillectomy skull with performed cuts (beige) and the virtual maxillectomy skull with planned cuts (green). **(B)** Is an inset of the zoomed in image of the superimposed skulls. The white arrows indicate an example of the distances that were measured using a cloud compare analysis to evaluate the mean linear deviation at the orbit cut.

the ESG and NSG groups, the CT scans obtained at each step were overlaid in Geomagics to align them with each other. Local multipoint alignment was used to manually indicate prominent features on the skulls for the software to then align the models. In the ESF group, only planned and performed cuts were evaluated. Planes of the cuts were generated for each of the aligned skulls and guides in 3-Matic using the three-point method, where a skilled user picked the points based on cut alignment. In addition to this, the surfaces of each of the cuts were marked manually in 3-Matic and exported as STL files, along with the planes.

Cloud Compare (open-source software), a 3D point cloud processing software, was used to calculate the deviation between the various generated planes and points on the cut surfaces obtained from CT scans superimposed (Figure 3), as reported in previous studies (20). For each cut, the surface STL was converted to a point cloud, and a primitive was generated by fitting a plane to the STL of the plane obtained from 3-Matic. The "distance to primitive" function was then used to obtain the distance of each of the points on the surface to the fit plane. Using this distribution, a histogram was generated and used for further analysis. Figure 4 shows an example of a heat map generated using the distances obtained for the zygomatic arch cut. The difference between the planned and performed cut were recorded as *total error* for each cut in all 3 treatment groups. The difference between the guided and performed cut was recorded for each cut only in the ESG and NSG groups, and considered *cutting error*. *The manufacturing/positioning errors* representing the difference between planned and guided cut was obtained by subtraction of the *cutting error* from the *total error* in the ESG and NSG groups.

2.8. Efficiency assessment

The portions of the maxillectomy procedures were divided in different time sequences to compare efficiency between ESG and ESF groups (reflecting on the use of the guide compared to freehand procedure) and between ESG and NSG groups (reflecting on the impact of surgical experience, while using the guide). The times required to model the corresponding head, and design and print each guide was also recorded as *CAD/CAM time*.

More specifically, for the ESF treatment group, the time to visually plan and mark the planned cuts on the cadaver head with a scalpel blade was recorded as *preparation time*. For the ESG and NSG groups, the *preparation time* was divided in *placement time* (time required to place the guide) and *securing time* (time required to secure the guide with screws), recorded separately.

The time to complete the osteotomy in the ESF treatment group was recorded as *freehand cut and removal time* and included the removal of the excised bone segment which clinically is performed concomitantly. In the ESG and NSG groups, the cutting and removal times were recorded separately as *guided cut* and *removal time* since the removal of the excised bone segment required prior removal of the guide.

The *total maxillectomy time* was recorded as the time from freehand planning or guide placement to osteotomy and removal of the excised bone segment. The surgical approach was not included in the time recorded. Accordingly, the *total maxillectomy time* for freehand subjects was an addition of *preparation* and *freehand cut and removal times.* The *total maxillectomy time* for guided subjects was the addition of the *placement, securing, guided cut, and guided removal times.*

2.9. Statistics

Based on a review of human literature regarding accuracy of 3D-printed custom-made surgical guides in maxillofacial, dental, and oncologic surgery, a mean linear deviation of 2 mm (+/−0.5) for both guided groups and 5 mm (+/−0.5) for ESF was expected (21–24). Using these expected values, a power analysis with an alpha of 0.05 and beta of 0.8 was performed and led to a minimum of 8 subjects per treatment group. Therefore, a total of 10 cadaver head sides was chosen as the number of subjects per group, which raised the total for 3 groups to 30 sides in 15 heads.

A Kruskal–Wallis test was performed to evaluate homogeneity across treatment groups in regards to head size.

The median and range was recorded for the total manufacturing time of the guides. The mean and standard deviations for times to plan freehand cuts, perform freehand cuts, place guide, secure guide, perform guided cuts, and remove the excised portions were recorded. The mean and standard deviations for linear deviation of each of the 5 cuts were calculated based on their absolute values, meaning that distances were examined without any consideration of which side of the designed cut they lay on. In cases where the performed cut intersected with the design cut, instead of the average then coming closer to 0 from having both

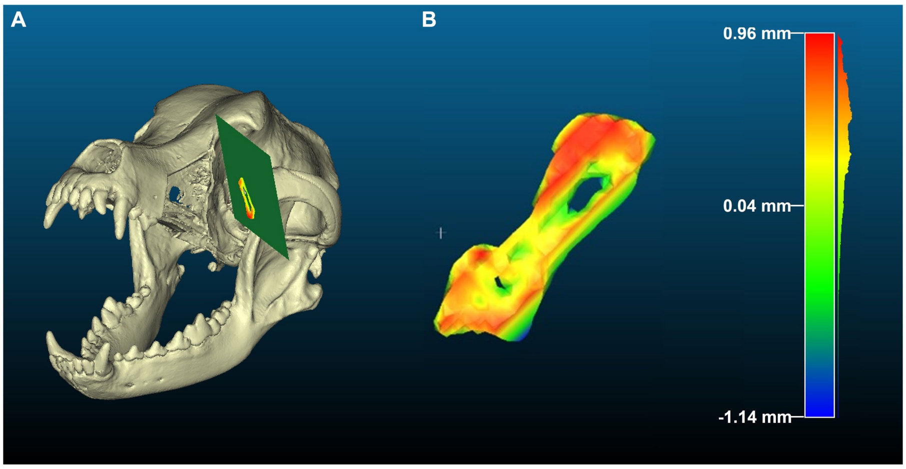


FIGURE 4
Heatmap generated using the "distance to primitive" function in cloud compare for cut 1 (orbit). **(A)** Demonstrates the planned cut plane in green and the performed cut in heatmap on the postoperative skull model. **(B)** Is a highlight of the cutting error observed along the orbit cut from planned to performed cut. In this case the cutting error was negative, or toward the defect, with the lowest to highest error pictured in warm (red) to cold (blue) colors, respectively. Scale bar indicates 15mm. Color bar units are in millimeters.

positive and negative values, the side was ignored to estimate how much deviation from the design plane occurred.

Standard deviations were similarly calculated from the average absolute deviations and were compared to give insight into the consistency of various cuts. Comparisons were done *via* t-test or, where the normality assumption failed, Wilcoxon rank sum test on the square root of the average absolute deviation or on the untransformed standard deviation. The square root transformation was applied to improve the normality of the data.

For the accuracy statistical analyses, two sets of data were run: the primary set with all heads and cuts included, and an *exclusion set*. To maximize test subjects with the limited cadavers available, maxillectomies were performed on the left and right sides of the head, as described previously. For some cadaver heads, the anatomic structure of the remaining head was altered slightly once both left and right caudal maxilla were removed, creating some inexactitude in the measurement of the linear deviation. Those heads were excluded from analysis in the *exclusion set* of data.

3. Results

3.1. Specimens

The cadaver heads ranged in length of 17–32 cm from the maxillary canine to the caudal aspect of the occipital bone. Two heads were brachycephalic, two dolicephalic, and the remainder of the heads were mesocephalic. There was no significant difference in the head lengths between the groups (p = 0.5157 from a Kruskal–Wallis test). The cadaver head lengths are summarized in Table 1.

TABLE 1 Summary of the cadaver head lengths.

	Median (minimum-maximum) (cm)	Mean (standard deviation) (cm)	Head classification (brachycephalic/ mesocephalic/ dolicephalic)
ESF	23 (17–32)	23 (±4.5)	0/9/1
ESG	22 (17–32)	22.4 (±4.6)	1/8/1
NSG	21 (17–24)	20.8 (±2.4)	1/9/0

The length was recorded from the most ventral aspect of the maxillary canine to the caudal aspect of the occipital bone. There was no significant difference in the head lengths between the groups (p = 0.5157 from a Kruskal–Wallis test).

3.2. Quality assessment

3.2.1. Ease of placement

Nine guides (9/10) were qualified as easy to place in the ESG group, and one guide (1/10) was graded as moderately easy. In the NSG group, 3/10 guides were considered easy to place, 6/10 moderately easy, and 1/10 difficult to place.

3.2.2. Guide failure

Five instances of guide failure were observed, 4/5 in the ESG group and 1/5 in the NSG group; none of the failures yielded low quality cuts. Failures of the guide were observed as follows: 2/5 loosen securing screws, 1/5 osteotome slot crack, 1/5 saw slot crack, and 1/5 extension of the saw mark out of the cutting groove.

3.2.3. Cut quality

Seven of thirty (23%) and 14/30 (47%) maxillectomy subjects had low or moderate quality cuts, respectively; distribution between groups is represented in Table 2. Of those with low or moderate quality cuts, 20/21 (95%) were related to the orbit cut. For the subject that had a moderate quality cut that did not occur with the osteotome, the jagged cut was located at the zygomatic arch. Nine subjects had incomplete cuts at the time of segment removal; 2 of those were observed in the ESF and ESG groups each, and 5 in the NSG group.

3.3. Accuracy assessment

3.3.1. Absolute linear deviation

The mean linear deviation from planned to performed cuts was overall lower in the ESG group compared to the ESF group, with a statistically significant difference for cuts 1 ($p = 0.032$), 2 ($p = 0.003$), 3 ($p < 0.001$), and 5 ($p = 0.045$) as shown in Table 3 and Figure 5. There was no significant difference in the mean linear deviation between ESG and NSG groups for any of the cuts. The greatest mean linear deviation from planned to performed cuts observed in the ESG, NSG, and ESF groups was 1.98 ± 0.81 mm, 3.19 ± 1.64 mm, and 5.46 ± 4.28 mm, respectively; all occurred at the orbit cut 5 (Table 3 and Figure 5).

The stepwise breakdown of the mean linear deviation observed with the use of the guide is represented in Figure 6. For all cuts combined, there was a trend toward the cutting error being more significant than the manufacturing/positioning error, however this was statistically significant for ESG only for cuts 2 and 5 ($p = 0.03$ and $p = 0.001$, respectively), and for NSG for cuts 2, 3, and 5 ($p = 0.021$, $p = 0.028$, and $p < 0.001$, respectively). The total error was partially corrected from the positioning/manufacturing error to the cutting error due to difference in directions (positive vs. negative) in 43 and 48% of cuts for the NSG and ESG groups, respectively.

3.3.2. Corrected values

Four heads were excluded in the *exclusion set* because of shifts in the skull anatomy post-maxillectomy suspected to be secondary to the removal of bilateral caudal maxilla. With those four heads excluded, there was no significant difference in mean linear deviation between the new values and the original data set.

3.3.3. Consistency

The values for standard deviation between the planned and performed cuts are shown in Table 4. The standard deviation from planned to performed cut was significantly lower for the ESG group than the ESF group for cut 2 ($p = 0.003$) and cut 3 ($p < 0.001$). There was no significant difference in standard deviation from planned to performed cut and from guided to performed cut for the NSG group compared to the ESG group for any of the cuts. There was no difference in frequency of deviation in one direction versus the other for all cuts.

TABLE 2 Quality of the performed cuts.

Treatment group	Cut quality		
	Low	Moderate	High
ESF	2/10	4/10	4/10
ESG	1/10	5/10	4/10
NSG	4/10	5/10	1/10

Quality of the cut was recorded as high (smooth and straight cuts for all cuts and all cuts intersecting within 2 mm of their individual demarcation), moderate (one or two irregular cuts and/or cuts extending 2–5 mm beyond their intersection), or low (more than two irregular cuts and/or cuts extending >5 mm beyond their intersection).

3.4. Efficiency assessment

3.4.1. Manufacturing time

The median total *CAD/CAM time* for the ESG and NSG groups was 10.7 h, with a range of 9 to 16.6 h per guide including a median printing time of 4.75 h (range, 3.4 to 10.5 h).

3.4.2. Preparation time

A significant difference was found in the *preparation time* between the ESG and ESF groups, with a median *preparation time* 3.38 time longer in the ESG group compared to the ESF group ($p < 0.001$) (Figure 7).

A significant difference was found in the *placement time* ($p = 0.004$) and the total *preparation time* ($p = 0.023$) between the ESG and NSG groups, with a median *placement time* and *preparation time* 2.3 and 1.2 time longer, respectively, for the NSG group compared to the ESG group (Figure 7). No significant difference was found in the median *securing time* between the ESG and NSG groups ($p = 0.13$).

3.4.3. Cut and removal time

A significant difference was found in the cut and removal time between the ESG and ESF groups ($p < 0.001$), with a median *guided cut and removal time* 1.9 time longer than the *freehand cut and removal time* for the ESG group compared to the ESF group (Figure 7).

A significant difference was found in the *guided cut time* ($p < 0.001$), the *guided removal time* ($p = 0.034$), and in the total *guided cut and removal time* ($p < 0.001$) between the ESG and NSG groups, with the NSG group having a median *guided cut time, guided removal time, and guided cut and removal time* 1.5, 1.8, and 1.6 time longer, respectively, than the ESG group.

3.4.4. Total maxillectomy time

A significant difference was found in the *total maxillectomy time* between the ESG and ESF groups ($p < 0.001$), with the ESG group having a median *total maxillectomy time* 2.3 time longer than the ESF group (Figure 7).

A significant difference was found in the *total maxillectomy time* between the ESG and NSG groups ($p < 0.001$), with the NSG group having a median *total maxillectomy time* 1.4 time longer than the ESG group (Figure 7).

TABLE 3 Summary values for average +/− SD absolute linear deviation from planned to performed cuts (mm).

Treatment group	Cut				
	1 (Zygomatic)	2 (Rostrolateral)	3 (Dorsolateral)	4 (Palatine)	5 (Orbit)
ESF	1.26 ± 0.76^{A}	1.32 ± 0.58^{A}	3.02 ± 1.72^{AB}	2.62 ± 1.77	5.46 ± 4.28^{A}
ESG	*0.61 (0.4, 0.73)^{A}	0.56 ± 0.36^{A}	0.78 ± 0.68^{A}	1.45 ± 0.72	1.98 ± 0.81^{A}
NSG	1.11 ± 0.7	0.82 ± 0.62	0.78 ± 0.46^{B}	*1.75 (1.19, 2.35)	3.19 ± 1.64

Superscript (A, B) indicates significant differences between average absolute deviation for $p < 0.05$. Results are reported as median with quartiles for variables with significant deviation from normality and noted with an asterisk.

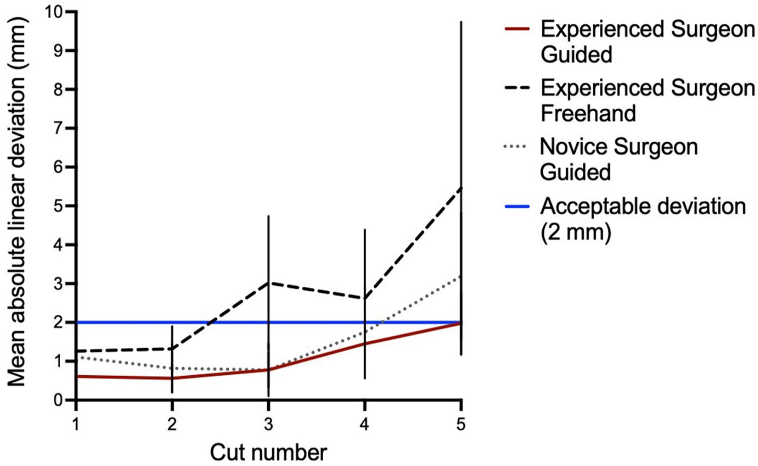


FIGURE 5
Linear deviation from planned to performed cut. All cuts for the experienced surgeon guided (ESG) fall within the acceptable deviation of 2mm. Most of the novice surgeon guided (NSG) cuts fall within the acceptable deviation. Three of five cuts for the experienced surgeon freehand (ESF) are beyond the acceptable deviation (>2mm).

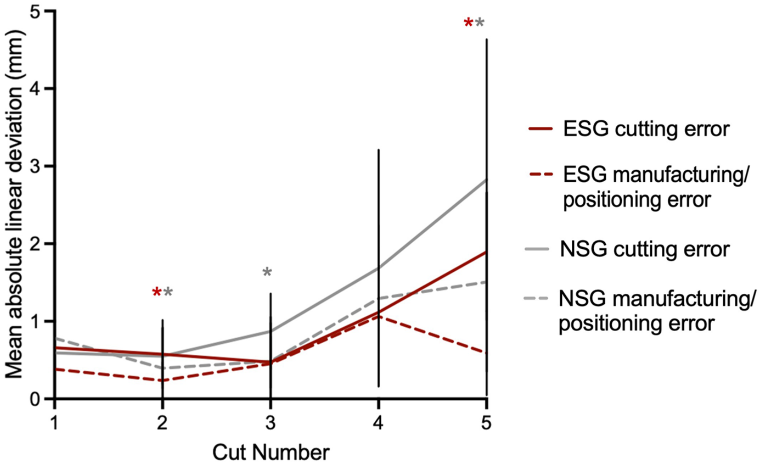


FIGURE 6
Origin of deviation observed with use of the guide. There is a trend toward the cutting error being more significant than the manufacturing/positioning error but that is not true for all cuts and only significant for the cuts with an asterisk. The asterisk indicates a significant difference between cutting error and manufacturing/positioning error with red for the experienced surgeon and gray for the novice. Although not statistically significant, there is a trend for the novice to introduce more cutting error than the experienced surgeon.

4. Discussion

In this study, we successfully designed, manufactured and used 20 3D-printed custom-made caudal maxillectomy guides in canine cadavers. We found that the use of the guide did improve the overall accuracy of the procedure when evaluating the outcomes of the experienced surgeon compared to conventional freehand technique, and allowed the novice surgeon to reach similar accuracy. The use of the guide did, however, decrease the efficiency of the procedure compared to conventional freehand technique, and this was more pronounced with the novice surgeon compared to the experienced surgeon. Therefore, we completed our objectives and partially accepted our first and second hypotheses.

To the authors' knowledge, this is the first veterinary study to specifically evaluate the accuracy of a custom-made 3D printed surgical guide for a caudal maxillectomy. A few veterinary studies have described the use of 3D printed cutting guides in combination with custom-made implants placement to reconstruct oncologic defects of the mandible, radius, tibia, and the skull (25–28). Improvements in accuracy have been reported with the use of surgical guides for pedicle screw placement in spinal surgery leading to a decreased risk of breaching the vertebral canal (29–31). In our study, an overall gain in accuracy was obtained with a global mean linear deviation improved from 2.74 mm to 1.08 mm with the use of the guide compared to the traditional freehand procedure for the experienced surgeon. This compares favorably with human literature

TABLE 4 Summary values for standard deviation from planned to performed cuts.

Treatment group	Cut				
	1	2	3	4	5
ESF	*0.12 (0.08, 0.16)	0.49 ± 0.47[AB]	1.25 ± 0.94[AB]	*0.15 (0.1, 0.41)[A]	3.28 ± 2.07
ESG	*0.08 (0.06, 0.14)	0.07 ± 0.05[B]	*0.04 (0.02, 0.09)[B]	*0.42 (0.19, 0.58)	*1.09 (0.41, 1.72)
NSG	0.1 ± 0.06	0.05 (0.05, 0.08)[A]	0.17 ± 0.15[A]	0.76 ± 0.58[A]	*0.85 (0.52, 2.57)

Superscript (A, B) indicates significant differences between average absolute deviation for $p < 0.05$. Results are reported as median with quartiles for variables. Significant deviation from normality is noted with an asterisk.

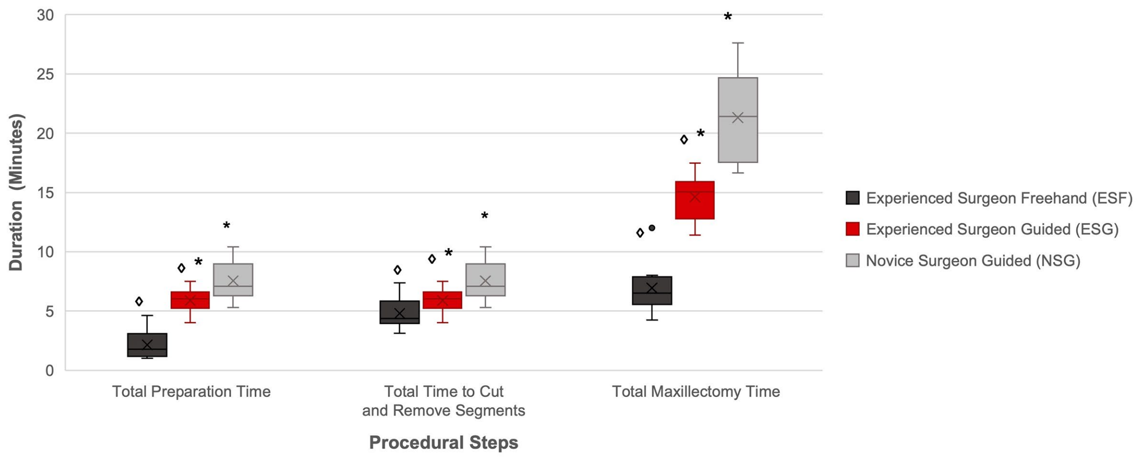


FIGURE 7
Comparison of total preparation time, cut and removal time, and total maxillectomy time per study group. The diamond and asterisks indicate significant differences (p<0.05). There was a significant difference between ESF and ESG, and NSG and ESG for the total preparation time, total time to cut and remove the segment, and total maxillectomy time.

reporting mean deviations of 1.17 mm and 2.49 mm from planned to achieved cuts with surgical guides, and ranges of 0.74 to 3.60 mm and 1.3 to 4.0 mm (11, 20, 23, 32). In comparison, the average deviation reported for freehand osteotomies when removing pelvic bone tumors in people is around 5 mm (13, 21). Most importantly, in our study, all cuts were within a mean linear deviation of 2 mm with the use of the guide while the most difficult cut along the orbit had a mean deviation superior to 5 mm without its use, similarly to previously reported freehand pelvic osteotomies (13, 21). This is particularly essential in an oncologic clinical scenario where the accuracy of oncologic margins could imply a difference between complete and incomplete tumor excision, leading to a higher risk of tumor recurrence and a potential need for adjuvant oncologic treatment. For most oral tumors, wide tumor excisions are planned with a margin of 1–2 cm of macroscopically normal tissue (33, 34). Based on those recommendations and previous papers evaluating accuracy of surgical guides, 2 mm was chosen as the threshold for a clinically acceptable surgical error (7, 20, 23, 32, 35, 36). Ultimately, achieving consistency in cutting accuracy is fundamental as one single largely deviated cut could be enough to lead to an incomplete excision. Our results show an improvement in the consistency for 2 out of 5 cuts with the use of the guide compared to the traditional freehand procedure. Overall, the use of the guide might help achieve more precise and more consistent osteotomy.

No significant difference in accuracy nor consistency was noted between the novice and the experienced surgeon suggesting that the 3D-printed custom-made surgical guide could provide assistance and reassurance to a less experienced surgeon performing a newly practiced procedure. Surgical approach and management of hemorrhage, however, are essential aspects of the procedure that are not exemplified by the guide. Few previous veterinary studies compared the accuracy obtained with the use of 3D-printed custom-made surgical guides from an experienced to a novice surgeon in corrective osteotomy and pedicle screw placement; no difference in accuracy was noted between the experienced and novice users in those scenarios (29, 37, 38). In the human literature, implant placement accuracy using 3D-printed custom-made surgical guides revealed no difference in accuracy between the experienced and novice surgeons in periodontal patients (39). This correlates to our findings and suggests that the guide represents a supportive surgical tool allowing a novice surgeon to perform a complex procedure without relying as intensely on traditional surgical mentorship. Additionally, the consistency of the procedure was similar between the novice and the experienced surgeon when using the cutting guide. This is in partial agreement with previous human studies as some report that custom guides allow novice surgeons to be as accurate as experienced surgeons and some suggests that experience still aids the user in obtaining higher accuracy even with the use of a guide (12, 39).

Part of our study design was also performed to measure the degree of error obtained at different steps of the guided procedure, as previous research has demonstrated that the total cumulative error is a sum of errors encountered during the CAD/CAM process, during the positioning of the implant, and during the actual cutting time (35, 40). The stepwise comparison between the planned, guided, and performed cuts revealed that the largest component of the total linear deviation originated from a cutting error, and suggests some existing micromotion after guide placement and during the performance of the osteotomy. The flexibility of the resin, the separation in 2 guide parts, the type of anatomic support (bone vs. mucosa vs. teeth), and the number of teeth included for support might all have played a role in this cumulative error (35, 40). Additionally, five of 20 guides showed some degree of failure (loosening of screws, osteotome or saw slot cracking, sawing outside of the groove) during use. Those failures may not represent relevant information to draw guidelines for guide use or manufacturing as the failure types reported were diverse and only one type of failure (screws loosening) occurred more than one time in this study. Overall, this is the first version of a caudal maxillectomy guide designed and tested in veterinary medicine; future prototyping might improve its outcome.

Finally, despite the noticeable improvement in accuracy reported in our study with use of the guide, the orbital cut carried the largest linear deviation in all groups and subjectively demonstrated a low to moderate cut quality in 20/30 (66%) of cases, all groups confounded. Overall, the cut quality along the orbit was higher for the experienced surgeon, with or without use of the guide, compared to the novice surgeon, which reflects on the difficulty of performing that part of the procedure. Its complexity is certainly related to the fact that the osteotomy is performed in a partially blind fashion, which is inherent to the procedure itself and the locoregional anatomy. One of the main purposes of our study was to facilitate that particular step by allowing the osteotome groove to guide the orbital cut, which was partially accomplished although our subjective impression remained mitigated. Improvement might come with some modifications in the guide design and use of different shapes and sizes of osteotomes.

Additionally, the overall efficiency of the procedure was lower compared to the traditional freehand technique. Therefore, our results compare unfavorably to human literature reviews of 3D-printed guides, which describe a decrease in intraoperative and operating room time in 80 and 46% of the cases, respectively (10, 41). Our results also compare unfavorably to a cadaveric veterinary study by Kim et al. evaluating the use of 3D-printed custom-made TPLO guides which reported a mean TPLO procedure time of 19 min with guide compared to 30 min without guide (42).

When breaking down the different procedural steps, the majority of the time difference observed for the experienced surgeon was found in the actual time required to perform the cut and remove the bone segment. This unfortunately also corresponds to the most critical step of the procedure where profuse hemorrhage can be encountered, thus the actual time of the procedure when an increase in efficiency would be beneficial. The choice made during our guide design to use screws instead of pins to improve guide stabilization might have been a trade off against efficiency, and replacing screws with pins could help placing and removing the guide faster. In addition, the cutting groove design did not allow to visualize the completion of the cuts at their intersections compared to the freehand procedure as the double cutting walls obscured the view; therefore, a larger number of incomplete cuts were noted in the guided groups, which was certainly responsible in part for the lack of efficiency. It is also possible that the efficiency demonstrated in the freehand group might have been overestimated by the use of cadavers. In live patients, surgeons generally experience profuse hemorrhage as the caudal maxillectomy is performed, unless the maxillary artery has already been ligated as previously described (43). Therefore, performing the maxillectomy and removing the segment in a live case scenario would likely take longer as blood obscures the field; the use of a guide could improve efficiency in that instance as the cuts could still be performed and finalized without absolute clear visualization.

Finally, the novice surgeon spent overall longer performing each step of the maxillectomy, and had a larger number of incomplete or subjectively low to moderate quality cuts. The increase in procedural time may have come from both the lack of familiarity with the procedure itself and with the use of a custom guide. Previous studies have demonstrated that experience with guides, even for the same individual, can improve the ability to position the guide (12). This was reflected as well in our efficiency and in our quality assessment as the placement of the guide was overall graded as more difficult for the novice group compared to the experienced surgeon. The cuts were also noted as incomplete in half of the cases for the novice group, which would have impacted the total time required to perform the cut and remove the maxillectomy segment. Therefore, although the use of a custom-made 3D printed guide might complement the surgical mentorship obtained, it certainly does not replace active supervision and might rather provide a comprehensive tool for initial practical exposure.

The main limitation of this study remains its cadaveric nature, and the fact that hemorrhage observed clinically may lower the efficiency and accuracy of the freehand procedure, therefore increasing the gap between the guided and traditional technique. The use of a guide may indeed aid in negating the effect of hemorrhage in a live patient because cuts could still be made even as blood obscures the field. Additionally, cadavers did not exhibit any pathology, and it is possible that the guide may not have the same fit with a large mass effect at the caudal maxilla. To accommodate for hypothetical large tumors, the guides were prophylactically designed with a large hollow center that would allow them to fit on the maxilla even in the instance where a tumor would surround the molars and/or last premolar teeth. Finally, the use and removal of left and right caudal maxilla to optimize cadaver usage and statistical power of our study design lead to four heads having a shift of the anatomy from an overall instability of the remaining skull. Even when excluding those four heads, however, the accuracy results were mostly identical and correlated with previous conclusions.

Another limitation of this study is the absence of comparison with a novice freehand group like previously reported by Bongers et al. (37) The goal of our study's group design was to evaluate if a guide improved the accuracy and efficiency of a caudal maxillectomy for an experienced surgeon (comparison ESF to ESG groups) and to evaluate if the guide could allow a novice surgeon to perform the procedure with similar accuracy and efficiency of an experienced surgeon (NSG to ESG groups). Further work could be considered with a novice freehand group, however we decided against that evaluation considering that a caudal maxillectomy is a complicated procedure that would not be expected from a novice surgeon without completing a general surgical training program. A novice freehand group, however, could be considered to determine if the guide helps improve the accuracy and/or efficiency of the procedure for a novice surgeon, and if the improvement obtained with the guide is proportional to the difference provided to the experienced surgeon.

In addition to the guide efficiency in terms of use, the manufacturing time for the guides was lengthy. Some guides were printed as the sole objects on the printer plate and some were printed 2–3 guides to a plate. The number of objects on a printer plate greatly influences the duration of a print and therefore, the total manufacturing time that was reported in this study may not be realistic for the total manufacturing time required for a single surgical guide printed for a clinical case.

In conclusion, our novel 3D-printed custom-made cutting guide appears to improve the accuracy of the caudal maxillectomy for an experienced surgeon, and allows a novice surgeon to perform the procedure with similar accuracy. However, it is not evident that the efficiency of the procedure is improved with the guide. In freehand caudal maxillectomies, the orbit cut is frequently considered the most challenging and, for the guided procedures, continues to lead to the most inaccuracy. There is a trend toward greater accuracy with the guide for the orbit cut, however, and further improvements to a maxillectomy guide could facilitate a more accurate procedure.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

Ethical review and approval was not required for the animal study because this study used cadavers from shelter animals euthanized for reasons unrelated to the study.

Author contributions

AC, MT, SN, and OH contributed to conception and design of the study. AC, MT, SN, and SK designed the guides. AC and MT performed the cadaver work. SN and SK performed the analysis. AC and MT wrote the first draft of the manuscript. SK wrote sections of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by 2020 NCSU CVM Extramural Funding Incentive Grant ($10000).

Acknowledgments

The authors would like to acknowledge James Robertson for aiding with statistical analysis and Jim Robey and Teri Critcher for aiding with numerous CT scans. John Joyner for capturing cadaveric photographic images and Jon Hash for collecting and storing cadaveric heads.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Goldschmidt SL, Bell CM, Hetzel S, Soukup J. Clinical characterization of canine acanthomatous ameloblastoma (CAA) in 263 dogs and the influence of postsurgical histopathological margin on local recurrence. *J Vet Dent*. (2017) 34:241–7. doi: 10.1177/0898756417734312

2. MacLellan RH, Rawlinson JE, Rao S, Worley DR. Intraoperative and postoperative complications of partial maxillectomy for the treatment of oral tumors in dogs. *J Am Vet Med Assoc*. (2018) 252:1538–47. doi: 10.2460/javma.252.12.1538

3. Sarowitz BN, Davis GJ, Kim S. Outcome and prognostic factors following curative-intent surgery for oral tumours in dogs: 234 cases (2004 to 2014): curative-intent surgery for oral tumours in dogs. *J Small Anim Pract*. (2017) 58:146–53. doi: 10.1111/jsap.12624

4. Schwarz PD, Withrow SJ, Curtis CR, Powers BE, Straw RC. Partial maxillary resection as a treatment for oral cancer in 61 dogs. *J Am Anim Hosp Assoc*. (1991) 27:617–24.

5. Verstraete FJM. Mandibulectomy and maxillectomy. *Vet Clin N Am Small Anim Pract*. (2005) 35:1009–39. doi: 10.1016/j.cvsm.2005.03.005

6. Bellanova L, Paul L, Docquier P-L. Surgical guides (patient-specific instruments) for pediatric tibial bone sarcoma resection and allograft reconstruction. *Sarcoma*. (2013) 2013:1–7. doi: 10.1155/2013/787653

7. Bosma SE, Wong KC, Paul L, Gerbers JG, Jutte PC. A cadaveric comparative study on the surgical accuracy of freehand, computer navigation, and patient-specific instruments in joint-preserving bone tumor resections. *Sarcoma*. (2018) 2018:1–9. doi: 10.1155/2018/4065846

8. Quinn-Gorham D, Javed Khan M. Thinking outside of the box: the potential of 3D printing in veterinary medicine. *J Vet Sci Technol*. (2016) 7. doi: 10.4172/2157-7579.1000360

9. Liang H, Ji T, Zhang Y, Wang Y, Guo W. Reconstruction with 3D-printed pelvic endoprostheses after resection of a pelvic tumour. *Bone Joint J*. (2017) 99-B:267–75. doi: 10.1302/0301-620X.99B2.BJJ-2016-0654.R1

10. Tack P, Victor J, Gemmel P, Annemans L. 3D-printing techniques in a medical setting: a systematic literature review. *Biomed Eng Online*. (2016) 15:115. doi: 10.1186/s12938-016-0236-4

11. Abou-ElFetouh A, Barakat A, Abdel-Ghany K. Computer-guided rapid-prototyped templates for segmental mandibular osteotomies: a preliminary report. *Int J Med Robot*. (2011) 7:187–92. doi: 10.1002/rcs.387

12. Cassetta M, Bellardini M. How much does experience in guided implant surgery play a role in accuracy? A randomized controlled pilot study. *Int J Oral Maxillofac Surg*. (2017) 46:922–30. doi: 10.1016/j.ijom.2017.03.010

13. Sallent A, Vicente M, Reverté MM, Lopez A, Rodríguez-Baeza A, Pérez-Domínguez M, et al. How 3D patient-specific instruments improve accuracy of pelvic bone tumour resection in a cadaveric study. *Bone Joint Res*. (2017) 6:577–83. doi: 10.1302/2046-3758.610.BJR-2017-0094.R1

14. Evans HE, de Lahunta A. *Miller's anatomy of the dog—E-book* Elsevier Health Sciences. United States: St. Louis, Missouri (2013).

15. Schoenebeck JJ, Ostrander EA. The genetics of canine skull shape variation. *Genetics*. (2013) 193:317–25. doi: 10.1534/genetics.112.145284

16. Cray M, Selmic LE, Kindra C, Abrams B, Story A, Hovis K, et al. Analysis of risk factors associated with complications following mandibulectomy and maxillectomy in dogs. *J Am Vet Med Assoc*. (2021) 259:265–74. doi: 10.2460/javma.259.3.265

17. Hernández-Irizarry R, Zendejas B, Ali SM, Lohse CM, Farley DR. Impact of resident participation on laparoscopic inguinal hernia repairs: are residents slowing us down? *J Surg Educ*. (2012) 69:746–52. doi: 10.1016/j.jsurg.2012.08.013

18. Puram SV, Kozin ED, Sethi R, Alkire B, Lee DJ, Gray ST, et al. Impact of resident surgeons on procedure length based on common pediatric otolaryngology cases: case length in pediatric cases. *Laryngoscope*. (2015) 125:991–7. doi: 10.1002/lary.24912

19. Lascelles BD, Thomson MJ, Dernell WS, Straw RC, Lafferty M, Withrow SJ. Combined dorsolateral and intraoral approach for the resection of tumors of the maxilla in the dog. *J Am Anim Hosp Assoc*. (2003) 39:294–305. doi: 10.5326/0390294

20. Ritacco LE, Milano FE, Farfalli GL, Ayerza MA, Luis Muscolo D, Aponte-Tinao LA. Accuracy of 3-D planning and navigation in bone tumor resection. *Orthopedics*. (2013) 36:e942–50. doi: 10.3928/01477447-20130624-27

21. Cartiaux O, Docquier P-L, Paul L, Francq BG, Cornu OH, Delloye C, et al. Surgical inaccuracy of tumor resection and reconstruction within the pelvis: an experimental study. *Acta Orthop*. (2008) 79:695–702. doi: 10.1080/17453670810016731

22. Ciocca L, Marchetti C, Mazzoni S, Baldissara P, Gatto MRA, Cipriani R, et al. Accuracy of fibular sectioning and insertion into a rapid-prototyped bone plate, for mandibular reconstruction using CAD-CAM technology. *J Cranio-Maxillofac Surg*. (2015) 43:28–33. doi: 10.1016/j.jcms.2014.10.005

23. Müller DA, Stutz Y, Vlachopoulos L, Farshad M, Fürnstahl P. The accuracy of three-dimensional planned bone tumor resection using patient-specific instrument. *Cancer Manag Res*. (2020) 12:6533–40. doi: 10.2147/CMAR.S228038

24. Wu C, Deng J-y, Li T, Tan L, Yuan D-c. Combined 3D printed template to guide iliosacral screw insertion for sacral fracture and dislocation: a retrospective analysis. *Orthop Surg*. (2020) 12:241–7. doi: 10.1111/os.12620

25. Bray JP, Kersley A, Downing W, Crosse KR, Worth AJ, House AK, et al. Clinical outcomes of patient-specific porous titanium endoprostheses in dogs with tumors of the mandible, radius, or tibia: 12 cases (2013–2016). *J Am Vet Med Assoc*. (2017) 251:566–79. doi: 10.2460/javma.251.5.566

26. James J, Oblak ML, Linden ARZ, James FMK, Phillips J, Parkes M. Schedule feasibility and workflow for additive manufacturing of titanium plates for cranioplasty in canine skull tumors. *BMC Vet Res*. (2020) 16:180. doi: 10.1186/s12917-020-02343-1

27. Séguin B, Pinard C, Lussier B, Williams D, Griffin L, Podell B, et al. Limb-sparing in dogs using patient-specific, three-dimensional-printed endoprosthesis for distal radial osteosarcoma: a pilot study. *Vet Comp Oncol*. (2020) 18:92–104. doi: 10.1111/vco.12515

28. Timercan A. Personalized 3D-printed endoprostheses for limb sparing in dogs: modeling and *in vitro* testing. *Med Eng Phys*. (2019) 71:17–29. doi: 10.1016/j.medengphy.2019.07.005

29. Bongers JJ, Wilkinson N, Kurihara M, Bridges JP, Baltzer W, Worth AJ. Accuracy of lumbosacral pedicle screw placement in dogs: a novel 3D printed patient-specific drill guide versus freehand technique in novice and expert surgeons. *Vet Comp Orthop Traumatol*. (2022) 35:381–9. doi: 10.1055/s-0042-1750433

30. Hamilton-Bennett SE, Oxley B, Behr S. Accuracy of a patient-specific 3D printed drill guide for placement of cervical transpedicular screws. *Vet Surg*. (2018) 47:236–42. doi: 10.1111/vsu.12734

31. Toni C, Oxley B, Clarke S, Behr S. Accuracy of placement of pedicle screws in the lumbosacral region of dogs using 3D-printed patient-specific drill guides. *Vet Comp Orthop Traumatol*. (2021) 34:053–8. doi: 10.1055/s-0040-1716840

32. Wong KC, Kumta SM, Geel NV, Demol J. One-step reconstruction with a 3D-printed, biomechanically evaluated custom implant after complex pelvic tumor resection. *Comput Aided Surg*. (2015) 20:14–23. doi: 10.3109/10929088.2015.1076039

33. Salisbury SK, Richardson DC, Lantz GC. Partial maxillectomy and premaxillectomy in the treatment of oral neoplasia in the dog and cat. *Vet Surg*. (1986) 15:16–26. doi: 10.1111/j.1532-950X.1986.tb00168.x

34. Tuohy JL, Selmic LE, Worley DR, Ehrhart NP, Withrow SJ. Outcome following curative-intent surgery for oral melanoma in dogs: 70 cases (1998–2011). *J Am Vet Med Assoc*. (2014) 245:1266–73. doi: 10.2460/javma.245.11.1266

35. Giacomo D, Giovanni A, da Silva JV, da Silva AM, Paschoal GH, Cury PR, et al. Accuracy and complications of computer-designed selective laser sintering surgical guides for flapless dental implant placement and immediate definitive prosthesis installation. *J Periodontol*. (2012) 83:410–9. doi: 10.1902/jop.2011.110115

36. Wong K-C, Sze K-Y, Wong IO-L, Wong C-M, Kumta S-M. Patient-specific instrument can achieve same accuracy with less resection time than navigation assistance in periacetabular pelvic tumor surgery: a cadaveric study. *Int J Comput Assist Radiol Surg*. (2016) 11:307–16. doi: 10.1007/s11548-015-1250-x

37. Bongers JJ, Wilkinson N, Kurihara M, Bridges JP, Baltzer W, Worth AJ. Accuracy of lumbosacral pedicle screw placement in dogs: a novel 3D printed patient-specific drill guide versus freehand technique in novice and expert surgeons. *Vet Comp Orthop Traumatol*. (2021) 34:053–8. doi: 10.1055/s-0040-1716840

38. Roh Y, Cho C, Ryu C, Lee J, Jeong S, Lee H. Comparison between novice and experienced surgeons performing corrective osteotomy with patient-specific guides in dogs based on resulting position accuracy. *Vet Sci*. (2021) 8:40. doi: 10.3390/vetsci8030040

39. de Wiele V, Gerlinde WT, Vercruyssen M, Coucke W, Temmerman A, Quirynen M. The accuracy of guided surgery via mucosa-supported stereolithographic surgical templates in the hands of surgeons with little experience. *Clin Oral Implants Res*. (2015) 26:1489–94. doi: 10.1111/clr.12494

40. Kholy KarimEl, Lazarin Rafael, Janner Simone F. M., Faerber Karin, Buser Ramona, Buser Daniel. (2019). "Influence of surgical guide support and implant site location on accuracy of static computer-assisted implant surgery." *Clin Oral Implants Res* 30 : 1067–1075. doi: 10.1111/clr.13520.

41. Rodby KA, Turin S, Jacobs RJ, Cruz JF, Hassid VJ, Kolokythas A, et al. Advances in oncologic head and neck reconstruction: systematic review and future considerations of virtual surgical planning and computer aided design/computer aided modeling. *J Plast Reconstr Aesthet Surg*. (2014) 67:1171–85. doi: 10.1016/j.bjps.2014.04.038

42. Kim J, Ko J, Kim J, Seo A, Eom K. Efficacy of a customized three-dimensional printing surgical guide for tibial plateau leveling osteotomy: a comparison with conventional tibial plateau leveling osteotomy. *Front Vet Sci*. (2021) 8:751908. doi: 10.3389/fvets.2021.751908

43. Carroll KA, Mathews KG. Ligation of the maxillary artery prior to caudal maxillectomy in the dog—a description of the technique, retrospective evaluation of blood loss, and cadaveric evaluation of maxillary artery anatomy. *Front Vet Sci*. (2020) 7:588945. doi: 10.3389/fvets.2020.588945

Frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 27 June 2023
DOI 10.3389/fbioe.2023.1186351

Check for updates

OPEN ACCESS

EDITED BY
Laura Cercenelli,
University of Bologna, Italy

REVIEWED BY
Marie-Noelle Giraud,
Université de Fribourg, Switzerland
Nattapon Chantarapanich,
Kasetsart University, Thailand

*CORRESPONDENCE
Erika Negrello,
e.negrello@smatteo.pv.it

†These authors have contributed equally to this work and share first authorship

RECEIVED 14 March 2023
ACCEPTED 13 June 2023
PUBLISHED 27 June 2023

CITATION
Pisani S, Mauri V, Negrello E, Mauramati S, Alaimo G, Auricchio F, Benazzo M, Dorati R, Genta I, Conti B, Ferretti VV, De Silvestri A, Pietrabissa A and Marconi S (2023), Assessment of different manufacturing techniques for the production of bioartificial scaffolds as soft organ transplant substitutes.
Front. Bioeng. Biotechnol. 11:1186351.
doi: 10.3389/fbioe.2023.1186351

COPYRIGHT
© 2023 Pisani, Mauri, Negrello, Mauramati, Alaimo, Auricchio, Benazzo, Dorati, Genta, Conti, Ferretti, De Silvestri, Pietrabissa and Marconi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Assessment of different manufacturing techniques for the production of bioartificial scaffolds as soft organ transplant substitutes

Silvia Pisani[1†], Valeria Mauri[2†], Erika Negrello[2*], Simone Mauramati[1], Gianluca Alaimo[3], Ferdinando Auricchio[3], Marco Benazzo[1,4], Rossella Dorati[5], Ida Genta[5], Bice Conti[5], Virginia Valeria Ferretti[6], Annalisa De Silvestri[6], Andrea Pietrabissa[2,7] and Stefania Marconi[3,8]

[1]Department of Otorhinolaryngology, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy, [2]SC General Surgery 2, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy, [3]Department of Civil Engineering and Architecture, University of Pavia, Pavia, Italy, [4]Department of Clinical, Surgical, Diagnostic and Pediatric Sciences, University of Pavia, Pavia, Italy, [5]Department of Drug Sciences, University of Pavia, Pavia, Italy, [6]SSD Biostatistica e Clinical Trial Center, Fondazione IRCCS Policlinico San Matteo, Pavia, Italy, [7]Department of Surgery, University of Pavia, Pavia, Italy, [8]Fondazione IRCCS Policlinico San Matteo, Pavia, Italy

Introduction: The problem of organs' shortage for transplantation is widely known: different manufacturing techniques such as Solvent casting, Electrospinning and 3D Printing were considered to produce bioartificial scaffolds for tissue engineering purposes and possible transplantation substitutes. The advantages of manufacturing techniques' combination to develop hybrid scaffolds with increased performing properties was also evaluated.

Methods: Scaffolds were produced using poly-L-lactide-co-caprolactone (PLA-PCL) copolymer and characterized for their morphological, biological, and mechanical features.

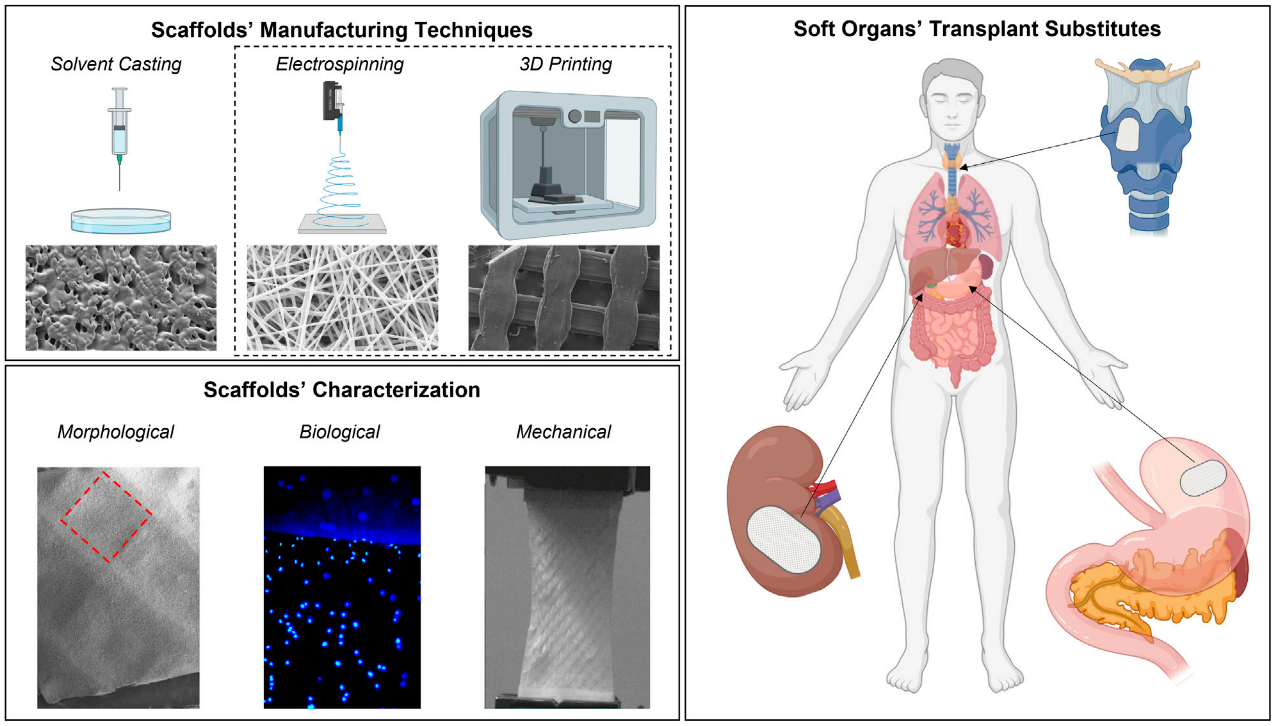
Results: Hybrid scaffolds showed the best properties in terms of viability (>100%) and cell adhesion. Furthermore, their mechanical properties were found to be comparable with the reference values for soft tissues (range 1–10 MPa).

Discussion: The created hybrid scaffolds pave the way for the future development of more complex systems capable of supporting, from a morphological, mechanical, and biological standpoint, the physiological needs of the tissues/organs to be transplanted.

Abbreviations: MTT, 3-(4,5-Dymethiltiazol-2-Y)-2,5 Diphenyltetrazolium Bromide; 3DP, 3D Printing; DAPI, 4′,6-Diamidino-2-Phenylindole); BOs, Bioartificial Organs; BSs, Bioartificial Scaffolds; CAD, Computer Aided Design; DMSO, Dimethyl sulfoxide; ES, Electrospinning; EtOH, Ethanol; ECM, Extracellular Matrix; FFF, Fused Filament Fabrication; GPC, Gel Permeation Chromatography; Tg°, Glass Transition Temperature; HNDF, Human Normal Dermal Fibroblast; HS, Hybrid Scaffolds; L, Length; Mn, Molecular Number; Mw, Molecular Weight; PLA-PCL, Poli-L-Lactide-Co-Caprolactone; PDI, Polydispersity Index; SEM, Scanning Electron Microscopy; SC, Solvent Casting; THF, Tetrahydrofuran; T, Thickness; W, Width; E, Young's Modulus.

KEYWORDS

bioartificial scaffolds, tissue engineering, 3D printing, electrospinning, transplantology, organ transplant, soft tissue regeneration



GRAPHICAL ABSTRACT

1 Introduction

Organ failure is still a leading cause of mortality worldwide, despite advances in surgical and transplantation techniques. The allogeneic transplant is still one of the most used strategies (more than 34,285 solid organ transplants were performed in the EU in 2019) (Vanholder et al., 2021). However, it has a series of problems such as the shortage of organs and donor cross-matching (Beyar, 2011). Indeed, it was reported by The World Health Organization (WHO) that only 10% of the worldwide need for organ transplantation is being met (Keeping Kidneys, 2012). To overcome the scarcity of donors, xenotransplantation techniques are being used (commonly from pigs) (Cooper et al., 2016; Zhang et al., 2019). However, this technique still features many limitations, including the most important one, a severe immune reaction concomitant with coagulation dysregulation and chronic inflammation; these adverse effects can nullify the transplant surgery and compromise a patient's life (Zhou et al., 2022).

Bioartificial organs (BOs) are emerging as a valid alternative to traditional organ transplantation able to obviate the lack of donors and avoid adverse reactions (Wang, 2019). Essentially, a BO is an engineered three-dimensional (3D) scaffold implanted or integrated into the human body, able to interact with surrounding living tissues with the aim of replacing an organ, promoting regeneration, and restoring its original functionality (Abu-Faraj, 2012; Ren and Ott, 2014).

The overarching vision is to provide better approaches for the development of bioartificial scaffold (BSs)-guided tissues to be applied to personalized therapy, organ replacement, or reconstruction as well as to *ex vivo* screening among treatment options for different human diseases (Oksdath et al., 2018). For BOs development, three-dimensional scaffolds alone are not enough for a correct and complete replacement of the damaged tissues, but it is preferred as the co-presence of suitable cells is able to make the repair process faster and more effective. In fact, the BSs geometry and architecture must mimic physiological tissue/organ and promote proper cell adhesion and proliferation to guarantee tissue/organ integration and functionality (Causa et al., 2007). As an example, fibroblast cells are widely used in tissue engineering thanks to their heterogeneous presence in numerous tissues. Depending on the anatomical region of origin, it has been demonstrated that fibroblasts have different gene expression profiles (Wong et al., 2007). Among them, dermal fibroblasts were demonstrated to have several functions such as increasing ECM components synthesis and deposition to promote proliferation and migration in response to stimuli (cytokines) and exhibiting autocrine and paracrine interaction (Stunova and Vistejnova, 2018). Therefore, the harmony and synergy between the 3D scaffold and the sown cells are fundamental to guarantee suitable chemical-physical, mechanical, and biological characteristics to mimic the physiological properties of the tissue to be replaced (Sultana, 2018; Okamoto, 2019). Crucial aspects must be considered to provide the right support for cell adhesion, growth, and/or differentiation on BSs; these include the choice of the polymer, the manufacturing technique, the geometry, and the consequent scaffolds' mechanical properties (Bacakova et al., 2011; Eltom et al., 2019). As for polymers, it is essential to choose biocompatible and non-toxic polymers such as Poly-lactic acid (PLA), Poly-caprolactone (PCL), Poly-lactic-co-glycolic acid (PLGA), or polyurethane (Pedersen et al., 2022). Many of these polymers are also biodegradable, therefore ensuring the tissue reformation and avoiding the generation of toxic degradation products; the main advantage of these polymers is that they do not have to be surgically removed once the tissue is grown and therefore do not require further interventions for the patient (Arif et al., 2019; Nikolova and Chavali, 2019).

Scaffold geometry provides biophysical signals that trigger cells' nucleus response (i.e., the regulation of gene expression) and modulates cells' behavior and functionality (Han et al., 2022). Geometrical parameters are also crucial to mimic the characteristics of the tissue/organ to be replaced and to improve cell proliferation: different cellular lines require different scaffold geometries and porosity percentages according to the function they have to perform (Chantarapanich et al., 2012; Loh and Choong, 2013).

Various techniques for the BSs production have been implemented over the years, starting from the simplest ones-such as solvent casting (SC)-up to the more advanced ones, such as electrospinning (ES) and 3D printing (3DP) (Murphy and Mikos, 2007; Wang, 2019; Morelli et al., 2022; Pien et al., 2022). SC technique is a scaffold manufacturing method that begins with the dissolution of a polymer in an organic solvent which afterwards hardens by exploiting solvent evaporation; a dry solid polymer scaffold with a porous network is left behind, although it is difficult to handle the desired pore shape and pore inter-connectivity of these types of scaffolds (Johnson et al., 2010; Dorati et al., 2017). ES is a fibers' fabrication technique which allows to obtain scaffolds consisting of a network of submicrometric-sized fibers. The collected electrospun fibers exhibit suitable properties for biomedical applications such as high surface area-to-volume ratio, small pore size, high porosity, and a structure able to mimic ECM (Pisani et al., 2018). More recently, 3DP technologies have emerged as versatile techniques to produce components to be used for a wide range of purposes such as anatomical models, surgical planning, training and simulation, medical devices, prosthetics prototyping, and regenerative medicine (Conti and Marconi, 2019). With Computer-Aided Design (CAD) programs, it is possible to define the geometry, the infill percentage, and the structure porosity according to 3D printers' resolution limits.

The purpose of the present work was to produce and characterize BSs manufactured through three different manufacturing techniques (SC, EL, and 3DP). Moreover, the advantages of combining EL and 3DP for the development of hybrid scaffolds (HS) useful for soft tissue replacement was evaluated. The combination of ES and 3DP technologies is already present with some examples in the literature. The important aspect highlighted in this work is that for the same used base material, obtained scaffolds have different properties due precisely to the manufacturing technique. Therefore, one of the main goals is to underline how the production technique has a strong influence on the final scaffolds' properties (both mechanical and biological).

TABLE 1 Main slicing parameters according to Ultimaker Cura nomenclature.

Slicing parameter (ultimaker cura)	Value
Layer Height	0.20 mm
Extrusion Width	0.35 mm
Wall Thickness	0.80 mm
Top Layers	4
Bottom Layers	4
Top/Bottom Pattern	lines
Infill Pattern	grid
Infill Overlap Percentage	10%
Flowrate	100%
Printing Temperature	200°C
Printing Temperature (Initial Layer)	200°C
Build Plate Temperature	80°C
Build Plate Temperature (Initial Layer)	80°C
Infill Speed	60 mm/s
Wall Speed	30 mm/s
Outer Wall Speed	30 mm/s
Inner Wall Speed	60 mm/s
Top/Bottom Speed	30 mm/s
Enable Retraction	True
Retract at Layer Change	False
Retraction Distance	6.50 mm
Retraction Speed	25 mm/s
Build Plate Adhesion Type	skirt
Brim Width	8 mm
Fan Speed	100%

BSs were made up by the same biopolymer (poli-L-lactide-co-caprolactone—PLA-PCL 70:30 ratio) that showed PLA prevalence with a suitable degradation rate (about 8 weeks *in vivo*) commensurate with the rate of tissue regeneration, while the PCL component gives plasticity. Moreover, PLA-PCL was chosen for its known bio-properties, such as biocompatibility and biodegradability, and because it has already been widely characterized in previous works by the authors, demonstrating excellent properties of soft tissue regeneration (Pisani et al., 2018; Pisani et al., 2020; Pisani et al., 2021). In the literature, the combination of Electrospinning (ES) and Fused Filament Fabrication (FFF) to produce copolymer PLA-PCL scaffolds usually refers to a hybrid scaffold made of a layer of 3D-printed PLA and a layer of electrospun PCL (or *vice versa*) (Pensa et al., 2019; Smith and Mele, 2021); on the contrary, here HS where both

ES and 3DP layers are made with the same PLA-PCL copolymer are proposed. The resulting BSs were compared in terms of morphological, mechanical, and biological properties and the influence of manufacturing technique on the final features was evaluated and compared. This proof of concept lays the foundations for the subsequent development of complete bioartificial soft body organs with the aim of providing replacement models for transplantation.

TABLE 2 Physical dimensions of patches produced with the different techniques.

	SC	ES	3DP	HS
L [mm]	25 ± 2	25 ± 1	20 ± 0.2	40 ± 2
T [mm]	2 ± 1	0.15 ± 0.03	0.4 ± 0.1	0.6 ± 0.2
W [mm]	15 ± 2	6 ± 1	10 ± 0.2	15 ± 0.5

2 Materials and methods

2.1 Scaffold manufacturing techniques

The SC manufacturing technique was used to prepare PLA-PCL (Resomer LC 703 S—Mw 160,000 Da-Evonik Nutrition & Care GmbH, Evonik Industries AG Relling-hauser Straße 1–11, 45128 Essen, Germany) film-scaffold. The copolymer was solubilized in 1-4 dioxane since this organic solvent has a boiling point of +101°C and a melting point of +12°C. These features limit the evaporation of the solvent during the preparation phase, ensure complete freezing of the solvent during the cooling phase of the system, and guarantee its complete sublimation during freeze-drying, facilitating the removal of the solvent from the finished product. The polymeric solution (10% w/v) was placed in a glass vial and subjected to magnetic stirring (1 h at RT) until completely dissolved. Once a solution was obtained, it was sonicated for 15 min (Sonica, Ultrasonic Cleaner. Soltec, Italy) to remove any air bubbles, which in the subsequent dripping phase could give rise to the formation of morphological heterogeneities, affecting the scaffold formation and reproducibility. Using a glass syringe (Gastight Syringes®/1 mL, Model 1001 LT SYR, Hamilton), 800 μL of the PLA-PCL solution in 1,4-dioxane was dropped into a Teflon mold 5 × 5 cm. After the dripping phase, the mold was frozen at −25°C for 5 h, followed by the freeze-drying process (Lyophilizer Lio-5P, Cinquepascal, Italy) at −48°C and 0.4 mbar for 12 h, to eliminate all the solvent residues and favor the formation of the dry polymeric film.

The same PLA-PCL copolymer was used to produce electrospun nanofibers. A 20% w/v polymeric solution in Methylene Chloride (MC) and N, N-Dimethylformamide (DMF) solvent blend 70: 30 ratio was electrospun following parameters optimized in previous works (Pisani et al., 2018; Conti and Marconi, 2019). The setup parameters were voltage (30 kV), flowrate (0.5 ml/h), needle-collector distance (15 cm), temperature (25 ± 3°C), and relative humidity (RH% = 30 ± 4%). Electrospinning apparatus NANON-01A equipped with a dehumidifier (MEEC instruments, MP, Pioltello, Italy) was used to produce electrospun polymeric fibers. PLA-PCL fibers were collected after 30 min spinning time using a metal plate collector (14 × 25 cm).

3DP scaffolds were produced using a PLA-PCL (70:30) copolymer filament (TreeD Filaments® company) through Fused Filament Fabrication (FFF) 3D printing technology. The filament had a calibrated diameter of 1.75 mm and was extruded - without any non-natural additive—using a LeapFrog HS® (Dutch LeapFrog Group). The printer was equipped with two extruders with nozzles of diameter 0.35 mm and a heated bed - with a printing area of 230 × 270 × 200 mm—that can reach a temperature of 90°C. The two extruders were equipped with a resistance heating them up to 250°C. The printing head was also equipped with two fans to cool the filament right before the extrusion chamber, to keep it rigid and able to push the heated polymer out of the nozzle.

The printer was instructed on how to deploy the material through a specific set of instructions, formatted according to the G-code language, which included information on the path, temperatures of the nozzles and the bed, printing speed, flow rate, etc. The G-code was computed thanks to the so-called slicing software, which started from virtual 3D geometry to be produced and generated the set of instructions to drive the production. In the present work, Ultimaker Cura 4.11.0 software was employed to generate the final G-code. To set-up the optimal printing strategy for the PLA-PCL copolymer, a preliminary investigation of printing parameters was required. On the one hand, PLA filament is one of the most common materials used in FFF processes and its printing parameters are therefore already well-known. On the other hand, little information is available in the literature about PCL printing parameters (printing temperatures (Darling and Sun, 2004), extrusion rates, and requirement of a ventilation system cooling off specimens during printing (Temple et al., 2014)), being less employed in FFF machine. Usually the two filaments (PLA and PCL) are printed separately, from two different nozzles, which are then combined layer-by-layer on the printing bed (Cheng et al., 2021; Espinosa and Moroni, 2021). In this work, the involved filament is unique and already includes both materials. PLA and PCL FFF printing parameters are widely known; on the contrary, the copolymer filament PLA-PCL 70:30 printing parameters had to be deeply studied and optimized for the specific purposes. Moreover, the production of a copolymer PLA-PCL filament is not a trivial task, due to the difficulty of producing a calibrated and regular filament with a diameter of 1.75 mm, starting from separate PLA and PCL granules. A change in PLA-PCL ratio would lead to different printing parameters, together with the possible use of various additives to facilitate the FFF scaffolds' production. The use of a copolymer filament allowed us to characterize—from the biological and mechanical viewpoints - patches made of the exact same material but produced through both standard (SC and ES) and innovative (3DP) manufacturing techniques.

Accordingly, preliminary tests were conducted to define optimal PLA-PCL copolymer printing parameters, starting from basic parameters—such as bed and extrusion temperatures—and then more advanced options—such as speeds, flow rate, extrusion width, retraction speed, and length. Different printing parameters could influence the mechanical properties of the final FFF sample (Chaitat et al., 2022); considering the specific involved cells and patches' geometry, the infill percentage parameter was considered as one of the most crucial aspects that could affect the final patches' biological and mechanical properties and was involved for further investigation and analyses. Rectangular 3DP samples (10 × 20 × 0.4–0.5 mm) with three different infill percentages, namely, 20%, 50%, and 100%, were then produced. The main printing parameters for PLA-PCL patches production are listed in Table 1.

After a preliminary testing on SC, ES, and 3DP scaffolds respectively, the obtained results showed that ES and 3DP techniques allowed the production of the most promising scaffolds. Consequently, the combination of these two technologies was also considered and tested for HS manufacturing.

HS were aimed to combine the advantages of the two production techniques: on the one hand, 3DP produces very complex and defined geometries and has high process reproducibility, while the electrospun nanofibers allow to obtain a network resembling the native extracellular matrix, thus ensuring better cell adhesion and proliferation. Using the same parameters cited above, 20% w/v PLA-PCL fibers were electrospun for 30 min on 3DP PLA-PCL scaffolds (20%, 50% and 100% infill) to achieve a composite HS. The number of fibers deposited on 3DP scaffolds was evaluated gravimetrically on the dry scaffolds using analytical balance (BCE64-1S, Sartorius AG, Göttingen German).

2.2 Morphological analysis

Zeiss EVO MA10 apparatus (Carl Zeiss, Oberkochen, Germany) was used to perform SEM analyses on polymeric SC, ES, 3DP, and HS scaffolds to evaluate electrospun fibers' dimensions and scaffolds porosity (%) per unit surface area. Analyses were performed at the following magnifications: 1.00 kX and 2.50 kX. The resulting images were processed by ImageJ software (an open-source image processing program designed for scientific multidimensional images) (Hotaling et al., 2015).

2.3 Gel permeation chromatography (GPC)

The chromatograph system consists of a guard column (Phenogel 10E 4˚A µm, 300 × 7.8 mm, Phenomenex, Milan, Italy) and three connected Ultrastyragel columns (7.7 × 250 mm each one with pores diameters of 104 ˚A, 103 ˚A and 500 ˚A), a pump (Varian 9,010), an infrared (IR) detector (Prostra 355 RI; Varian), and software used to process the data relating to MW (Galaxie Workstation, ver.1.8 Single-Instrument, Varian). Samples for GPC analysis were prepared by dissolving scaffolds (SC, ES and 3DP) in tetrahydrofuran (THF) at a concentration of 1–2 mg/mL. THF solutions were filtered through a 0.45 mm filter (Millipore, Massachusset, United States) and injected in the GPC apparatus; the GPC eluent was tetrahydrofuran at a flow rate of 1 mL/min. The resulting data of molecular weight (MW), molecular number (Mn), and polydispersity index (PDI) are expressed as an average of five parallel samples.

2.4 Mechanical characterization

Scaffolds' mechanical characterization was performed through uniaxial tensile tests using an MTS Insight Testing Systems (MTS System Corporation). Tests were carried out at 10 mm/min crosshead speed using a 250N load cell. Testing parameters were selected according to the ASTM D638 standard, considered the most suitable for the materials under examination. Geometrical parameters—namely, specimens' length (L), thickness (T), and width (W)—of the different types of scaffolds are reported in Table 2.

Scaffolds' stiffness was retrieved through linear interpolation of the initial part of stress-strain curves, using least square method (LSM) defined as the Young's modulus (E). Reliability was assessed computing the maximum R2 factor of the linear interpolation of σ-ε curves for each sample (≥0.95).

Given the different size and thickness of the samples made with different techniques, results of mechanical characterization have been normalized to obtain comparable data.

2.5 Cell seeding and viability

Scaffolds' cytocompatibility and cell adhesion were evaluated for SC, ES, 3DP, and HS 48 h after cells sowing, as defined by standard ISO 10993-5 (ISO 10993-5). Human Normal Dermal Fibroblast (HNDF) passage 5 (P.5) were used (100.000 cell/scaffold) for scaffolds cellularization. Scaffolds were fixed in CellCrown™ (CellCrown™ 12 NX Scaffdex, Tampere, Finland) supports, suitable for 12-multiwell, and sterilized before cell seeding. CellCrown™ was used to avoid samples floating in the cell culture medium which can cause cells to fall off the scaffold during cellularization. Scaffolds' sanitization was performed by dipping scaffolds in an 85% v/v EtOH solution for 20 min followed by 15 min in a 70% v/v EtOH solution. Scaffolds were then washed with sterile PBS supplemented with 2% penicillin/streptomycin (P/S) and left under a laminar hood with UV irradiation overnight. A protocol for scaffolds' sterilization was already used in previous work demonstrating that no scaffolds were damaged (Pisani et al., 2021).

Engineered scaffolds were maintained in cell culture medium (DMEM 10%v/v Fetal Bovine Serum) at 37°C and 5% CO_2 for 48 h before analysis. The cell viability percentage was evaluated using underwent 3-(4,5-dymethiltiazol-2-y)-2,5 diphenyltetrazolium bromide (MTT) (Merck KGaA, Darmstadt, Germany) test. Cell culture media was removed from samples and the engineered scaffold was washed twice with PBS (pH 7.4). Subsequently, 300 µL of MTT solution (5 mg/mL in PBS) was added to each scaffold sample and fresh PBS was added to guarantee the total immersion of the samples. The scaffolds were maintained for 2 h and 30 min of incubation at 37°C and 5% of CO_2, and then MTT solution was removed paying attention not to suck precipitated formazan salts.

The engineered scaffolds were withdrawn and transferred into a vial to be solubilized with 1 mL THF under magnetic stirring for 1 h to completely dissolve the polymeric matrix and lyse cellular membranes of cells adhered that gave a purple coloring proportional to cell viability. Instead, 12-multiwell bottoms (from

TABLE 3 SC, EL, 3DP, and HS scaffold information obtained after the manufacturing process.

Scaffold manufacturing technique	PLA-PCL 70:30	Production time	Surface porosity	Pore surface area
	[mg]	[min]	[%]	[mm^2]
SC	80 ± 5.32	1440	0.24 ± 0.05	0.00092 ± 0.00015
ES	50 ± 3.23	30	3.94 ± 1.12	0.00472 ± 0.0021
3DP 20% infill	120 ± 0.25	5	38.41 ± 4.45	2.3 ± 0.3
3DP 50% infill	230 ± 0.15	7	11.7 ± 2.42	0.2 ± 0.03
3DP 100% infill	350 ± 3.30	15	-	-
HS 20% infill	150 ± 0.41	5 + 30	4.14 ± 1.06	0.085 ± 0.1
HS 50% infill	250 ± 5.05	7 + 30	6.92 ± 1.02	0.069 ± 0.04
HS 100% infill	380 ± 1.50	15 + 30	8.28 ± 1.31	0.1399 ± 0.1

where scaffolds were withdrawn) were treated with 1 mL DMSO and left under soaking for 1 h to lyse cellular membranes of cells that did not adhere to scaffolds. HNDF (100.000 cell/bottom) was used as control, treated with MTT solution for 2 h and 30 min, washed, and then solubilized with 1 mL DMSO. The solutions obtained by dissolution of the engineered scaffolds, from 12-multiwell bottoms and HNDF control, were spectrophotometrically analyzed at 570 nm wavelength in a quartz cell (6705 UV/Vis Spectrophotometer—Single cell holder JENWAY) to obtain Abs values and correlate them to cell viability.

DAPI (4′,6-diamidino-2-phenylindole) (Thermo Fisher Scientific Inc., Waltham, Massachusetts, US) is a fluorescent nuclear staining useful for highlighting the nuclei of the cells and detect cells on the scaffold surface. After 48 h incubation, scaffolds were washed with PBS and 0,4% v/v glutaraldehyde solution was added and left for 10 min to fix the adhered cells. Scaffolds were washed twice with PBS and TRITON X 0,1% v/v was added and left for 10 min in order to permeabilize the cell membranes. A further washing step with PBS was performed and scaffolds were treated with DAPI 300 nM solution over-night at 4°C. The samples were analyzed with fluorescence microscope (Leica DM IL LED with ebq 50 ac-L) and images obtained were processed by ImageJ software.

Cell visualization on scaffolds' surfaces was performed using SEM analysis. After 48 h incubation, scaffolds were washed with PBS and dehydrated using ethanol (EtOH) solutions at increasing concentrations (30% v/v, 70% v/v, 80% v/v, and 100% v/v), for 10 min each passage. Samples were then washed with a 50: 50 mixture of dry EtOH 100% and hexamethyldisilane (HDMS) for 20 min. Before SEM analysis, samples were left under a laminar hood to remove solvent residues and then made conductive with surface deposition of a thin layer of gold in an Argon atmosphere. SEM images were further processed by ImageJ software to characterize engineered scaffolds.

2.6 Statistical analysis

The results collected in this experimental study were presented as mean, standard deviation (SD), and 95% Confidence Interval (95%CI). All experiments were carried out in triplicate, unless otherwise stated. To take into account the correlation between technical triplicates (non-independence between observations), we fitted clustered regression models, using robust Sandwich estimator for the calculation of standard errors. The comparison of polymer molecular weight and molecular number for SC, ES, and 3DP techniques compared to the raw materials was evaluated by t-test for paired data.

Statistical analyses were performed using Stata 17 (StataCorp. 2021. Stata Statistical Software: Release 17. College Station, TX: StataCorp LLC). A p-value less than 0.05 was considered statistically significant.

3 Results

3.1 Morphological analysis

Results in terms of scaffolds' weight, manufacturing time, and surface porosity resulting from the three different manufacturing techniques are reported in Table 3.

SC scaffolds showed a lower porosity compared to the ES ones, with an average pore size of 0.92 ± 0.15 μm^2 and 4.72 ± 2.1 μm^2 respectively.

To calculate the porosity of the 3DP scaffolds, the area of the square voids left by the filaments' deposition was computed (Figures 1C,D). The porosity and average pore size progressively decreased from the lowest to the highest infill percentage; 3DP 100% infill resulted full of printed materials, i.e., 0 porosity. The single 3DP filament composing the 3DP scaffolds was in all cases pore-free and smooth.

HSs showed homogeneous fibers covering on 3DP scaffolds with an average weight increase of 25 ± 4.62 mg compared to the respective 3DP scaffold. The weight increase was comparable for all HS taking as reference their corresponding 3DP scaffolds. As concerns scaffolds' porosity percentage and pore surface area, ES layer was considered.

SEM images were obtained for all scaffolds. SC scaffolds (Figure 1A) have shown undefined geometry and rough surface with some pores of variable size and random position generated by solvent sublimation during the manufacturing process. ES scaffolds (Figure 1B) are characterized by the

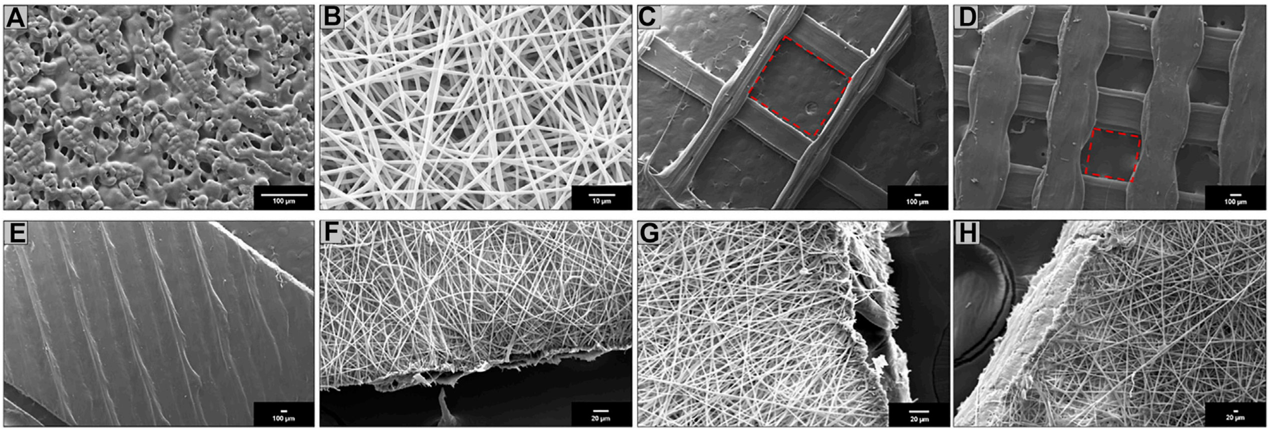


FIGURE 1
Scanning electron microscopy (SEM) of **(A)** SC; **(B)** ES; **(C)** 3DP 20% infill; **(D)** 3DP 50% infill; **(E)** 3DP 100% infill; **(F)** HS 20% infill; **(G)** HS 50% infill; and **(H)** HS 100% infill PLA-PCL 70:30 scaffolds.

presence of homogeneous nanofibers (745 ± 23 nm) forming an interconnected network with a high surface area and porosity. 3DP scaffolds obtained with three different infill, 20% (Figure 1C), 50% (Figure 1D), and 100% (Figure 1E), have shown polymeric filaments with a full, smooth, and non-porous structure; the porosity of the scaffold is controlled by the distance between the deposited filaments.

HS showed complete coverage and adhesion between the 3D scaffolds and electrospun fibers at all infills: 20% (Figure 1F), 50% (Figure 1G), and 100% (Figure 1H). Moreover, fibers electrospun on the 3DP scaffold maintained nanometer size range (775 ± 31 nm).

3.2 Gel permeation chromatography (GPC)

Results obtained shown how all scaffold production techniques lead to a reduction in polymer molecular weight compared to the raw materials (Figure 2). In the case of SC and ES techniques, PLA-PCL powder was solubilized in solvents which were then evaporated to obtain the dry scaffolds, while in the case of 3DP the PLA-PCL wire was softened with heat ($T > Tg^{\circ}$) to be extruded.

In all cases, between the raw material and the respective scaffold obtained there was a reduction in the molecular weight of 26% ± 3.7%.

3.3 Mechanical characterization

Results obtained by uniaxial tensile tests were reported in Figures 3, 4. ES exhibited E = 6.21 ± 3.45 MPa while 3DP scaffolds showed higher values (E = 20.90 ± 6.11 MPa and E = 77.26 ± 23.58 MPa for the 20% infill and 50% infill scaffolds respectively). SC scaffolds showed lower E values (E = 4.56 ± 1.47 MPa), indicating that they are less stiff with respect to the previous ones.

HS 20% infill and HS 50% infill scaffolds showed values of E = 7.39 ± 3.85 MPa and E = 34.77 ± 6.37 MPa respectively. Stress-strain curve for 3DP 100% and HS 100% infill was not reported in the graph because their E values resulted out of scale if compared to the other ones (E > 100 MPa). These scaffolds resulted too rigid and stiff and – for these reasons – not suitable to mimic soft tissues' mechanical properties.

The Young's modulus values obtained from the tensile tests performed on the different scaffolds shown how even if the same material is used, the production technique influenced by the mechanical properties of the final structure. In fact, the production technique rules the resulting sub-millimeter microstructure (see Figure 1) which governs the mechanical response. All scaffolds have E values that fit within the physiological mechanical strength values reported in the literature for different soft (E ≈ 1 MPa) and hard tissues (E > 10 MPa) (Akhtar et al., 2011; Guimarães et al., 2020). Specifically, SC, ES, 3DP 20% infill, and its hybrid equivalent scaffolds can find their application in muscle and cartilage tissue regeneration. On the other hand, the 3DP 50% and 100% infill scaffolds are more suitable for bone regeneration, where much greater mechanical stiffness and strength is required.

3.4 Cell seeding and viability

Results of cells seeding and viability % were reported in Figure 5.

The cell viability test showed good cell viability for cells seeded on SC (83% ± 21%) and ES (127% ± 15%) scaffolds. Almost no cells were found at the bottom of the multi-well, indicating that the scaffolds were able to retain the cells on their surface. The fact that the cell viability value of the ES scaffolds is higher than the control is justified as the adhesion and growth of the cells seeded on the scaffold and on the 2D control (multi-well plate) can be different due to the diversity of the support. In this case, there was better adhesion to the three-dimensional polymeric scaffolds than to the multi-well, justifying slightly higher vitality values. For the 3DP 20% infill and

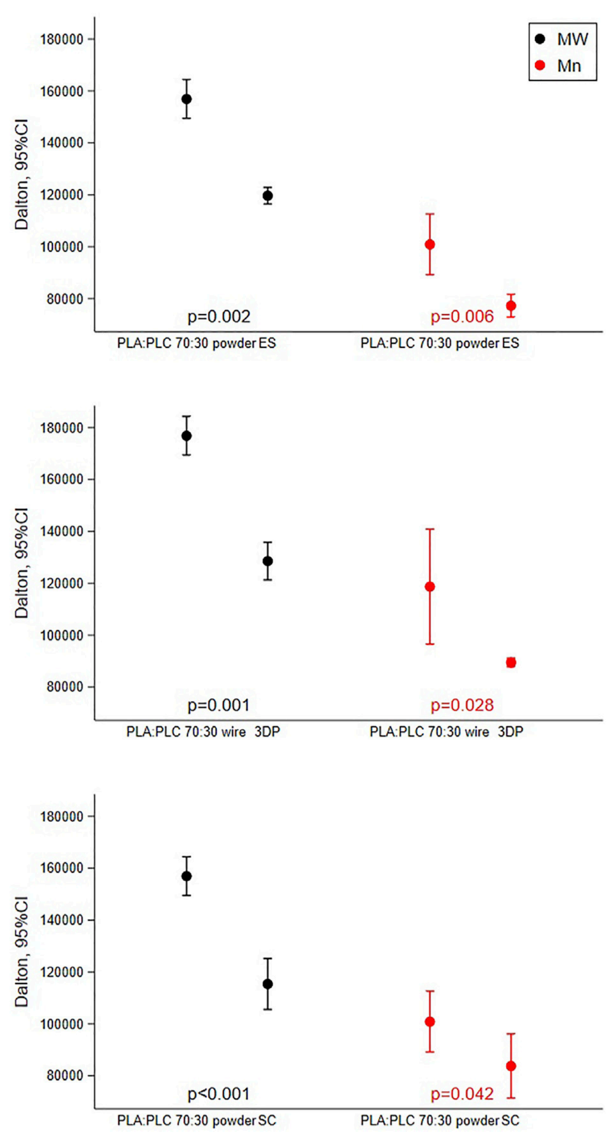


FIGURE 2
Results of GPC analysis reporting mean with 95%CI of molecular weight and molecular number performed on PLA-PCL 70:30 powder, wire, and scaffolds obtained by SC, ES, and 3DP.

50% infill scaffolds, most of the cells remained viable, but deposited at the bottom of the multi-well. This happened because the high porosity created by the geometry of the scaffolds does not retain the cells that fall by gravity and do not adhere to the scaffolds. The 3DP 100% infill scaffold demonstrated another problem; the absence of porosity and the smooth surface do not allow good cell adhesion (Pisani et al., 2020; Pisani et al., 2022). Increase in cell viability and adhesion was achieved in HS that exploit the porous texture of electrospun nanofibers combined to the rigidity of the 3DP scaffold. This is demonstrated to regulate cell behavior since substrate stiffness affects many different processes, such as cell growth, migration, and differentiation (Breuls et al., 2008).

Biological characterization was also performed by fluorescent microscopy through nuclear DAPI staining (Figures 6, 7). Images obtained reflect cell viability data in which a higher number of cells appears on ES scaffolds than on SC or 3DP scaffolds (both 20% and 50% infill). 3DP 100% infill scaffolds show some cells on the surface, but with few clusters. All the HSs showed a high cell density and complete cellularization of the electrospun surfaces with the initial formation of clusters.

Confirmation of the presence of cells on the scaffolds was obtained from SEM images (Figure 8). Also, in this case there are more cells on the EL than on SC scaffolds. Only a few isolated cells were detected on 3DP 20% and 50% infill scaffolds. 3DP 100% infill scaffolds showed few living cells, which tend to detach from the scaffolds because of their smoother surface that avoids correct cell adhesion and proliferation. Otherwise, HSs show good surface cellularization with large spaces to allow further cell proliferation.

4 Discussion

The demonstrated improved performance of HSs produced through the combination of different techniques (EL and 3DP) is a step forward for current investigations. First of all–by analyzing the time-quality balance of the different production methods—it emerges that SC is a very long manufacturing technique requiring more than 24h, and not at all functional for the development of BSs due to scaffolds' poor reproducibility and non-uniformity that compromise final outcomes. On the contrary, ES and 3DP techniques are characterized by short production times and a high process reproducibility. ES produces uniform and dimensionally homogeneous nanofibers which can be employed alone as scaffolds. Moreover, nanofibers offer the possibility to be used as a "functional coating material", exploiting their high surface area over volume ratio to improve cell adhesion (Kishan and Cosgriff-Hernandez, 2017). 3DP technique—based on a defined CAD design—can be modified and adjusted in order to perfectly adapt to the user needs, offering high customization possibilities in the perspective of personalized precision medicine purposes (Vaz and Kumar, 2021). Moreover, important to consider is the scaffolds' resolution that the different technologies can implement. The obtainable minimum thickness of scaffolds with the 3D technique at our disposal has been created to adapt to the values of mechanical properties compatible with soft tissue. On the other hand, the EL scaffolds, which mainly have a biological role in this case, as the nanofibrous texture improves cell adhesion, were electrospun for a suitable time to allow deposition both on the metal collector and on the polymeric 3D-printed scaffold.

Promising results have emerged from the integration of the two techniques: HS with different infill percentages (20% and 50%) show an improved surface porosity than the respective 3DP scaffolds. Particularly, in the case of the 100% 3DP scaffolds with no porosity, the addition of nanofibers leads to a suitable surface porosity able to guarantee cells' adhesion. These values obtained for the HS scaffolds were more suitable to be used as supports for cell cultures and to allow the passage of nutrients/waste (Yadav et al., 2021). From the biological point of view, cell viability studies also demonstrated that HS scaffolds (20%, 50%, and 100% infill) show not only a higher viability than the respective 3DP and ES scaffolds, but also a better ability to retain cells on their surfaces. In fact, 3DP 20% and 50% scaffolds' higher porosity did not allow the cells to adhere to the

10.3389/fbioe.2023.1186351

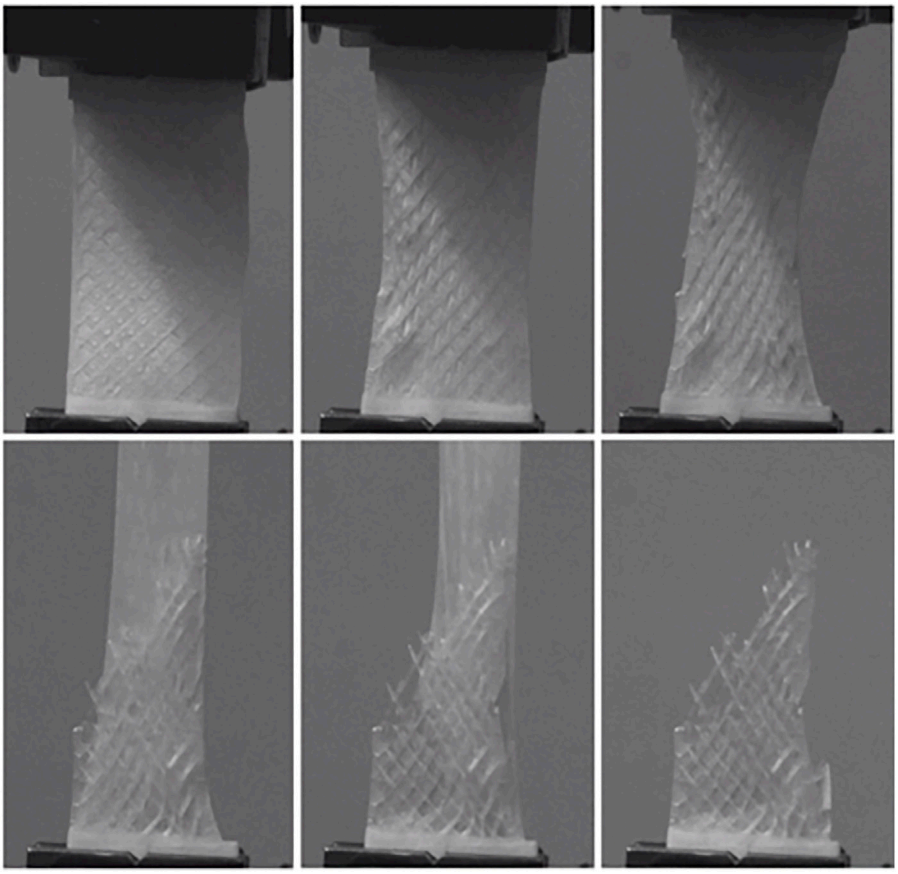


FIGURE 3
Different frames of uniaxial tensile test performed on HS scaffolds (20% infill).

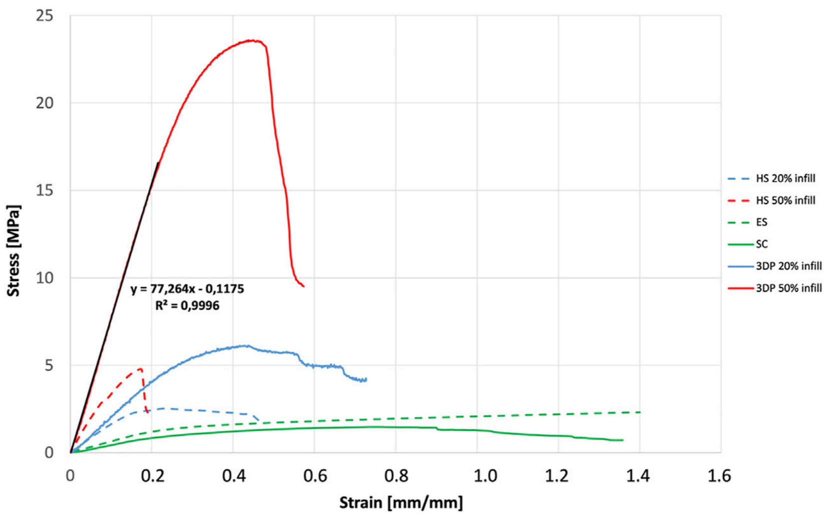


FIGURE 4
Mechanical characterization: stress-strain curves and slope resulting from the mechanical characterization of SC, ES, 3DP, and HS scaffolds.

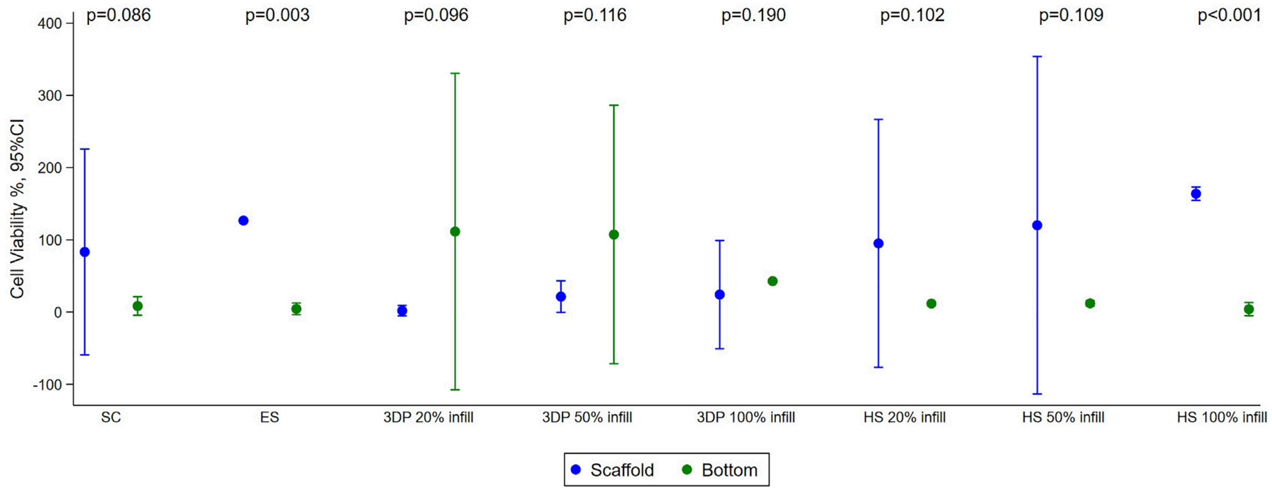


FIGURE 5
Results of cell viability test performed with HNDF (Human normal dermal fibroblast) 48 h after sowing on SC, ES, 3DP, and HS scaffolds. Mean with 95%CI are shown.

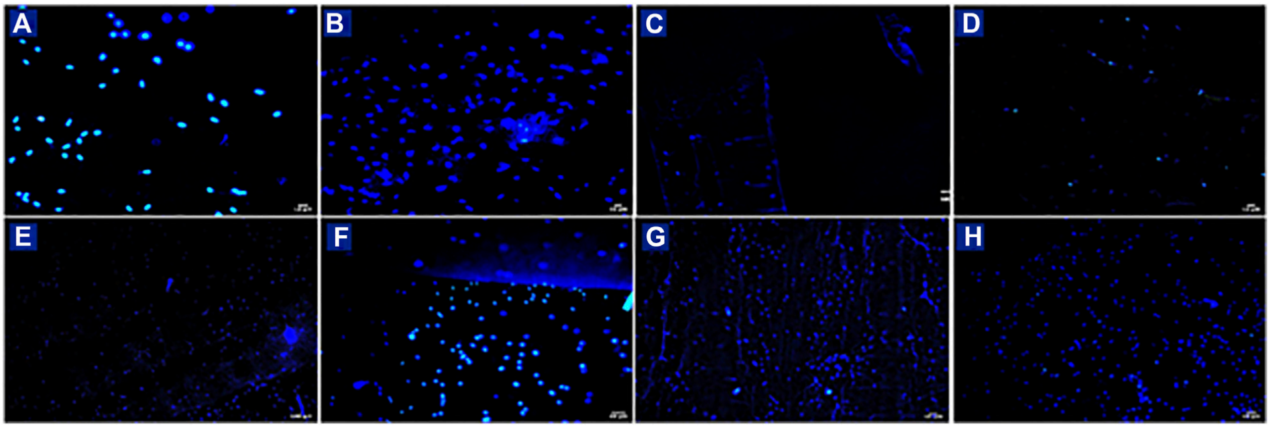


FIGURE 6
DAPI staining of HNDF 48 h after sowing on **(A)** SC; **(B)** ES; **(C)** 3DP 20% infill; **(D)** 3DP 50%; **(E)** 3DP 100% infill; **(F)** HS 20% infill; **(G)** HS 50% infill; and **(H)** HS 100% infill scaffolds.

scaffold surface, making them fall to the multi-well bottom. On the other hand, the lack of porosity of 3DP 100% scaffolds did not guarantee suitable cell adhesion and growth, showing lower cell viability values. Both problems can be overcome by the combination between 3DP and ES, with the ES substrate improving the HS biological surface properties. It is also interesting to note that the performance of ES scaffolds is improved by the integration with 3DP support. It is therefore clear that HS scaffolds offer better biological performance than the ones produced through the two manufacturing techniques individually. Moreover, cell viability on HS scaffolds increased by increasing the infill of the 3DP scaffold. Moreover, when PLA-PCL was used to print the 100% 3DP scaffolds, the mechanical properties resulted were too high to be used as a support for soft tissue replacement and regeneration. Thus, an important aspect on which future research will focus will be to evaluate the use of less stiff materials for the development of high infill 3DP scaffolds for soft tissue regeneration.

From the biomechanical point of view, it was noted that the HS (20% and 50% infill) show lower Young's modulus values (E) than the corresponding 3DP scaffolds: the high stiffness being one of the main problems of 3DP scaffolds, the integration of ES nanofibers offers the possibility to tune this parameter and overcome mechanical limitations. The measured values are compatible with the Young's moduli values of some human soft tissues: in particular,

10.3389/fbioe.2023.1186351

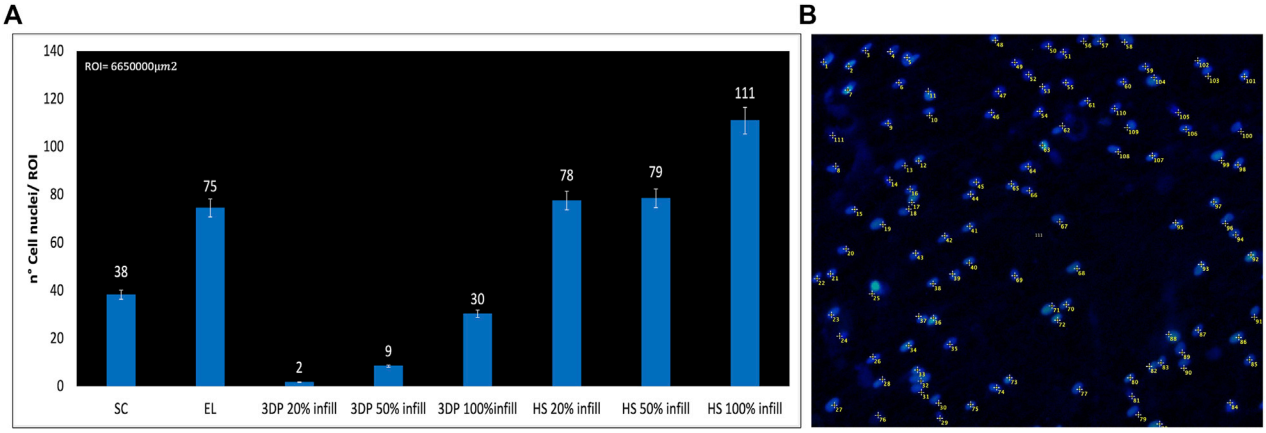


FIGURE 7
(A) Number of DAPI-stained cell nuclei for the region of interest (ROI); **(B)** Example of DAPI nucleus counting on HS 100% infill scaffolds.

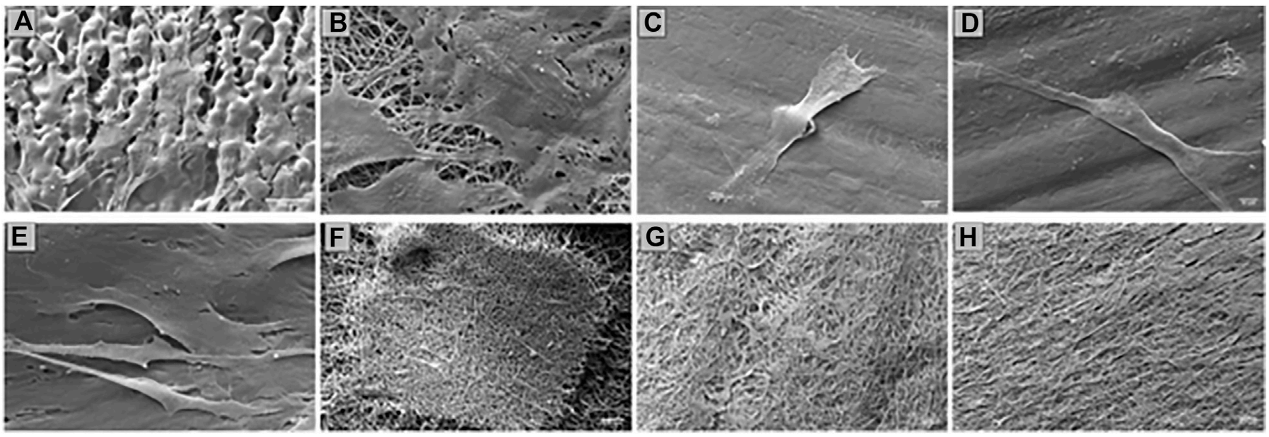


FIGURE 8
SEM of HNDF 48 h after sowing on **(A)** SC; **(B)** ES; **(C)** 3DP 20% infill; **(D)** 3DP 50% infill; **(E)** 3DP 100% infill; **(F)** HS 20% infill; **(G)** HS 50% infill; and **(H)** HS 100% infill scaffolds.

possible applications for the proposed PLA-PCL HS could be the replacement of the aorta (2.0–3.0 MPa), myocardium (2.0–4.0 MPa), nerves (5 MPa), cartilage (5.7–6.2 MPa), and ligaments (25–93 MPa) (Guimarães et al., 2020; Kwon et al., 2020). Moreover, other studies highlighted that, especially in biological materials, mechanical properties are more controlled by the nano-micro-structure. At the nanostructure level of ECM proteins (10^{-9}–10^{-6} m) stiffness is higher (10–10^{3} MPa) compared to the tissue (10^{-5}–10^{-3} m) and organ level (10^{-3}–10 m), which show a stiffness ranging from 10^{-3} and 10^{2} MPa (Akhtar et al., 2011). Because scaffolds also have to maintain mechanical strength and integrity at the nanoscale level unless new tissue regeneration takes place, it is acceptable that Young's modulus of BSs is higher compared to the target tissue/organ (Yadav et al., 2021). However, for future applications, modulation of the E values could be adapted to the physiological values of other tissues/organs, their architecture, and individual needs, and modified according to the use of materials, architecture-geometry, and the presences of selected cells on the scaffold (Sultana, 2018; Pisani et al., 2022).

5 Conclusion

The proposed work stands as a "proof of concept" to validate the integrated use of two of the most innovative and scalable scaffolds' production techniques, namely, electrospinning and 3D printing. Preliminary results obtained from this work defined some crucial aspects that are certainly very interesting for future development of more advanced BOs. HSs can be modulated as a function of i) used materials, ii) 3D construct micro-geometry, and iii) electrospinning nano-topography, to be adaptable to different soft tissue/organ types. BSs are therefore a valid option for further studies and advanced development of BOs could be used as alternatives to organ transplantation.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

Conceptualization, methodology, formal analysis, investigation, writing—original draft, writing—review and editing, and visualization, SP, VM, and SM; methodology and investigation, GA, EN, and SM; methodology, investigation, writing—original draft, and writing—review and editing, RD, IG, BC, and FA; resources, writing—review and editing, supervision, MB and AP; statistical analyses, ADS and VVF; all authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by the Italian Ministry of Health, RC-2021-n.986-rcr2021i-24 "3D-hybrid engineered tubular bioscaffold for esophageal tissue regeneration: from *in vitro* to *in vivo* validation" and by the Italian ministry of Health RC-2019-945-rcr2019i2-17 "Patient-specific 3D-printing-based Simulation Platform for Live-Donor Robotic Nephrectomy and Transplantation".

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abu-Faraj, Z. O. (2012). Handbook of research on biomedical engineering education and advanced bioengineering learning: Interdisciplinary concepts: Interdisciplinary concepts. *Med. Inf. Sci. Reference.* doi:10.4018/978-1-4666-0122-2

Akhtar, R., Sherratt, M. J., Cruickshank, J. K., and Derby, B. (2011). Characterizing the elastic properties of tissues. *Mater. Today* 14 (3), 96–105. doi:10.1016/S1369-7021(11)70059-1

Arif, U., Haider, S., Haider, A., Khan, N., Alghyamah, A. A., Jamila, N., et al. (2019). Biocompatible polymers and their potential biomedical applications: A review. *Curr. Pharm. Des.* 25 (34), 3608–3619. doi:10.2174/1381612825999191011105148

Bacakova, L., Filova, E., Parizek, M., Ruml, T., and Svorcik, V. V. (2011). Modulation of cell adhesion, proliferation and differentiation on materials designed for body implants. *Biotechnol. Adv.* 29 (6), 739–767. doi:10.1016/j.biotechadv.2011.06.004

Beyar, R. (2011). Challenges in organ transplantation. *Rambam Maimonides Med. J.* 2 (2), e0049. doi:10.5041/RMMJ.10049

Breuls, R. G., Jiya, T. U., and Smit, T. H. (2008). Scaffold stiffness influences cell behavior: Opportunities for skeletal tissue engineering. Open Orthop. *J.* 2, 103–109. doi:10.2174/1874325000802010103

Causa, F., Netti, P. A., and Ambrosio, L. (2007). A multi-functional scaffold for tissue regeneration: The need to engineer a tissue analogue. *Biomaterials* 28 (34), 5093–5099. doi:10.1016/j.biomaterials.2007.07.030

Chaitat, S., Chantarapanich, M., and Wanchat, S. (2022). Effects of the 3DP process parameters on mechanical properties of polylactic acid part used for medical purposes. *Rapid Prototyp. J.* 28 (1), 143–160. doi:10.1108/RPJ-01-2021-0014

Chantarapanich, N., Puttawibul, P., Sucharitpwatskul, S., Jeamwatthanachai, P., Inglam, S., and Sitthiseripratip, K. (2012). Scaffold library for tissue engineering: A geometric evaluation. *Comput. Math. Methods Med.*, 2012, 1–14. doi:10.1155/2012/407805

Cheng, C. H., Shie, M. Y., Lai, Y. H., Foo, N. P., Lee, M. J., and Yao, C. H. (2021). Fabrication of 3D printed poly (lactic acid)/polycaprolactone scaffolds using TGF-β1 for promoting bone regeneration. *Polymers* 13, 3731. doi:10.3390/polym13213731

Conti, M., and Marconi, S. (2019). Three-dimensional printing for biomedical applications. *Int. J. Artif. Organs* 42 (10), 537–538. doi:10.1177/0391398819860846

Cooper, D. K., Dou, K. F., Tao, K. S., Yang, Z. X., Tector, A. J., and Ekser, B. (2016). Pig liver xenotransplantation: A review of progress toward the clinic. *Transplantation* 100 (10), 2039–2047. doi:10.1097/TP.0000000000001319

Darling, A. L., and Sun, W. W. (2004). 3D microtomographic characterization of precision extruded poly-ϵ-caprolactone scaffolds. *J. Biomed. Mater. Res. Part B Appl. Biomaterials* 70B (2), 311–317. doi:10.1002/jbm.b.30050

Dorati, R., De Trizio, A., Marconi, S., Ferrara, A., Auricchio, F., Genta, I., et al. (2017). Design of a bioabsorbable multilayered patch for esophagus tissue engineering. *Macromol. Biosci.* 17 (6), 1600426. doi:10.1002/mabi.201600426

Eltom, A., Zhong, G., and Muhammad, A. (2019). Scaffold techniques and designs in tissue engineering functions and purposes: A review. *Adv. Mater. Sci. Eng.*, 2019, 1–13. doi:10.1155/2019/3429527

Espinosa, S. C., and Moroni, L. (2021). Janus 3D printed dynamic scaffolds for nanovibration-driven bone regeneration. *Nat. Commun.* 12, 1031. doi:10.1038/s41467-021-21325-x

Guimarães, C. F., Gasperini, L., Marques, A. P., and Reis, R. L. (2020). The stiffness of living tissues and its implications for tissue engineering. *Nat. Rev. Mater.* 5 (5), 351–370. doi:10.1038/s41578-019-0169-1

Han, P., Gomez, G. A., Duda, G. N., Ivanovski, S., and Poh, P. S. P. (2022). Scaffold geometry modulation of mechanotransduction and its influence on epigenetics. *Acta Biomater.* 2022, 259–274. doi:10.1016/j.actbio.2022.01.020

Hotaling, N. A., Bharti, K., Kriel, H., and Simon, C. G., Jr. (2015). DiameterJ: A validated open-source nanofiber diameter measurement tool. *Biomaterials* 61, 327–338. doi:10.1016/j.biomaterials.2015.05.015

Johnson, T., Bahrampourian, R., Patel, A., and Mequanint, K. K. (2010). Fabrication of highly porous tissue-engineering scaffolds using selective spherical porogens. *Bio-Medical Mater. Eng.* 20, 107–118. doi:10.3233/BME-2010-0621

Keeping kidneys (2012). Keeping kidneys. *Bull. World Health Organ* 90 (10), 718–719. doi:10.2471/BLT.12.021012

Kishan, A. P., and Cosgriff-Hernandez, E. M. (2017). Recent advancements in electrospinning design for tissue engineering applications: A review. *J. Biomed. Mater Res. A* 105 (10), 2892–2905. doi:10.1002/jbm.a.36124

Kwon, J., Ock, J., and Kim, N. N. (2020). Mimicking the mechanical properties of aortic tissue with pattern-embedded 3D printing for a realistic phantom. *Mater. (Basel)* 13 (21), 5042. doi:10.3390/ma13215042

Loh, Q. L., and Choong, C. (2013). Three-dimensional scaffolds for tissue engineering applications: Role of porosity and pore size. *Tissue Eng. Part B Rev.* 19 (6), 485–502. doi:10.1089/ten.TEB.2012.0437

Morelli, S., Piscioneri, A., Salerno, S., and De Bartolo, L. (2022). Hollow fiber and nanofiber membranes in bioartificial liver and neuronal tissue engineering. *Cells Tissues Organs* 211 (4), 447–476. doi:10.1159/000511680

Murphy, M. B., and Mikos, A. G. (2007). "Chapter twenty-two - polymer scaffold fabrication," in *Principles of tissue engineering*. Editors R. Lanza, R. Langer, and J. Vacanti Third Edition (Burlington: Academic Press), 309–321.

Nikolova, M. P., and Chavali, M. S. (2019). Recent advances in biomaterials for 3D scaffolds: A review. *Bioact. Mater* 4, 271–292. doi:10.1016/j.bioactmat.2019.10.005

Okamoto, M. (2019). "The role of scaffolds in tissue engineering," in *Handbook of tissue engineering scaffolds: Volume one*. Editors M. Mozafari, F. Sefat, and A. Atala (Woodhead Publishing), 23–49. 9780081025642.

Oksdath, M., Perrin, S. L., Bardy, C., Hilder, E. F., DeForest, C. A., Arrua, R. D., et al. (2018). Review: Synthetic scaffolds to control the biochemical, mechanical, and geometrical environment of stem cell-derived brain organoids. *Apl. Bioeng.* 2 (4), 041501. doi:10.1063/1.5045124

Pedersen, D. D., Kim, S., and Wagner, W. R. (2022). Biodegradable polyurethane scaffolds in regenerative medicine: Clinical translation review. *J. Biomed. Mater Res. A* 110 (8), 1460–1487. doi:10.1002/jbm.a.37394

Pensa, N. W., Curry, A. S., Bonvallet, P. P., Bellis, N. F., Retting, K. M., Reddy, M. S., et al. (2019). 3D printed mesh reinforcements enhance the mechanical properties of electrospun scaffolds. *Biomaterials Res.* 23 (22), 22. doi:10.1186/s40824-019-0171-0

Pien, N., Palladino, S., Copes, F., Candiani, G., Dubruel, P., Van Vlierberghe, S., et al. (2022). Tubular bioartificial organs: From physiological requirements to fabrication processes and resulting properties. A critical review. *Cells Tissues Organs* 211 (4), 420–446. doi:10.1159/000519207

Pisani, S., Calcaterra, V., Croce, S., Dorati, R., Bruni, G., Genta, I., et al. (2022). Shape memory engineered scaffold (SMES) for potential repair of neural tube defects. *React. Funct. Polym.* 173 (105223), 105223–115148. doi:10.1016/j.reactfunctpolym.2022.105223

Pisani, S., Croce, S., Chiesa, E., Dorati, R., Lenta, E., Genta, I., et al. (2020). Tissue engineered esophageal patch by mesenchymal stromal cells: Optimization of electrospun patch engineering. *Int. J. Mol. Sci.* 21 (5), 1764. doi:10.3390/ijms21051764

Pisani, S., Dorati, R., Conti, B., Modena, T., Bruni, G., and Genta, I. (2018). Design of copolymer PLA-PCL electrospun matrix for biomedical applications. *React. Funct. Polym.* 124, 77–89. doi:10.1016/j.reactfunctpolym.2018.01.011

Pisani, S., Dorati, R., Genta, I., Benazzo, M., Conti, B., and Prina Mello, A. A. (2021). A study focused on macrophages modulation induced by the polymeric electrospun matrices (EL-Ms) for application in tissue regeneration: *In vitro* proof of concept. *Int. J. Pharm.* 603, 120712. doi:10.1016/j.ijpharm.2021.120712

Ren, X., and Ott, H. C. (2014). On the road to bioartificial organs. *Pflugers Arch.* 466 (10), 1847–1857. doi:10.1007/s00424-014-1504-4

Smith, J. A., and Mele, E. (2021). Electrospinning and additive manufacturing: Adding three dimensionality to electrospun scaffolds for tissue engineering. *Front. Bioeng. Biotechnol.* 9, 674738. doi:10.3389/fbioe.2021.674738

Stunova, A., and Vistejnova, L. (2018). Dermal fibroblasts - a heterogeneous population with regulatory function in wound healing. *Cytokine Growth Factor Rev.* 39, 137–150. ISSN 1359-6101. doi:10.1016/j.cytogfr.2018.01.003

Sultana, N. (2018). "Mechanical and biological properties of scaffold materials," in *Functional 3D tissue engineering scaffolds*. Editors Y. Deng and J. Kuiper (Woodhead Publishing), 1–21. 9780081009796.

Temple, J. P., Hutton, D. L., Hung, B. P., Huri, P. Y., Cook, C. A., Kondragunta, R., et al. (2014). Engineering anatomically shaped vascularized bone grafts with hASCs and 3D-printed PCL scaffolds. *J. Biomed. Mater Res. A* 102 (12), 4317–4325. doi:10.1002/jbm.a.35107

Vanholder, R., Domínguez-Gil, B., Busic, M., Cortez-Pinto, H., Craig, J. C., Jager, K. J., et al. (2021). Organ donation and transplantation: A multi-stakeholder call to action. *Nat. Rev. Nephrol.* 17 (8), 554–568. doi:10.1038/s41581-021-00425-3

Vaz, V. M., and Kumar, L. (2021). 3D printing as a promising tool in personalized medicine. *AAPS PharmSciTech* 22 (1), 49. doi:10.1208/s12249-020-01905-8

Wang, X. (2019). Bioartificial organ manufacturing technologies. *Cell Transpl.* 28 (1), 5–17. doi:10.1177/0963689718809918

Wong, T., McGrath, J. A., and Navsaria, H. (2007). The role of fibroblasts in tissue engineering and regeneration. *Br. J. Dermatology* 156, 1149–1155. doi:10.1111/j.1365-2133.2007.07914.x

Yadav, P., Beniwal, G., and Saxena, K. K. (2021). A review on pore and porosity in tissue engineering. *Mater. Today Proc.* 44, 2623–2628. doi:10.1016/j.matpr.2020.12.661

Zhang, X., Li, X., Yang, Z., Tao, K., Wang, Q., Dai, B., et al. (2019). A review of pig liver xenotransplantation: Current problems and recent progress. *Xenotransplantation* 26 (3), e12497. doi:10.1111/xen.12497

Zhou, Q., Li, T., Wang, K., Zhang, Q., Geng, Z., Deng, S., et al. (2022). Current status of xenotransplantation research and the strategies for preventing xenograft rejection. *Front. Immunol.* 13, 928173. doi:10.3389/fimmu.2022.928173

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 04 August 2023
DOI 10.3389/fbioe.2023.1204643

Check for updates

OPEN ACCESS

EDITED BY
Sudheesh K. Shukla,
Shobhit University, India

REVIEWED BY
Giovanni Rubeis,
Karl Landsteiner University of Health Sciences, Austria
Rajiv Dutta,
Shobhit University, India

*CORRESPONDENCE
Julie Urlings,
julie.urlings@radboudumc.nl

†These authors have contributed equally to this work and share senior authorship

RECEIVED 12 April 2023
ACCEPTED 10 July 2023
PUBLISHED 04 August 2023

CITATION
Urlings J, Abma I, Aquarius R, Aalbers M, Bartels R, Maal T, Henssen D and Boogaarts J (2023), Augmented reality—The way forward in patient education for intracranial aneurysms? A qualitative exploration of views, expectations and preferences of patients suffering from an unruptured intracranial aneurysm regarding augmented reality in patient education.
Front. Bioeng. Biotechnol. 11:1204643.
doi: 10.3389/fbioe.2023.1204643

COPYRIGHT
© 2023 Urlings, Abma, Aquarius, Aalbers, Bartels, Maal, Henssen and Boogaarts. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Augmented reality—The way forward in patient education for intracranial aneurysms? A qualitative exploration of views, expectations and preferences of patients suffering from an unruptured intracranial aneurysm regarding augmented reality in patient education

Julie Urlings[1,2,3]*, Inger Abma[4], René Aquarius[1], Marlien Aalbers[1], Ronald Bartels[1], Thomas Maal[2], Dylan Henssen[3†] and Jeroen Boogaarts[1†]

[1]Department of Neurosurgery, Radboud University Medical Centre, Nijmegen, Netherlands, [2]3D Lab Radboudumc, Radboud University Medical Centre, Nijmegen, Netherlands, [3]Department of Medical Imaging, Radboud University Medical Centre, Nijmegen, Netherlands, [4]IQ Healthcare, Radboud Institute of Health Sciences, Radboud University Medical Center, Nijmegen, Netherlands

Objectives: The goal of this project is to explore the views, expectations and preferences of patients with an unruptured intracranial aneurysm regarding the use of AR in patient education.

Methods: To gain an in-depth understanding of the patients' perspective, a face-to-face interview study was conducted using an interview protocol with a predefined topic list. All interviews were audio-recorded and transcribed verbatim afterwards. Transcripts were analyzed using thematic content analyses. Coding was performed using Atlas.ti software.

Results: Seventeen interviews were conducted. The views, expectations and preferences of patients regarding patient education with AR could be subdivided into 15 categories, which could be grouped into 4 general themes: 1) experiences with current patient education, 2) expectations of AR in patient education, 3) opportunities and limitations of AR, and 4) out-of-hospital use of an AR application. Patients' expectations were predominantly positive regarding improving patients' understanding of their medical situation and doctor-patient communication.

Discusssion: This study suggests that patients with unruptured intracranial aneurysms are open to receive patient education regarding their disease with AR. Patients expect that AR models can help patients with intra-cranial aneurysms better understand their disease, treatment options and risks. Additionally, patients expect AR could improve doctor-patient communication.

KEYWORDS

augmented reality, mixed reality, patient education, interview study, qualitative research

1 Introduction

An intracranial aneurysm is a dilation in the wall of an artery of the brain. This increases the chances of the blood vessel rupturing, which can cause severe bleeding. However, aneurysm treatment is associated with possible complications. Therefore, the risk of spontaneous rupture has to be balanced against the risks of procedural complications. Clinical decision making is therefore complicated and for patients with unruptured, often asymptomatic, intracranial aneurysms, the process of information transfer needs to be in-depth and detailed. Physicians have the responsibility to inform patients about the aneurysm itself, the different treatment options (e.g., endovascular coiling, open surgical clipping) and the associated risks. This process of clinical decision making can be challenging. Survey studies reported that there was low agreement between patients and neurosurgeons regarding the "best" treatment option for each individual patient (King et al., 2005; Saito et al., 2012). Furthermore, almost no agreement with regard to the understanding of treatment options and corresponding risks has been reported (Saito et al., 2012) (King et al., 2005). Patients estimated much higher risks of stroke or death from surgical clipping, endovascular embolization, or no intervention compared with the estimates offered by their neurosurgeons (King et al., 2005). These results illustrate that important discrepancies exist between the perceived risks and benefits as estimated by neurosurgeons and those estimated by patients. Patient education with innovative 3D visualization techniques might be able to overcome these discrepancies.

AR is a form of 3D technology which overlays a computer-generated image on a user's view of the real world, providing additional data and context (Barsom et al., 2016). Although AR is already being studied and used in the education of students and residents (Kamphuis et al., 2014; Zhu et al., 2014; Pelargos et al., 2017), its benefits in the context of patient education are mostly unknown (Urlings et al., 2022). Theoretically, the use of AR might add to standard methods of information transfer as AR has the ability to simulate events on top of reality, creating a hybrid immersive learning environment. This could facilitate the development of skills, such as problem solving, critical thinking and communicating (Dunleavy et al., 2008). Additionally, studies on AR for anatomy education suggest that using AR applications decreases cognitive load in students (i.e., the amount of working memory resources used) (Iordache et al., 2012; Di Serio ÁIbáñez and Kloos, 2013; Kucuk et al., 2016; Henssen et al., 2020). In a conference paper of Jakl et al. (2020) participants thought that an AR system as a complementary tool for medical patient education could lead to an improved understanding of the content of a medical consultation. Considering these benefits of AR for students and residents, we hypothesize that AR could also be beneficial for patient education.

However, before implementing AR in the education of unruptured intracranial aneurysm patients, it is important to ask these patients if and how they would like to see AR used in the clinic. Patients' views on the current patient education process and their expectancies of the usage of AR, can provide valuable insights into effective AR implementation. Previous work has shown that including patients through qualitative research can result in identifying facilitators and barriers from the patient perspective, and positive and negative effects of new educational tool usage (van de Belt et al., 2018; Hilt et al., 2020).

This qualitative study aimed to provide more insight in the expectations and wishes of patients suffering from unruptured intracranial aneurysms on AR in patient education. These insights could provide indications for AR in patient education, and thus form a basis for developing a suitable AR application and for further research.

2 Materials and methods

2.1 Study design

To obtain an in-depth understanding of the subject and to explore patients' views, expectations and preferences, a qualitative research design was used. Ethical approval was not required for this type of study under Dutch law, and an exemption was obtained by the local Medical Ethics Committee "CMO Regio Arnhem-Nijmegen" (registration number: 2020-7206). Written informed consent was obtained from all participants. Data were collected using semi structured interviews to obtain a nuanced understanding of the patients' expectations and wishes concerning AR in patient education.

Based on a literature study, an interview protocol was constructed based on a topic list. In collaboration with a radiologist and neurosurgeons and 3D technicians, the topic list was finetuned. The interview protocol provided structure for the interview to ensure that all necessary topics were covered. Interviews were conducted from May 2021 until December 2021.

The interviews were conducted by two trained researchers (JU and/or D.H.). Each participant was interviewed once, at a time and place suitable for the patient. The setting of this interview was informal and was conducted as a conversation. Based on patients' preferences, close relatives attended the interview and were allowed to assist patients during the interviews. The researchers made explicit that they had no involvement in the medical care of the participants.

Interviews were conducted face-tot-face or by use of video call programs due to the COVID-19 pandemic and its subsequent restrictions (e.g., social distancing). All interviews were audio-recorded and transcribed verbatim after conducting the interview. Additional interviews were conducted until data saturation was suspected (i.e., no new topics emerged during the interviews), after which two additional interviews were held to confirm this.

Open-ended questions were used as starting points for the interview. The first part of the interview consisted of three topics: 1) general/demographics, 2) experiences with current patient education, and 3) context (friends and family). Then, to clearly illustrate augmented reality technology to the participants, two

TABLE 1 Timeline of each interview including main interview topics and accompanying interview questions.

Interview topics	Main questions
General/demographics	- Questions on age, work and occupational status
	- Can you tell me about your aneurysm?
	- How were you diagnosed?
	- Which treatment did you receive?
Experiences with current patient education	- How did your physician educate you about your aneurysm?
	- What was your opinion about the information you received from your physician?
	- Did your physician use visualization tools?
	- Did you fully understand the information you received from your physician?
	- What was the most important information you received from your physician?
Context (family and friends)	- Did you talk about your condition with family and/or friends?
First video is shown	
First reaction to AR	- What are your first reactions on AR after watching this video?
	- Did you already have experience with AR?
	- Do you have any questions after watching the first video?
Second video is shown	
First reaction to AR	- What are your first reactions on AR after watching this video?
Expected value of using 3D prints during consultation	- Do you think there are advantages of using AR in patient education?
	- Do you think there are disadvantages of using AR in patient education?
	- Do you think AR could change something about the subjects that are being discussed in the consultation room?
Wishes for AR	- If you look back on the education that you received, do you think you would have liked the education to have been with the use of AR?
	- Besides the features of AR that you saw in the video, are there other things you would like to see in AR?
	- When and where would you like to use AR?
	- On which devise would you like to see AR?
The AR model	- What did you think of the AR model that was used in the video?
Context (friend and family)	- Do you think AR could help you talk about your condition with family and friends?
Additional experiences and needs	- Do you have additional topics you want to discuss?
	- Do you have any questions?

videos were shown during the interviews. These videos can be viewed in Supplementary Material SA, B. The first video gave a general idea on what AR looks like. The second video showed an AR aneurysm model. The interview guide used after watching the first video consisted of the topic 1) first reaction to AR. After watching the second video, the guide consisted of six major topics: 1) first reactions to AR, 2) expected value of using 3D prints during consultation, 3) wishes for AR, 4) the AR model, 5) context (friend and family), 6) additional experiences and needs. The main topics/questions are shown in Table 1. Patients were encouraged to express their own opinions and experiences freely. Clarification was asked regularly to ensure that answers given were understood correctly. When new topics emerged from the interviews, they were added to the topic list.

2.2 Participants and recruitment

Patients were eligible for inclusion if they had a diagnosis for an intracranial aneurysm or they were already treated for their intracranial aneurysm. Patients were identified through the outpatient clinic of the Radboud university medical center, Nijmegen, Netherlands, or the weekly neurovascular interdisciplinary meeting of the first 3 months of 2021 of this hospital. Patients were first approached by a neurosurgeon or nurse practitioner. If patients were interested to participate in this project, their contact details were passed on to the researchers who then provided more information on the study. To seek in-depth information from a wide range of patients, purposive sampling was conducted based on gender, age and educational background. Only patients with a sufficient understanding of the Dutch language were included in the study.

TABLE 2 Patient characteristics.

	Participants (Brown et al., 2019)
Gender	
- Female (n)	11
- Male (n)	6
Median age	65 years (20–75 years)
Occupational status	
- Working (n)	9
- Unable to work (n)	5
- Retired (N)	3
Treatment	
- Minimal invasive (Coiling/surpass flow diverter)	10
- Clipping	1
- Watchfull waiting	5
- Clipping and watchfull waiting	1
Level of education	
- Bachelor's degree	7
- No Bachelor's degree	10

TABLE 3 Themes and categories.

Themes	Categories
Experiences with current patient education	- Received education
	- Asking questions
	- Visualization tools
	- Points of improvement
Expectations of AR in patient education	- Improving patients' understanding
	- Emotional confrontation
Opportunities and limitations of AR	- Wishes used AR model
	- Explaining relate complaints
	- Showing treatment options in AR
	- Patients' needs
	- Necessity
	- Concerns regarding the use of AR
	- Timing
Out- of- hospital use	- AR at home
	- AR for relatives

2.3 Data analysis

The data from these interviews were analysed using thematic content analysis. This was done using Atlas. ti software version 22 Windows (http://atlasti.com; ATLAS.ti Scientific Software Development GmbH, Berlin, Germany). The first interviews were independently coded by two authors (JU and DH). The assigned codes were compared and discussed until consensus was reached on the codes for the codebook. Analysis took place via an inductive iterative process using the constant comparative method. This means that data analysis started after the first interview was completed. As the study continued, codes derived from the previous interview(s) were used as a starting point for coding the next interview, adding additional codes where and whenever needed.

JU and DH also started axial coding in which codes were linked together and combined into categories. This process continued throughout the coding of the subsequent interviews, in which JU analysed the interviews first, adding new codes if needed, after which DH checked the coded interviews for agreement. Together, they adapted the codebook throughout this process and created an overview of the code categories and themes. Finally, the findings were discussed with IA and JB, and the final categories and themes were decided upon.

3 Results

3.1 Patient characteristics

Seventeen patients with an unruptured intracranial aneurysm were interviewed. Patient characteristics are summarized in Table 2.

3.2 Results of the analysis

The patients' expectations and wishes on AR were categorized into 15 categories. These were sorted into 4 themes: 1) experiences with current patient education, 2) expectations of AR in patient education, 3) opportunities and limitations of AR, and 4) out of hospital use of an AR application. Details are summerized in Table 3.

3.2.1 Current patient education

This theme describes experiences patients have with current patient education, which can provide valuable insights into effective AR implementation.

3.2.1.1 Experiences

Participants mentioned that creating clarity and providing assurance was the most important goal of patient education. They indicated that this could be achieved by providing complete, though concise and understandable information. Additionally, it was mentioned that it was very important for patients to be able to trust their physician and that the education is approached positively, i.e., not just focusing on the risks. This is important because emphasizing the risks of an intracranial aneurysm could cause patients to feel overly anxious.

When discussing which information, they valued the most, interviewees reported that detailed information about aneurysm location, size and origin was important to them. The participants described they needed this information to understand the explanation of different treatment options, the advantages and disadvantages, the associated risks and the possible consequences of rupture or treatment failure.

It was expressed by participants that they had forgotten a great amount of information that was given during the first consultation. The reasons that people gave for this memory loss were pre-existent memory problems, being nervous and/or the excessive amount of information given during their first visit. Interviewees told us that they therefore learned most about their aneurysm by reading information provided by their physician or information they found online.

Participants stated that they found it difficult to formulate the right questions during the first consultation due to shock and the amount of information provided at this first consultation. Therefore, most questions arose after their hospital visit.

3.2.1.2. Visualization tools

Two tools were described by the interviewees which were used to help them visualize the aneurysm: radiological data and a sketch made by the physician. A mentioned disadvantage of the radiological data was that participants found it too difficult to understand. They mentioned that they needed a lot of explanation from their physician to fully comprehend what they saw when looking at scans. One patient even stated that imaging was interesting for physicians, but not for patients.

> "A neurologist explained to me that there were at least two aneurysms in there and she also showed them on a scan but well, all you see are a few gray spots. You can't do much with that. (P14)"

Patients expressed that they were more satisfied with the sketches made by the physician. The reasons patients gave for this preference was that these sketches gave them more insight in their diagnosis and possibilities for treatment in a simple way. Additionally, the physician showing a stent when explaining treatment options was well appreciated by patients.

> "They also showed it on the MRI or CT, but that is less clear. It is more difficult to understand. With a sketch, it was just a blood vessel with a ball on top and then he drew some coils in it. (P4)"

3.2.2 Expectations regarding AR

3.2.2.1 Improving patients' understanding

The biggest benefit patients expected from AR was that it could increase patients' understanding of the location and size of their aneurysm. In addition, patients believed AR could be a valuable addition to help explain treatment options and the associated risks. They expected that, especially the attractive visualization used in AR would help patients better understand the information given by their physician. This might also make the information easier to remember. Moreover, patients expected that increased understanding could also lead to patients making more well-educated decisions on their treatment. Participants believed this visualization and the increased understanding that could come with it would make it easier to live at ease with their intracranial aneurysm.

> "Then I can see what they're really talking about. Because I try to understand what they are telling me, but if you really just have a clear image, it's just a lot easier. Then it becomes easier to remember and you can also explain it to others more easily. (P9)"

It was also expressed that seeing an aneurysm in AR would give more insight in possible complaints due to their aneurysm.

> 'Very clear, nice to see. You understand it a little better. I suppose that if I could see this of my own head, then I would know precisely where my aneurysm is located. So, when I feel pain elsewhere, I don't have to be afraid. But then, if I feel pain on the location of the aneurysm, I'd think, oh, could there be something wrong there? (P10)"

With regard to AR in the consultation room, views were that the use of AR would be suitable in a consultation with their physician. Patients differed in their view whether using AR would lead to discussing extra topics during a consultation. One view was that the use of AR would not change what is being discussed in the consultation room. The contrasting view was that, because patients can see exactly what the physician is talking about, AR could lead to patients asking more questions.

3.2.2.2 Emotional confrontation

Before viewing the video of the AR aneurysm model, patients expected that viewing their aneurysm in AR could be emotionally confronting and even scary. However, after watching the AR video, none of the participants expressed feelings of this kind. Patients did not find the AR model scary because it was clear and informative and because it did not look too 'real'. Having a good doctor-patient relationship was expected to make it less scary to watch the AR model, emphasizing the guiding role of the physician.

3.2.3 Opportunities and limitations AR

3.2.3.1 Wishes regarding the used AR model

Generally, patients thought the AR aneurysm model used in this study was clear and that it gave insight in the anatomy and the location of the aneurysm. Participants were happy with an innovation like this in patient education and mentioned they would like to see a personalized AR model. The adaptability of AR was seen as an advantage. Unlike with a sketch, with AR you can easily remove and add different layers of the model, which provides patients with a more detailed explanation. However, several points for improvement and wishes for AR emerged during the interviews. Concerning the current model, patients stated that the aneurysm in the model should be presented larger or more striking. Thereby, it was stated that more details should be incorporated in the AR model, such as cranial nerves, white matter and the thickness of blood vessels. Additionally, it was suggested that a brain should be incorporated in the AR model to make it clearer where the blood vessels are in relation to the brain. Also indicating the front and back of the model could be helpful.

3.2.3.2 Explaining related complaints

When receiving information using AR, patients stated that the model should be used to explain potential causes of the aneurysm and causes of complaints provoked by their aneurysm or perceived treatment. One patient stated that they would like their physician to use AR every consult to update them on the state of their aneurysm.

3.2.3.3 Showing treatment options with AR

Participants stated that they would like to see their treatment options depicted in AR. Some patients stated they would like AR to be used to create more insight in treatment risks, whereas others disclosed that treatment risks should not be visualized in AR because it would be too confronting.

3.2.3.4 Concerns regarding the use of AR

Concerns were expressed regarding the accuracy of AR. Participants were afraid details would get lost when transferring data from a scan to an AR model or that the model will show a distorted reality.

Another problem that was raised was that patients need time to get used to the new technology, which could be difficult at an older age. Proper guidance when using AR and a physician well-trained in using the device were therefore considered necessary by patients. Additionally, patients specifically stated that when using AR, medical education provided by a specialist will always stay necessary, preferably by a physician.

> "My physician could explain it in great detail with a drawing, or at least well enough, and that gave a soothing feeling. You know that the physician understands everything about it and that he will fix it. That feeling should stay if you're going to do it that way. (P4)"

Views on preferred AR device were divided. One view was that using a phone to view AR would make is easy to watch AR at home and prevent having to purchase a new device. Several patients were inconclusive about which AR device they would prefer, because they did not have any experience in using AR. When discussing the potential disadvantages of AR in patient education, patients stated that although they would like education using AR, they were afraid it would not work in practice, because it would consume too much time during the consultation.

3.2.3.5 Patients' needs

Several perspectives on patients' need for AR arose during the interviews. Patients told us they expected that it depends on the patient's coping strategy whether someone wants education with AR or not. For example, some patients just do not want to know everything about their condition. It was suspected that AR might be too confronting when a patient has a more severe diagnosis or needs to undergo a treatment with more risks. It was specifically expressed that therefore you should always ask the patient whether they would like to see an AR model. Additionally, it was stated that AR would be most favorable for patients with a low understanding of their condition.

3.2.3.6 Necessity

The necessity of AR in patient education was questioned during the interviews. Participants reported that AR gave them more insight in their condition, but that it was not necessary for patient education. Additionally, it was stated that AR would be more beneficial for conditions that are more difficult to understand than having an aneurysm.

3.2.3.7 Timing

Participants advised not to provide AR when a patient first hears about having an aneurysm, but at a later stage. This way, patients would have time to let the news sink in before further patient education takes place with AR. However, not all patients agreed: one patient stated that he would not mind receiving AR patient education in the acute moment.

3.2.4 Out of hospital use

3.2.4.1 AR at home

Several patients mentioned that they would like to be able to watch an AR model of their aneurysm at home, for example, by using an AR-application on a phone or tablet. This would allow them to rewatch information in a comfortable environment with relatives. One patient expressed that they wanted to gather as much information at home as possible. A suggestion that was made by participants was that an AR recording at home should include the physician's explanation that they received at their hospital visit.

3.2.4.2 AR for relatives

AR could also be used for relatives. A reason patients mentioned to do this was that using AR could make talking about aneurysms easier for patients. Another reason was that it would be beneficial that one would not have to explain everything repeatedly, because they could instead show their relatives using AR. Reasons mentioned against using AR for relatives were that AR could make relatives uncomfortable or that relatives would not be interested in AR.

4 Discussion and conclusion

4.1 Discussion

This study explored the views, expectations and preferences with regard to the use of AR in educating patients about their intracranial aneurysms. In general, participants expressed that AR could be useful in this setting, particularly with regard to improving patients' understanding of their medical situation. Also, AR was suggested to improve the communication between doctor and patient. It was preferred to use AR later in the educational process (i.e., not directly after receiving the diagnosis) and most interviewees saw opportunities regarding the use of AR at home and in explaining their disease to their relatives. Nevertheless, interviewees disclosed that AR could be too time-consuming during consultation and that a physician trained in using AR is necessary. Also, AR was believed to be emotionally confronting, although none of the participants experienced this after viewing the here described AR system.

The few previous studies that have been carried out on the use of AR in patient education, rarely involved patients in the process of designing and implementing AR applications (Domhardt et al., 2015; Calle-Bustos et al., 2017; Azman et al., 2019; Brown et al., 2019; Calle-Bustos et al., 2019; Wake et al., 2019; Bray et al., 2020; House et al., 2020; Sezer et al., 2020; Tait et al., 2020). For a recent review, see (Urlings et al., 2022). End-users often stress misalignments among their problems and the solutions that technology systems aim to solve (Calvillo-Arbizu et al., 2019). Including patient views early on in the process of design of patient education is important, because physicians and other healthcare workers might not always be able to judge what is important to patients (Hilt et al., 2020). Our current study is the first research that has been performed to explore the views of patients with unruptured intracranial aneurysms regarding AR for educational purposes.

The previous studies comprised patients suffering from a diverse spectrum of chronic diseases (e.g., prostate cancer, diabetes mellitus, multiple sclerosis, epilepsy). These studies showed benefits of AR such as (perceived) knowledge gain and increased patient satisfaction (Urlings et al., 2022). These benefits are similar to the benefits of AR as suspected by the patients interviewed in our study. More high-quality studies are needed to conclude whether AR truly has those beneficial effects on patient education and whether it could thereby improve doctor-patient communication.

Additionally, inter-individual differences in visuospatial abilities significantly impacted student performance when working with AR (Moro et al., 2020).

Furthermore, it was noted that patient-specific factors, such as conditions like strabismus that affect three-dimensional vision, could affect the choice of the most suitable AR device for patient education (Jakl et al., 2020). Nevertheless, whether spatial abilities and other co-variates play a significant role when working with AR in patient education remains elusive and should be investigated in future studies.

Patients in this study emphasized the ongoing necessity of medical education provided by a specialist, preferably a physician, when augmented reality (AR) is utilized. This finding corresponds with the existing literature on the use of visualization tools for patient education. Previous research on alternative methods such as videotaped instructions or computer-aided information systems has produced mixed results. However, existing evidence shows that providing a combination of spoken and written or visual information is best (Thomson et al., 2001; Kessels, 2003).

Furthermore, patients expressed their preference to utilize AR later in the educational process as this would give patients the time to process their diagnosis before further patient education takes place. We know evidence exists, indicating that attentional narrowing occurs when events are perceived as stressful or emotional (Kessels, 2003). In such situations, the central message, such as "you have an aneurysm in your brain," becomes the primary focus, leading to limited attention towards other provided information. Consequently, any remaining information, perhaps about treatment options and risks, is not processed and stored into memory and therefore cannot be recalled. Considering this, providing AR patient education later in the educational process could me more beneficial (Kessels, 2003).

Similar to the expectations of patients in our study that AR could help in communicating with relatives, it was found that an AR intervention helped parents to talk about a planned procedure with their children and that it significantly decreased anxiety in parents whose children were undergoing invasive procedures (Bray et al., 2020). In another study relatives expressed their preference of an AR application over a physical model and chose it as the future standard tool for patient education (House et al., 2020). The latter contrasts with our finding that some patients feared AR could make relatives uncomfortable or that relatives would not be interested in AR. Future research should examine the exact value of using AR between patients and relatives.

Finally, patients in our study expressed their worries about needing time to get used to the new technology and that this could be difficult at an older age. Proper guidance when using AR and a physician well-trained in using the device were considered necessary. When looking at the existing literature however, most studies reported a high usability and likability of the AR applications used (Calle-Bustos et al., 2017; Brown et al., 2019; Calle-Bustos et al., 2019; Wake et al., 2019; Bray et al., 2020; House et al., 2020; Tait et al., 2020). These studies comprised patients with an age range between 8 and 63 years. Five of these studies offered participants training or assistance in using AR (16, 17, 21, 22, 24). Therefore, the use of AR might not be problematic for patients with an older age as long as there is proper guidance.

4.1.1 Strengths and limitations

This study is the first to present the expectations and wishes concerning the use of AR for patient education among patients with unruptured intracranial aneurysms. It has been shown in other aspects of medical treatment that the patient's expectations and wishes can be different from those of physicians (Hilt et al., 2020). The identification of positive and negative expectations and patients' wishes as perceived by individual participants allows us to improve the use of AR models for patient education. The results form a basis for future quantitative studies on the effect of AR in patient education for these patients. In these studies, it is necessary to determine whether the use of these models truly contributes to patients' understanding of their disease, the procedure, and risks compared with the use of 2D imaging alone. Another strength of this study is the role of the researchers as an independent party in relation to the treatment process. This enabled participants to speak openly.

To seek in-depth information from a wide range of patients purposive sampling was conducted in this study. As a result, the participant group comprised patients with varying genders, ages, and educational backgrounds. It is worth noting that there were more female participants than male participants, which aligns with the epidemiology of unruptured intracranial aneurysms, as these are more commonly observed in females (Vlak et al., 2011; Brown and Broderick, 2014).

A limitation of this study was that all study participants were treated in the same hospital located in the Netherlands. Patients might perceive AR differently depending on the health resources in a country. This partially limits the transferability of these findings to other patient populations.

Another limitation is that we opted for videos to illustrate AR instead of an AR application or AR-glasses due to practical considerations. This made it possible to conduct the interviews online, which was necessitated by the interviews being conducted during the COVID-19 pandemic. It is possible that this resulted in a less immersive experience compared to an application or AR glasses.

4.2 Conclusion

This study suggests that patients with unruptured intracranial aneurysms are open to receive patient education regarding their disease with AR. Patients expect that AR models can help patients with intra-cranial aneurysms better understand their disease, treatment options and risks. Additionally, patients expect AR could improve doctor-patient communication. The views, expectations and preferences of patients identified in this study can contribute to improving information provision and communication using AR applications by providing insights into patients' perceptions.

4.3 Practice implications

Future studies on AR in patient education should take these expectations and wishes into account and evaluate the extent to which the use of AR models could positively influence the quality of patient education.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the CMO Regio Arnhem-Nijmegen. The patients/participants provided their written informed consent to participate in this study.

Author contributions

Conception: JU, DH, and IA; Data analysis: JU and DH; Supervision: DH and JB. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2023.1204643/full#supplementary-material

References

Azman, A., Poyade, M., and Overell, J. (2019). Towards a more user-friendly medication information delivery to people living with multiple sclerosis: A case study with alemtuzumab. *Adv. Exp. Med. Biol.* 1120, 67–82. doi:10.1007/978-3-030-06070-1_6

Barsom, E. Z., Graafland, M., and Schijven, M. P. (2016). Systematic review on the effectiveness of augmented reality applications in medical training. *Surg. Endosc.* 30 (10), 4174–4183. doi:10.1007/s00464-016-4800-6

Bray, L., Sharpe, A., Gichuru, P., Fortune, P. M., Blake, L., and Appleton, V. (2020). The acceptability and impact of the xploro digital therapeutic platform to inform and prepare children for planned procedures in a hospital: Before and after evaluation study. *J. Med. Internet Res.* 22 (8), e17367. doi:10.2196/17367

Brown, H. M., Collins, C. E., Bucher, T., and Rollo, M. E. (2019). Evaluation of the effectiveness and usability of an educational portion size tool, ServARpreg, for pregnant women. *J. Hum. Nutr. Diet.* 32 (6), 719–727. doi:10.1111/jhn.12660

Brown, R. D., and Broderick, J. P. (2014). Unruptured intracranial aneurysms: Epidemiology, natural history, management options, and familial screening. *Lancet Neurology* 13 (4), 393–404. doi:10.1016/s1474-4422(14)70015-8

A. M. Calle-Bustos, M. C. Juan, F. Abad, and R. Mollá, editors. An augmented reality app for therapeutic education and suitable for mobile devices with different features. 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT); 15-18 July 2019, New York, IEEE.

Calle-Bustos, A. M., Juan, M. C., Garcia-Garcia, I., and Abad, F. (2017). An augmented reality game to support therapeutic education for children with diabetes. *PLoS One* 12 (9), e0184645. doi:10.1371/journal.pone.0184645

Calvillo-Arbizu, J., Roa-Romero, L. M., Estudillo-Valderrama, M. A., Salgueira-Lazo, M., Aresté-Fosalba, N., Del-Castillo-Rodríguez, N. L., et al. (2019). User-centred design for developing e-Health system for renal patients at home (AppNephro). *Int. J. Med. Inf.* 125, 47–54. doi:10.1016/j.ijmedinf.2019.02.007

Di Serio ÁIbáñez, M. B., and Kloos, C. D. (2013). Impact of an augmented reality system on students' motivation for a visual art course. *Comput. Educ.* 68, 586–596. doi:10.1016/j.compedu.2012.03.002

Domhardt, M., Tiefengrabner, M., Dinic, R., Fotschl, U., Oostingh, G. J., Stutz, T., et al. (2015). Training of carbohydrate estimation for people with diabetes using mobile augmented reality. *J. Diabetes Sci. Technol.* 9 (3), 516–524. doi:10.1177/1932296815578880

Dunleavy, M., Dede, C., and Mitchell, R. (2008). Affordances and limitations of immersive participatory augmented reality simulations for teaching and learning. *J. Sci. Educ. Technol.* 18 (1), 7–22. doi:10.1007/s10956-008-9119-1

A. Jakl, A. Lienhart, C. Baumann, A. Jalaeefar, A. Schlager, L. Schöffer, et al. (2020). "Enlightening patients with augmented reality," *IEEE conference on virtual reality and 3D user interfaces* (USA: VR).

Henssen, D., van den Heuvel, L., De Jong, G., Vorstenbosch, M., van Cappellen van Walsum, A. M., Van den Hurk, M. M., et al. (2020). Neuroanatomy learning: Augmented reality vs. Cross-sections. *Anat. Sci. Educ.* 13 (3), 353–365. doi:10.1002/ase.1912

Hilt, A. D., Mamaqi Kapllani, K., Hierck, B. P., Kemp, A. C., Albayrak, A., Melles, M., et al. (2020). Perspectives of patients and professionals on information and education after myocardial infarction with insight for mixed reality implementation: Cross-sectional interview study. *JMIR Hum. Factors* 7 (2), e17147. doi:10.2196/17147

House, P. M., Pelzl, S., Furrer, S., Lanz, M., Simova, O., Voges, B., et al. (2020). Use of the mixed reality tool "VSI Patient Education" for more comprehensible and imaginable patient educations before epilepsy surgery and stereotactic implantation of DBS or stereo-EEG electrodes. *Epilepsy Res.* 159, 106247. doi:10.1016/j.eplepsyres.2019.106247

Iordache, D. D., Pribeanu, C., and Balog, A. (2012). Influence of specific AR capabilities on the learning effectiveness and efficiency. *Stud. Inf. Control* 21 (3), 233–240. doi:10.24846/v21i3y201201

Kamphuis, C., Barsom, E., Schijven, M., and Christoph, N. (2014). Augmented reality in medical education? *Perspect. Med. Educ.* 3 (4), 300–311. doi:10.1007/s40037-013-0107-7

Kessels, R. P. (2003). Patients' memory for medical information. *J. R. Soc. Med.* 96 (5), 219–222. doi:10.1177/014107680309600504

King, J. T., Jr., Yonas, H., Horowitz, M. B., Kassam, A. B., and Roberts, M. S. (2005). A failure to communicate: Patients with cerebral aneurysms and vascular neurosurgeons. *J. Neurol. Neurosurg. Psychiatry* 76 (4), 550–554. doi:10.1136/jnnp.2004.051649

Kucuk, S., Kapakin, S., and Goktas, Y. (2016). Learning anatomy via mobile augmented reality: Effects on achievement and cognitive load. *Anat. Sci. Educ.* 9 (5), 411–421. doi:10.1002/ase.1603

Moro, C., Birt, J., Stromberga, Z., Phelps, C., Clark, J., Glasziou, P., et al. (2020). Virtual and augmented reality enhancements to medical and science student physiology and anatomy test performance: A systematic review and meta-analysis. *Anat. Sci. Educ.* 14, 368–376. doi:10.1002/ase.2049

Pelargos, P. E., Nagasawa, D. T., Lagman, C., Tenn, S., Demos, J. V., Lee, S. J., et al. (2017). Utilizing virtual and augmented reality for educational and clinical enhancements in neurosurgery. *J. Clin. Neurosci.* 35, 1–4. doi:10.1016/j.jocn.2016.09.002

Saito, M., Takahashi, Y., Yoshimura, Y., Shima, A., Morita, A., Houkin, K., et al. (2012). Inadequate communication between patients with unruptured cerebral aneurysms and neurosurgeons. *Neurol. Med. Chir. (Tokyo)* 52 (12), 873–877. doi:10.2176/nmc.52.873

Sezer, S., Piai, V., Kessels, R. P. C., and Ter Laan, M. (2020). Information recall in pre-operative consultation for glioma surgery using actual size three-dimensional models. *J. Clin. Med.* 9 (11), 3660. doi:10.3390/jcm9113660

Tait, A. R., Connally, L., Doshi, A., Johnson, A., Skrzpek, A., Grimes, M., et al. (2020). Development and evaluation of an augmented reality education program for pediatric research. *J. Clin. Transl. Res.* 5 (3), 96–101. doi:10.18053/jctres.05.202003.002

Thomson, A. M., Cunningham, S. J., and Hunt, N. P. (2001). A comparison of information retention at an initial orthodontic consultation. *Eur. J. Orthod.* 23 (2), 169–178. doi:10.1093/ejo/23.2.169

Urlings, J., Sezer, S., Ter Laan, M., Bartels, R., Maal, T., Boogaarts, J., et al. (2022). The role and effectiveness of augmented reality in patient education: A systematic review of the literature. *Patient Educ. Couns.* 105, 1917–1927. doi:10.1016/j.pec.2022.03.005

van de Belt, T. H., Nijmeijer, H., Grim, D., Engelen, L. J. L. P. G., Vreeken, R., van Gelder, M. M. H. J., et al. (2018). Patient-specific actual-size three-dimensional printed models for patient education in glioma treatment: First experiences. *World Neurosurg.* 117, e99–e105. doi:10.1016/j.wneu.2018.05.190

Vlak, M. H., Algra, A., Brandenburg, R., and Rinkel, G. J. (2011). Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: A systematic review and meta-analysis. *Lancet Neurol.* 10 (7), 626–636. doi:10.1016/s1474-4422(11)70109-0

Wake, N., Rosenkrantz, A. B., Huang, R., Park, K. U., Wysock, J. S., Taneja, S. S., et al. (2019). Patient-specific 3D printed and augmented reality kidney and prostate cancer models: Impact on patient education. *3D Print Med.* 5 (1), 4. doi:10.1186/s41205-019-0041-3

Zhu, E., Hadadgar, A., Masiello, I., and Zary, N. (2014). Augmented reality in healthcare education: An integrative review. *PeerJ* 2, e469. doi:10.7717/peerj.469

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 17 November 2023
DOI 10.3389/fbioe.2023.1302524

Check for updates

OPEN ACCESS

EDITED BY
Adrian Elmi-Terander,
Stockholm Spine Center, Sweden

REVIEWED BY
Shireen Y. Elhabian,
The University of Utah, United States
Xiaojun Chen,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Wang Li,
wang.l@cqut.edu.cn

†These authors have contributed equally to this work

RECEIVED 26 September 2023
ACCEPTED 01 November 2023
PUBLISHED 17 November 2023

CITATION
Lv J, Zhang L, Xu J, Li W, Li G and Zhou H (2023), Automatic segmentation of mandibular canal using transformer based neural networks.
Front. Bioeng. Biotechnol. 11:1302524.
doi: 10.3389/fbioe.2023.1302524

COPYRIGHT
© 2023 Lv, Zhang, Xu, Li, Li and Zhou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Automatic segmentation of mandibular canal using transformer based neural networks

Jinxuan Lv†, Lang Zhang†, Jiajie Xu, Wang Li*, Gen Li and Hengyu Zhou

School of Pharmacy and Bioengineering, Chongqing University of Technology, Chongqing, China

Accurate 3D localization of the mandibular canal is crucial for the success of digitally-assisted dental surgeries. Damage to the mandibular canal may result in severe consequences for the patient, including acute pain, numbness, or even facial paralysis. As such, the development of a fast, stable, and highly precise method for mandibular canal segmentation is paramount for enhancing the success rate of dental surgical procedures. Nonetheless, the task of mandibular canal segmentation is fraught with challenges, including a severe imbalance between positive and negative samples and indistinct boundaries, which often compromise the completeness of existing segmentation methods. To surmount these challenges, we propose an innovative, fully automated segmentation approach for the mandibular canal. Our methodology employs a Transformer architecture in conjunction with cl-Dice loss to ensure that the model concentrates on the connectivity of the mandibular canal. Additionally, we introduce a pixel-level feature fusion technique to bolster the model's sensitivity to fine-grained details of the canal structure. To tackle the issue of sample imbalance and vague boundaries, we implement a strategy founded on mandibular foramen localization to isolate the maximally connected domain of the mandibular canal. Furthermore, a contrast enhancement technique is employed for pre-processing the raw data. We also adopt a Deep Label Fusion strategy for pre-training on synthetic datasets, which substantially elevates the model's performance. Empirical evaluations on a publicly accessible mandibular canal dataset reveal superior performance metrics: a Dice score of 0.844, click score of 0.961, IoU of 0.731, and HD95 of 2.947 mm. These results not only validate the efficacy of our approach but also establish its state-of-the-art performance on the public mandibular canal dataset.

KEYWORDS
mandibular canal, transformer, feature fusion, segmentation, CBCT

1 Introduction

The mandibular canal (MC) is a tubular anatomical structure situated within the mandible and chiefly houses the inferior alveolar nerve (IAN) and associated vasculature (Agbaje et al., 2017). This nerve shares a critical relationship with the third molar (Rai et al., 2014). Any insult to the MC can lead to adverse outcomes such as patient discomfort, acute pain, or even facial paralysis (Al-Juboori et al., 2014). Therefore, precise segmentation of the MC from imaging modalities is instrumental for clinicians to appreciate the spatial relationship between the MC and adjacent anatomical landmarks, thereby minimizing the risk of iatrogenic nerve injury during surgical interventions (Li et al., 2022). Owing to the cumbersome and error-prone nature of manual delineation, automated segmentation of the MC from radiological images has emerged as a focal point in dental research (Usman et al., 2022).

With the advent of advanced deep learning techniques, neural network-based segmentation of oral structures has shown significant progress (Cui et al., 2022; Fontenele et al., 2023). However, the segmentation of the mandibular canal still falls short when compared to other anatomical structures. The primary challenges are multifaceted. First, the mandibular canal occupies a minute fraction of the overall CBCT image, which can lead the neural network to prioritize the background over the target foreground. Second, the low contrast of CBCT images makes it difficult to distinguish the mandibular canal from surrounding tissues, often resulting in blurred or indistinct boundaries. Traditional segmentation approaches such as region growing, level set, thresholding, and model matching have proven insufficient for overcoming these obstacles (Kainmueller et al., 2009; Abdolali et al., 2017). U-Net-based architectures have exhibited excellent performance across various domains since their introduction. Nonetheless, they often lack the capability to provide holistic information, causing them to neglect the topological structure of the mandibular canal during segmentation tasks (Jaskari et al., 2020; Lahoud et al., 2022). In recent years, Transformer-based encoder-decoder frameworks [e.g., TransUNet (Chen et al., 2021), UNETR (Hatamizadeh et al., 2022a), UNETR++ (Shaker et al., 2022)] have emerged, demonstrating promising results (Li et al., 2023). These Transformer-based methodologies utilize a global mechanism to capture features over long distances, addressing the limitations of CNN-based networks. However, the existing Transformer-based segmentation methods predominantly focus on larger organs, and they still do not provide effective solutions for segmenting the mandibular canal, which has smaller voxel sizes.

To address the aforementioned challenges, we have enhanced the Swin-UNetR model specifically for mandibular canal segmentation. We also incorporate a pixel-level feature fusion module to augment the model's capability to discern finer details of the mandibular canal. To mitigate the severe class imbalance between the foreground and background, as well as the low contrast prevalent in mandibular canal data, we introduce a cropping technique grounded in mandibular foramen localization and a contrast enhancement strategy based on Contrast-Limited Adaptive Histogram Equalization (CLAHE). Given the topological continuity of the mandibular canal, we employ clDice as the model's loss function. Moreover, to improve model robustness, we propose a straightforward yet effective deep label fusion technique that capitalizes on the sparse data in the dataset. Our main contributions can be summarized as follows:

(1) We introduce an enhanced Transformer-based segmentation network tailored for mandibular canal segmentation, offering a novel avenue for accurate segmentation of this complex structure.
(2) We proposed a pixel-level feature fusion module to improve the model's detail perception ability, and improve the model's segmentation accuracy and convergence speed.
(3) We introduce a cropping method that autonomously localizes the mandibular and mental foramina, coupled with an image contrast enhancement strategy, as preprocessing steps to address the challenges of category imbalance and unclear mandibular canal boundaries. Furthermore, our depth expansion technique is used to generate fused label datasets, enhancing the model's robustness.

The remainder of the paper is structured as follows: Section 2 reviews related work in mandibular canal segmentation. Section 3 provides a comprehensive description of our proposed method. Section 4 discusses the materials and implementation details. In Section 5, we present the results along with comparative analyses. Section 6 contains the analysis and discussion of our work. Finally, Section 7 concludes the paper.

2 Related work

In the fisrt chapter, we delineated the broader research context, current state of the field, and specifically emphasized the importance and challenges associated with mandibular canal segmentation. In the subsequent chapters, we will delve deeper into the historical development of various mandibular canal segmentation techniques. These methods can be broadly categorized based on the underlying technology into traditional image processing techniques, CNN (Convolutional Neural Network)-based approaches, and Transformer-based segmentation methods.

2.1 Traditional image processing-based segmentation method

To address the clinical issue of solely relying on manual segmentation of the mandibular canal by dental professionals, Kainmueller et al. (2009) proposed an automated segmentation technique that combines the Dijkstra tracking algorithm with the Statistical Shape Model (SSM). This method successfully reduced the average error to 1.0mm, achieving a level of automation. Subsequently, Kim et al. (2010) presented a segmentation strategy that employs 3D panoramic volume rendering (VR) and texture analysis. Their approach captured variations in the curvature of the mandibular canal using a line tracing algorithm. Furthermore, threshold-based segmentation technologies have seen some advancements. Moris et al. (2012) employed a thresholding technique to identify the mandibular and mental foramina and then used template matching technology to recursively calculate the optimal path between them, leveraging strong prior knowledge to achieve effective segmentation results. Building on Mori et al.'s work, Onchis-Moaca et al. (2016) enhanced template matching technology by using the anisotropic generalized Hough transform of the Gabor filter, significantly improving computational efficiency. However, these methods suffer from excessive reliance on prior knowledge and limited generalizability. On the other hand, to tackle the low contrast of CBCT images, Abdolali et al. (2017) innovatively employed low-rank matrix decomposition to enhance image quality, thereby increasing the visibility of the mandibular canal in the shape model. Similarly, Wei and Wang (2021) utilized windowing and

K-means clustering algorithms for data enhancement to improve the mandibular canal's visibility and subsequently deployed two-dimensional linear tracking coupled with tetranomial fitting for segmentation. In summary, traditional segmentation methods either depend excessively on prior knowledge or require significant manual intervention, leading to pronounced human-induced biases.

2.2 CNN-based segmentation method

In recent years, CNN-based segmentation methods have achieved significant advancements in mandibular canal segmentation. Jaskari et al. (2020) first employed a Fully Convolutional Network (FCN) for this task, achieving a Dice coefficient of 0.57 and thus substantiating the efficacy of CNN approaches in this domain. Following this, Kwak et al. (2020) utilized thresholding technology for rapid mandibular canal localization, converting the full-volume image into a 2D slice sequence. They then employed SegNet and 3D UNet models for 2D slice-level and 3D volume-level segmentation, respectively. However, their approach did not adequately consider the structural information of the mandibular canal. To address this gap, Widiasri et al. (2022) segmented 3D images into 2D slices and utilized the Dental-Yolo algorithm for feature detection. This method computed the dimensions between the alveolar bone and the mandibular canal, allowing the model to acquire rich positional information. Additionally, to enhance segmentation accuracy and mitigate computational limitations, researchers have proposed generalized hierarchical frameworks (Lahoud et al., 2022; Verhelst et al., 2021). For instance, Verhelst et al. (2021) initially downsampled images to reduce resolution, retained only patches with foreground classes, and employed a 3D UNet in conjunction with the Marching Cubes algorithm for smoothing and segmentation. However, this method necessitates some manual input and struggles with samples that have indistinct mandibular canal boundaries. To counteract the issue of blurred boundaries, Faradhilla et al. (2021) introduced a Double Auxiliary Loss (DAL) in the loss function to make the network more attentive to the target area and its boundaries, achieving a high Dice accuracy of 0.914 on their private dataset. To combat class imbalance, Du et al. (2022) innovatively introduced a pre-processing step involving centerline extraction and region growing to identify the mandibular canal's location. They used a fixed point as a reference to crop a localized region around the mandibular canal, thereby minimizing the impact of background samples. Despite the successes of these methods, they generally sacrifice rich global information during training, leading to a loss of structural integrity in the segmented mandibular canal.

2.3 Transformer-based segmentation method

In the realm of medical imaging, Transformer-based techniques have garnered considerable attention, finding applications across a range of tasks including segmentation, recognition, detection, registration, reconstruction, and enhancement (Li et al., 2023; Dosovitskiy et al., 2020). One key advantage of the Transformer architecture over Convolutional Neural Networks (CNNs) is its robust capability for global perception, allowing for a more effective understanding of global contextual information and capturing long-range dependencies. Many Transformer-based approaches have been adapted for segmentation tasks involving major human organs, and have yielded impressive results (Liu and Shen, 2022; Pan et al., 2022). For instance, Wang et al. (2021) introduced the UCTransNet model, which for the first time incorporated the Transformer into the channel dimension. By leveraging feature fusion and multi-scale channel attention, the model optimized the information integration between low- and high-dimensional spaces (Chen et al., 2021). Hatamizadeh et al. (2022a) then proposed the UNETR model, which employed the Vision Transformer (ViT) as the encoding layer. This model leveraged the Transformer's strong global modeling capabilities to achieve excellent performance on multi-organ segmentation datasets. To address the UNETR model's relatively weaker performance in capturing local details, Hatamizadeh et al. (2022b) introduced the Swin UNETR segmentation model. This variant ensured a global receptive field while also giving ample consideration to local details, and it has shown promising results in tasks such as brain tumor segmentation. Specifically in the context of mandibular canal segmentation, Jeoun et al. (2022) introduced the Canal-Net, a continuity-aware context network designed to help the model understand the spatial structure of the mandibular canal. This approach achieved a Dice coefficient of up to 0.87. These outcomes provide compelling evidence to suggest that the Transformer's strong context-aware capabilities could be particularly effective for mandibular canal segmentation tasks. However, it is worth noting that research in Transformer-based mandibular canal segmentation is still in its nascent stages. Recognizing the unique challenges and characteristics of mandibular canal segmentation, we sought to improve upon the Swin UNETR model. Our modified approach has yielded promising segmentation results, underscoring the potential utility of Transformer-based architectures in this domain.

3 Methods

3.1 Data preprocessing

Considering the impact of preprocessing on model performance, we employ a comprehensive set of preprocessing steps to address existing challenges in CBCT imagery and thereby enhance the segmentation accuracy of the mandibular canal. The specific process is shown in Figure 1. The rectangular box in Figure 1 highlights the changes in the mandibular canal.

3.1.1 Volume cropping

The proportion of voxels representing the mandibular canal in the entire CBCT image is exceedingly small, exacerbating the class imbalance between foreground and background. This imbalance adversely affects the model's segmentation performance, as illustrated in Figure 1A. To address this challenge, we introduce a cropping technique based on the localization of the mandibular foramen. This approach aims to identify the largest connected domain of the mandibular canal by locating the jaw foramen, as depicted in Figure 1B. This method locates the positional

10.3389/fbioe.2023.1302524

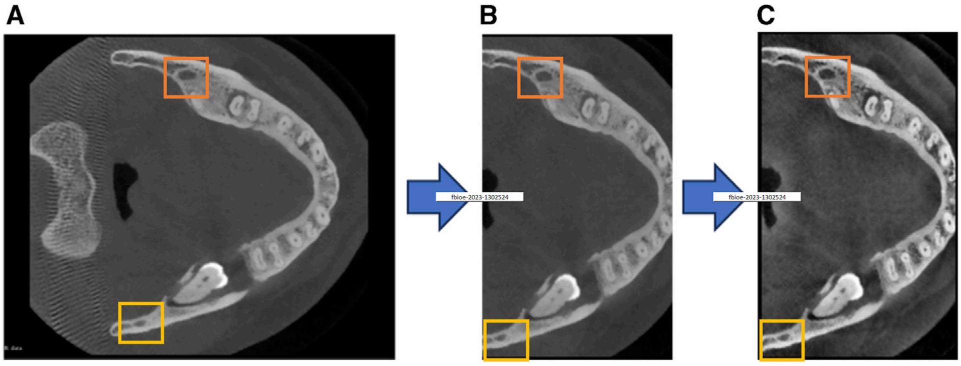


FIGURE 1
Our proposed preprocessing process, **(A)** represents the original image, **(B)** represents the cropped image, and **(C)** represents the contrast-enhanced image.

relationship between the mandibular foramen and chin foramen in the labeled information, and then maps this positional relationship to the original image for cropping. Following this preprocessing step, the total number of voxels is reduced by approximately 60%. This reduction not only enhances the model's convergence speed and segmentation accuracy but also minimizes hardware resource consumption.

3.1.2 Contrast enhancement

In CBCT imaging, the gray values of the mandibular canal and surrounding tissues are often similar, which obscures the boundary of the mandibular canal. Further complicating the matter, some CT devices may produce images with low resolution and blurriness, making it difficult to differentiate the mandibular canal from adjacent structures. To overcome these challenges, we employ Contrast-Limited Adaptive Histogram Equalization (CLAHE) to enhance image contrast, thereby improving the model's segmentation performance. This enhancement is demonstrated in Figure 1C.

3.2 Mandibular canal segmentation network structure

We employ the aforementioned preprocessing techniques on the CBCT images and use them as input for the segmentation network. In the encoder portion of the network, a 4-layer Swin Transformer serves as the feature extractor. This architecture leverages the Transformer's robust capability for global modeling, allowing it to focus more effectively on the overall structural features of the mandibular canal, compared to traditional CNN-based feature extractors. The decoder part of the network adheres to the conventional U-Net decoding structure. In this design features extracted by the encoder are connected to the decoder via skip connections at each scale. At each stage of the encoder i the output features are reshaped to size $\frac{H}{2^i} \times \frac{W}{2^i} \times \frac{D}{2^i}$, which are then fed into a residual module consisting of two $3 \times 3 \times 3$ convolutional layers. Subsequently, the feature map is upsampled by 2 times using a deconvolution layer, and is concatenated with the output of the previous layer and fed into the residual module. Finally, the output of the residual module is sent to the DRC module to achieve pixel-level feature fusion with the previous layer features. The final segmentation result is calculated by using a $1 \times 1 \times 1$ convolutional layer and a sigmoid activation function. It restores the spatial dimensions of the feature map through a series of five upsampling operations, as shown in Figure 2.

To further improve the network's ability to perceive the details of the mandibular canal, we introduce a feature fusion strategy of element-by-element addition and use the DRC (Deep Residual Convolution) module for each decoding layer to further extract features, as shown in Figure 3B. Comparing with traditional CNN structure, as shown in Figure 3A, this module is mainly composed of two branches: the first branch consists of a 1×1 convolution, the second branch consists of two 1×1 and a 3×3 convolution, and to improve the expressiveness of the convolution, we perform normalization and ReLU activation operations after each convolution operation. The output of the DRC module can be expressed as:

$$DRC = F(X, Y)_L + F(X, Y)_R \tag{1}$$

among them, X represents the input data, $F(X, Y)_L$ represents the output of the left branch, and $F(X, Y)_R$ represents the output of the right branch. The extracted features are fused layer by layer at the pixel level to obtain the fused feature map $F(x, y)$:

$$F(x, y) = F_n(x, y) + DRC(F_{n-1}(x, y)) \tag{2}$$

where $F(x, y)$ represents the pixel position in the feature map, and n represents the nth decoder layer. Through this fusion strategy, the model can learn more information from different feature map.

3.3 Deep Label fusion

To optimally leverage our set of 256 sparsely labeled data, we introduce an innovative approach for pseudo-label generation. Initially, the model is trained using densely annotated data, after which it generates pseudo-labels for the 256 sparsely annotated samples. Compared to the original sparse labels, these pseudo-labels offer a richer semantic context but may lack adequate connectivity. To address this limitation, we implement an intelligent label fusion

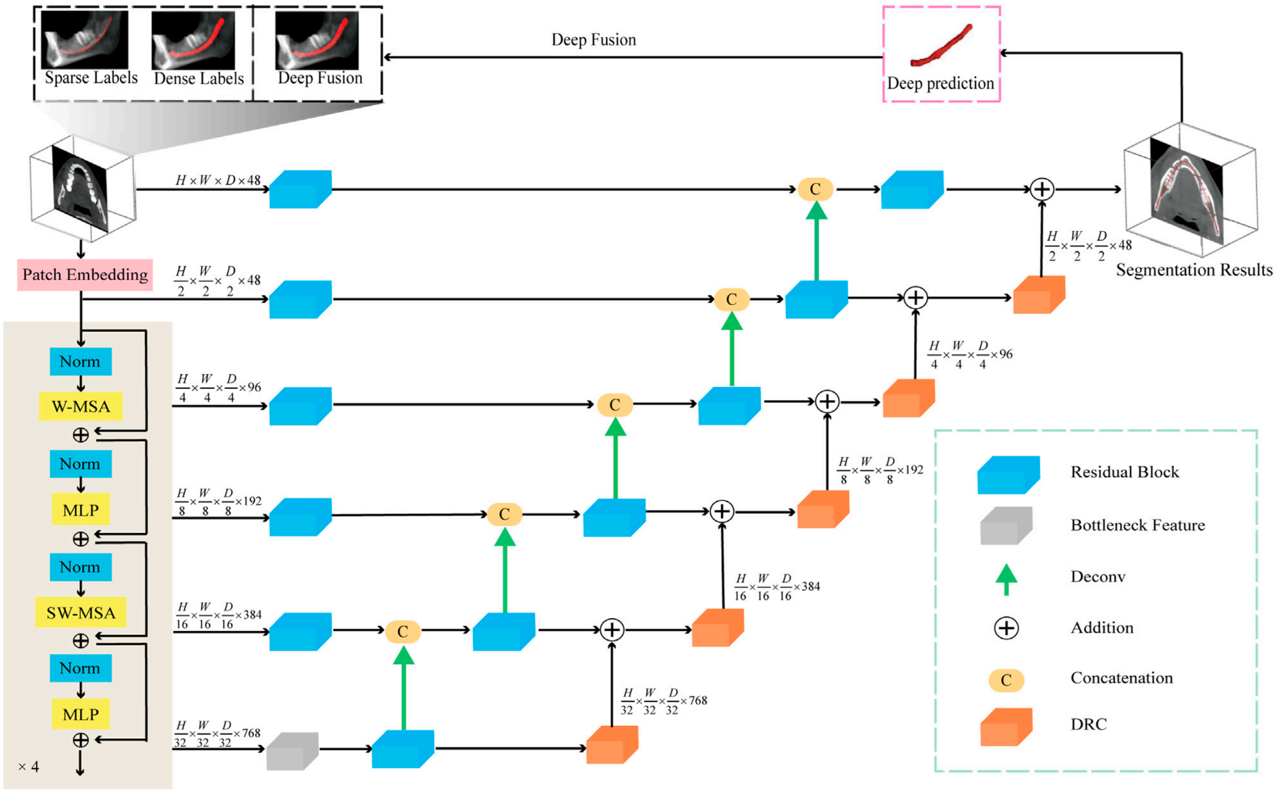


FIGURE 2
The network diagram used in this article consists of a Transformer encoding module, a decoding module, and a feature fusion module. In addition, the model accepts three types of labels: sparse label, dense label, and deep fusion label.

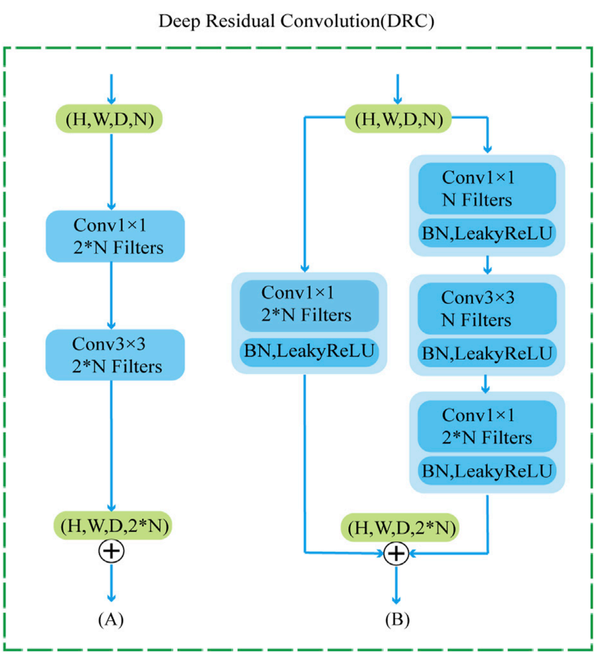


FIGURE 3
DRC module structure diagram, where **(A)** represents the traditional convolution module **(B)** represents our proposed DRC module.

algorithm. This method first integrates instance level features of circular extended labels and newly generated pseudo labels through interpolation. More specifically, the pseudo-labels contribute valuable semantic insights, while the circular extension labels provide precise boundary delineation. We have coined this method "Deep Label Fusion," and employ it to create an augmented dataset for this study. Utilizing this extended dataset for pre-training the prediction model led to a notable improvement in the Dice metric, particularly when compared to the performance achieved with the original set of 256 circularly extended labels.

3.4 Loss function analysis

The primary objective of the loss function in medical image segmentation tasks is to quantify the discrepancy between the predicted segmentation outcomes and the ground-truth labels. Given that the mandibular canal is a tubular structure, its connectivity is a crucial consideration. In 2021, Shit et al. introduced a loss function designed to take into account both vessel topology and connectivity, known as centerline Dice (clDice). This function is computed based on the intersection between the segmentation output and the extracted cartilage scaffold. Importantly, clDice is adept at evaluating the connectivity of tubular anatomical features. In our research, we employ clDice as the loss function for training the network. The expression for the clDice function is as follows:

TABLE 1 Comparison of results of different segmentation methods.

Test	Methods	Training set	HD	IoU	clDice#	Dice
1	Jaskari et al. (2020)	Cir. Exp.	—	0.39	—	0.56
2	Ours	Cir. Exp.	7.844	**0.405**	0.845	**0.573**
3	Cipriano et al., (2022a)	3D Ann.	—	0.61	—	0.75
4	Usman et al., (2022)	3D Ann.	—	0.79	—	0.77
5	3D UNet	3D Ann.	16.048	0.558	0.809	0.709
6	nn-UNet	3D Ann.	6.363	0.665	**0.935**	0.796
7	UNetR	3D Ann.	8.027	0.569	0.823	0.722
8	Swin-Unet (Cao et al.)	3D Ann.	7.072	0.482	0.733	0.640
9	Ours	3D Ann.	**5.002**	**0.692**	0.933	**0.815**

Bold represents the optimal result. # is the measurement standard for tubular structure proposed by Shit et al. (2021).

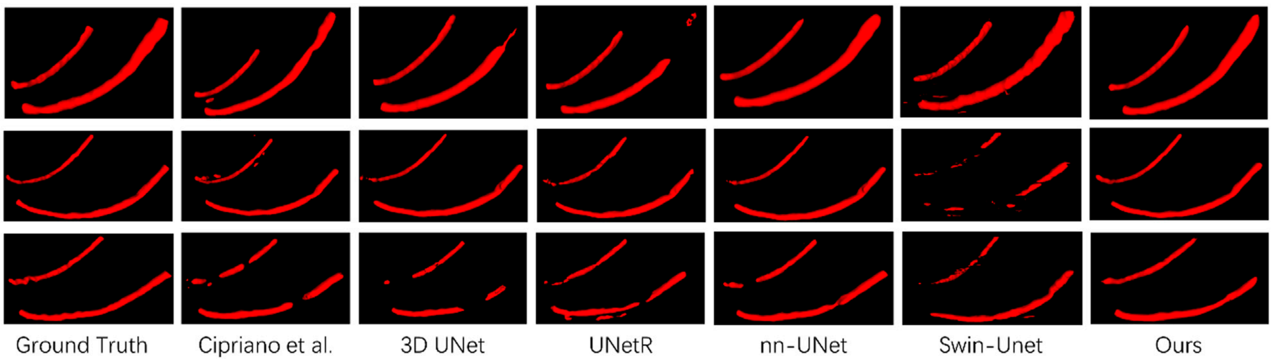


FIGURE 4
Comparison of visualization results of different segmentation methods.

TABLE 2 Quantitative analysis results of different preprocessing methods on model performance.

Input images	HD_{95} (mm)	IoU	clDice	Dice
Original	6.355	0.656	0.912	0.788
Volume Cutting	5.008	0.684	0.927	0.810
Contrast Enhancement	**5.000**	**0.692**	**0.933**	**0.815**

Bold values are reports the optimal result.

$$T_p(S_P, V_L) = \frac{|S_P \cap V_L|}{|S_P|}, \tag{3}$$

$$T_s(S_L, V_P) = \frac{|S_L \cap V_P|}{|S_L|}, \tag{4}$$

$$L_{clDice} = -2 \times \frac{T_p(S_P, V_L) \times T_s(S_L, V_P)}{T_p(S_P, V_L) \times T_s(S_L, V_P)}, \tag{5}$$

among them, V_L and V_P refer to the predicted results and real labels of network segmentation, respectively, S_P and S_L refer to the soft skeleton extracted from V_L and V_P, respectively, $Tp(S_P, V_L)$ refers to the topological accuracy, $Ts(S_L, V_P)$ is the topological sensitivity. L_{clDice} is the harmonic mean of the above two metrics to focus on object connectivity. The total loss function L_{total} combines Dice Loss and clDice Loss, the formula is as follows:

$$L_{total} = (1 - \lambda)L_{Dice} + \lambda L_{clDice}, \tag{6}$$

where λ is a scaling factor.

4 Data and implementation details

4.1 Data

The CBCT dataset utilized in this study was supplied by Cipriano et al. (2022b) and exists in two versions: old and new. The old dataset comprises 91 3D densely annotated primary datasets and 256 2D sparsely annotated auxiliary datasets. This primary dataset is further divided into 68 training sets, 8 validation sets, and 15 test sets. The spatial resolutions of these CBCT scans range from 148 × 272 × 334 to 171 × 423 × 462, featuring a voxel size of 0.3 × 0.3 × 0.3 mm^3. Conversely, the new dataset consists of 153 densely annotated primary datasets and 290 sparsely annotated auxiliary datasets. The spatial resolution in this new version ranges from 148 × 265 × 312 to 178 × 423 × 463. Additionally, the training set in this new version has been expanded to include 130 datasets. To

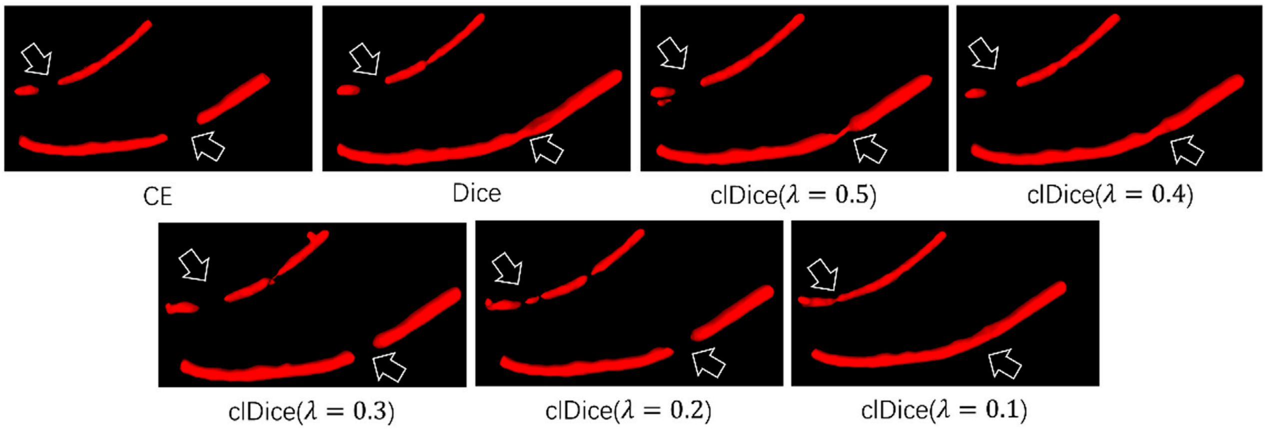


FIGURE 5
Visual segmentation results obtained by training the network with different loss functions.

maintain a fair and rigorous comparison with other studies, all comparative analyses were conducted using the old dataset. Moreover, to demonstrate the cutting-edge nature of our research, we also conducted verifications using the new dataset version.

4.2 Experimental details

Our experiments are implemented using NVIDIA Tesla V100S in the PyTorch and MONAI deep learning libraries. During the preprocessing, we processed the raw data offline by the proposed jaw-foramen localization-based volume cropping method and image contrast enhancement method. During the training process, Diceloss and clDiceloss are used as the loss function of the model, the Adam optimizer with momentum (μ = 0.99) is used, the initial learning rate is set to 0.0001, and the learning rate is automatically adjusted using cosine annealing, and the batch size is set to 1, and the number of iterations of the model is uniformly set to 500. In our experiments, to reduce memory usage, we use a 96 × 96 × 96 sliding window with a stride of 48 to crop the original CBCT image, and then feed the cropped image into the network for training. After outputting the predicted patch, we restore the output predicted patch to the original image size by stitching.

4.3 Evaluation indicators

In the test phase, we use four commonly used evaluation indicators for segmentation tasks to evaluate the performance of the model: Dice coefficient (Dice), Intersection Over Union (IOU), Hausdorff distance (HD):

4.3.1 Dice coefficient (Dice)

The Dice coefficient is a set similarity measurement function, which is usually used to calculate the similarity between two samples, and the value range is [0, 1].

$$Dice = \frac{2TP}{FP+2TP+FN} \tag{7}$$

4.3.2 Intersection over union (IOU)

The IOU indicator calculates the overlap rate of predicted results and real results, that is, the ratio of their intersection and union.

$$IOU = \frac{A \cap B}{A \cup B} \tag{8}$$

4.3.3 Hausdorff distance (HD)

The HD indicator is a metric used to measure the similarity or difference between two sets.

$$H(A, B) = \max\{h(A, B), h(B, A)\} \tag{9}$$

where TP represents true positives, TN represents true negatives, FP represents false positives, FN represents false negatives, A represents the set of true labels, and B represents the set of predicted segmentations.

5 Results

5.1 Evaluation of result

To prove the effectiveness of our proposed mandibular canal segmentation method, we conducted a performance evaluation. The specific results are as follows: only trained on 91 dense data, average Dice = 0.815, average IoU = 0.69, average clDice = 0.93, the average HD95 = 5 mm, all evaluation indicators have proved the excellence of our proposed mandibular canal segmentation method. In addition, to prove the advanced nature of our proposed method, we also compared and analyzed it with existing methods, and the specific comparison results are shown in Table 1. From the comparison results, it can be seen that the improvement of the segmentation method we proposed is very significant. Compared with the most advanced method that also uses

TABLE 3 Quantitative results of different feature fusion methods.

Test	HD_{95} (mm)	IoU	clDice	Dice
Baseline	10.320	0.629	**0.92**	0.769
Baseline+C	9.550	0.642	0.894	0.777
Baseline+DRC	**6.355**	**0.656**	0.912	**0.788**

Among them, C means to use the traditional convolution module, and DRC means to use the deep residual convolution module. Bold values are reports the optimal result.

TABLE 4 Quantitative results of training models with different loss functions.

Loss function	HD_{95} (mm)	IoU	clDice	Dice
Cross-Entropy (CE)	8.207	0.660	0.907	0.790
Dice	7.828	0.665	0.888	0.795
clDice (λ= 0.5)	9.392	0.656	0.909	0.787
clDice (λ= 0.4)	6.958	0.680	0.927	0.806
clDice (λ= 0.3)	9.550	0.660	0.907	0.789
clDice (λ= 0.2)	5.005	0.681	0.927	0.806
clDice (λ= 0.1)	**5.002**	**0.692**	**0.933**	**0.815**

Bold values are reports the optimal result.

91 densely labeled data, our Dice index has increased by 4.5%. In addition, using 256 sparse data for pre-training and our proposed deep label fusion strategy for training, the Dice index reached 0.824 and 0.840, respectively. The specific visual comparison results are shown in Figure 4.

5.2 Ablation experiment

5.2.1 Preprocessing

In Section 3.2, we proposed two data preprocessing methods, namely, the cropping method based on jaw hole positioning and the contrast enhancement method based on Contrast-Limited Adaptive Histogram Equalization. Figure 4 shows our proposed preprocessing method in detail. We have analyzed the effectiveness of the two proposed methods, and the specific results are shown in Table 2. It can be seen from the table that the Dice index has increased by 2.7% after data preprocessing.

5.2.2 Feature fusion

To deeply study the impact of our proposed feature fusion strategy on the performance of mandibular canal segmentation, we conduct a series of ablation experiments and summarize the experimental results in Table 3. The results show that the feature fusion strategy plays an important role in the mandibular canal segmentation task. The Dice coefficient using this feature fusion strategy is 0.788, which is significantly improved compared to the case where this module is not used. In addition, we conduct a comparative analysis of the proposed DRC module and traditional convolution. This design can effectively enhance the representation ability of the model and help to further optimize our proposed feature fusion strategy to improve segmentation performance.

5.2.3 Loss function analysis

In our work, we use clDice Loss as the loss function to train the network. In order to prove that clDice Loss is helpful in improving the mandibular canal segmentation effect, we compared clDice Loss with Cross-Entropy (CE) Loss and Dice Loss. The specific results are shown in Table 4. We found that using clDice Loss as the loss function achieved the optimal Dice coefficient of 0.815. In addition, we also compared the impact of different hyperparameters λ in clDice Loss on model segmentation performance. The specific visual comparison results are shown in Figure 5. The broken part of the mandibular canal in the segmentation result is marked with arrows. From the figure, it can be clearly seen that the segmentation result has the best connectivity when λ = 0.1.

5.2.4 Deep label fusion

In Section 3.3, we proposed a deep label fusion method, which generated 256 fused labels on a sparse dataset, forming a new deep fusion dataset. This deep fusion dataset contains richer semantic information compared to sparse datasets, which can better guide model training. To verify the effectiveness of this method, we conducted a performance evaluation, and the specific evaluation results are shown in Table 5. From the table, it can be seen that the Dice index has been significantly improved when using our proposed deep fusion label for pre training. When using 3D Ann data for training, using our method for pre training is 1.6% higher than using circular extended data for pre training Dice index.

6 Discussion

In this study, we propose a Transformer-based method for automatic segmentation of the mandibular canal, which is capable of simultaneously focusing on local fine-grained details and global semantic information of the mandibular canal to segment the mandibular canal with highly consistent accuracy across the entire CBCT image. We validate the method on the largest mandibular canal segmentation dataset, and the evaluation index Dice coefficient exceeds previous research methods.

Due to the low contrast in CBCT images and the close similarity in grayscale values between the mandibular canal and surrounding tissues, neural networks face difficulties in effectively distinguishing the boundaries of the mandibular canal (Waltrick et al., 2013). Furthermore, the mandibular canal constitutes only a minute fraction of the total CBCT image volume, leading to a pronounced class imbalance between foreground and background. This imbalance causes the network to disproportionately focus on background features (Dai et al., 2023). To address these challenges, we employ two pre-

TABLE 5 Analysis of training results using different labels.

Test	Pre-training set	Training set	HD	IoU	clDice	Dice
1	—	3D Ann.	5.002	0.692	0.933	0.815
2	Cir. Exp.	3D Ann.	4.061	0.704	0.947	0.824
3	Deep fusion	3D Ann.	3.213	0.727	0.960	0.840
4	Deep fusion	ToothFairy	**2.947**	**0.731**	**0.961**	**0.844**

Notes: 3D Ann indicates densely labeled data, Cir. Exp. represents circle expansion data, Deep fusion represents synthetic data sets, and ToothFairy represents new version data sets. Bold values are reports the optimal result.

processing techniques aimed at mitigating issues related to blurred boundaries and class imbalances: a cropping method for automated localization of the mandibular and mental foramina, and Contrast-Limited Adaptive Histogram Equalization (CLAHE) for image contrast enhancement. By eliminating extraneous information and enhancing the contrast between the mandibular canal and its surrounding tissue, these techniques facilitate precise localization and segmentation of the mandibular canal. The efficacy of this contrast enhancement approach has also been successfully validated in two-dimensional microscopic images as per Wu et al. (2022). As demonstrated in Section 5, the implementation of these two preprocessing methods results in an approximate 4% improvement in Dice coefficient performance.

Secondly, to maintain the connectivity of the segmented mandibular canal, we incorporated the Transformer architecture into the segmentation task. This enables the network to learn both the local fine-grained details and the global semantic information pertinent to the mandibular canal (Liu et al., 2021). Given the small volumetric proportion of the mandibular canal in the CBCT images, we introduced a pixel-level feature fusion strategy to augment the network's segmentation performance. The deployment of the Deep Residual Convolution (DRC) module further enriches the network's perception of intricate details. Previous studies have substantiated the efficacy of feature fusion strategies in enhancing the segmentation performance for small and indistinct targets (Dai et al., 2023). Our empirical tests show that the feature fusion module not only improves segmentation performance but also accelerates model convergence, reducing training time by as much as 50%. This acceleration is likely attributed to the enhanced perceptual capabilities conferred by the module, partially ameliorating the slow convergence typically associated with Transformer models. Regarding the loss function, we employed the centerline Dice (clDice) loss function to better account for the tubular topology of the mandibular canal (Shit et al., 2021). As evidenced in Section 5, there was a 1% increase in the Dice coefficient, corroborating the effectiveness of this method in improving the segmentation of tubular structures. Figure 5 clearly demonstrates enhanced connectivity in the segmentation results, an outcome of clDice loss function's calculation of connectivity disparities between segmented outcomes and the extracted cartilage scaffolding. This quantification allows the network to focus more on ensuring connectivity in the segmentation results, thereby significantly enhancing the morphological integrity of the mandibular canal's tubular structure. Similar findings are reported in Huang et al. (2022) and Pan et al. (2021). Additionally, we leveraged sparse existing data to generate an augmented dataset through our proposed Deep Label Fusion technique. Compared to pre-training on the circle-extension dataset, our method resulted in a 1.6% increase in the Dice coefficient, reaching a score of 0.840. When validated on the new version of the ToothFairy dataset, the Dice coefficient was further improved to 0.844.

In the segmentation performance of CBCT images, our method achieved the highest performance on the public mandibular canal dataset (Cipriano et al., 2022b). Although this research work achieved the best segmentation results overall, there are still certain limitations. First of all, compared with the CNN network, the convergence speed of this network needs to be improved. Secondly, because the pixel changes of the mandibular canal at the mandibular and mental foramen are not obvious, the segmentation effect of the mandibular canal at the head and tail of the foramen is poor. Therefore, we will focus on the first and last features in future research to further improve the accuracy of the model.

7 Conclusion

In this study, we introduce a Transformer-based method for the robust segmentation of the mandibular canal. Our approach adeptly addresses key challenges, including morphological preservation of the mandibular canal, class imbalance, and ambiguous boundaries, subsequently achieving substantial improvements in segmentation metrics. Firstly, we employ Contrast-Limited Adaptive Histogram Equalization (CLAHE) to enhance image contrast, substantially ameliorating the low-contrast issues inherent to original CBCT scans. This step results in a notable increase in the model's segmentation accuracy. Secondly, we implement an image cropping strategy founded on mandibular foramen localization. This alleviates the class imbalance issue and substantially reduces extraneous background information, streamlining the segmentation process. Further, we introduce a specialized pixel-level feature fusion module known as the Deep Residual Convolution (DRC). This module not only amplifies the model's sensitivity to fine details in smaller targets such as the mandibular canal but also accelerates the convergence speed of the model, partially mitigating the known slow-convergence issue associated with Transformer architectures. To improve the topological integrity of the segmented mandibular canal, we utilize the centerline Dice (clDice) loss function. This forces the network to concentrate on maintaining the connectivity of the segmented structures, enhancing the morphological integrity of the mandibular canal. Lastly, we deploy a Deep Label Fusion technique to mine further information from the original, sparsely-annotated dataset. This step significantly bolsters the model's segmentation performance. Our method was rigorously evaluated on a publicly available mandibular canal dataset. The empirical results demonstrate that our proposed segmentation approach outperforms existing methods, underscoring its strong potential for application in the domain of mandibular canal segmentation.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

JL: Writing–original draft. LZ: Writing–original draft. JX: Writing–review and editing. WL: Writing–review and editing. GL: Writing–review and editing. HZ: Writing–review and editing.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN202201164), the Chongqing University of Technology Research and Innovation Team Cultivation Program (Grant No. 2023TDZ012), and Chongqing Graduate Student Research Innovation Project (Grant No. CYS23698).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdolali, F., Zoroofi, R. A., Abdolali, M., Yokota, F., Otake, Y., and Sato, Y. (2017). Automatic segmentation of mandibular canal in cone beam CT images using conditional statistical shape model and fast marching. *Int. J. Comput. assisted radiology Surg.* 12, 581–593. doi:10.1007/s11548-016-1484-2

Agbaje, J. O., de Casteele, E. V., Salem, A. S., Anumendem, D., Lambrichts, I., and Politis, C. (2017). Tracking of the inferior alveolar nerve: its implication in surgical planning. *Clin. Oral Investig.* 21, 2213–2220. doi:10.1007/s00784-016-2014-x

Al-Juboori, M., Al-Wakeel, H., SuWen, F., and Yun, C. M. (2014). Mental foramen location and its implication in dental treatment plan. *World J. Med. Med. Sci. Res.* 2 (3), 35–42.

Cao, H., et al. Swin-unet: unet-like pure transformer for medical image segmentation." Available: http://arxiv.org/abs/2105.05537.

Chen, J., et al. (2021). Transunet: transformers make strong encoders for medical image segmentation. *arXiv preprint arXiv:2102.04306.*

Cipriano, M., Allegretti, S., Bolelli, F., Di Bartolomeo, M., Pollastri, F., Pellacani, A., et al. (2022b). Deep segmentation of the mandibular canal: a new 3D annotated dataset of CBCT volumes. *IEEE Access* 10, 11500–11510. doi:10.1109/ACCESS.2022.3144840

Cipriano, M., Allegretti, S., Bolelli, F., Pollastri, F., and Grana, C. (2022a). "Improving segmentation of the inferior alveolar nerve through deep label propagation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*,China, IEEE, 21137–21146.

Cui, Z., Fang, Y., Mei, L., Zhang, B., Yu, B., Liu, J., et al. (2022). A fully automatic AI system for tooth and alveolar bone segmentation from cone-beam CT images. *Nat. Commun.* 13 (1), 2096. doi:10.1038/s41467-022-29637-2

Cui, Z., Li, C., and Wang, W. (2019). ToothNet: automatic tooth instance segmentation and identification from cone beam CT images *IEEE/CVF conference on computer vision and pattern recognition (CVPR).* Long Beach, CA, USA: IEEE, 6361–6370. doi:10.1109/CVPR.2019.00653

Dai, S., Zhu, Y., Jiang, X., Yu, F., Lin, J., Yang, D., et al. (2023). TD-Net: trans-Deformer network for automatic pancreas segmentation. *Neurocomputing* 517, 279–293. doi:10.1016/j.neucom.2022.10.060

Dosovitskiy, A., et al. (2020). *An image is worth 16x16 words: transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.*

Du, G., Tian, X., and Song, Y. (2022). Mandibular canal segmentation from CBCT image using 3D convolutional neural network with scSE attention. *IEEE Access* 10, 111272–111283. doi:10.1109/access.2022.3213839

Faradhilla, Y., Arifin, A. Z., Suciati, N., Astuti, E. R., Indraswari, R., and Widiasri, M. (2021). Residual fully convolutional network for mandibular canal segmentation. *Int. J. Intelligent Eng. Syst.* 14 (6), 208–219. doi:10.22266/ijies2021.1231.20

Fontenele, R. C., Gerhardt, M. d. N., Picoli, F. F., Van Gerven, A., Nomidis, S., Willems, H., et al. (2023). Convolutional neural network-based automated maxillary alveolar bone segmentation on cone-beam computed tomography images. Clin. Oral Implants Res. *P. Clr.* 34, 565–574. doi:10.1111/clr.14063

Hatamizadeh, A., et al. (2022a). "Unetr: transformers for 3d medical image segmentation," in *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, USA, IEEE, 574–584.

Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H., Xu, D., and Swin, UNETR: Swin transformers for semantic segmentation of brain tumors in MRI images, 2022b. Available: http://arxiv.org/abs/2201.01266

Huang, S., Li, J., Xiao, Y., Shen, N., and Xu, T. (2022). RTNet: relation transformer network for diabetic retinopathy multi-lesion segmentation. *IEEE Trans. Med. Imaging* 41 (6), 1596–1607. doi:10.1109/TMI.2022.3143833

Jaskari, J., Sahlsten, J., Järnstedt, J., Mehtonen, H., Karhu, K., Sundqvist, O., et al. (2020). Deep learning method for mandibular canal segmentation in dental cone beam computed tomography volumes. *Sci. Rep.* 10 (1), 5842. doi:10.1038/s41598-020-62321-3

Jeoun, B.-S., Yang, S., Lee, S. J., Kim, T. I., Kim, J. M., Kim, J. E., et al. (2022). Canal-Net for automatic and robust 3D segmentation of mandibular canals in CBCT images using a continuity-aware contextual network. *Sci. Rep.* 12 (1), 13460. doi:10.1038/s41598-022-17341-6

Kainmueller, D., Lamecker, H., Seim, H., Zinser, M., and Zachow, S. (2009). "Automatic extraction of mandibular nerve and bone from cone-beam CT data," in Medical image computing and computer-assisted intervention–MICCAI 2009: 12th international conference September 20-24, 2009, (London, UK: Springer), 76–83.

Kim, G., Lee, J., Lee, H., Seo, J., Koo, Y. M., Shin, Y. G., et al. (2010). Automatic extraction of inferior alveolar nerve canal using feature-enhancing panoramic volume rendering. *IEEE Trans. Biomed. Eng.* 58 (2), 253–264. doi:10.1109/TBME.2010.2089053

Kwak, G. H., Song, J. M., Park, H. R., Jung, Y. H., Cho, B. H., et al. (2020). Automatic mandibular canal detection using a deep convolutional neural network. *Sci. Rep.* 10 (1), 5711. doi:10.1038/s41598-020-62586-8

Lahoud, P., Diels, S., Niclaes, L., Van Aelst, S., Willems, H., Van Gerven, A., et al. (2022). Development and validation of a novel artificial intelligence driven tool for accurate mandibular canal segmentation on CBCT. *J. Dent.* 116, 103891. doi:10.1016/j.jdent.2021.103891

Li, J., Chen, J., Tang, Y., Wang, C., Landman, B. A., and Zhou, S. K. (2023). *Transforming medical imaging with Transformers? A comparative review of key properties, current progresses, and future perspectives*. China, Medical image analysis.102762

Li, Y., et al. (2022). *Association of the inferior alveolar nerve position and nerve injury: a systematic review and meta-analysis*. Germany, Healthcare, MDPI, 1782.

Liu, Z., et al. (2021). Swin transformer: hierarchical vision transformer using shifted windo. Available at: http://arxiv.org/abs/2103.14030.

Liu, Z., and Shen, L. (2022). Medical image analysis based on transformer: a review. *arXiv preprint arXiv:2208.06643*.

Moris, B., Claesen, L., Sun, Y., and Politis, C. (2012). Automated tracking of the mandibular canal in CBCT images using matching and multiple hypotheses methods *Fourth international conference on communications and Electronics (ICCE)*. Germany, IEEE, 327–332.

Onchis-Moaca, D., Zappalá, S., Goția, S. L., Real, P., and Pricop, M. (2016). Detection of the mandibular canal in orthopantomography using a Gabor-filtered anisotropic generalized Hough transform. *Pattern Recognit. Lett.* 83, 85–90. doi:10.1016/j.patrec.2015.12.001

Pan, L., Zhang, Z., Zheng, S., and Huang, L. (2021). MSC-net: multitask learning network for retinal vessel segmentation and centerline extraction. *Appl. Sci.* 12 (1), 403. doi:10.3390/app12010403

Pan, S., et al. (2022). "CVT-Vnet: convolutional-transformer model for head and neck multi-organ segmentation," in *Medical imaging 2022: computer-aided diagnosis* China, (SPIE), 914–921.

Rai, R., Shrestha, S., and Jha, S. (2014). Mental foramen: a morphological and morphometrical study. *Int. J. Health Biomed. Res.* 2, 144–150.

Shaker, A., Maaz, M., Rasheed, H., Khan, S., Yang, M.-H., and Khan, F. S. (2022). *UNETR++: delving into efficient and accurate 3D medical image segmentation*. *arXiv preprint arXiv:2212.04497*.

Shit, S., et al. (2021). "clDice-a novel topology-preserving loss function for tubular structure segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, IEEE, China, 16560–16569.

Usman, M., Rehman, A., Saleem, A. M., Jawaid, R., Byon, S. S., Kim, S. H., et al. (2022). Dual-stage deeply supervised attention-based convolutional neural networks for mandibular canal segmentation in CBCT scans. *Sens* 22 (24), 9877. doi:10.3390/s22249877

Verhelst, P.-J., Smolders, A., Beznik, T., Meewis, J., Vandemeulebroucke, A., Shaheen, E., et al. (2021). Layered deep learning for automatic mandibular segmentation in cone-beam computed tomography. *J. Dent.* 114, 103786. doi:10.1016/j.jdent.2021.103786

Waltrick, K. B., De Abreu Junior, M. J. N., Corrêa, M., Zastrow, M. D., and D'Avila Dutra, V. (2013). Accuracy of linear measurements and visibility of the mandibular canal of cone-beam computed tomography images with different voxel sizes: an *in vitro* study. *J. Periodontology* 84 (1), 68–77. doi:10.1902/jop.2012.110524

Wei, X., and Wang, Y. (2021). Inferior alveolar canal segmentation based on cone-beam computed tomography. *Med. Phys.* 48 (11), 7074–7088. doi:10.1002/mp.15274

Widiasri, M., Arifin, A. Z., Suciati, N., Fatichah, C., Astuti, E. R., Indraswari, R., et al. (2022). Dental-yolo: alveolar bone and mandibular canal detection on cone beam computed tomography images for dental implant planning. *IEEE Access* 10, 101483–101494. doi:10.1109/access.2022.3208350

Wu, Y., et al. (2022). Blood vessel segmentation from low-contrast and wide-field optical microscopic images of cranial window by attention-gate-based network IEEE/CVF conference on computer vision and pattern recognition workshops (CVPRW). Jun. 2022, New Orleans, LA, USA: IEEE, 1863–1872. doi:10.1109/CVPRW56347.2022.00203

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Systematic Review
PUBLISHED 22 November 2023
DOI 10.3389/fbioe.2023.1276338

Check for updates

OPEN ACCESS

EDITED BY
Run Zhang,
The University of Queensland, Australia

REVIEWED BY
Cosima Prahm,
University of Tübingen, Germany
Wen Qi,
Polytechnic University of Milan, Italy

*CORRESPONDENCE
Naqash Nasir,
naqash.nasir2@unibo.it
Laura Cercenelli,
laura.cercenelli@unibo.it

†These authors have contributed equally to this work and share first authorship

‡These authors have contributed equally to this work and share last authorship

RECEIVED 11 August 2023
ACCEPTED 03 November 2023
PUBLISHED 22 November 2023

CITATION
Nasir N, Cercenelli L, Tarsitano A and Marcelli E (2023), Augmented reality for orthopedic and maxillofacial oncological surgery: a systematic review focusing on both clinical and technical aspects.
Front. Bioeng. Biotechnol. 11:1276338.
doi: 10.3389/fbioe.2023.1276338

COPYRIGHT
© 2023 Nasir, Cercenelli, Tarsitano and Marcelli. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Augmented reality for orthopedic and maxillofacial oncological surgery: a systematic review focusing on both clinical and technical aspects

Naqash Nasir[1,2*†], Laura Cercenelli[3*†], Achille Tarsitano[1,4‡] and Emanuela Marcelli[3‡]

[1]Oral and Maxillo-Facial Surgery Unit, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy, [2]Orthopedics and Traumatology Department, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy, [3]eDIMES Lab-Laboratory of Bioengineering, Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy, [4]Maxillofacial Surgery Unit, Department of Biomedical and Neuromotor Science, University of Bologna, Bologna, Italy

This systematic review offers an overview on clinical and technical aspects of augmented reality (AR) applications in orthopedic and maxillofacial oncological surgery. The review also provides a summary of the included articles with objectives and major findings for both specialties. The search was conducted on PubMed/Medline and Scopus databases and returned on 31 May 2023. All articles of the last 10 years found by keywords augmented reality, mixed reality, maxillofacial oncology and orthopedic oncology were considered in this study. For orthopedic oncology, a total of 93 articles were found and only 9 articles were selected following the defined inclusion criteria. These articles were subclassified further based on study type, AR display type, registration/tracking modality and involved anatomical region. Similarly, out of 958 articles on maxillofacial oncology, 27 articles were selected for this review and categorized further in the same manner. The main outcomes reported for both specialties are related to registration error (i.e., how the virtual objects displayed in AR appear in the wrong position relative to the real environment) and surgical accuracy (i.e., resection error) obtained under AR navigation. However, meta-analysis on these outcomes was not possible due to data heterogenicity. Despite having certain limitations related to the still immature technology, we believe that AR is a viable tool to be used in oncological surgeries of orthopedic and maxillofacial field, especially if it is integrated with an external navigation system to improve accuracy. It is emphasized further to conduct more research and pre-clinical testing before the wide adoption of AR in clinical settings.

KEYWORDS

augmented reality, mixed reality, orthopedics oncology, maxillofacial oncology, surgery, tracking

1 Introduction

Augmented reality (AR) is a technology that allows the fusion of digital content into the real environment. The achieved augmented continuum is a virtual world in which virtual objects are overlaid on real elements, in the surrounding actual environment (Azuma et al., 2001).

The first AR system using a Head Mounted Display (HMD) was developed by Sutherland in 1968 (Feiner, 2002). Since its discovery, AR technology has been utilized by experts in many areas; such as entertainment, sports, gaming, retail, and also medicine. Indeed, the recent technological advancements in headsets and computer hardware resulted in many companies, especially in the entertainment sector, investing in AR devices which have become increasingly available and accessible. Therefore employed also in health-related applications particularly in the surgical fields.

When applied to surgery, the AR allows to improve the user's perceptual and comprehensive ability by projecting three-dimensional underlying anatomy directly onto the user's retina (via HMDs) or on a display screen.

AR in surgery has enormous potential to help the surgeon in identifying tumor locations, delineating the planned dissection planes, and reducing the risk of injury to invisible structures. Therefore, using AR in the operating room (OR) could be helpful in performing surgical tasks in a more accurate way. HMDs are particularly beneficial for AR surgical applications since they intrinsically provide the surgeon with an egocentric viewpoint, and offer improved ergonomics if compared to traditional computer-assisted surgical systems. This allows surgeons to concentrate on the task at hand without having to turn their heads away from the surgical field to constantly look at imaging monitor. The most ambitious goal in surgery is to use AR for intraoperative navigation. This involves taking data from preoperative imaging and using anatomical anchors in the operating field to register the two representations in real time.

Registration is an important step in computer-assisted surgical navigation in order to correlate the virtual content and the real surgical scene. In this context, the registration error can be defined as; the measurement of how much the virtual objects displayed in AR appear incorrectly positioned relative to the real environment.

For virtual-to-real surgical scene registration, AR systems typically use a camera coupled to a device marker; such as QR code, anchored to the patient (marker-based registration). Another option is marker-less registration which includes a combination of location data (from Global Positioning System), inertial measurement unit (IMU) data, and computer vision to track image features such as scene depth, the object surface, and object edges (Venkatesan et al., 2021).

The core of the registration modality is tracking, which means to determine and follow the position and orientation of an object with respect to some reference coordinate system over time.

Over the past decade, with the advent of multimodal and high-detailed 4D medical imaging (Bradley, 2008), numerous surgical specialties have integrated AR into their surgical workflow, namely; neurosurgery (Cannizzaro et al., 2022), urological surgery (Bianchi et al., 2021; Schiavina et al., 2021; Roberts et al., 2022), ophthalmology (Li et al., 2021), gastrointestinal endoscopy (Mahmud et al., 2015), cardiovascular surgery (Rad et al., 2022), spinal surgery (Molina et al., 2021a), breast surgery (Gouveia et al., 2021), and thyroid surgery (Lee et al., 2020). Some authors utilized AR to perform lateral skull-based surgery for cerebellopontine angle tumor (Schwam et al., 2021) and some used it in open hepatic surgery (Golse et al., 2021). Moreover, AR has also been employed in procedures such as perforator flap transfer (Jiang et al., 2020) and percutaneous nephrolithotomy (Ferraguti et al., 2022).

Orthopedic and Maxillofacial surgeries have been pioneers in the use of AR in a surgical setting (Barcali et al., 2022).

These two surgeries may represent very promising fields for the future clinical implementation of AR, since they are based on bony hard tissues which make it easier to have fixed references, i.e., bony structures, to be used for ensuring an accurate virtual-to-real scene registration between preoperative (virtual) and intraoperative (real) views.

Regarding orthopedic surgery, Alexander et al. (2020) formulated a 3D augmented reality system for the placement of acetabular component during total hip arthroplasty (THA) and found it to be more precise and faster than standard fluoroscopic guidance. Similarly, Ogawa et al. (2018) found AR to be more accurate when comparing it to conventional goniometer for acetabular cup placement during THA. In 2019, Tsukada et al. (2019) conducted an *in vitro* study on sawbone models for employing AR during total knee arthroplasty and concluded that the system provided accurate measurements for tibial bone resection. Consequently, in 2021, the same authors, formulated prospective cohort study on 72 patients. They emphasized that AR-assisted navigation to resect distal femur is more precise than the conventional method (Tsukada et al., 2021).

Augmented reality and its tools have emerged as a new paradigm also in spinal surgeries. Many authors have validated the use of AR navigation for the precise placement of pedicle screw (Elmi-Terander et al., 2018; Elmi-Terander et al., 2019; Gibby et al., 2019; Dennler et al., 2020) and some compared its accuracy with free-hand approach (Elmi-Terander et al., 2020). In 2021, Molina et al. (2021b) conducted the first human trial of using an FDA approved AR-HMD (X-vision Spine System, Augmedics) and demonstrated its clinical and technical accuracy in spine surgery.

In the context of oral and cranio-maxillofacial surgery, AR applications are of increasing interest and adoption (Badiali et al., 2020).

Sharma et al. (2021) proposed a marker-less AR navigation system algorithm with greater precision and faster processing time for jaw surgery. Similar to the article on marker-less image registration for jaw experiments published by Wang et al. (2019), this study demonstrated its clinical viability through minimal registration error and processing time.

Some other experiences of marker-less AR navigation have been reported for assisting the harvesting of periosteum pedicle flap and osteomyocutaneous fibular flap in head and neck reconstruction (Battaglia et al., 2020), as well as for guiding osteotomies in pediatric cranio-facial surgery (Ruggiero et al., 2023).

Similarly, recent studies in dental implantology have demonstrated the efficacy of AR for displaying dynamic navigation systems (Pellegrino et al., 2019; Shrestha et al., 2021). Ma et al. (2019) proposed an AR-assisted navigation with cone beam computed tomography (CBCT) registration method to attain the desired dental implant precision. They compared the navigation method to physician's experience and concluded that AR guidance had better outcomes in terms of mean target error and mean angle error. Budhathoki et al. (2020) emphasized the use of AR navigation to visualize deep-seated anatomy, narrow areas and to provide positioning of surgical instruments to avoid positioning error complications during jaw surgery. Moreover, Gao et al. (2019) employed AR in mandibular split osteotomy. Same as, Pietruski et al. (2019) who incorporated AR navigation and cutting guides for mandibular osteotomies in 2019 and concluded that this technology can enhance the surgeon's perception and hand-eye coordination during mandibular resection and reconstruction procedures.

Although AR technology has a long history in orthopedic and maxillofacial surgery, a complete analysis of its clinical and technical application on oncological cases is still lacking.

In this literature review, we provide the comprehensive up-to-date overview on current clinical applications of AR in orthopedic oncology and maxillofacial oncology, pointing out its benefits and current limitations. Moreover, we also elucidate the different technological aspects of AR used in each of these experiences to give an insight on how AR can be administered in oncological surgical scenarios.

2 Materials and methods

2.1 Searching criteria

This systematic literature review was conducted on PubMed/Medline and Scopus databases using the terms, "Augmented Reality" AND "Orthopedic oncology," "Mixed Reality" AND "Orthopedic oncology," "Augmented Reality" AND "Maxillofacial" AND "oncology," "Mixed Reality" AND "Maxillofacial" AND "oncology" and "Augmented Reality" AND "Head and Neck" AND "cancer." The same search was also attempted using the term "Cranio-maxillofacial" AND "oncology" in the place of "Maxillofacial" AND "oncology," from which, however, no additional results were obtained with respect to what was already found. Searches for both specialty domains were done separately and 2 independent users performed search until 31 May 2023. Relevant articles of only last 10 years were included in this review paper. Manual search was also done in references of papers to see missing of any relevant paper. PRISMA-guidelines were kept in mind while preparing this review article.

The SPIDER (Sample, Phenomena of Interest, Design, Evaluation, Research type) method was used to construct the suitable research question: "Can augmented reality be considered a beneficial tool in orthopedic and maxillofacial oncological fields in achieving surgical accuracy?"

This review addresses this question by focusing on both clinical and technical aspects of AR in these two surgical disciplines, as well as on reporting current limitations and benefits.

Due to the qualitative and mixed-method nature of the included articles and the heterogeneity of the data, the term "evaluation" was left intentionally broad.

We performed the study selection based on the following inclusion criteria: studies on augmented reality applications in either orthopedic or maxillofacial field, only focused on oncology; studies reporting applications on different targets (e.g., phantoms, cadavers, animals and patients). All articles with either quantifiable or qualitative outcomes on augmented reality confined to both specialties with case reports were included.

Exclusion criteria were the following: articles "not relevant" (i.e., not related to augmented reality, not strictly related to oncological surgery, not related to orthopedic or maxillofacial surgery); articles with language of publication other than English; theses; conference papers; editorials; book chapters; review articles (as Review articles typically do not include sufficient specifics regarding the recommended solutions and are also considered as secondary source, therefore they cannot be used in data extraction process).

2.2 Data extraction and analysis

All the search articles available till 31 May 2023 were screened by title first and then abstract.

The authors, date of publication, study design, and data from the eligible articles were tabulated in Microsoft Excel® (Microsoft Corporation, WA). All articles meeting the inclusion criteria were read carefully and stratified following two parallel perspectives: a clinical one, i.e., focusing on the specific surgical application, the type of study (on phantoms, on cadavers, on animals, on humans), the anatomical region of interest, the virtual information provided to surgeon, and a technical one, i.e., the registration/tracking modality the type of AR display, the achieved registration and surgical accuracy).

Key findings of each article were also stated in given tables in the Results section for both orthopedic and maxillofacial specialties, and also depicted in bar histograms. However, meta-analysis could not be performed due to heterogeneity of literature. All these findings were validated by a second independent investigator to ensure the correct data acquisition and selection of the appropriate relevant literature.

Due to the fact that the included study designs exhibited a significant level of variability, as is often observed in the case of new technologies, they are developed individually with distinct features. Consequently, conventional approaches for evaluating the risk of bias were not suitable for use in this context. The authors generally evaluated and assessed the risk of bias to be low or negligible for data description, but it could be high for the analysis of the effectiveness of approaches used in these studies. Furthermore, none of the articles included in the review refers to a specific methodological protocol.

3 Results

The initial search of the PubMed/Medline and Scopus databases was completed on 31 May 2023, and all available articles were scrutinized using the above-described criteria.

In the following paragraphs, an analytic overview of the selected papers and their classification were presented for both orthopedic oncology and maxillofacial oncology.

As depicted in flowchart (Figure 1), the databases search for orthopedic oncology returned 89 results while manual search yielded 4 publications. Fifty-nine articles remained after removing duplicates based on titles and abstracts. According to the inclusion criteria, only 9 of the 59 articles were included in this review. Other articles were excluded for the following reasons: "not-relevant" (n = 27), review articles/book chapters/editorials/conference papers (n = 22), non-English articles (n = 1).

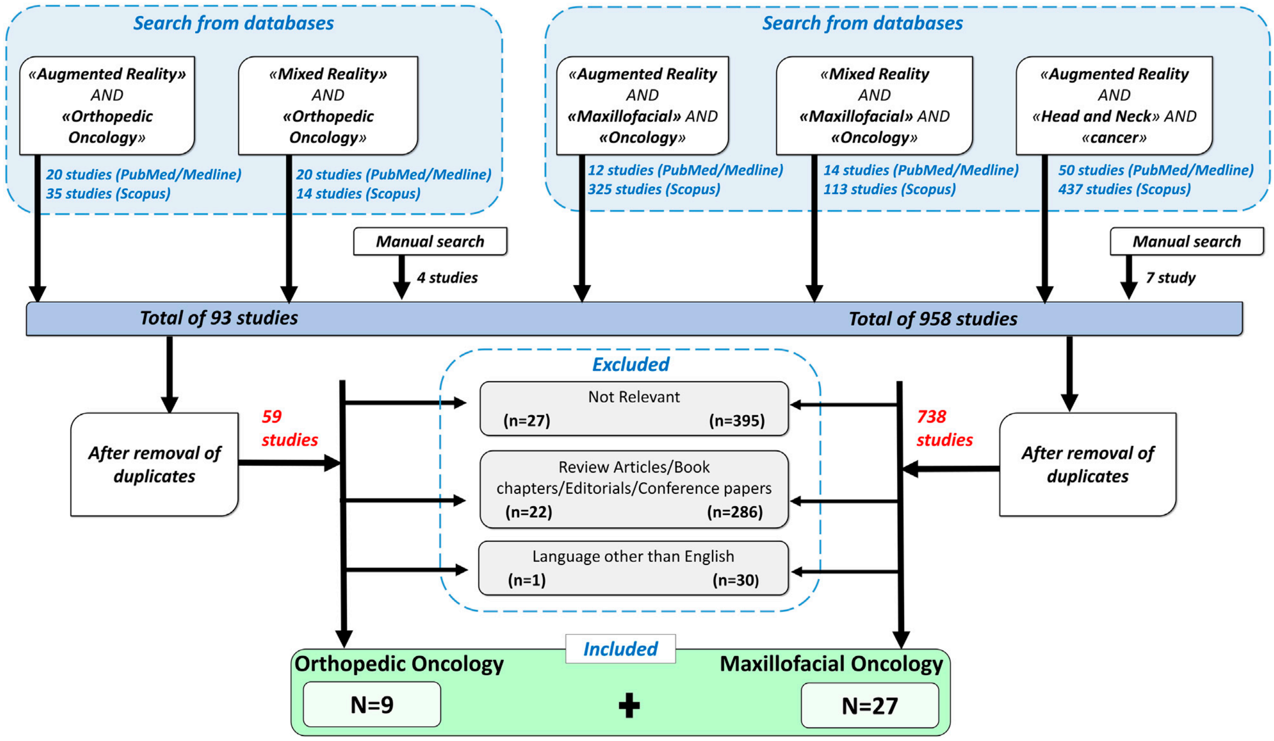


FIGURE 1
A flowchart showing inclusion and exclusion criteria used for the search, and the resulting selected papers.

Similarly, for CMF oncology, 951 articles were found through PubMed/Medline and Scopus search and 7 from manual search. Out of total 958, 738 articles remained after removal of duplicates. Based on above defined criteria, 27 publications were included and others were eliminated for the following reasons: "not-relevant" ($n = 395$), review articles/book chapters/editorials/conference papers ($n = 286$), non-English language ($n = 30$).

For the included articles, both clinical and technical aspects are summarized separately in given tables for both specialties (Tables 1–4). For the clinical aspects, we classified the papers according to: the specific surgical application, the number of cases involved, the involved anatomical region, the type of study (i.e., on phantom, on cadaver, on patient), the virtual information provided to augment the surgeon's view (Tables 1, 3). For the technical aspects, we considered the type of AR display and device, the used registration/tracking modality, as well as the achieved registration error and surgical accuracy (measured in mm) (Tables 2, 4). For each included article we provided in a separate table a brief description of the study and the major outcomes (Tables 5, 6).

We also show in bar histograms the breakdown of the studies according to some of the most interesting aspects above mentioned (Figures 2, 3).

3.1 Orthopedic oncology

For orthopedic oncology (Cho et al., 2017; Choi et al., 2017; Cho et al., 2018; Moreta-Martinez et al., 2018; Abdel Al et al., 2020; García-Sevilla et al., 2021a; Molina et al., 2021c; Moreta-Martinez et al., 2021; Pose-Díez-de-la-Lastra et al., 2022), four studies were conducted on phantoms ($n = 1$) or both on phantoms and prospectively on patients ($n = 3$). Three studies were conducted on animal cadavers, whereas, two studies directly employed AR on patients in OR. In the context of AR display and registration/tracking modality, five studies used screen-based display and eight studies opted for a marker-based registration/tracking modality. Regarding the anatomical region of interest, pelvis is the most involved anatomical region in the selected studies ($n = 4$).

The main outcomes reported in the included articles referred to the registration error of the AR systems, the AR-guided surgical accuracy in performing tumor resection compared to preoperative planning and/or to standard procedures (i.e., manual measurements), the placement error in positioning surgical guides or patient-specific implant under AR assistance.

Despite the fact that the application of AR in the field of orthopedic oncology has been relatively limited compared to maxillofacial oncology to date, the included studies demonstrate the potential future significance of AR technology in this surgical field.

3.2 Maxillofacial oncology

In the field of maxillofacial oncology (Scolozzi and Bijlenga, 2017; Battaglia et al., 2019; Gsaxner et al., 2019; Pepe et al., 2019; Kim et al., 2020; García-Sevilla et al., 2021b; Gsaxner et al., 2021; Meng et al., 2021; Ochandiano et al., 2021; Sahovaler et al., 2021; Scherl et al., 2021; Sugahara et al., 2021; Tel et al., 2021; Ceccariglia

TABLE 1 Distribution of studies by clinical aspects of AR in Orthopedic Oncology.

Reference	Authors	Year	Target of study	Virtual content	Specific surgical application	Anatomical region	No. of cases
Moreta-Martinez et al. (2021)	Moreta et al.	2021	Phantom and Patient	Bones, tumors, cutting planes	2 myxofibrosarcoma,1 liposarcoma, 1 Ewings sarcoma, 1 fibrous dysplasia, 1 undifferentiated pleomorphic sarcoma	Femur, thigh, calf, pelvis, shoulder	6
Cho et al. (2017)	Choi et al.	2017	Animal cadaver	Coloured template of tumor with normal bone and safety margin	Simulated femur tumor	Femur	Total = 123 AR = 82 con = 41
Cho et al. (2018)	Choi et al.	2018	Animal cadaver	Tumor, surgical plane and safety margin	Simulated pelvic tumor	Pelvis	Total = 36 AR = 18, con = 18
Choi et al. (2017)	Choi et al.	2017	Animal cadaver	Tumor, safety margin, resection and saw plane	Simulated pelvic tumor	Pelvis	Total = 60, AR = 30, con = 30
Molina et al. (2021c)	Molina et al.	2021	Patient	Tumor and osteotomy trajectory	L1 chondroma	Spine	1
Abdel Al et al. (2020)	Abdel et al.	2020	Patient	MRI images with tumor	Soft tissue sarcoma	Foot	1
García-Sevilla et al. (2021a)	Garcia-Sevilla et al.	2021	Phantom	Bones and PSI	Simulated Pelvic tumor	Pelvis	6
Moreta-Martinez et al. (2018)	Moreta et al.	2018	Phantom and Patient	Skin, bone and tumor	Ewings sarcoma	Tibia-fibula	1
Pose-Díez-de-la-Lastra et al. (2022)	Pose-Díez-de-la-Lastra et al.	2022	Phantom and Patient	Bone, tumor and surgical guide	Extraosseous Ewing's sarcoma and Undifferentiated pleomorphic sarcoma	Tibia-fibula, shoulder	2

et al., 2022; Cercenelli et al., 2022; Chan H et al., 2022; Gao et al., 2022; Han et al., 2022; Modabber et al., 2022; Shi et al., 2022; Tang et al., 2022; Winnand et al., 2022; Yang et al., 2022; Necker et al., 2023; Prasad et al., 2023; Shaofeng et al., 2023; Zhao et al., 2023), in the majority of articles ($n = 14$) the proposed AR systems were tested on phantoms, and 8 on patients only. Five researcher groups carried out pre-clinical research on phantoms before using AR on patients. The marker-less registration approach ($n = 14$) and wearable AR displays, i.e., HMDs ($n = 14$), were utilized in the majority of the research studies. These two investigating factors made up 52% of the total studies. Mandible, being the most involved anatomical area of interest for AR implementation, comprises of 52% of total studied areas ($n = 14$).

The following are the primary findings that are covered in the articles: the registration error, the AR-guided surgical accuracy in performing tumor resection or flap harvesting compared to preoperative planning, and the AR-guided surgical accuracy compared to the more conventional use of 3D printed cutting guides. It is interesting to mention that, some studies in maxillofacial oncology have also incorporated external tracking navigation systems into augmented reality to improve accuracy and spatial relationships. The articles included have shown the importance of AR and its future perspectives in this field. Nevertheless, the majority of the studies underlined the need for additional research before clinical application.

4 Discussion

In the set of orthopedic oncology articles, all studies utilized marker for registration except 1 (Abdel Al et al., 2020), which opted for marker-less registration. Mean registration error found to be less than 3 mm where measured (not all articles reported the registration error) excepting the one case of complex shoulder phantom (Pose-Díez-de-la-Lastra et al., 2022), where RMSE was slightly more than 3 mm with HoloLens 2.

Out of nine articles, only three randomized controlled trial studies (RCT) were reported (performed on animal cadavers). Mean AR-assisted resection was less than 2 mm as compared to conventional approach (e.g., manual resection) which had mean resection error more than 2.5 mm in these studies (Cho et al., 2017; Choi et al., 2017; Cho et al., 2018).

Included studies of orthopedic oncology also reported two cases, both achieved intended outcomes using AR without any complications (Abdel Al et al., 2020; Molina et al., 2021c). Moreover, one study demonstrated a better positioning of surgical guide in simulated pelvic tumor through AR as compared to freehand method (García-Sevilla et al., 2021a).

Regarding usability, only one study conducted a questionnaire survey with patients and surgeons, and the results turned out to be satisfactory (Moreta-Martinez et al., 2021). Ergonomics of the most popular AR devices, HoloLens 1 and 2, were discussed by Pose-Díez-

TABLE 2 Distribution of studies by technical aspects of AR in orthopedic oncology.

Reference	Authors	Year	AR display type	AR device	Registration/ Tracking modality	No. of cases	AR registration error	AR surgical accuracy
Moreta-Martinez et al. (2021)	Moreta et al.	2021	Screen-based	Smartphone iphone 6	Marker-based (cubic marker)	6	Mean registration error: 2.80 ± 0.98 mm	n.a
Cho et al. (2017)	Choi et al.	2017	Screen-based	Tablet (Surface Pro 3)	Marker-based (ArUCo code)	Total = 123 AR = 82 conv = 41	n.a	Mean resection error (difference between the obtained and the planned surgical margin): AR-assisted resection: 1.71 ± 0.25 mm; Conventional resection (manual measurement): 2.64 ± 0.5 mm
Cho et al. (2018)	Choi et al.	2018	Screen-based	Tablet (Surface Pro 3)	Marker-based (ArUCo code)	Total = 36 AR = 18 conv = 18	n.a	Mean resection error (difference between the obtained and the planned surgical margin):AR-assisted resection: 1.59 ± 4.13 mm; Conventional resection (based on CT image): 4.55 ± 9.7 mm
Choi et al. (2017)	Choi et al.	2017	Screen-based	Tablet (Surface Pro 3)	Marker-based (ArUCo code)	Total = 60 AR = 30 conv = 30	Mean registration error: 0.58 ± 0.22 mm	Mean resection error (i.e., difference between the obtained and the planned surgical margin): AR-assisted resection: 0.15 ± 1.02 mm; Conventional resection (manual measurement): 2.89 ± 4.30 mm
Molina et al. (2021c)	Molina et al.	2021	HMD	X-Vision Augmedics	Marker-based	1	n.a	n.a
Abdel Al et al. (2020)	Abdel el at	2020	screen-based	Smartphone (Samsung Galaxy A5)	Marker-less	1	n.a	n.a
García-Sevilla et al. (2021a)	Garcia-Sevilla et al.	2021	Screen-based/ HMD	Smartphone/ HoloLens 2	Marker-based	6	n.a.	Median MOD (Maximum Osteotomy Deviation) between planned and real osteotomy planes (for realistic phantom): AR-assisted positioning of surgical guide (smartphone): 1.54 mm; AR-assisted positioning of surgical guide (HoloLens2): 1.84 mm; Freehand positioning: 3.37 mm
Moreta-Martinez et al. (2018)	Moreta et al.	2018	HMD	HoloLens 1	Marker-based	1	(Phantom): Registration error (Root Mean Square Error (RMSE) in AR point localization): 2.90 mm (Clinical): n.a.	n.a.

(Continued on following page)

TABLE 2 (*Continued*) Distribution of studies by technical aspects of AR in orthopedic oncology.

Reference	Authors	Year	AR display type	AR device	Registration/ Tracking modality	No. of cases	AR registration error	AR surgical accuracy
Pose-Díez-de-la-Lastra et al. (2022)	Pose-Díez-de-la-Lastra et al.	2022	HMD	HoloLens 1 and 2	Marker-based	2	Registration error (RMSE in AR point localization): −2.16 mm (HoloLens2–Leg phantom); −2.83 mm (HoloLens1–Leg phantom); −3.11 mm (HoloLens2–shoulder phantom);	n.a.

de-la-Lastra et al. (2022) while comparing them on two orthopedic cases. They indicated that HoloLens 2 has superior ergonomics (score 4 out of 5) as compared to HoloLens 1 (score 2.84 out of 5).

The primary outcomes of the included articles in orthopedic segment demonstrated the utility of AR and showed even better results than conventional procedures in some cases. However, some of the limitations of the orthopedic segment is that few articles discuss the usability and ergonomics of this technology. Only three out of nine articles are randomized controlled trials (RCTs), making it difficult to derive a statistical comparison value from the current studies. Furthermore, only four articles provided surgical accuracy (Cho et al., 2017; Choi et al., 2017; Cho et al., 2018; García-Sevilla et al., 2021a) and only four provided registration error (Choi et al., 2017; Moreta-Martinez et al., 2018; Moreta-Martinez et al., 2021; Pose-Díez-de-la-Lastra et al., 2022) showing the limitation of this research.

Over all, with consideration of sufficient accuracy achieved in terms of registration error and surgical accuracy during surgical simulations, it can be said that AR is beneficial and helpful in orthopedic oncological surgeries. However, due to limited literature in orthopedic oncology till date, further testing is recommended.

In maxillofacial oncology, out of eight marker-based studies, only two reported AR regisration error in terms of fiducial registration error (<1 mm) (Sahovaler et al., 2021; Chan H et al., 2022). Marker-less registration error (where measured in articles) ranges from less than 1 mm to upto 2 cm (Gsaxner et al., 2019; Gsaxner et al., 2021; Cercenelli et al., 2022; Shi et al., 2022; Yang et al., 2022). This shows that some articles did not achieve the desired precision in AR registration through the marker-less approach.

Marker-based AR surgical accuracy was reported in five articles and it was demonstrated in terms of deviation from pre-planned surgical resection. Mean surgical resection error was measured to be less than 3 mm among these articles (Kim et al., 2020; García-Sevilla et al., 2021b; Chan H et al., 2022; Gao et al., 2022; Zhao et al., 2023). Conversely, seven articles reported surgical accuracy using the marker-less approach. The accuracy ranges from 0.49 to 2.77 mm in six articles (Ceccariglia et al., 2022; Cercenelli et al., 2022; Modabber et al., 2022; Shi et al., 2022; Tang et al., 2022; Winnand et al., 2022), whereas, one study had a maximum surgical error of 6.08 mm which did not meet surgical requirement (Yang et al., 2022).

Two studies employed both marker-less and marker-based registration, the latter being used mainly in the case of lack of surface features easily recognizable (Shaofeng et al., 2023) or to introduce a calibration procedure aimed at improving registration accuracy (Pepe et al., 2019). Particularly in (Shaofeng et al., 2023) the surgical accuracy, in terms of distance deviation between the planned osteotomies and postoperative cuts performed under AR guidance, was measured in the range of 0.88–2.01 mm. Manual alignment for virtual-to-real registration was performed in five articles (Meng et al., 2021; Scherl et al., 2021; Han et al., 2022; Necker et al., 2023; Prasad et al., 2023) and only two of them reported errors in terms of overlaying AR images. One study showed average relocation error of 4 mm ± 3.9 mm with HoloLens 2 (Prasad et al., 2023), whereas the other measured mean registration error of 1.3 cm with HoloLens 1 (Scherl et al., 2021). Even though it is hard to comment based on only two results, it seems that chances of error are more in manual registration as compared to marker-less or marker-based approach.

Two of the three case reports on maxillofacial oncological surgery were qualitative, while one case of fibrous histiocytoma had quantified findings with an overall mean discrepancy of 2.77 ± 1.29 mm using AR (Kim et al., 2020). According to qualitative studies, the incorporation of mixed reality during pre-surgical and intra-operative phases allows for precise surgical outcomes and is helpful for lesion identification and determination of its extension (Scolozzi and Bijlenga, 2017; Sugahara et al., 2021).

A total of seven studies have examined the utilization and precision of AR in the context of flap harvesting for mandibular restoration (Battaglia et al., 2019; Meng et al., 2021; Cercenelli et al., 2022; Han et al., 2022; Modabber et al., 2022; Winnand et al., 2022; Zhao et al., 2023). Among these, two of them conducted a comparison between marker-less AR guidance and cutting guides in the context of iliac crest harvesting (Modabber et al., 2022; Winnand et al., 2022). The results indicated that cutting guides exhibited superior precision compared to AR navigation in terms of, both distance and angular deviation from pre-determined trajectories. However, the distance deviations were less than 2.7 mm in AR group. In contrast, Zhao et al. (2023) employed a marker-based methodology to assess the fibular flap, yielding an average distance deviation of 1.22 ± 0.12 mm. One possible explanation for this phenomenon is that the marker-based approach tends to exhibit more accuracy in comparison to the

TABLE 3 Distribution of studies by clinical aspects of AR in maxillofacial oncology.

Reference	Authors	Year	Target of study	Virtual content	Specific surgical application	Anatomical region	No. of cases
Chan H et al. (2022)	Chan et al.	2022	Phantom	Tumors and cutting planes	Simulated Maxillary tumors	Maxilla	5
Sahovaler et al. (2021)	Sahovaler et al.	2021	Phantom	Tumors and cutting planes	Simulated Sinonasal tumors	Sinonasal	4
Shi et al. (2022)	Jiafeng shi et al.	2022	Phantom	Mandible and osteotomy line	Mandibular tumor	Mandible	1
Ceccariglia et al. (2022)	Ceccariglia et al.	2021	Patient	Skin, tumor with surrounding normal bone and cutting planes	1 maxillary squamous cell carcinoma,1 osteomyelitis of left jaw, 1 osteomyelitis of left mandible	Maxilla, mandible	3
Ochandiano et al. (2021)	Ochandiano et al.	2022	Phantom and Patient	Mandible with teeth, implant position and angulation, splint (tracker)	8 mandibular, 1 maxilla,1 tongue and 1 hard palate	Mandible, maxilla, tongue, hard palate	7
Gsaxner et al. (2019)	Gsaxner et al.	2019	Phantom and Patient	Bone and tumor mass	Head and neck cancers (not specified)	Head and neck (not specified)	8
Pepe et al. (2019)	Pepe et al.	2019	Phantom	Tumor, reference markers, slice of PET-CT scan	Simulated Head and neck tumor	Head and neck (not specified)	1
García-Sevilla et al. (2021b)	García-Sevilla et al.	2022	Phantom and Patient	Bone, tumor, surgical resection margin, splint for tracking	Adenoid cystic carcinoma	Hard palate	1
Tel et al. (2021)	Tel et al.	2021	Phantom and Patient	Skull with maxillary sinuses, muscles and fat, tumor, infraorbital nerve, optic nerve	Inferior orbital compartment tumors (3 cavernous hemangioma, 1 neurofibroma, 1 schwannoma)	Orbit	5
Gsaxner et al. (2021)	Gsaxner et al.	2021	Phantom	Skull and tumor, orthogonal slices of imaging in anatomical planes	Head and neck carcinoma (not specified)	Head and neck (not specified)	1
Sugahara et al. (2021)	Sugahara et al.	2021	Patient	Tumor, nasal cavity and maxillary sinus with surrounding normal skull	Maxillary calcifying odontogenic cyst	Maxilla	1
Shaofeng et al. (2023)	Shaofeng Liu et al.	2023	Phantom and Animal	Mandible, teeth, resection lines, surgical saw, fixation screw	Bilateral mandibular simulated tumor	Mandible	9 model trials, 12 animal trials
Prasad et al. (2023)	Prasad BA et al.	2023	Human Cadaver	Simulated cancer	20 different simulated Head and neck cancers	Head and neck (not specified)	20
Gao et al. (2022)	Gao et al.	2022	Phantom	Skull and virtual recontouring plan	Craniofacial fibrous dysplasia	Maxilla	5
Kim et al. (2020)	Jin Kim et al.	2020	Patient	Skull, tumor and cutting planes	Malignant fibrous histiocytoma of the maxilla	Maxilla	1
Yang et al. (2022)	Yang et al.	2022	Patient	Mandible with tumor and surgical planes	Mandibular tumor	Mandible	4
Tang et al. (2022)	Tang et al.	2022	Patient	Skull, tumor, surgical planes, probe	Maxillary and mandibular tumor	Maxilla, mandible	7
Scolozzi and Bijlenga (2017)	Scolozzi et al.	2017	Patient	Skull and tumor	Pleomorphic adenoma of lacrimal gland	Orbit	1
Cercenelli et al. (2022)	Cercenelli et al.	2022	Phantom	Bone, vessels, skin, osteotomy line	Skin paddle harvesting in osteomyocutaneous fibular flap	Mandible	1
Scherl et al. (2021)	Scherl et al.	2021	Patient	Surface of the face, mandible, masseter, parotid gland, tumor	Parotid tumor	Parotid Gland	6

(Continued on following page)

TABLE 3 (*Continued*) Distribution of studies by clinical aspects of AR in maxillofacial oncology.

Reference	Authors	Year	Target of study	Virtual content	Specific surgical application	Anatomical region	No. of cases
Necker et al. (2023)	Necker et al.	2022	Human Cadaver	Mandible, tumor	Mandibular tumor	Mandible	1
Han et al. (2022)	Lin et al.	2022	Phantom and Patient	Arteries, veins, bone tissues	Mandibular ameloblastoma	Mandible	1
Zhao et al. (2023)	Zhao et al.	2022	Cadaver	Fibula, osteotomy lines	Mandible (application: fibular flap harvesting)	Mandible	7
Modabber et al. (2022)	Modabber et al.	2021	Cadaver	Iliac crest with planned osteotomies	Mandibular tumor (application: iliac crest harvesting)	Mandible	10
Meng et al. (2021)	Meng et al.	2021	Cadaver	Fibula, fibular flap with cutting planes	Mandible (application: fibula flap harvesting)	Mandible	1
Battaglia et al. (2019)	Battaglia et al.	2021	Patient	Skin, fibula, cutting guides, arteries	Mandible (application: fibular flap harvesting)	Mandible	3
Winnand et al. (2022)	Winnand et al.	2022	Phantom	Iliac crest with planned osteotomies	Mandible (application: iliac crest graft harvesting)	Mandible	10

marker-less approach in relation to lower limb bones. This might be attributed to the specific shape and contour of these bones, which pose challenges for marker-less registration techniques.

It is interesting to outline that five of the studies utilized external navigator system with mixed reality. Four of them incorporated external navigation into mixed reality to enhance accuracy and spatial relationships (Scolozzi and Bijlenga, 2017; Gao et al., 2022; Tang et al., 2022; Yang et al., 2022) and one study compared AR and optical tracking system (OTS) accuracy (García-Sevilla et al., 2021b).

García-Sevilla et al. (2021b) compared AR and OTS for surgical navigation and concluded that they have similar accuracy with errors below 1 mm. Gao et al. (2022) used HoloLens and OTS on five patients of cranio-fibrous dysplasia and the mean registration error across all cranium models was 1.036 ± 0.081 mm. Tang et al. (2022) conducted a study to evaluate efficacy and accuracy of mixed reality, which is augmented by surgical navigator, on seven patients of maxilla and mandibular tumors. The mean deviation from pre-defined osteotomy plane was 1.68 ± 0.92 mm (set target error was 2 mm). However, in the study by Yang et al. (2022), cross linking of mixed reality and optical navigator did not produce clinically required accuracy (maximum error was 6.08 mm), but authors highlighted that this system enhanced spatial experience and work efficiency. Furthermore, in one case report, researchers utilized tracked-microscope-based AR system and they emphasized that this system was helpful in identifying and determining the extension of pleomorphic adenoma of lacrimal gland (Scolozzi and Bijlenga, 2017).

Feasibility and usability studies were conducted in two articles in maxillofacial oncology. In one study, the standard ISO-9241/110 feasibility questionnaire based on the 5-point Likert Scale was conducted on both medical staff and AR experts, and overall feedback turned out to be positive (Pepe et al., 2019). In another article by Gsaxner et al. (2021), AR usability evaluation received a SUS of 74.8 ± 15.9 (>68 indicates above average) with the 5-point Likert scale of 4.5 ± 0.7 out of 5 (5 representing extremely positive).

In addition, AR guided simulations scored higher as compared to virtual unguided resections in mental demand, performance, effort and frustration in preclinical study conducted on phantoms by Chan H et al. (2022).

Some clinicians find the AR system simple to learn and use, which improves their decision-making skills (Gsaxner et al., 2021), whereas, some authors stressed on the importance of pre-clinical education in managing the technology's steep learning curve (Ochandiano et al., 2021).

Despite of the fact that a few studies did not achieve sufficient accuracy in terms of registration error and surgical accuracy in maxillofacial oncology, the overall outcomes seem to have a positive impact of AR in maxillofacial oncological surgeries. For instance, according to Chan H et al. (2022), AR guided resection improved negative margin and had more similarities with pre-planned cutting planes. Similarly, Shaofeng et al. (2023) came to conclusion that AR system accuracy is similar to that of surgical guide while testing it on mandibular tumor and fibular reconstruction. They also emphasized that it enhances surgeon's hand-eye coordination in executing surgeries. Ceccariglia et al. (2022) found the discrepancy of under 2 mm between AR projected osteotomy and customized cutting guide osteotomy. Shi et al. (2022) concluded that AR navigation can effectively display and guide the surgical path and helps in achieving desired results. Furthermore, Sahovaler et al. (2021) showed advantage of AR over unguided simulations. However, some pressing concerns like limited depth perception and time required for auto-registration were also mentioned in some studies (Pepe et al., 2019; Gsaxner et al., 2021; Modabber et al., 2022).

Similar to orthopedic oncology, maxillofacial oncology section also lacks in many aspects. Feasibility questionnaire survey and ergonomics are discussed in only few articles. The different methods of evaluation in articles limited the ability to provide quantifiable results. Additionally, only 10 out of 27 articles reported on registration error and only 14 reported on surgical accuracy. Conversely, further research is emphasized pre-clinically before implementing AR in operating rooms.

We should advocate for the development of a technique that is uniform and consistent in order to investigate this new technology

TABLE 4 Distribution of studies by technical aspects of AR in maxillofacial oncology.

Reference	Authors	Year	AR display type	AR device	Registration/ Tracking modality	No. of cases	AR registration error	AR surgical accuracy
Chan H et al. (2022)	Chan et al.	2022	Projector-based	Portable high-definition projector (PicoPro, Celluon Inc.)	Marker-based	5	Fiducial registration error of ≤1 mm	AR-guided osteotomies had high similarity with the preplanned, with interclass correlation index (ICC) close to 1 (0.893) in "adequate" (5–15 mm) margins
Sahovaler et al. (2021)	Sahovaler et al.	2021	Projector-based	Portable high-definition projector (PicoPro, Celluon Inc.)	Marker-based	4	Fiducial registration error of ≤1 mm	n.a.
Shi et al. (2022)	Jiafeng shi et al.	2022	HMD	HiAR G200 AR Glasses	Marker-less	1	Surface matching error: 0.64 ± 0.28 mm	Mean surgical resection error: 0.49 ± 0.37 mm
Ceccariglia et al. (2022)	Ceccariglia et al.	2021	HMD	HoloLens 2	Marker-less	3	n.a.	Surgical resection error (deviation between AR-projected osteotomy and the one planned and performed with cutting guide: <2 mm
Ochandiano et al. (2021)	Ochandiano et al.	2022	Screen -based	Smartphone (iphone 6)	Marker-based	7	n.a.	n.a.
Gsaxner et al. (2019)	Gsaxner et al.	2019	HMD	HoloLens 1	Marker-less	8	Target registration error: 9.2 ± 1.5 mm	n.a.
Pepe et al. (2019)	Pepe et al.	2019	HMD	HoloLens 1	Marker-less Marker-based	1	Mean registration error along (x, y, z): Marker-less: (3.3 ± 2.3, −4.5 ± 2.9, −9.3 ± 6.1) mm; Marker-based (7.0 ± 2.1, −12.5 ± 2.5, −19.0 ± 2.0) mm	n.a.
García-Sevilla et al. (2021b)	García-Sevilla et al.	2022	Screen based	Smartphone (Iphone 6)	Marker-based	1	n.a.	Median surgical resection error: 0.40 mm
Tel et al. (2021)	Tel et al.	2021	Screen based	Smartphone (Iphone 12 Pro)	Marker-less	5	n.a.	n.a.
Gsaxner et al. (2021)	Gsaxner et al.	2021	HMD/PC	HoloLens 1/PC	Marker-less	1	Registration error between a few millimeters of up to 2 cm	n.a.
Sugahara et al. (2021)	Sugahara et al.	2021	HMD	HoloLens 1	Marker-based (attach to splint)	1	n.a.	n.a.
Shaofeng et al. (2023)	Shaofeng Liu et al.	2023	HMD	HiAR G200 AR glasses	Marker-less Marker-based	9 model trials, 12 animal trials		Model Trials Distance deviations: 1.62 ± 0.38 mm (mandible) 1.86 ± 0.43 mm (fibula); 1.67 ± 0.70 mm (fixation screws) Angular deviations: 3.68 ± 0.71° (mandible); 5.48 ± 2.06° (fibula); 7.50 ± 1.39° (fixation screws) Animal trials Distance deviations: 0.93 ± 0.63 mm (condyle); 2.01 ± 2.49 mm (mandibular angle); 1.41 ±

(Continued on following page)

TABLE 4 (*Continued*) Distribution of studies by technical aspects of AR in maxillofacial oncology.

Reference	Authors	Year	AR display type	AR device	Registration/ Tracking modality	No. of cases	AR registration error	AR surgical accuracy
								0.61 mm (mandible); 0.88 ± 0.22 mm (tibiofibular bones) Angular deviations: 6.81 ± 2.21° (mandible); 6.47 ± 3.03° (tibiofibular bones)
Prasad et al. (2023)	Prasad BA et al.	2023	HMD	HoloLens 2	Manual	20	Average relocation error of 4 mm ± 3.9 mm	n.a.
Gao et al. (2022)	Gao et al.	2022	HMD	HoloLens 1	Marker-based	5	n.a.	Mean surgical error: 1.036 ± 0.081 mm
Kim et al. (2020)	Jin Kim et al.	2020	Screen-based/ HMD	Self-developed AR viewer (Vive pro, monitor)	Marker-based	1	n.a.	Mean surgical error: 2.77 ± 1.29 mm
Yang et al. (2022)	Yang et al.	2022	HMD	HoloLens 1	Marker-less	4	Registration error <1 mm	Maximum surgical error: 6.08 mm
Tang et al. (2022)	Tang et al.	2022	HMD	HoloLens 2	Marker-less	7	n.a.	Mean surgical error: 1.68 ± 0.92 mm
Scolozzi and Bijlenga (2017)	Scolozzi et al.	2017	Screen-based	Microscope-based AR system	Marker-less	1	n.a.	n.a.
Cercenelli et al. (2022)	Cercenelli et al.	2022	Screen-based HMD	Tablet HoloLens 2	Marker-less	1	Registration errors ranging between 1–5 mm	Accuracy: 2.0 mm (100% success rate with HoloLens; 97% with tablet)
Scherl et al. (2021)	Scherl et al.	2021	HMD	HoloLens 1	Manual	6	Mean error of the alignment: 1.3 cm	n.a.
Necker et al. (2023)	Necker et al.	2022	HMD	HoloLens 2	Manual	1	n.a.	n.a.
Han et al. (2022)	Lin et al.	2022	HMD	HoloLens 2	Manual	1	n.a	n.a
Zhao et al. (2023)	Zhao et al.	2022	Screen-based	Monitor	Marker-based	7	n.a	Length difference: 1.18 ± 0.84 mm, Angular deviation: 5.45 ± 1.47°, Volume overlap rate: 95.31% ± 2.09%, Average surface distance: 1.22 ± 0.12 mm.
Modabber et al. (2022)	Modabber et al.	2021	Projector-based	ML750ST, Optoma projector	Marker-less	10	n.a	Angulation of the osteotomy plane: AR group: 14.99 ± 11.69° Cutting guides group: 8.49 ± 5.42° Osteotomy plane distance: AR group: 2.65 ± 3.32 mm Cutting guides: 1.47 ± 1.36 mm
Meng et al. (2021)	Meng et al	2021	HMD	HoloLens 1	Manual	1	n.a	Mean location of the fibular osteotomies: 2.11 ± 1.31 mm Angular deviation of the fibular segments: 2.85° ± 1.97° Intergonial angle distances: 7.24 ± 3.42 mm
Battaglia et al. (2019)	Battaglia et al.	2019	Screen-based	Tablet	Marker-less	3	n.a	n.a

(Continued on following page)

TABLE 4 (*Continued*) Distribution of studies by technical aspects of AR in maxillofacial oncology.

Reference	Authors	Year	AR display type	AR device	Registration/ Tracking modality	No. of cases	AR registration error	AR surgical accuracy
Winnand et al. (2022)	Winnand et al.	2021	Projector-based	ML750ST, Optoma projector	Marker-less	10	n.a	Average discrepancy in osteotomy plane angulation AR: 10.21 ± 7.22° Cutting guides: 6.98 ± 4.70 Mean variations between the osteotomy planes and the planned trajectories: AR: 2.29 ± 1.98 mm Cutting guide: 1.32 ± 1.00 mm

and make it possible to conduct meta-analyses for future investigations. This stage is crucial for gathering data to support the use of AR in oncological procedures in both disciplines. Additionally, we stress the importance of including external navigation in AR in future experiments in order to enhance the precision and depth perception of this infant, yet useful technology. Moreover, through the use of standardized questionnaires, SUS, and a 5-point Likert scale, feasibility and ergonomics should be evaluated.

We observed from the collected papers that an important aspect of AR implementation is three-dimensional printing (3D printing), also referred to as rapid prototyping, which is typically used for obtaining pre-operative patient-specific phantoms replicating the anatomical structures of interest. These phantoms are typically used in the papers to perform the surgical task preoperatively under AR guidance, as well as to evaluate both the registration error and surgical accuracy (Moreta-Martinez et al., 2018; Gsaxner et al., 2019; Pepe et al., 2019; García-Sevilla et al., 2021a; García-Sevilla et al., 2021b; Gsaxner et al., 2021; Moreta-Martinez et al., 2021; Ochandiano et al., 2021; Sahovaler et al., 2021; Tel et al., 2021; Chan H et al., 2022; Gao et al., 2022; Pose-Díez-de-la-Lastra et al., 2022; Shi et al., 2022; Shaofeng et al., 2023). When using a marker-based tracking approach, the reference marker which is designed to fit in a unique position on the patient, is produced by 3D printing (Moreta-Martinez et al., 2021; Cho et al., 2017; Choi et al., 2017; Cho et al., 2018; Moreta-Martinez et al., 2018; García-Sevilla et al., 2021b; Ochandiano et al., 2021; Sugahara et al., 2021; Pose-Díez-de-la-Lastra et al., 2022; Shaofeng et al., 2023). Moreover, in some cases (Moreta-Martinez et al., 2018; Ceccariglia et al., 2022) patient-specific surgical guides used for comparative evaluation with AR guidance on surgical accuracy are manufactured via 3D printing. Finally, some studies clearly suggest to use AR and 3D printing in combination to improve surgical efficacy, accuracy, and patients experience (García-Sevilla et al., 2021a; Moreta-Martinez et al., 2021).

4.1 Limitations of AR in surgery

Despite the fact that AR is a growing technology, it is not without limitations and complications. Surgeons should be well aware of the limitations of augmented reality in surgery, including technical challenges, limited field of view which limits the amount of virtual content available to the user, high implementation costs and limited user experience. In addition, they should consider how these limitations may impact the accuracy and efficacy of AR systems, as well as the surgical outcomes. Before incorporating AR into clinical practice, it is essential to execute a comprehensive analysis of its viability and benefits.

For instance, the viewing distance and angle of commercially available HMDs, such as HoloLens 2, are not optimized for use in surgery since the focus distance is suboptimal for medical procedures that are typically carried out at arm's length and with the head bowed to observe the operative field (Wong et al., 2022).

To the authors' knowledge, today only two "surgery-specific" headsets are available for AR-based intraoperative guidance: the X-vision Spine System by Augmedics, which received FDA approval (https://www.augmedics.com/), and the VOSTARS system, still under investigation (https://www.vostars.eu/). VOSTARS is promising new wearable AR system designed as a hybrid Optical-See-Through (OST)/Video-See-Through (VST) HMD capable to offer a highly advanced navigation tool for maxillofacial surgery and other open surgeries. An early prototype of the VOSTARS system (Ruggiero et al., 2023) has been already evaluated in phantom tests and demonstrated a sub-millimetric accuracy (0.5÷1 mm) in the execution of high-precision maxillofacial tasks (Cercenelli et al., 2020; Condino et al., 2020).

The issue of depth perception is another challenge that surgeons have to consider when applying AR technology during surgical procedures (Sielhorst et al., 2006). Surgeons must accurately gauge the distance between their instruments and the intended targets for AR surgery to be successful. However, accurate distance estimation during AR-assisted surgery is complicated by the fact that tools and target landmarks are 3D-rendered (Choi et al., 2016).

In order to implement AR in surgery, complex technical solutions, including medical-grade software and hardware systems, are required. For example, consumer-grade computer systems are suboptimal for displaying high-quality 3D rendered objects, and HMDs have a limited battery life (2–3 h), which can result in technical issues such as system failure, calibration errors, and latency. These obstacles may limit the accuracy of AR systems and result in surgical complications (Wong et al., 2022).

In addition, surgeons using AR-HMDs must contend with the limited field of view, restricted binocular field and projection size (Lareyre et al., 2021).

TABLE 5 Brief description of each article with aim and major outcomes, on Orthopedic Oncology.

Reference	Author, date	Brief description with aim and outcomes
Moreta-Martinez et al. (2021)	Moreta et al., 2021	Aim: Authors introduced a surgical workflow to orthopedic oncology by combining 3D printing and a smartphone-based AR application. A 3D-printed reference marker was used for virtual-to-real patient registration. The system was experienced on six patient-specific phantoms and in two clinical cases. This system was evaluated in terms of visualization accuracy and usability during the whole surgical workflow Results/Conclusion: Phantom experiments provided a visualization accuracy (i.e., registration error) below 3 mm. Positive feedback was obtained from surgeons and patients
Cho et al. (2017)	Choi et al., 2017	Aim: Authors conducted an experiment on pig femurs using a tablet-based AR navigation system to evaluate the accuracy of AR-assisted oncological surgeries compared to conventionally performed surgeries. To simulate a bone tumor in the pig femur, a cortical window was made in the diaphysis and bone cement was inserted. A total of 164 surgical removals using the AR technique and 82 resections using the conventional procedure (i.e., manual measurement as per routine clinical practice) were performed Results/Conclusion: The mean resection error (i.e., the difference between the obtained and the planned surgical margin) was 1.71 mm (0–6) and 2.64 mm (0–11) for AR-based and conventional interventions, respectively ($p < 0.05$). In AR-based navigation resection, 90.2% of the times, the surgical margin of 10 mm was successfully attained, compared to only 70.7% in conventionally performed surgeries. The authors concluded that the accuracy of tumor resection achieved with the proposed AR-based navigation was satisfactory
Cho et al. (2018)	Choi et al., 2018	Aim: Authors conducted animal trials showing that AR navigation may prove useful in pelvic tumor resections. Researchers injected bone cement into the acetabular dome of 36 cadaver pig pelvises to compare the AR navigation technique and the conventional navigation technique (based on CT image). Each technique was assigned with 18 tumor pelvises to operate on with a safety margin of 1 cm Results/Conclusion: With marker-based AR technology, the mean margin of resection was 1.59 ± 4.13 mm, whereas in the conventional navigation it was 4.55 ± 9.7 mm. 100% AR-assisted resections had errors <6 mm as compared to 78% resections done using conventional method, showing the attainability of planned margins using AR. The authors further emphasized that more *in vivo* trials are needed before adopting it for clinical trials
Choi et al. (2017)	Choi et al., 2017	Aim: This study was aimed at achieving the margin of cancer resection in 60 porcine acetabular regions using AR navigation method and conventional method based on manual measurement (30 pelvises were assigned for each technique). A tablet PC was used for AR visualization and cuboidal reference markers were used for tracking. 3D dilation technique was used in simulated tumor model, and separation of this 3D dilation and actual tumor was deemed as safety margin. The target margin of resection was set at 10 mm. The surgical resection plane, determined by this safety margin, was adopted in accordance with the direction of cutting saw in real time, and minimum distance between the cutting saw and resection plane was measured Results/Conclusion: The mean fiducial registration error for anatomical landmarks was 0.58 ± 0.22 mm. After resection, analysis showed the measured resection margin to be 9.85 ± 1.02 mm for AR navigational resection as compared to 7.11 ± 4.30 mm with conventional resection. The reduction in standard deviation from 4.30 to 1.02 mm demonstrates the high precision of surgical resection margins using the AR method. With the conventional method, 1/4 of surgical procedures resulted in invasion of the 5 mm tolerance margin, while none of the surgeries using the AR method resulted in such invasion. Hence, AR navigation is proved to be more accurate for guiding surgical resection of complex bone tumors
Molina et al. (2021c)	Molina et al., 2021	Aim: Authors presented a first-ever case report using AR-assisted spinal surgery on L1 chondroma to perform *en bloc* wide osteotomy via posterior approach Clinical symptoms: A 69-year-old man experienced lower back pain and paresthesis 6 months prior to surgery. He had a history of lymphoma for which he received chemotherapy and radiation.He has been symptom-free since 2010 Investigations: Imaging by MRI revealed an interosseous lesion with spinal extension at L1 level. And chondroma was diagnosed through a biopsy Results/Conclusion: The minimal invasive *en bloc* procedure was performed with precise lumbar osteotomy and screw placement using AR-HMD devices and flippable tracker to avoid any line-of-sight obliteration. Postoperatively, there were no reports of long-term complications, and pathology confirmed a negative resection margin. Therefore, AR guidance assisted in achieving the intended navigational trajectories for the placement of pedicle screws
Abdel Al et al. (2020)	Abdel el at, 2020	Aim: Authors published a case report of a 39-year-old male with a deep-seated impalpable soft tissue (synovial fluid) sarcoma of the medial aspect of the left foot. The resection was done using smartphone-based AR guided MRI images to outline sarcoma superficially Clinical Symptoms: Two years ago, the patient complained of pain on the medial side of the left sole without a relevant medical history. This non-radiating pain

(Continued on following page)

TABLE 5 (*Continued*) Brief description of each article with aim and major outcomes, on Orthopedic Oncology.

Reference	Author, date	Brief description with aim and outcomes
		intensified while walking, and topical analgesics had no effect Investigations: Imaging and biopsy were performed to detect the mass, and synovial sarcoma was diagnosed. The subsequent physical examination was unremarkable, and MRI revealed a bilobed lesion on the medial aspect of the left foot. In addition, the PET-CT scan revealed a small, mildly hypermetabolic lesion, but no distant metastases Results/Conclusion: A patient-specific AR application was devised to superimpose MRI images on foot landmarks. The tumors borders were outlined on skin using sagittal MRI images. Using AR-guided pre-operative markings for tumor localization, *en bloc* excision was performed complication-free, except for decreased sensation in the medial aspect of the foot. Furthermore, follow-up imaging ruled out any infection, recurrence, and metastasis. It was emphasized that this smartphone-based AR application is suitable for small and fixed tumor
García-Sevilla et al. (2021a)	Garcia-Sevilla et al. 2021	Aim: Authors conducted a pre-clinical study demonstrating the importance of AR to guide patient-specific instruments (PSI) placement in pelvic tumor resection. Six different ilium tumor scenarios were simulated for evaluation. In addition, six pairs of PSI were designed, resulting in two PSI's for each case scenario. Experiment was devised using a smartphone and HoloLens 2 to test the system's accuracy and compared it to freehand method. Two varieties of phantoms were utilized for assessment. i.e., conventional plastic pelvic bone (no silicone layer) and realistic phantom (silicone layer). System accuracy was analyzed based on PSI transformation from its intended position to its actual one, which provided the maximum osteotomy deviation (MOD) Results/Conclusion: The median values for MOD with smartphone or HoloLens 2 are found to be less than 2 mm in silicone phantom and below 1 mm in conventional plastic bone, whereas the median values for PSI placed with freehand were 3.37 mm for realistic phantom and 1.70 mm for non-silicone one. There was significant difference between freehand method and using either smartphone or HoloLens ($p < 0.001$). For phantom's comparison, high errors were observed in all cases with silicone phantom ($p < 0.001$). In the end, authors found encouraging evidence that AR has the ability to overcome the current constraints of PSIs in a straightforward and efficient manner
Moreta-Martinez et al. (2018)	Moreta et al., 2018	Aim: Authors demonstrated an AR approach for Extra-osseous Ewing sarcoma (EES) of distal limb. Using patient-specific tools with marker attached, pre-surgical simulation was done on patient-specific phantom replicating Ewing's sarcoma before testing during actual surgical intervention. Two sets of 3D models were printed. One was used for validation purposes (including conical holes in design) and the other for surgical application. For system precision, evaluation was done on the basis of surgical guide placement error and error in AR point localization for visualization. For that purpose, conical holes on both phantoms and surgical guide were used as a reference point for registration and error measurement. The measurement was accomplished using an optical tracking system (Polaris) Result/Conclusion: Surgical guide placement error and AR visualization error was measured in average root-mean square and valued 1.87 and 2.90 mm, respectively among all 3 repetition attempts. In addition, the complete workflow was implemented by an expert surgeon on actual surgical field using HoloLens, and satisfactory alignment was achieved. The authors believe that the developed pre-surgical system will pave the way for the creation of user-friendly AR systems that can be used in the medical industry for training, simulation, and guidance
Pose-Díez-de-la-Lastra et al. (2022)	Pose-Díez-de-la-Lastra et al., 2022	Aim: In this study, researchers evaluated the accuracy of HoloLens 1 and HoloLens 2 using two orthopedic oncological cases. For this purpose, they acquired patient-based phantom simulating EES of distal leg. Secondly, they examined the accuracy of HoloLens 2 on more complex case of right shoulder undifferentiated pleomorphic sarcoma. And finally, researchers employed HoloLens 2 on a real patient with a pleomorphic sarcoma of the right shoulder in operating room to evaluate its technical and ergonomic aspects. Additionally, patient-specific surgical guides were also designed to accommodate AR registration markers. Using an optical tracking system, the AR projection error was analysed by documenting the positions of AR spheres on a phantom surface. The procedure was repeated three times by three researchers Results/Conclusion: Leg phantom (EES) showed greater accuracy with HoloLens 2 with RMSE of 2.16 mm as compared to HoloLens 1 (RMSE = 2.833 mm). Significant difference among devices was noted ($p < 0.05$). In a case of pleomorphic sarcoma of the shoulder, the RMSE for HoloLens 2 was 3.108 mm. The increased error by HoloLens 2 in second case was due to larger size and dimensions of phantom. Lastly, HoloLens 2 were tested on actual patient with pleomorphic sarcoma and the surgeon praised the enhanced ergonomics of this HMD. In addition, a survey was conducted to compare both head gears and the HoloLens 1 received a score of 2.84 out of 5 as compared to 4 for the HoloLens 2 indicating the superior ergonomics of this device. In conclusion, HoloLens 2 had better results in terms of both accuracy and ergonomics when tested on orthopedic oncological cases

TABLE 6 Brief description of each article with aim and major outcomes, on Maxillofacial Oncology.

Reference	Author, date	Brief description with aim and outcomes
Chan H et al. (2022)	Chan et al., 2022	Aim: Authors conducted a preclinical study on phantoms to address the issue of margin control using AR. 5 phantom models with maxillary tumors were created and 5 resident surgeons carried out both AR-guided simulated resection and virtual unguided resections for comparison. 115 osteotomies were performed virtually and comparison was done on basis of intralesional cuts (≤0 mm), close (>0 mm and ≤5 mm), adequate (>5 mm and ≤15 mm) and exceeding distance (>15 mm) from tumor Results/Conclusion: Registration error was less than 1 mm for AR application. In context of surgical accuracy, Intratumor margin was 0% in AR-guided resection versus 1.9% in unguided simulation. Close margin also showed low percentage of 0.8% and 7.9% in AR-guided and unguided resections, respectively. In both cases, *p*-value was <0.0001. With a *p*-value of 0.018, the percentage was greater for the AR-guided simulation at 25.3% as opposed to 18.6% in the unguided simulation for "adequate" resection. No differences were noted for excessive margins. In addition, AR-guided simulations scored higher in mental demand, performance, effort, and frustration. According to authors, AR-guided resections had more similarities with pre-planned surgical planes. Consequently, they concluded that AR methodology improves negative margin through more precise rendering of preplan cutting planes
Sahovaler et al. (2021)	Sahovaler et al., 2021	Aim: Researchers conducted a preclinical study to compare approach of AR and intraoperative navigation (IN) on sinonasal malignancies removal. five surgeons performed simulations of virtual cuts to compare AR approach and advance IN approach on four tumor models. Unguided, AR, IN and AR+IN simulations were performed and statistically compared. Making intratumor cuts the key outcome, the others "close, adequate, and excessive distances" from tumor were also analyzed in percentages. Additionally, screening timing was calculated based on the information from gaze tracker headset Results/Conclusion: AR application registration error was <1 mm. Out of 335 cuts, percentage of intratumoral cuts were 20.7%, 9.4%, 1.2% and 0% for the unguided, AR, IN, and AR+IN simulations, respectively ($p < 0.0001$) showing the advantage of AR over unguided simulation. IN approach decreases intratumoral cuts as compared to AR alone approach. Whereas, combination of both AR and IN did not improve intratumoral rate significantly (*p*-value 0.5). The screening timing in unguided, AR, IN, and AR+IN turned out to be 55.5%, 0%, 78.5%, and 61.8%, respectively ($p < 0.001$). The screening time and workload score (NASA-TLX questionnaire Score) in AR+IN approach improves as compared to IN alone approach. Hence, the authors concluded that AR navigation improves open sinonasal tumors resections as well as overcome the attention-deteriorating-screening problem of IN. However, more works needs to be done on this application before clinical implementation
Shi et al. (2022)	Jiafeng shi et al., 2022	Aim: In this article, authors conducted a study based on marker-less tracking on mandibular edge for resection of benign maxillofacial tumor to avoid the complications caused by guiding plate. Before surgery, they replicated the 3D model for diseased bone for pre-surgical simulation and access the lines of resection of tumor Results/Conclusion: They analyzed the marker-less surface registration error and turned out to be 0.6453 ± 0.2826 mm (<1 mm), affected by system error and impact was ignored surgically. Surgical error was assessed using an experimental AR system and found out to be 0.4858 ± 0.3712 mm (<1 mm). The authors concluded that AR-guided marker-less navigation can effectively display and guide the surgical path. Therefore, it helps in achieving the desired results and has a positive impact on doctors
Ceccariglia et al. (2022)	Ceccariglia et al., 2021	Aim: In this study authors demonstrated the application of marker-less AR registration for removal of maxillofacial tumors, and performed the resectional surgeries on three patients suffering from oral tumors using AR. Two males and one female patients with the mean age of 56 years underwent seven group of osteotomies in total. These osteotomies were analyzed by comparing corticotomy lines drawn by AR guidance and customized cutting guides Results/Conclusion: The difference of under 2 mm was noted between AR projected osteotomy and customized cutting guide osteotomy, hence showing that marker-free AR navigation is achievable. However, the authors also emphasized that further research is needed to be done on marker-less facial registration for maxillofacial tumors resection despite being considered safe
Ochandiano et al. (2021)	Ochandiano et al., 2022	Aim: Authors published an article emphasizing the role of 3D printing, virtual surgical planning (VSP) and augmented reality in head and neck tumors ablation and dental implants. They included 11 patients. Out of which, 8 suffered from mandibular, 1 tongue, 1 maxilla and 1 hard palate carcinoma. Total of 56 implants were inserted, but 6 of them were withdrawn from data analysis due to unavoidable intra operative complications. Using intraoperative infrared optical navigation (for the first four patients), surgeons virtually planned and transferred the prosthetically driven dental implant placement to the patient. Finally, they used a combination of conventional static teeth supported 3D-printed acrylic guide stent, intraoperative dynamic navigation, and AR for final intraoperative verification for other seven patients. Differences in implants coronal, apical, and angular location between preoperative 3D planning and guided intraoperative placement were quantified Results/Conclusion: Due to jig registration instability, initial simple infrared navigated cases achieved low accuracy. Dynamic navigation cases that follow achieved 1–1.5 mm insertion point deviation with the help of highly stable acrylic static guides used as reference and to register markers. Hence, Image-guided surgery, 3D printing, and AR technology could be used to precisely plan, implantation and reconstructive surgeries. The authors went on to stress the importance of pre-clinical education in managing the technology's steep learning curve

(Continued on following page)

10.3389/fbioe.2023.1276338

TABLE 6 (*Continued*) Brief description of each article with aim and major outcomes, on Maxillofacial Oncology.

Reference	Author, date	Brief description with aim and outcomes
Gsaxner et al. (2019)	Gsaxner et al., 2019	Aim: Authors employed marker-free inside-out image to face registration for augmented reality in head and neck oncological cases. For that, they 3D printed eight phantom heads of subjects with tumors in this region and subsequently tested the application's viability on human subject. For accuracy estimation, target registration error (TRE) was measured through five different landmarks and compared it with optical tracking system Results/Conclusion: TRE was repeated 10 times on each phantom and human subject reporting a mean TRE of 9.2 ± 1.5 mm, in comparison to error of high-precision optical tracking system of 3.9 ± 1.8 mm in translation and 4.9 ± 2.4° in rotation. Even though authors did not achieve the desired precision required for this type of intervention due to certain restrictions, they still believe that AR is a viable tool for these interventions
Pepe et al. (2019)	Pepe et al., 2019	Aim: In this article, authors employed a marker-less approach for a head and neck tumor demonstration on a 3D PET CT scan-acquired model. In model, *x* and *y* axes were marked along the horizontal and vertical axis of the tumor mass for evaluation purposes, whereas the z-axis depicts the direction of gaze. Red points were added to facial landmarks to measure the registration error against the green virtual points from the AR application, and an experienced user repeated the measurements four times. Automatic registration error was analyzed with and without user calibration Results/Conclusion: Errors with user calibration (markers) along the y, x, and z-axes were to be −4.5 ± 2.9 mm, 3.3 ± 2.3 mm and −9.3 ± 6.1 mm, respectively. These errors had nearly doubled values without user calibration (without markers) along these axes (−12.5 ± 2.5 mm, 7.0 ± 2.1 mm and − 19.0 ± 2.0 mm), indicating that user calibration guarantees a more precise registration. However, the most anticipated difficulty during this study appears to be physician's limited depth perception, which restricted him from viewing the 3D model from certain angles. Moreover, the standard ISO-9241/110 feasibility questionnaire based on 5-point Likert Scale was conducted and overall feedback turned out to be positive
García-Sevilla et al. (2021b)	García-Sevilla et al., 2022	Aim: In this study, authors compared three different navigation techniques for *en bloc* resection of exophytic invasive adenoid cystic carcinoma of the hard palate. Preoperative simulation and accuracy of these surgical navigations based on pre-defined surgical margins were estimated on phantom and each simulation was repeated three times Results/Conclusion: Initially, an optical tracking system (OTS) facilitating registration using screws (five screws were attached to maxilla before image acquisition) was examined for surgical guidance. The median deviation from the desired surgical margin was 0.57 mm, but with lowest variations (IQR of 0.24 mm). Whereas, OTS enabling registration with surgical guide (splint), had median and IQR both higher than screw registration. AR navigation, on the other hand, had the lowest median of 0.40 mm, but with high variation exhibiting IQR of 0.89 mm. However, no statistical difference was noted in terms of accuracy (errors below 1 mm). But, owing to low error variations, OTS with screw registration was considered for real surgical intervention, giving the fiducial point registration errors of 0.77, 0.93 and 0.81 mm over three different repetitions. Additionally, AR was also tested in OR for qualitative study. Authors concluded that these navigational approaches are accurate and convenient for minimal invasive and conservational approaches
Tel et al. (2021)	Tel et al., 2021	Aim: The authors conducted a qualitative investigation into the use of computer-assisted pre-operative surgical planning, virtual endoscopy, and AR navigation for the resection of various histopathological masses in the inferior orbital compartment via transantral approach. Five patients with disease-related main signs and symptoms were included. Pre-surgical simulation was conducted on patient-specific 3D-printed model Results/Conclusion: Ultimately, surgery was performed under transmaxillary navigation guidance using virtual models to verify the correct surgical positioning throughout all essential stages of endoscopic transantral approach. Excisions were performed through a transantral approach on all patients without intraoperative or long-term postoperative complications. It was determined that each step of virtual planning can be replicated with relative precision during actual surgery. In the context of AR implementation, accurate tracking and overlaying were performed on each patient-specific phantom for preoperative simulation, highlighting its role during preoperative study, training, and actual surgery
Gsaxner et al. (2021)	Gsaxner et al., 2021	Aim: This article emphasized on practicability and utility of a designed AR system for head and neck tumor simulation and clinical settings. Eleven healthcare personnel with extensive experience in the head and neck region were recruited. During the initial segment of training, participants received a comprehensive introduction and demonstration of the HoloLens system using a patient phantom. The researchers were then requested to evaluate the AR system on a healthy volunteer with a simulated tumor in the head and neck region to demonstrate a clinical case during the testing phase Results/Conclusion: Registration error was noted between a few millimeters to 2 cm. After completing both phases (training and demonstration), physicians participated in system feasibility questionnaires and an informal interview. The AR usability evaluation received a System Usability score (SUS) of 74.8 ± 15.9 (>68 indicates above average) with a 5-point Likert scale of 4.5 ± 0.7 out of 5 (with "1" representing extremely negative and "5" representing extremely positive). According to users, however, registration precision and the time required for auto-registration were pressing concerns. In conclusion, clinicians find this AR system simple to learn and use, which improves their decision-making skills

(Continued on following page)

10.3389/fbioe.2023.1276338

10.3389/fbioe.2023.1276338

TABLE 6 (*Continued*) **Brief description of each article with aim and major outcomes, on Maxillofacial Oncology.**

Reference	Author, date	Brief description with aim and outcomes
Sugahara et al. (2021)	Sugahara et al., 2021	Aim: Authors presented a case report on the resection of maxillary calcifying odontogenic cyst and reconstruction using mixed reality Clinical Symptoms: 27 years old female had a gingival swelling in 2010 and left untreated due to the abstinence of discomfort. Later in 2016, during routine dental care, an x-ray revealed radiolucent area but again disregarded against the medical advice. Next year, upon displacement of left anterior maxillary teeth, patient was referred to our department Investigations: Physical examination revealed left nasal ala deformity along with intraoral and left upper lip swelling. The Electric pulp test showed vital reaction in this region but no pain upon pressure Pre-requisites: Partial resection of left maxilla from the lateral incisor to the second pre-molar with iliac cancellous bone and marrow grafting was planned. Virtual planning and estimation of graft quantity to fill post-resection cavity were accomplished prior to surgery. 3D model with all details on tumor and its resection was printed. During surgery, tumor and surrounding anatomical structures were virtually visualized by three surgeons using HoloLens Results/Conclusion: The resection was successfully executed according to plan and post-surgical void was reconstructed using iliac cancellous bone graft. At the 18 months follow-up, there were no complications or relapse. In conclusion, the incorporation of mixed reality during pre-surgical and intra-operative phases allows for accurate and secure surgery
Shaofeng et al. (2023)	Shaofeng Liu et al., 2023	Aim: Authors tested the feasibility and accuracy of marker-less contour-based AR registration for simulated mandibular tumor and fibula-based reconstruction on phantom models and subsequently compared this AR based system accuracy with surgical guides in animal trials. Virtual tumor tissue of right and left mandible was designed, and 9 mandibular and fibular models were 3D printed. Based on the trigger detection algorithm, virtual-real registration of osteotomy instruments for tracking and calibration of instrument's angle was also accomplished. After virtual-real scene registration of instrument and lesion, mandibular resection, screw fixation and fibular reconstruction were accomplished under AR guidance and post-surgical CBCT data was analyzed for deviations from pre-operative planning Results/Conclusion: For model trials, the distance and angular deviation for mandibular osteotomy surface were 1.62 ± 0.38 mm and 3.68 ± 0.71°, respectively. And for fixation screws, it was 1.67 ± 0.70 mm and 7.50 ± 1.39°, respectively. In contrast, the distance and angular deviation of reconstructed fibular osteotomy were 1.86 ± 0.43 mm and 5.48 ± 2.06°, respectively. No statistical difference was noted between pre-operative planning and post-surgical analysis ($p < 0.001$). For animal trials, 12 New Zealand rabbits were recruited and divided into six pairs of AR and surgical guide groups. The animals were scanned again for error analysis in order compare deviation between AR group and surgical guide group. For bilateral condylar outer poles, distance deviations for AR and surgical guide groups were 0.93 ± 0.63 mm and 0.81 ± 0.30 mm, respectively ($p = 0.68$). For bilateral mandibular posterior angle, they were 2.01 ± 2.49 mm and 2.89 ± 1.83 mm, respectively ($p = 0.50$). Whereas, for mandibular osteotomy surface, distance deviation for both groups were 1.41 ± 0.61 mm and 1.21 ± 0.18 mm, respectively ($p = 0.45$) and the angular deviation were 6.81 ± 2.21° and 6.11 ± 2.93°, respectively ($p = 0.65$). The distance deviation for reconstructed tibiofibular osteotomy surfaces were 0.88 ± 0.22 mm and 0.84 ± 0.18 mm ($p = 0.70$), whereas, angular deviations were 6.47 ± 3.03° and 6.90 ± 4.01° ($p = 0.84$), respectively. There was no statistical difference between two groups ($p > 0.05$). In conclusion, AR system is feasible with accuracy similar to surgical guide. It also enhances surgeon's hand-eye coordination during surgeries. However, more testing is required prior to clinical implementation
Prasad et al. (2023)	Prasad BA et al., 2023	Aim: This article demonstrated the feasibility and accuracy of AR-guided re-resections of head, neck and oral cavity cancers. 20 different dissections were performed on 3 cadavers by 2 head and neck surgeons. Before simulating cancer resection, two sutures were placed: one at the edge of the resection site(S) and other at adjacent resection bed(R). After resection, 2 fiducial markers were used to pinpoint the precise location of stitch R, which was then removed. Using a HMD, surgeons manually align the AR hologram to resection bed. Stitch S on the hologram was used as a guide to relocate stitch R and position a new stitch R′. The distance between R and R′ was measured for accuracy Results/Conclusion: The statistical analysis showed a mean relocation error of 4.0 ± 3.9 mm. However, errors pertaining to maxillary and mandibular resections differed significantly from those associated with other resections (10.7 vs. 2.8 mm; $p < 0.01$). After evaluating its applicability on cadavers, the authors emphasized on applying this technology to patients in the operating room
Gao et al. (2022)	Gao et al., 2022	Aim: Authors conducted a study aimed at treatment of craniofacial fibrous dysplasia with AR navigation and analyze its usability in cranio-maxillofacial surgery. Randomly selected data from five patients with craniofacial fibrous dysplasia was utilized for virtual planning and 3D printing. As a reference for estimating the bony resection, the normal contralateral side of the cranium was mirrored on the diseased side. The virtual recontouring plan was then superimposed on a 3D skull model. For AR registration, a marker fixed to patient's dental model is tracked by HoloLens and Optical Tracking System (OTS), which was also used to track the surgical drill in real time. For accuracy measurement, a post-operative 3D model was superimposed onto a pre-operative surgical plan. Using software for 3D analysis, discrepancies were measured Results/Conclusion: The mean error across all five recontoured cranium models using AR guidance was 1.036 ± 0.081 mm. It was concluded that the AR treatment modality for craniofacial fibrous dysplasia is both effective and safe

(Continued on following page)

10.3389/fbioe.2023.1276338

TABLE 6 (*Continued*) Brief description of each article with aim and major outcomes, on Maxillofacial Oncology.

Reference	Author, date	Brief description with aim and outcomes
Kim et al. (2020)	Jin Kim et al., 2020	Aim: In this case report, virtual planning and AR guidance were employed for the resection of a fibrous histiocytoma involving maxillary sinuses on left side Clinical symptoms: A 39-year-old Korean presented with facial deformity and discomfort in the left maxillary and palate region Investigations: Physical examination revealed no cervical lymphadenopathy and imaging showed radiolucent lesion. A CT scan revealed an expansile mass obstructing the left maxillary sinus entirely and extending into the nasal cavity Pre-requisites: Four osteotomies were planned for the total maxillectomy according to pre-surgical virtual planning. The AR viewer could track forehead marker and overlay images on actual patient. Additionally, a patient-specific polycaprolactone mesh implant was 3D printed for reconstruction of the orbital floor. For intervention, a Modified Weber Ferguson incision was made and AR guided surgery was performed using AR viewer. Osteotomies were performed and post-operative CT scan was collected to assess discrepancies between the pre-planned osteotomies and the actual osteotomies performed Results/Conclusion: The discrepancies in frontomaxillary, pterygomaxillary, pre-maxillary and zygomaticomaxillary osteotomies were 4.99, 2.25, 1.95, and 1.88 mm, respectively, with overall mean of 2.77 ± 1.29 mm. In conclusion, total maxillectomy performed under the AR guidance is appeared to be a powerful tool in applied surgery
Yang et al. (2022)	Yang et al., 2022	Aim: The authors examined the effect of cross-linking a mixed reality display device (HoloLens) and an optical navigator, analysing its accuracy with enhanced spatial relationship and graphical conversion on four patients with mandibular tumor. Imaging data was gathered for pre-operative 3D reconstruction and planning, and it was compared with 1 month postoperative data via error distribution map (3DMeshMetric) for error analysis Results/Conclusion: Using six marker points (anatomical landmarks), point registration was performed with a controlled registration error of less than 1 mm. In all cases, the surgical error analysis revealed a maximum error of 6.08 mm, with 4.79 mm in the majority of areas. Due to variations in chin surface and fibular morphology for reconstruction, the error is confined to the area of the chin in all cases. However, the authors highlighted that despite the fact this system enhances spatial experience and work efficiency, the overall accuracy still does not meet clinical requirements
Tang et al. (2022)	Tang et al., 2022	Aim: The authors conducted a retrospective study on seven patients (four with a maxillary tumor and three with a mandibular tumor) to evaluate the efficacy and accuracy of mixed reality augmented by surgical navigation. Virtual surgical plan with osteotomy planes was designed following image acquisition and fusion. Moreover, integration of mixed reality and surgical navigator was accomplished through IGT-link port for the transmission of image data between workstations. Using a HMD (HoloLens), osteotomy lines were distinguished by predetermined reference points using a navigation probe, and the spatial relationships between the probe and approaching structures were displayed on the user's retina Results/Conclusion: The target error was set at 2 mm and intra-operative frozen section biopsies were performed to guarantee negative resection. A CT scan was ordered 1 week after surgery. Both pre-operative and post-operative 3D models were analyzed for accuracy. In total of 13 groups of osteotomy planes, the mean deviation from pre-defined osteotomy plane was 1.68 ± 0.92 mm and 80.16% of mean deviations were within 3 mm. Whereas, in context of maxillary and mandibular tumor, the mean deviations were 1.60 ± 0.93 mm and 1.86 ± 0.93 mm, respectively. During the follow-up period, no significant complications or recurrences were observed in any patients. Although considered safe and effective during oral and maxillofacial surgeries, the authors emphasized the need for additional research on surgical navigation and mixed reality application
Scolozzi and Bijlenga (2017)	Scolozzi et al., 2017	Aim: In this case report, researchers utilized tracked-microscope-based AR system for excision of pleomorphic adenoma of lacrimal gland Clinical Symptoms: In 2015, a 42 years old female was presented with facial asymmetry and an asymptomatic, slow growing mass in the upper lid of her left eye, resulting in restricted movement and diminished visual acuity Investigations: CT imaging showed multilobulated mass in the left lacrimal gland as well as erosion of upper and lateral orbital walls. Later, biopsy confirmed pleomorphic adenoma Results/Conclusion: The en-bloc excision was performed using AR and neuronavigation. The authors emphasized that this system was beneficial for identification, extension of lesion as well as en-bloc resection. At 2 years imaging follow-up, no recurrence was detected
Cercenelli et al. (2022)	Cercenelli et al., 2022	Aim: Authors evaluated the achievable registration accuracy and the success rate in performing an AR-guided skin paddle incision in fibular flap harvesting for mandibular reconstruction. They performed the experiment on a patient-specific 3D printed phantom and two display solutions (tablet and HoloLens 2) were compared Results/Conclusion: On average, the marker-less AR protocol showed comparable registration errors (ranging within 1–5 mm) for tablet-based and HoloLens-based solution. In 97% and 100% of cases, the AR-guided task was performed with an accuracy of ±2 mm (error margin of 4 mm), for tablet and HoloLens, respectively. The authors concluded that the proposed marker-less AR protocol can be suitable for assisting skin paddle harvesting in clinical setting

(Continued on following page)

TABLE 6 (*Continued*) Brief description of each article with aim and major outcomes, on Maxillofacial Oncology.

Reference	Author, date	Brief description with aim and outcomes
Scherl et al. (2021)	Scherl et al., 2021	Aim: Authors reported the first ever trial of AR-assisted surgery for parotid tumor ablation on six patients. Segmentation was done using MRI images and 3D hologram is manually overlaid on patient's skin to estimate the position of tumor border. Alignment accuracy was measured through pre-defined landmarks using electromagnetic navigation device Results/Conclusion: The mean error of alignment was 1.3 cm (0.5–2.1). Statistical difference was noted among central and peripheral structures showing better accuracy centrally ($p = 0.0059$). No long-term complication was noted in any AR-guided case. However, authors emphasized that further work need to be done on skin surface registration and alignment to improve accuracy
Necker et al. (2023)	Necker et al., 2022	Aim: Authors introduced a technique for flap sizing after mandible tumor resection on cadaver. A resected mandibular tumor was 3D scanned using smartphone and annotated with colours for orientation. The 3D scanned virtual specimen was displayed through AR-HMD Results/Conclusion: The 3D hologram was manually placed back in its original site accurately. After adjusting the size, the hologram was then overlaid on flap harvesting site for reconstruction planning. Authors concluded that this technique could assist in fibular flap reconstruction of mandibular tumor
Han et al. (2022)	Lin et al., 2022	Aim: Authors proposed an integrated approach in mandibular reconstruction based on 3 technologies: 3D printing, mixed reality (MR), Robotic-Assisted navigation (RAN). MR was used to output the visualized project and matched the anatomical 3D reconstruction model in reality. The 3D plate was printed for surgical guidance. RAN was used to guide and position the vascularized fibula autograft and the immediate dental implantation Results/Conclusion: Authors concluded that constructed MR, 3D, and RAN technologies assist each other to make the surgery more accurate and minimally invasive
Zhao et al. (2023)	Zhao et al., 2022	Aim: Authors reported a cadaver study to investigate a novel method of fibula free flap (FFF) osteotomy based on AR technology. AR-based surgical navigation was used to guide the FFF osteotomy and these fibular segments were used to reconstruct a defective mandible model. After reconstruction, all segments were scanned by CT and osteotomy accuracy was evaluated by measuring the length and angular deviation between the virtual plan and the final result, and the volume overlap rate and average surface distance between the planned and obtained reconstruction Results/Conclusion: The length difference, angular deviation, volume overlap rate and average surface distance were 1.18 ± 0.84 mm, 5.45 ± 1.47°, 95.31% ± 2.09%, and 1.22 ± 0.12 mm, respectively
Modabber et al. (2022)	Modabber et al., 2021	Aim: Authors proposed AR guided solution for harvesting grafts for mandibular reconstruction and demonstrated its clinical application and surgical accuracy. They used human cadavers to demonstrate a projector-based marker-less AR setup for harvesting iliac crest grafts and compared it to a cutting guides procedure. A total of 10 iliac crests from 5 cadavers were used, and each iliac crest was randomly assigned to either harvested using AR guidance or using cutting guides. The transplants were digitized using CBCT and accuracy was measured in terms of angles, distances, and volumes of the actual and intended osteotomies Results/Conclusion: Both AR and cutting guides effectively replicated the virtually projected transplant volume with precision. Nevertheless, there was a significant difference ($p = 0.018$) in the cumulative angulation of the osteotomy plane between the AR group and the group using cutting guides (14.99 ± 11.69° vs. 8.49 ± 5.42°). The results indicate that the accuracy of AR-guided navigation in terms of cumulative osteotomy plane distance was lower (2.65 ± 3.32 mm) compared to the cutting guides (1.47 ± 1.36 mm), however this difference was not statistically significant. Furthermore, more research was recommended before clinical implementation
Meng et al. (2021)	Meng et al., 2021	Aim: Authors reported a cadaver study to investigate the feasibility of the application of MR in mandible reconstruction with fibula flap. AR was used to guide the osteotomy and shaping of the fibular bone. After fixing the fibular segments using the titanium plate, all segments underwent a CT examination. The planned and the actual postoperative fibula osteotomies were compared in terms of angular deviation of fibular segments, and intergonial angle distances. To evaluate the accuracy of MR technique, the distance between the postoperative actual cutting edge and preoperative osteotomy plan of each fibular segment's was calculated Results/Conclusion: The mean location of the fibular osteotomies, angular deviation of the fibular segments, and intergonial angle distances were 2.11 ± 1.31 mm, 2.85° ± 1.97°, and 7.24 ± 3.42 mm, respectively

(Continued on following page)

10.3389/fbioe.2023.1276338

TABLE 6 (*Continued*) **Brief description of each article with aim and major outcomes, on Maxillofacial Oncology.**

Reference	Author, date	Brief description with aim and outcomes
Battaglia et al. (2019)	Battaglia et al., 2019	Aim: Authors reported a case series of 3 consecutive patients who underwent mandibular reconstruction using AR-assisted fibular free flap harvesting. Using an app installed on a tablet, the virtual-to-real registration is performed according to a shape recognition system of the leg of the patient, rendering in real time a superimposition of the anatomy of the bony, vascular, and skin of the patient and also the surgical planning of the reconstruction Results/Conclusion: Accuracy of AR overlay was verified visually by the surgeons, who believe that AR can be a prospective improving technology for mandibular complex reconstruction
Winnand et al. (2022)	Winnand et al., 2021	Aim: The objective of this research is to examine the usability, visual perception, and accuracy of projector-based marker-less navigation of AR guided harvesting of the iliac crest, in comparison to harvesting using cutting guides for facial abnormalities on phantoms. A random selection of $5^{-\Delta\Delta CT}$ scans was made to enroll a total of 10 iliac crests. Two commonly shaped transplants (box shaped and hockey stick shaped) were virtually planned. Each of the 10 iliac crests was replicated (3D printed) four times for the purpose of testing. Following the surgical removal, the transplants underwent digitization using CBCT. The accuracy of the measurements was assessed by quantifying the volume, distance, and angular deviation Results/Conclusion: In the context of volume rendering, the osteotomies supported by AR demonstrated higher levels of accuracy compared to those performed using a cutting guide ($p = 0.057$). Therefore, a statistically significant difference was seen in the accuracy of osteotomies for box-shaped transplants, but no significant difference was observed for hockey stick-shaped transplants. In general, the average discrepancy in osteotomy plane angulation from the virtual design was found to be $10.21 \pm 7.22°$ with the assistance of AR, and $6.98 \pm 4.70°$ when utilizing cutting guides. A notable disparity between the methods was observed, with a statistically significant difference ($p = 0.004$). Similarly, the mean differences between the osteotomy planes and the planned trajectories were determined to be 2.29 ± 1.98 mm for AR and 1.32 ± 1.00 mm for cutting guides. This analysis revealed a significant difference between the two methods ($p = 0.002$). Therefore, due to certain limitations in AR-guided harvesting further work on research was recommended

Virtual-to-real scene registration with HMDs is another major issue when using AR for surgical guidance and simulation, resulting in inaccurate identification of the deep anatomical structure in question. Due to the fact that these display devices were not devised for medical purposes, their technical characteristics are less suited for surgical procedures (Badiali et al., 2020).

Surely, a marker-less registration, i.e., without the use of fiducial markers or trackers anchored to the patient, is highly preferable in surgery, however it is not always feasible for certain surgeries. As also emerged from our analysis, in maxillofacial oncological surgery a marker-less based approach seems to be more viable since the edges of anatomical parts (e.g., mandible or skull) are more accessible and trackable during surgery; conversely, in orthopedic field the intraoperative recognition of bone edges may be more difficult.

During surgery, a surgeon using an AR headset may endure discomfort, weariness, eye strains, and headache. Furthermore, it is possible that the surgeon's ability to focus on the surgical operation field might get impaired due to the visualization of augmented information. However, in a study conducted on simulation sickness, in 2018, authors showed that out of 142 HMD's users from various fields, only few experienced mild discomfort (Vovk et al., 2018).

Because this technology is high priced and requires significant investment to initiate and implement in hospitals, it is not widely accessible in all medical settings. Consequently, the data associated with augmented reality in the surgical field are rather preliminary and require further testing and analysis.

The AR application has a steep and costly learning curve. It requires expertise, and most surgeons are unfamiliar with AR utilization. So they must collaborate with biomedical engineers to implement this technology in the operating room. Personnel must endure time-consuming hands-on training in order to implement AR in hospital surgical setups.

When information needs to be electronically distributed to several departments in order to make patient-specific tools for AR, patient data privacy is another concern that must be addressed. When dealing with sensitive information pertaining to patients, certain national regulations should be followed in order to protect and secure both their safety and their privacy.

4.2 Future of AR

Augmented reality is being recognized as a promising application for enhancing the outcomes and standard of care for orthopedic and maxillofacial oncology patients. Especially in complex oncological cases that require a strategic planning for execution and comprehension, it is emphasized that surgeons should consider using augmented reality where applicable, in combination with 3D printing. However, implementation of AR and its tools in surgical cases and healthcare has certain shortcomings which can be improved with future advancement in technology. Depending on the complexity of cases, process from procuring CT scan/MRI images to 3D printing of pre-surgical phantoms and developing a patient-specific AR application take days, sometimes even months including preoperative surgical planning and simulation training. Surgeons and biomedical engineers must work together to refine and successfully execute the procedure. Therefore, as future perspective, AR cannot be used in

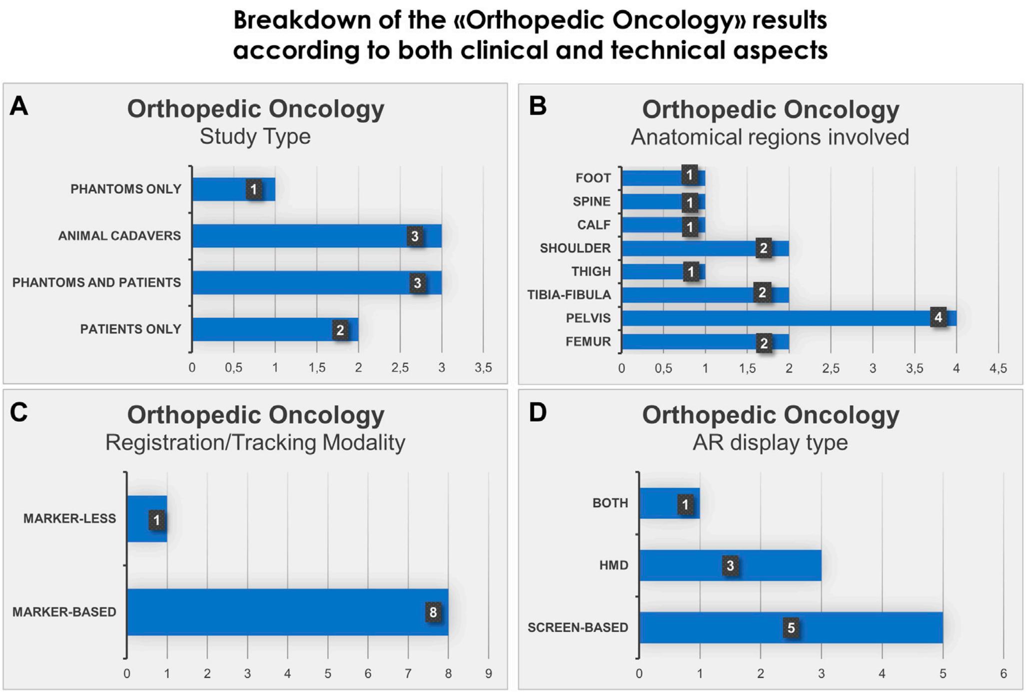


FIGURE 2
Bar histograms depicting types of study **(A)**, anatomical regions involved in the surgery **(B)**, registration/tracking modality **(C)** and AR display type **(D)**, for Orthopedic Oncology.

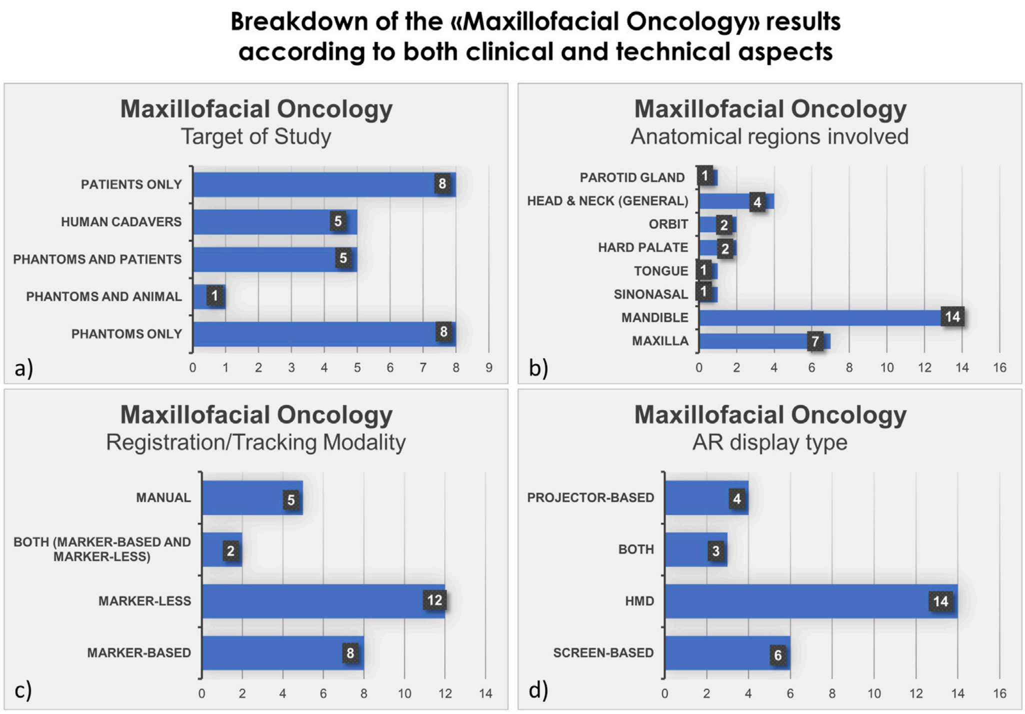


FIGURE 3
Bar histograms depicting types of study **(A)**, anatomical regions involved in the surgery **(B)**, registration/tracking modality **(C)** and AR display type **(D)**, for Maxillofacial Oncology.

emergency cases and can only be advised in elective surgeries. Despite these facts, we continue to believe that AR technology has the potential to revolutionize the conventional methods of oncological surgical procedures and overcome all of these limitations. Augmented reality is expected to help in better understanding of tumor anatomy and in plan the resection accordingly. This will contribute to improve the outcomes and standard of care by limiting the recurrence rate while attaining the desired surgical margins with accuracy.

5 Conclusion

Currently, augmented reality is one of the most innovative technology in the field of surgery, particularly orthopedic and maxillofacial surgery. Due to its real-time visualization of preoperative images and planning directly on the patient, AR can be particularly beneficial for oncological surgeries in both fields to achieve the desired surgical accuracy. In oncological surgery, the AR allows to overcome some limitations of conventional computer-assisted surgical navigation, such as the surgeon's attention shift from the operative field to view the navigation monitor, as well as to avoid the lead-time in manufacturing 3D-printed cutting guides. Indeed, AR should be used as a complementary tool to other computer-assisted technologies, as suggested by our literature review: particularly for maxillofacial oncology, surgeons have begun to incorporate external navigation systems into AR to track the surgical probe or instruments, to further improve accuracy and spatial relationships.

Even though AR technology is still in its infancy and has certain limitations, the current outcomes of its application in both disciplines are promising to support its clinical use. Certain concerning aspects still remain, related to image-to-patient registration and surgical accuracy. In the present review, we attempted to identify a range of registration error and surgical accuracy based on results from both surgical domains. Although it is difficult to derive general ranges due to the involvement of various anatomical regions and different complexity for each domain it can be observed that AR resection error exhibited greater accuracy compared to conventional un-guided resection, hence successfully attaining the desired goals without any associated complications. Additionally, in some studies the AR navigation showed comparable accuracy with pre-planned virtual cutting planes and with customized cutting guides.

We believe that the still limiting technical aspects on registration and surgical accuracy can be improved and overcome with further development in hardware and software used for AR. For that, industry-academic partnerships are essential to advance the technology, in conjunction with clinical studies to assess its benefits and role in the clinical practice.

In conclusion, although AR is seen to have the capacity to enhance surgical efficiency and ensure patient's safety, further search needs to be done pre-clinically in order to improve its accuracy and before achieving its wide adoption in clinical settings.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

NN: Conceptualization, Data curation, Methodology, Writing–original draft, Writing–review and editing. LC: Conceptualization, Data curation, Methodology, Writing–original draft, Writing–review and editing. AT: Resources, Supervision, Writing–review and editing. EM: Resources, Supervision, Writing–review and editing.

Funding

The authors declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdel Al, S., Chaar, M. K. A., Mustafa, A., Al-Hussaini, M., Barakat, F., and Asha, W. (2020). Innovative surgical planning in resecting soft tissue sarcoma of the foot using augmented reality with a smartphone. *J. Foot Ankle Surg.* 59 (5), 1092–1097. doi:10.1053/j.jfas.2020.03.011

Alexander, C., Loeb, A. E., Fotouhi, J., Navab, N., Armand, M., and Khanuja, H. S. (2020). Augmented reality for acetabular component placement in direct anterior total hip arthroplasty. *J. Arthroplasty* 35 (6), 1636–1641.e3. doi:10.1016/j.arth.2020.01.025

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Comput. Graph. Appl.* 21 (6), 34–47. doi:10.1109/38.963459

Badiali, G., Cercenelli, L., Battaglia, S., Marcelli, E., Marchetti, C., Ferrari, V., et al. (2020). Review on augmented reality in oral and cranio-maxillofacial surgery: toward "surgery-specific" head-up displays. *IEEE Access* 8, 59015–59028. doi:10.1109/access.2020.2973298

Barcali, E., Iadanza, E., Manetti, L., Francia, P., Nardi, C., and Bocchi, L. (2022). Augmented reality in surgery: a scoping review. *A Scoping Rev.* 12 (14), 6890. doi:10.3390/app12146890

Battaglia, S., Badiali, G., Cercenelli, L., Bortolani, B., Marcelli, E., Cipriani, R., et al. (2019). Combination of CAD/CAM and augmented reality in free fibula bone harvest. *Plastic Reconstr. Surg. Glob. open* 7 (11), e2510. doi:10.1097/gox.0000000000002510

Battaglia, S., Ratti, S., Manzoli, L., Marchetti, C., Cercenelli, L., Marcelli, E., et al. (2020). Augmented reality-assisted periosteum pedicled flap harvesting for head and neck reconstruction: an anatomical and clinical viability study of a galeo-pericranial flap. *J. Clin. Med.* 9 (7), 2211. doi:10.3390/jcm9072211

Bianchi, L., Chessa, F., Angiolini, A., Cercenelli, L., Lodi, S., Bortolani, B., et al. (2021). The use of augmented reality to guide the intraoperative frozen section during robot-assisted radical prostatectomy. *Eur. Urol.* 80 (4), 480–488. doi:10.1016/j.eururo.2021.06.020

Bradley, W. G. (2008). History of medical imaging. *Proc. Am. Philosophical Soc.* 152 (3), 349–361.

Budhathoki, S., Alsadoon, A., Prasad, P. W. C., Haddad, S., and Maag, A. (2020). Augmented reality for narrow area navigation in jaw surgery: modified tracking by detection volume subtraction algorithm. *Int. J. Med. robotics + computer assisted surgery MRCAS* 16 (3), e2097. doi:10.1002/rcs.2097

Cannizzaro, D., Zaed, I., Safa, A., Jelmoni, A. J. M., Composto, A., Bisoglio, A., et al. (2022). Augmented reality in neurosurgery, state of art and future projections. A systematic review. *Frontiers in surgery* 9, 864792. doi:10.3389/fsurg.2022.864792

Ceccariglia, F., Cercenelli, L., Badiali, G., Marcelli, E., and Tarsitano, A. (2022). Application of augmented reality to maxillary resections: a three-dimensional approach to maxillofacial oncologic surgery. *Maxillofacial Oncologic Surgery* 12 (12), 2047. doi:10.3390/jpm12122047

Cercenelli, L., Babini, F., Badiali, G., Battaglia, S., Tarsitano, A., Marchetti, C., et al. (2022). Augmented reality to assist skin paddle harvesting in osteomyocutaneous fibular flap reconstructive surgery: a pilot evaluation on a 3D-printed leg phantom. *Front. Oncol.* 11, 804748. doi:10.3389/fonc.2021.804748

Cercenelli, L., Carbone, M., Condino, S., Cutolo, F., Marcelli, E., Tarsitano, A., et al. (2020). The wearable VOSTARS system for augmented reality-guided surgery: preclinical phantom evaluation for high-precision maxillofacial tasks. *Journal of clinical medicine* 9 (11), 3562. doi:10.3390/jcm9113562

Chan H, H. L., Sahovaler, A., Daly, M. J., Ferrari, M., Franz, L., Gualtieri, T., et al. (2022). Projected cutting guides using an augmented reality system to improve surgical margins in maxillectomies: a preclinical study. *Oral oncology* 127, 105775. doi:10.1016/j.oraloncology.2022.105775

Cho, H. S., Park, M. S., Gupta, S., Han, I., Kim, H. S., Choi, H., et al. (2018). Can augmented reality Be helpful in pelvic bone cancer surgery? An *in vitro* study. *Clinical orthopaedics and related research* 476 (9), 1719–1725. doi:10.1007/s11999.0000000000000233

Cho, H. S., Park, Y. K., Gupta, S., Yoon, C., Han, I., Kim, H. S., et al. (2017). Augmented reality in bone tumour resection: an experimental study. *Bone Joint Res* 6 (3), 137–143. doi:10.1302/2046-3758.63.bjr-2016-0289.r1

Choi, H., Cho, B., Masamune, K., Hashizume, M., and Hong, J. (2016). An effective visualization technique for depth perception in augmented reality-based surgical navigation. *The international journal of medical robotics + computer assisted surgery MRCAS* 12 (1), 62–72. doi:10.1002/rcs.1657

Choi, H., Park, Y., Lee, S., Ha, H., Kim, S., Cho, H. S., et al. (2017). A portable surgical navigation device to display resection planes for bone tumor surgery. *Minimally Invasive Therapy & Allied Technologies* 26 (3), 144–150. doi:10.1080/13645706.2016.1274766

Condino, S., Fida, B., Carbone, M., Cercenelli, L., Badiali, G., Ferrari, V., et al. (2020). Wearable augmented reality platform for aiding complex 3D trajectory tracing. *Sensors (Basel).* 20 (6), 1612. doi:10.3390/s20061612

Dennler, C., Jaberg, L., Spirig, J., Agten, C., Götschi, T., Fürnstahl, P., et al. (2020). Augmented reality-based navigation increases precision of pedicle screw insertion. *Journal of Orthopaedic Surgery and Research* 15 (1), 174. doi:10.1186/s13018-020-01690-x

Elmi-Terander, A., Burström, G., Nachabé, R., Fagerlund, M., Ståhl, F., Charalampidis, A., et al. (2020). Augmented reality navigation with intraoperative 3D imaging vs fluoroscopy-assisted free-hand surgery for spine fixation surgery: a matched-control study comparing accuracy. *Sci Rep* 10 (1), 707. doi:10.1038/s41598-020-57693-5

Elmi-Terander, A., Burström, G., Nachabe, R., Skulason, H., Pedersen, K., Fagerlund, M., et al. (2019). Pedicle screw placement using augmented reality surgical navigation with intraoperative 3D imaging: a first in-human prospective cohort study. *Spine* 44 (7), 517–525. doi:10.1097/brs.0000000000002876

Elmi-Terander, A., Nachabe, R., Skulason, H., Pedersen, K., Söderman, M., Racadio, J., et al. (2018). Feasibility and accuracy of thoracolumbar minimally invasive pedicle screw placement with augmented reality navigation technology. *Spine* 43 (14), 1018–1023. doi:10.1097/brs.0000000000002502

Feiner, S. K. (2002). Augmented reality: a new way of seeing. *Scientific American* 286 (4), 48–55. doi:10.1038/scientificamerican0402-48

Ferraguti, F., Farsoni, S., and Bonfè, M. (2022). Augmented reality and robotic systems for assistance in percutaneous nephrolithotomy procedures: recent advances and future perspectives. *Recent Advances and Future Perspectives* 11 (19), 2984. doi:10.3390/electronics11192984

Gao, Y., Lin, L., Chai, G., and Xie, L. (2019). A feasibility study of a new method to enhance the augmented reality navigation effect in mandibular angle split osteotomy. *Journal of cranio-maxillo-facial surgery official publication of the European Association for Cranio-Maxillo-Facial Surgery* 47 (8), 1242–1248. doi:10.1016/j.jcms.2019.04.005

Gao, Y., Liu, K., Lin, L., Wang, X., and Xie, L. (2022). Use of augmented reality navigation to optimise the surgical management of craniofacial fibrous dysplasia. *British Journal of Oral and Maxillofacial Surgery* 60 (2), 162–167. doi:10.1016/j.bjoms.2021.03.011

García-Sevilla, M., Moreta-Martinez, R., García-Mato, D., Arenas de Frutos, G., Ochandiano, S., Navarro-Cuéllar, C., et al. (2021b). Surgical navigation, augmented reality, and 3D printing for hard palate adenoid cystic carcinoma en-bloc resection: case report and literature review. *Frontiers in oncology* 11, 741191. doi:10.3389/fonc.2021.741191

García-Sevilla, M., Moreta-Martinez, R., García-Mato, D., Pose-Diez-de-la-Lastra, A., Pérez-Mañanes, R., Calvo-Haro, J. A., et al. (2021a). Augmented reality as a tool to guide PSI placement in pelvic tumor resections. *Sensors (Basel)* 21 (23), 7824. doi:10.3390/s21237824

Gibby, J. T., Swenson, S. A., Cvetko, S., Rao, R., and Javan, R. (2019). Head-mounted display augmented reality to guide pedicle screw placement utilizing computed tomography. *International journal of computer assisted radiology and surgery* 14 (3), 525–535. doi:10.1007/s11548-018-1814-7

Golse, N., Petit, A., Lewin, M., Vibert, E., and Cotin, S. (2021). Augmented reality during open liver surgery using a markerless non-rigid registration system. *Journal of gastrointestinal surgery official journal of the Society for Surgery of the Alimentary Tract* 25 (3), 662–671. doi:10.1007/s11605-020-04519-4

Gouveia, P. F., Costa, J., Morgado, P., Kates, R., Pinto, D., Mavioso, C., et al. (2021). Breast cancer surgery with augmented reality. *The Breast* 56, 14–17. doi:10.1016/j.breast.2021.01.004

Gsaxner, C., Pepe, A., Li, J., Ibrahimpasic, U., Wallner, J., Schmalstieg, D., et al. (2021). Augmented reality for head and neck carcinoma imaging: description and feasibility of an instant calibration, markerless approach. *Computer Methods and Programs in Biomedicine* 200, 105854. doi:10.1016/j.cmpb.2020.105854

Gsaxner, C., Pepe, A., Wallner, J., Schmalstieg, D., and Egger, J. (2019). "Markerless image-to-face registration for untethered augmented reality in head and neck surgery," in 22nd International Conference, Shenzhen, China, October 2019, 236–244.

Han, J. J., Sodnom-Ish, B., Eo, M. Y., Kim, Y. J., Oh, J. H., Yang, H. J., et al. (2022). Accurate mandible reconstruction by mixed reality, 3D printing, and robotic-assisted navigation integration. *The Journal of craniofacial surgery* 33 (6), e701–e706. doi:10.1097/SCS.0000000000008603

Jiang, T., Yu, D., Wang, Y., Zan, T., Wang, S., and Li, Q. (2020). HoloLens-based vascular localization system: precision evaluation study with a three-dimensional printed model. *Journal of medical Internet research* 22 (4), e16852. doi:10.2196/16852

Kim, H.-J., Jo, Y.-J., Choi, J.-S., Kim, H.-J., Park, I.-S., You, J.-S., et al. (2020). Virtual reality simulation and augmented reality-guided surgery for total maxillectomy: a case report. *A Case Report* 10 (18), 6288. doi:10.3390/app10186288

Lareyre, F., Chaudhuri, A., Adam, C., Carrier, M., Mialhe, C., and Raffort, J. (2021). Applications of head-mounted displays and smart glasses in vascular surgery. *Annals of Vascular Surgery* 75, 497–512. doi:10.1016/j.avsg.2021.02.033

Lee, D., Yu, H. W., Kim, S., Yoon, J., Lee, K., Chai, Y. J., et al. (2020). Vision-based tracking system for augmented reality to localize recurrent laryngeal nerve during robotic thyroid surgery. *Scientific Reports* 10 (1), 8437. doi:10.1038/s41598-020-65439-6

Li, T., Li, C., Zhang, X., Liang, W., Chen, Y., Ye, Y., et al. (2021). Augmented reality in ophthalmology: applications and challenges. *Front Med (Lausanne)* 8, 733241. doi:10.3389/fmed.2021.733241

Ma, L., Jiang, W., Zhang, B., Qu, X., Ning, G., Zhang, X., et al. (2019). Augmented reality surgical navigation with accurate CBCT-patient registration for dental implant placement. *Medical & biological engineering & computing* 57 (1), 47–57. doi:10.1007/s11517-018-1861-9

Mahmud, N., Cohen, J., Tsourides, K., and Berzin, T. (2015). Computer vision and augmented reality in gastrointestinal endoscopy: figure 1. *Gastroenterology Report* 3, 179–184. doi:10.1093/gastro/gov027

Meng, F. H., Zhu, Z. H., Lei, Z. H., Zhang, X. H., Shao, L., Zhang, H. Z., et al. (2021). Feasibility of the application of mixed reality in mandible reconstruction with fibula flap: a cadaveric specimen study. *Journal of stomatology, oral and maxillofacial surgery* 122 (4), e45–e49. doi:10.1016/j.jormas.2021.01.005

Modabber, A., Ayoub, N., Redick, T., Gesenhues, J., Kniha, K., Möhlhenrich, S. C., et al. (2022). Comparison of augmented reality and cutting guide technology in assisted harvesting of iliac crest grafts - a cadaver study. *Annals of anatomy = Anatomischer Anzeiger official organ of the Anatomische Gesellschaft* 239, 151834. doi:10.1016/j.aanat.2021.151834

Molina, C. A., Dibble, C. F., Lo, S.-fL., Witham, T., and Sciubba, D. M. (2021c). Augmented reality–mediated stereotactic navigation for execution of *en bloc* lumbar spondylectomy osteotomies. *Journal of Neurosurgery Spine* 34 (5), 700–705. doi:10.3171/2020.9.spine201219

Molina, C. A., Dibble, C. F., Lo, S. L., Witham, T., and Sciubba, D. M. (2021a). Augmented reality-mediated stereotactic navigation for execution of *en bloc* lumbar spondylectomy osteotomies. *J Neurosurg Spine* 34, 700–705. doi:10.3171/2020.9.spine201219

Molina, C. A., Sciubba, D. M., Greenberg, J. K., Khan, M., and Witham, T. (2021b). Clinical accuracy, technical precision, and workflow of the first in human use of an augmented-reality head-mounted display stereotactic navigation system for spine surgery. *Operative neurosurgery (Hagerstown, Md)* 20 (3), 300–309. doi:10.1093/ons/opaa398

Moreta-Martinez, R., García-Mato, D., García-Sevilla, M., Pérez-Mañanes, R., Calvo-Haro, J., and Pascau, J. (2018). Augmented reality in computer-assisted interventions based on patient-specific 3D printed reference. *Healthc Technol Lett* 5 (5), 162–166. doi:10.1049/htl.2018.5072

Moreta-Martinez, R., Pose-Díez-de-la-Lastra, A., Calvo-Haro, J. A., Mediavilla-Santos, L., Pérez-Mañanes, R., and Pascau, J. (2021). Combining augmented reality and 3D printing to improve surgical workflows in orthopedic oncology: smartphone application and clinical evaluation. *Sensors (Basel, Switzerland)* 21 (4), 1370. doi:10.3390/s21041370

Necker, F. N., Chang, M., Leuze, C., Topf, M. C., Daniel, B. L., and Baik, F. M. (2023). Virtual resection specimen interaction using augmented reality holograms to guide margin communication and flap sizing. *Otolaryngology–head and neck surgery* 169, 1083–1085. doi:10.1002/ohn.325

Ochandiano, S., García-Mato, D., Gonzalez-Alvarez, A., Moreta-Martinez, R., Tousidonis, M., Navarro-Cuellar, C., et al. (2021). Computer-assisted dental implant placement following free flap reconstruction: virtual planning, CAD/CAM templates, dynamic navigation and augmented reality. *Frontiers in oncology* 11, 754943. doi:10.3389/fonc.2021.754943

Ogawa, H., Hasegawa, S., Tsukada, S., and Matsubara, M. (2018). A pilot study of augmented reality technology applied to the acetabular cup placement during total hip arthroplasty. *J Arthroplasty* 33 (6), 1833–1837. doi:10.1016/j.arth.2018.01.067

Pellegrino, G., Mangano, C., Mangano, R., Ferri, A., Taraschi, V., and Marchetti, C. (2019). Augmented reality for dental implantology: a pilot clinical report of two cases. *BMC Oral Health* 19 (1), 158. doi:10.1186/s12903-019-0853-y

Pepe, A., Trotta, G. F., Mohr-Ziak, P., Gsaxner, C., Wallner, J., Bevilacqua, V., et al. (2019). A marker-less registration approach for mixed reality-aided maxillofacial surgery: a pilot evaluation. *Journal of digital imaging* 32 (6), 1008–1018. doi:10.1007/s10278-019-00272-6

Pietruski, P., Majak, M., Światek-Najwer, E., Żuk, M., Popek, M., Mazurek, M., et al. (2019). Supporting mandibular resection with intraoperative navigation utilizing augmented reality technology - a proof of concept study. *Journal of cranio-maxillofacial surgery* 47 (6), 854–859. doi:10.1016/j.jcms.2019.03.004

Pose-Díez-de-la-Lastra, A., Moreta-Martinez, R., García-Sevilla, M., García-Mato, D., Calvo-Haro, J. A., Mediavilla-Santos, L., et al. (2022). HoloLens 1 vs. HoloLens 2: improvements in the new model for orthopedic oncological interventions. *Oncological Interventions* 22 (13), 4915. doi:10.3390/s22134915

Prasad, K., Miller, A., Sharif, K., Colazo, J. M., Ye, W., Necker, F., et al. (2023). Augmented-reality surgery to guide head and neck cancer Re-resection: a feasibility and accuracy study. *Annals of Surgical Oncology* 30, 4994–5000. doi:10.1245/s10434-023-13532-1

Rad, A. A., Vardanyan, R., Lopuszko, A., Alt, C., Stoffels, I., Schmack, B., et al. (2022). Virtual and augmented reality in cardiac surgery. *Brazilian journal of cardiovascular surgery* 37 (1), 123–127. doi:10.21470/1678-9741-2020-0511

Roberts, S., Desai, A., Checcucci, E., Puliatti, S., Taratkin, M., Kowalewski, K. F., et al. (2022). Augmented reality" applications in urology: a systematic review. *Minerva urology and nephrology* 74 (5), 528–537. doi:10.23736/s2724-6051.22.04726-7

Ruggiero, F., Cercenelli, L., Emiliani, N., Badiali, G., Bevini, M., Zucchelli, M., et al. (2023). Preclinical application of augmented reality in pediatric craniofacial surgery: an accuracy study. *An Accuracy Study* 12 (7), 2693. doi:10.3390/jcm12072693

Sahovaler, A., Chan, H. H. L., Gualtieri, T., Daly, M., Ferrari, M., Vannelli, C., et al. (2021). Augmented reality and intraoperative navigation in sinonasal malignancies: a preclinical study. *Frontiers in oncology* 11, 723509. doi:10.3389/fonc.2021.723509

Scherl, C., Stratemeier, J., Rotter, N., Hesser, J., Schönberg, S. O., Servais, J. J., et al. (2021). Augmented reality with HoloLens® in parotid tumor surgery: a prospective feasibility study. *ORL* 83 (6), 439–448. doi:10.1159/000514640

Schiavina, R., Bianchi, L., Chessa, F., Barbaresi, U., Cercenelli, L., Lodi, S., et al. (2021). Augmented reality to guide selective clamping and tumor dissection during robot-assisted partial nephrectomy: a preliminary experience. *Clinical genitourinary cancer* 19 (3), e149–e155. doi:10.1016/j.clgc.2020.09.005

Schwam, Z. G., Kaul, V. F., Bu, D. D., Iloreta, A. C., Bederson, J. B., Perez, E., et al. (2021). The utility of augmented reality in lateral skull base surgery: a preliminary report. *Am J Otolaryngol* 42 (4), 102942. doi:10.1016/j.amjoto.2021.102942

Scolozzi, P., and Bijlenga, P. (2017). Removal of recurrent intraorbital tumour using a system of augmented reality. *British Journal of Oral and Maxillofacial Surgery* 55 (9), 962–964. doi:10.1016/j.bjoms.2017.08.360

Shaofeng, L., Yunyang, L., Bingwei, H., Bowen, D., Zhaoju, Z., Jiafeng, S., et al. (2023). Mandibular resection and defect reconstruction guided by a contour registration-based augmented reality system: a preclinical trial. *Journal of Cranio-Maxillofacial Surgery* 51, 360–368. doi:10.1016/j.jcms.2023.05.007

Sharma, A., Alsadoon, A., Prasad, P. W. C., Al-Dala'in, T., and Haddad, S. (2021). A novel augmented reality visualization in jaw surgery: enhanced ICP based modified rotation invariant and modified correntropy. *Multimedia Tools and Applications* 80 (16), 23923–23947. doi:10.1007/s11042-021-10787-2

Shi, J., Liu, S., Zhu, Z., Deng, Z., Bian, G., and He, B. (2022). Augmented reality for oral and maxillofacial surgery: the feasibility of a marker-free registration method. *The international journal of medical robotics + computer assisted surgery MRCAS* 18 (4), e2401. doi:10.1002/rcs.2401

Shrestha, L., Alsadoon, A., Prasad, P. W. C., AlSallami, N., and Haddad, S. (2021). Augmented reality for dental implant surgery: enhanced ICP. *The Journal of Supercomputing* 77 (2), 1152–1176. doi:10.1007/s11227-020-03322-x

Sielhorst, T., Bichlmeier, C., Heining, S. M., and Navab, N. (2006). "Depth perception – a major issue in medical AR: evaluation study by twenty surgeons," in *Medical image computing and computer-assisted intervention – miccai 2006* (Berlin, Germany: Springer).

Sugahara, K., Koyachi, M., Koyama, Y., Sugimoto, M., Matsunaga, S., Odaka, K., et al. (2021). Mixed reality and three dimensional printed models for resection of maxillary tumor: a case report. *Quantitative imaging in medicine and surgery* 11 (5), 2187–2194. doi:10.21037/qims-20-597

Tang, Z.-N., Hu, L.-H., Soh, H. Y., Yu, Y., Zhang, W.-B., and Peng, X. (2022). Accuracy of mixed reality combined with surgical navigation assisted oral and maxillofacial tumor resection. *Front Oncol* 11, 715484. doi:10.3389/fonc.2021.715484

Tel, A., Arboit, L., Sembronio, S., Costa, F., Nocini, R., and Robiony, M. (2021). The transantral endoscopic approach: a portal for masses of the inferior orbit—improving surgeons' experience through virtual endoscopy and augmented reality. *Front Surg* 8, 715262. doi:10.3389/fsurg.2021.715262

Tsukada, S., Ogawa, H., Nishino, M., Kurosaka, K., and Hirasawa, N. (2019). Augmented reality-based navigation system applied to tibial bone resection in total knee arthroplasty. *Journal of experimental orthopaedics* 6 (1), 44. doi:10.1186/s40634-019-0212-6

Tsukada, S., Ogawa, H., Nishino, M., Kurosaka, K., and Hirasawa, N. (2021). Augmented reality-assisted femoral bone resection in total knee arthroplasty. *JBJS Open Access* 6 (3), e21.00001. doi:10.2106/jbjs.oa.21.00001

Venkatesan, M., Mohan, H., Ryan, J. R., Schürch, C. M., Nolan, G. P., Frakes, D. H., et al. (2021). Virtual and augmented reality for biomedical applications. *Cell reports Medicine* 2 (7), 100348. doi:10.1016/j.xcrm.2021.100348

Vovk, A., Wild, F., Guest, W., and Kuula, T. (2018). "Simulator sickness in augmented reality training using the Microsoft HoloLens," in Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, April, 2018, 1–9.

Wang, J., Shen, Y., and Yang, S. (2019). A practical marker-less image registration method for augmented reality oral and maxillofacial surgery. *International journal of computer assisted radiology and surgery* 14 (5), 763–773. doi:10.1007/s11548-019-01921-5

Winnand, P., Ayoub, N., Redick, T., Gesenhues, J., Heitzer, M., Peters, F., et al. (2022). Navigation of iliac crest graft harvest using markerless augmented reality and cutting guide technology: a pilot study. *The international journal of medical robotics + computer assisted surgery MRCAS* 18 (1), e2318. doi:10.1002/rcs.2318

Wong, K. C., Sun, Y. E., and Kumta, S. M. (2022). Review and future/potential application of mixed reality technology in orthopaedic oncology. *Orthopedic research and reviews* 14, 169–186. doi:10.2147/orr.s360933

Yang, R., Li, C., Tu, P., Ahmed, A., Ji, T., and Chen, X. (2022). Development and application of digital maxillofacial surgery system based on mixed reality technology. *Front. Surg.* 8, 719985. doi:10.3389/fsurg.2021.719985

Zhao, R., Zhu, Z., Shao, L., Meng, F., Lei, Z., Li, X., et al. (2023). Augmented reality guided in reconstruction of mandibular defect with fibular flap: a cadaver study. *Journal of stomatology, oral and maxillofacial surgery* 124 (2), 101318. doi:10.1016/j.jormas.2022.10.017

Frontiers in Bioengineering and Biotechnology

TYPE Review
PUBLISHED 22 February 2024
DOI 10.3389/fbioe.2024.1329263

Check for updates

OPEN ACCESS

EDITED BY
Laura Cercenelli,
University of Bologna, Italy

REVIEWED BY
Shujun Wang,
Hong Kong Polytechnic University, Hong Kong SAR, China
Jiong Zhang,
University of Southern California, Los Angeles, United States

*CORRESPONDENCE
Zhentao Zhu,
jshayyzzt@163.com
Weihua Yang,
benben0606@139.com
Qiang Wang,
wangqiang1981@wmu.edu.cn

RECEIVED 28 October 2023
ACCEPTED 12 February 2024
PUBLISHED 22 February 2024

CITATION
Chen N, Zhu Z, Yang W and Wang Q (2024), Progress in clinical research and applications of retinal vessel quantification technology based on fundus imaging.
Front. Bioeng. Biotechnol. 12:1329263.
doi: 10.3389/fbioe.2024.1329263

COPYRIGHT
© 2024 Chen, Zhu, Yang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Progress in clinical research and applications of retinal vessel quantification technology based on fundus imaging

Naimei Chen[1], Zhentao Zhu[1]*, Weihua Yang[2]* and Qiang Wang[3]*

[1]Department of Ophthalmology, Huaian Hospital of Huaian City, Huaian, China, [2]Department of Ophthalmology, Shenzhen Eye Hospital, Jinan University, Shenzhen, China, [3]Department of Ophthalmology, Third Affiliated Hospital, Wenzhou Medical University, Ruian, China

Retinal blood vessels are the only directly observed blood vessels in the body; changes in them can help effective assess the occurrence and development of ocular and systemic diseases. The specificity and efficiency of retinal vessel quantification technology has improved with the advancement of retinal imaging technologies and artificial intelligence (AI) algorithms; it has garnered attention in clinical research and applications for the diagnosis and treatment of common eye and related systemic diseases. A few articles have reviewed this topic; however, a summary of recent research progress in the field is still needed. This article aimed to provide a comprehensive review of the research and applications of retinal vessel quantification technology in ocular and systemic diseases, which could update clinicians and researchers on the recent progress in this field.

KEYWORDS

retinal vasculature, artificial intelligence, ocular diseases, systemic diseases, retinal vessel quantification

1 Introduction

Recent advancements in ophthalmic imaging technology have led to challenges in monitoring disease progression during early and advanced stages, largely due to its limited use for quantitative analysis of diseases; it has been mainly used for qualitative analysis of diseases. The rise in computational power, imaging data expansion, and ethical system improvement has led to a rapid development of AI in medicine (Ting et al., 2021). The retinal vessel quantification technology has been advanced, providing new opportunities for quantitative analysis of diseases. The retinal blood vessels are the only blood vessels that can be directly observed without using an invasive approach; their structures and functions are similar to those of the systemic vascular system. Quantitative research on retinal blood vessels could provide insights into the conditions of cardiovascular, cerebrovascular, and systemic blood vessels because they are components of the body's circulatory system (Dumitrascu and Koronyo-Hamaoui, 2020). The AI-based retinal vessel quantification technology can help clinicians and researchers study ocular and systemic diseases, benefiting disease prevention, diagnosis, and treatment (Gadde et al., 2016; Fuchs et al., 2022; Fu et al., 2023). A few articles have reviewed this technology; however, summarizing its recent progress is still needed. This article aimed to provide a comprehensive review of the research and applications of retinal vessel quantification technology in ocular and systemic diseases. As shown in Table 1.

TABLE 1 Summary of key points of the article.

Retinal vessel quantification has become a hot area of research for common eye and related systemic diseases due to its high specificity and sensitivity
Clinical research and applications of retinal vessel quantification in eye diseases
Clinical research and applications of retinal vessel quantification in systemic diseases
Limitations and solutions in clinical research and applications of retinal vessel quantification
Researching and applying retinal vessel quantification to ocular and systemic diseases is promising

2 Retinal vessel quantification technology

Retinal vessel quantification is a method to measure and analyze retinal blood vessels in the fundus through imaging technology, and extensive research has been conducted in clinical Retinal vessel quantification based on AI using optical coherence tomography angiography (OCTA). OCTA utilizes specific algorithms to image only moving red blood cells in the vessels, providing three-dimensional (3D) visualization and quantitative analysis of the blood flow status across different vascular membrane layers of the retina and choroid (Lommatzsch, 2020; Koutsiaris et al., 2023). With the continual advancement in AI algorithms and ophthalmic imaging technology, several quantitative models for retinal blood vessels have been established, with high levels of sensitivity and specificity (Nguyen et al., 2013; Nunez do Rio et al., 2020; Comin et al., 2021; Liefers et al., 2021; Zhou et al., 2022; Nardini et al., 2023). Retinal vessel quantification is crucial in studying ocular and systemic diseases. Repeated quantitative analyses of vascular parameters such as vascular diameter, vascular fractal dimension (FD), vascular angle, vascular density (VD), retinal non-perfusion (RNP), and foveal avascular zone (FAZ) can assist in the comprehension of changes in disease occurrence, progression, and treatment (Pournaras and Riva, 2013; Lee et al., 2018; Ramos et al., 2019; Alibhai et al., 2020; Kadomoto et al., 2021).

Quantitative analysis of the normal retina allows for an in-depth examination of the retinal vascular structure, benefiting the diagnosis and analysis of vascular abnormalities. The density of the deep retinal vascular plexus is higher than that of the superficial retinal vascular plexus. The retina is divided into four sectors centered on the FAZ. The vascular density of the inferior region, whether deep or superficial, is higher than that of the other regions (temporal, superior, and nasal), and age has no significant effect on VD (Gadde et al., 2016). Macular vascular parameters are related to sex and age, and the FD and VD of blood vessels in males are significantly higher than those in females; in contrast, no significant differences exist in vessel curvature, and the entire macular region's FD, VD, and average vascular diameter exhibit negative correlations with age (Feng et al., 2021). Based on the measurement of the retinal vascular parameters of individuals aged 50 and over, the greater the FAZ area and non-circular index, the lower the average capillary density and superficial vascular density, the poorer visual acuity. These vascular parameters might be used as biomarkers for predicting the visual acuity of those aged 50 and over (Li et al., 2023).

Retinal vessel quantification technology has the potential to foster research on relevant ailments and provide novel methods for managing systemic and ocular health. Precise quantitative analysis of the retinal vasculature can provide insight into the pathophysiological mechanisms of ocular diseases and early warning and monitoring of systemic diseases as shown in Table 2.

3 Clinical research and applications in ocular disorders

3.1 Diabetic retinopathy

Diabetic retinopathy (DR) is a prevalent microvascular complication of diabetes and a frequent cause of blindness; the disease triggers vascular variations by damaging capillary cells like the vascular endothelial cells and pericytes, leading to local ischemia. DR progression leads to an increase in the number of microaneurysms (Wu et al., 2014; Jiang et al., 2020), vascular curvature (Lee et al., 2018), RNP area (Baxter et al., 2019; Alibhai et al., 2020; Kim et al., 2021), FAZ area (Mihailovic et al., 2019; Ratra et al., 2021; Meng et al., 2022), and non-circular index and a decrease in VD (Durbin et al., 2017). The RNP area shows a positive correlation with the emergence of neovascularization (Yu et al., 2020). These vascular parameters can function as reference indicators for the onset and progression of DR and predictive indicators for disease transformation. Certain parameters, including VD, vascular length, and FAZ area, could undergo changes prior to the onset of visual impairment in individuals with DR, and they offer substantial value in facilitating the detection of DR lesions at an early stage (Zhu et al., 2019). Among all parameters, there was a statistically significant difference in vascular curvature between patients with non-DR and mild non-proliferative diabetic retinopathy (NPDR), especially within the 1.5-mm area of the superficial layer of the retina (Lee et al., 2018). Quantitative indicators of retinal blood vessels can serve as potential biomarkers for DR staging (Xu et al., 2019; Chua et al., 2020; Boned-Murillo et al., 2021; Xu et al., 2021). Borrelli and colleagues attempted to utilize 3D analysis of OCTA to create a 3D image of the retinal vasculature and used a global threshold algorithm to procure two vascular parameters, 3D vascular volume, and 3D perfusion density, to assess the status of macular ischemia in patients with NPDR; the smaller the 3D vascular volume and 3D perfusion density, the more severe the macular ischemia (Borrelli et al., 2020). After thorough examinations, the reliability of 3D analysis for evaluating the state of retinal blood vessels has been confirmed, indicating broad potential applications (Borrelli et al., 2020). Retinal vessel quantification assesses the treatment prognosis of patients with DR, and it quantifies the alterations in retinal neovascularization pre-

TABLE 2 The importance of quantifying retinal vessels in ocular and systemic diseases.

Ocular diseases	Systemic diseases
DR: Early diagnosis of disease, staging of disease severity, prediction of disease transformation, and assessment of treatment efficacy	Cerebrovascular diseases: Monitoring the risk of disease occurrence, alternative invasive or expensive examinations, and early diagnosis of diseases
RVO: Prediction of the risk of neovascularization and complications, selection of timing for PRP intervention, staging of disease severity, and evaluation of treatment effectiveness	Hypertension: Monitoring the risk of disease occurrence and predicting the risk of complications
Glaucoma: Detection of optic nerve injury, evaluation of visual function prognosis, and detection indicators of disease progression	CAD: Monitoring the risk of disease occurrence
EAMD: Detection indicators for disease progression	SCD: Monitoring the risk of disease occurrence
Other eye diseases: Selection of surgical methods and monitoring of disease progression	Fabry disease: Monitoring the risk of disease occurrence and predicting the risk of complications

and post-treatment with anti-vascular endothelial growth factor (anti-VEGF) therapy to evaluate the susceptibility of patients with DR to VEGF (Hu et al., 2019); helping informed decisions for follow-on treatment. Pan-retinal photocoagulation (PRP) can effectively reverse retinal DR-caused ischemia while maintaining the integrity of macular microvascular structure. The treatment effect can be effectively evaluated by quantifying the retinal VD and FAZ areas after PRP in patients with DR (Abdelhalim et al., 2022; Sariyildiz et al., 2023).

3.2 Retinal vein occlusion

Retinal vein occlusion (RVO) is the second most common retinal vascular disease that can cause blindness following DR. It happens when one or more veins in the retina are blocked or obstructed. The retina, located at the posterior of the eye, detects light and transmits signals to the brain for visual perception. Venous obstruction interferes with normal blood flow on the retina, potentially resulting in vision impairments. The impact of RVO on the deep capillary plexus (DCP) of the retina is 1.77–1.84 times that of the shallow capillary plexus (SCP) (Kim et al., 2020). The larger the area of RNP, the greater the risk of neovascularization. Quantifying the area of RNP can effectively assess the risk of neovascularization (Kadomoto et al., 2021). Neovascularization and associated neovascular glaucoma are common complications of RVO that can cause serious damage to a patient's vision and the eyeball itself (Hayreh, 2021). PRP can effectively prevent and treat neovascularization, which could benefit from quantifying the area of the RNP to select the timing of PRP use. The severity of macular ischemia increases as the VD in the macular area reduces and the RNP area expands, accurately quantifying the VD and the RNP area can be used to grade macular ischemia in RVO patients, and accurately quantifying the VD and the RNP areas can be used to grade macular ischemia in patients with RVO and assess disease severity and prognosis (Ouederni et al., 2019; Tang et al., 2021; Yeung et al., 2021). Huang et al., 2022 classified patients with branch vein occlusion (BRVO) into reactive and refractory groups based on their responses to anti-VEGF treatment, which was determined by semi-automatic quantitative fluorescein angiography to measure the amount of fluorescein leakage around and near the fovea of the macular area in patients with BRVO. The refractory group demonstrated more severe leakage than the reactive group; thus, this technique could effectively predict the efficacy of anti-VEGF treatment for evaluating treatment feasibility (Huang et al., 2022).

3.3 Glaucoma

Glaucoma is an irreversible condition that can cause blindness. Primary angle closure glaucoma is the most common type of glaucoma in China, which occurs when the angle between the iris and cornea in the front chamber of the eye becomes narrow or even closes completely, preventing the flow of aqueous humor. If the anterior chamber angle is narrow, the normal outflow of aqueous humor is impeded, causing an elevation in intraocular pressure, resulting in optic nerve damage and eventual vision loss. Based on the quantification of the retinal blood vessels in patients suffering from primary angle closure, the microvessel density around the optic disc reduces, despite the absence of any changes in the thickness of the retinal nerve fiber layer and the ganglion cell complex, the microvascular density around the optic papilla is a sensitive indicator of changes in intraocular pressure and a predictive and monitoring parameter for the onset of glaucoma nerve changes (Miguel et al., 2021; Nascimento E Silva et al., 2021; Wang et al., 2021). These observations highlight the potential utility of adjusting target intraocular pressure based on changes in microvessel density. Van Melkebeke et al., 2018 showed that the microvascular density surrounding the optic disc is a prognostic indicator of visual function in patients with glaucoma. Kromer et al., 2019 used OCTA to measure macular blood flow density in patients with open-angle glaucoma and found that macular blood flow density was significantly decreased in glaucoma patients compared to the healthy population. A noteworthy correlation exists between the density of blood flow in the macula and the visual field (Yarmohammadi et al., 2017), presenting an opportunity to repeatedly evaluate glaucoma progression by measuring macular blood flow density.

3.4 Wet age-related macular degeneration

Wet age-related macular degeneration (wAMD) is a common cause of irreversible visual impairment, typically affecting the central vision of the eye. It is usually caused by abnormal vascular growth beneath the retina, resulting in macular region damage and fluid exudation; the incidence rate of wAMD has increased significantly in recent years (Stahl, 2020). Gao et al. quantitatively analyzed the RNP areas of the extrafoveal superficial vascular complex (SVC), intermediate capillary plexus (ICP), and deep capillary plexus (DCP) in patients with wAMD and healthy individuals, ruling out related interfering factors. They found that these patients had larger areas of RNP in SVC, ICP, and DCP (Gao et al., 2022). Hence, it is evident that the area of the RNP is closely related to the onset and progression of wAMD; measuring the area of the RNP can be used in clinical practice to predict and monitor the onset and progression of wAMD.

Quantitative analysis of retinal effusion secondary to retinal angiopathy facilitates evaluating disease progression and treatment responses in patients with retinal effusion like RVO, wAMD, and diabetes macular edema (Farinha et al., 2020; Schmidt-Erfurth et al., 2020; Fuchs et al., 2022; Michl et al., 2022; Muste et al., 2022; Reiter and Schmidt-Erfurth, 2022; Coulibaly et al., 2023). Wu et al., 2021 proposed a new optimized segmentation and quantification algorithm for neovascularization based on OCTA, which has higher accuracy and effectively monitors changes in neovascularization. It can be useful in follow-up monitoring during diagnosing and treating ischemic ophthalmopathy and systemic diseases.

3.5 Other ocular diseases

Iris VD is decreased shortly after refractive surgery, and the densities of superficial and deep retinal blood vessels do not recover within 3 months of surgery (Olcay et al., 2015). The small incision corneal stromal lenticule extraction (SMILE) is more significantly reduced than femtosecond laser *in situ* keratomileusis (FS-LASIK), which might be related to an abnormal elevation in intraocular pressure during the surgical process (Olcay et al., 2015). Thus, it is necessary to consider the impact of vascular changes and the selection of surgical approaches for patients requiring refractive surgery in clinical practice (Cui et al., 2022). Retinal vessel quantification can monitor and analyze vascular changes in the occurrence, development, and treatment of eye diseases, such as retinitis pigmentosa (Wang et al., 2019; Lu et al., 2022), type 2 macular telangiectasia (Chidambara et al., 2016; Pauleikhoff et al., 2019; Pauleikhoff et al., 2022), familial retinal arteriolar tortuosity (Saraf et al., 2019), Behcet's disease (Türkcü et al., 2020), optic disc drusen (Leal-González et al., 2020), and retinopathy of prematurity (Cabrera et al., 2021).

4 Clinical research and applications in systemic diseases

4.1 Cerebrovascular diseases

Retinal and cerebral blood vessels come from the internal carotid artery and interconnect and influence each other. Thus, abnormalities in retinal blood vessels are often accompanied by abnormalities in cerebral blood vessels, and abnormal changes in retinal blood vessels are associated with stroke, vascular cognitive impairment, and dementia (Frost et al., 2017; Cabrera DeBuc et al., 2018; Cabrera DeBuc et al., 2020; Dumitrascu and Koronyo-Hamaoui, 2020). Widespread retinal arteriolar stenosis is linked to a high risk of disabling dementia (Jinnouchi et al., 2017), which could serve as an effective biomarker for populations with disabling dementia and one of the traditional screening indicators. The APOE ε4 allele is among the genetic factors most closely linked with Alzheimer's disease, and individuals carrying the APOE ε4 allele have a greater risk of developing Alzheimer's disease. Carriers of the APOE ε4 allele have significantly higher vascular width ratios than normal individuals, and measuring the ratio of retinal vessel width is an alternative approach to the invasive examination of APOE ε4 (Frost et al., 2017). Brain white matter volume reduction and enlargement of the inferior lateral ventricle adversely affect brain function, potentially leading to cognitive impairments, motor dysfunction, epilepsy, and visual and speech issues, depending on the extent, location, and cause of the reduction. Brain white matter volume is typically assessed and evaluated using MRI, and a larger diameter of retinal venules is associated with a smaller brain white matter volume (Ikram et al., 2013). Decreased densities of retinal SCP and perfusion are associated with the enlargement of the inferior lateral ventricle (Yoon et al., 2019). As an alternative to MRI, assessing and monitoring changes in brain white matter volume and the inferior lateral ventricle can be achieved by measuring retinal venule diameter, retinal SCP density, and perfusion density. Early abnormalities in brain microvasculature, that cannot be detected by head MRI, are closely associated with changes in retinal vasculature. Therefore, quantifying retinal vascular changes can be used to assess brain vascular function, achieving early diagnosis and intervention (London et al., 2013; Wardlaw et al., 2013).

4.2 Hypertension

Hypertension is a common cardiovascular disease. As blood pressure becomes unstable or consistently rises, the risk of developing diseases, such as heart, retinal, stroke, and kidney diseases, also increases. Hypertension typically results in an elevated ratio of small artery length to diameter and decreased terminal branch arteries. Alterations in the retina often signal potential disease risks in other target organs; thus, quantifying retinal vascular parameters could help evaluate the risk of hypertension complications (Hughes et al., 2006; Leclaire et al., 2021). High blood pressure elevates the pressure on blood vessels, leading to weakened and hardened elasticity of retinal arterioles, compressing the veins, and resulting in decreased blood flow and gradual thinning of the veins, known as Gunn's sign (Wigdahl et al., 2015). Furthermore, the pressure of arterioles at the intersection can be transmitted to the veins, resulting in the characteristic S-shaped appearance; the Salus sign is a term used to describe a phenomenon at the intersection of retinal arteries and veins, and quantifying the Gunn and Salus signs could predict hypertensive retinopathy and RVO development (Wigdahl et al., 2015). Bringing together deep learning and OCTA could more precisely quantify the retinal vascular structure, it is feasible to precisely forecast the onset and development risk of hypertension and its complications based on the changes in retinal vascular structure. (Tan et al., 2022).

4.3 Coronary artery disease

Coronary artery disease (CAD) is a cardiovascular disease occurring when the coronary artery (one of the main blood vessels supplying the heart) narrows or becomes blocked. CAD can trigger angina (chest pain) and myocardial infarction (severe damage to heart muscles), endangering the patient's life. The high incidence and mortality rates of CAD imply early detection, and intervention are imperative. Fu et al. used deep learning technology to quantify retinal vascular parameters from color fundus photography of 57,947 participants without CAD; after approximately 11 years of follow-up, the FD of blood vessels had decreased, and the reduction in the number of arterial and small venous segments and arterial and venous bone densities was closely associated with an increased risk of subsequent CAD (Fu et al., 2023). Hence, detecting retinal vascular parameters could help predict CAD (Shokr et al., 2021).

4.4 Sickle cell disease

Sickle cell disease (SCD), also called sickle cell anemia, is a prevalent hereditary blood disease characterized by the deformation of red blood cells into sickle or curved shapes, different from the round shape of normal red blood cells. Such abnormal red blood cells are susceptible to adhesion and blockage within blood vessels, leading to ischemic damage. When OCTA was used to determine retinal capillary perfusion, dynamic changes were observed in retinal capillary perfusion in patients with SCD compared with healthy individuals (Zhou et al., 2021). Retinal ischemia and hypoxia can result in various complications. Non-invasive dynamic monitoring of retinal capillary perfusion in patients with SCD allows for effective evaluation of the state of systemic blood vessels, benefiting the early detection and treatment of the disease and assessing treatment efficacy.

4.5 Other systemic diseases

Fabry disease is a rare genetic disorder where glycolipids accumulate in multiple tissues and cells due to deficiency or reduced activity of the enzyme cleavage lipase, resulting in damage to multiple organs and systems. Quantifying retinal blood vessels demonstrate a negative correlation between retinal VD and myocardial damage associated with fabry disease (Cennamo et al., 2020). Primary nephrotic syndrome (PNS) is typically linked to dysfunction in glomerular filtration. The VD and blood flow perfusion density in the macular areas of patients with PNS significantly decreased compared with healthy individuals and negatively correlated with urinary protein levels (Yao et al., 2022). Thus, retinal blood vessels can be quantified to monitor and analyze the onset and progression of associated systemic illnesses and their complications and to assess disease alterations throughout treatment.

5 Limitations and solutions

Limitations exist with the retinal vessel quantification technology. This technology has limited generalization for large clinical applications, and trust and acceptance of algorithm outcomes vary among physicians and patients (Ting et al., 2019). Retinal vessel quantification techniques have some limitations, including inconsistent data quality, high equipment costs, inconsistent algorithms, and insufficient baseline data. To address these issues, we propose the following improvements: Firstly, we suggest improving reproducibility and reducing costs by enhancing imaging techniques and standardizing data acquisition and image processing processes. Secondly, we recommend developing international standards and guidelines to ensure algorithmic consistency and comparability. Thirdly, we propose establishing better baseline data through big data analytics to support clinical applications. Lastly, we suggest adopting data encryption and privacy protection measures to ensure the security of patient data. To enhance the acceptance of AI algorithms among physicians and patients, it is important to further develop explainable AI algorithms., Limitations and solutions of retinal vessel quantification are shown in Table 3. Fostering inter-team cooperation, improving technology, and increasing collaboration with clinicians and patients would help overcome these limitations, increasing its clinical applications and benefiting disease diagnosis and treatment.

6 Application trends and prospects

The development of modern technologies, dataset expansion, improvement in doctor-patient acceptance, data privacy and ethics, and increased financial support will render retinal vessel quantification to have broader application prospects in disease research. By observing changes in the retinal blood vessels of diverse populations, researchers evaluate the effectiveness of disease treatment protocols and study pathophysiologic changes in disease to more accurately diagnose and treat disease and develop drugs (Al-Shabrawey, 2023; Middel et al., 2023; Yucel Gencoglu et al., 2023). Prospects and trends of retinal vessel quantification in clinical disease are shown in Table 4.

Retinal vessel quantification will continue to play an important role in clinical research. It would improve the effectiveness of disease diagnosis, treatment, and prevention and promote the further development of medical science by combining advanced technologies and data analysis methods.

7 Conclusion

The quantification of retinal vessels has important clinical research and application values, mainly reflected in the following aspects. First, it can be used for early diagnosis of diseases. Doctors can detect early signs of DR, glaucoma, CAD, and other diseases by quantifying retinal vessels, allowing for early intervention and treatment to reduce permanent damage. Second, it can be used to monitor disease progression. Retinal vessel quantification can be used to monitor disease progression, determine the effectiveness of treatment, and adjust treatment plans in a timely manner in patients with eye and related systemic diseases. Third, it can be used for personalized treatment. By understanding each patient's retinal vascular status, physicians can develop more personalized treatment plans to improve treatment efficacy and reduce side effects. Fourth, it can assist in replacing relevant clinical examinations. Retinal vessel quantification can effectively assist in

TABLE 3 Limitations and solutions of retinal vessel quantification.

Limitations	Solutions
Data quality is inconsistent	Improving imaging technology, standardizing data acquisition and image processing, and improving repeatability
The equipment used for quantification is complex and expensive, which is not conducive to dissemination	Reducing the cost of equipment and promoting technology use in remote areas
The inconsistency of retinal vessel quantification algorithms leads to variable results	Developing international standards and guidelines to ensure consistency and comparability of retinal vessel quantification algorithms
The lack of sufficient baseline data for comparison with individual retinal vascular diseases may limit its clinical applications	By collecting large-scale retinal images and clinical data, baseline data can be better established; this will help improve disease diagnosis and treatment
Patient privacy and safety issues	Taking data encryption and privacy measures to ensure the security and legality of patient data
Doctors and patients have a low acceptance of AI "black box" algorithms	Further development of interpretable AI algorithms

TABLE 4 Prospects and trends of retinal vessel quantification in clinical disease.

Personalized medicine: As technology advances, the quantification of retinal blood vessels will become a part of personalized medicine. Doctors can develop personalized treatment plans based on the specific retinal characteristics of each patient to improve treatment efficacy
Predictive medicine: AI can help doctors predict eye and systemic health risks by analyzing large-scale retinal image data and cases. This predictive medicine helps with early intervention and disease prevention
New drug development: Retinal vessel quantification can be used to evaluate the therapeutic effects of new drugs on diseases, speeding up the development of new drugs and providing more treatment options for patients
Identification of disease subtypes: Vascular quantification can help identify subtypes of different diseases, leading to more accurate diagnosis and treatment
Remote monitoring: The remote acquisition and analysis of retinal images allow doctors to monitor patients regularly without frequent clinic visits. This is particularly beneficial for managing long-term chronic conditions
Multi-field applications in ocular and systemic diseases: Retinal vessel quantification has applications in ophthalmology and plays a role in various medical fields, such as systemic health management and cardiovascular disease research

replacing some traumatic and costly clinical examinations with repeatability, helping clinicians evaluate disease occurrence, development, and treatment effectiveness. Fifth, it can be used for studying disease mechanisms and pathogenesis, providing a foundation for new treatment methods and drug development.

Overall, retinal vessel quantification has broad research application prospects in the diagnosis, treatment, and research of ocular and systemic diseases, helping prevent disease onset and development and better protect and maintain the physical health of patients (Keskinbora and Güven, 2020).

Author contributions

NC: Writing–original draft. ZZ: Writing–review and editing. WY: Writing–review and editing. QW: Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the Public Benifit Technology Research Program of Zhejiang Science and Technology Department (GC21H120001), Shenzhen Fund for Guangdong Provincial High-level Clinical Key Specialties (SZGSP014) and Sanming Project of Medicine in Shenzhen (SZSM202311012).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abdelhalim, A. S., Abdelkader, M. F. S. O., Mahmoud, M. S. E., and Mohamed Mohamed, A. A. (2022). Macular vessel density before and after panretinal photocoagulation in patients with proliferative diabetic retinopathy. *Int. J. Retina Vitr.* 8, 21. doi:10.1186/s40942-022-00369-1

Alibhai, A. Y., De Pretto, L. R., Moult, E. M., Or, C., Arya, M., McGowan, M., et al. (2020). Quantification of retinal capillary nonperfusion in diabetics using wide-field optical coherence tomography angiography. *Retina* 40, 412–420. doi:10.1097/IAE.0000000000002403

Al-Shabrawey, M. (2023). Methods of studying retinal vessels in health and diseases. *J. Vis. Exp.* 192. doi:10.3791/65008

Baxter, S. L., Ashir, A., Nguyen, B. J., and Nudleman, E. (2019). Quantification of retinal nonperfusion associated with posterior segment neovascularization in diabetic retinopathy using ultra-widefield fluorescein angiography. *Ophthalmic Surg. Lasers Imaging Retina* 50, 86–92. doi:10.3928/23258160-20190129-04

Boned-Murillo, A., Albertos-Arranz, H., Diaz-Barreda, M. D., Orduna-Hospital, E., Sánchez-Cano, A., Ferreras, A., et al. (2021). Optical coherence tomography angiography in diabetic patients: a systematic review. *Biomedicines* 10, 88. doi:10.3390/biomedicines10010088

Borrelli, E., Sacconi, R., Querques, L., Battista, M., Bandello, F., and Querques, G. (2020). Quantification of diabetic macular ischemia using novel three-dimensional optical coherence tomography angiography metrics. *J. Biophot.* 13, e202000152. doi:10.1002/jbio.202000152

Cabrera, M. T., Chia, T., Wallace, D. K., Ulrich, J. N., Freedman, S. F., Ding, L., et al. (2021). Short-term computer-assisted quantification of plus disease after treatment of TYPE 1 retinopathy of prematurity with intravitreal bevacizumab or retinal laser photocoagulation. *Retin. Cases Brief. Rep.* 15, 314–319. doi:10.1097/ICB.0000000000000794

Cabrera DeBuc, D., Feuer, W. J., Persad, P. J., Somfai, G. M., Kostic, M., Oropesa, S., et al. (2020). Investigating vascular complexity and neurogenic alterations in sectoral regions of the retina in patients with cognitive impairment. *Front. Physiol.* 11, 570412. doi:10.3389/fphys.2020.570412

Cabrera DeBuc, D., Somfai, G. M., Arthur, E., Kostic, M., Oropesa, S., and Mendoza Santiesteban, C. (2018). Investigating multimodal diagnostic eye biomarkers of cognitive impairment by measuring vascular and neurogenic changes in the retina. *Front. Physiol.* 9, 1721. doi:10.3389/fphys.2018.01721

Cennamo, G., Montorio, D., Santoro, C., Cocozza, S., Spinelli, L., Di Risi, T., et al. (2020). The retinal vessel density as a new vascular biomarker in multisystem involvement in Fabry disease: an optical coherence tomography angiography study. *J. Clin. Med.* 9, 4087. doi:10.3390/jcm9124087

Chidambara, L., Gadde, S. G., Yadav, N. K., Jayadev, C., Bhanushali, D., Appaji, A. M., et al. (2016). Characteristics and quantification of vascular changes in macular telangiectasia type 2 on optical coherence tomography angiography. *Br. J. Ophthalmol.* 100, 1482–1488. doi:10.1136/bjophthalmol-2015-307941

Chua, J., Sim, R., Tan, B., Wong, D., Yao, X., Liu, X., et al. (2020). Optical coherence tomography angiography in diabetes and diabetic retinopathy. *J. Clin. Med.* 9, 1723. doi:10.3390/jcm9061723

Comin, C. H., Tsirukis, D. I., Sun, Y., and Xu, X. (2021). Quantification of retinal blood leakage in fundus fluorescein angiography in a retinal angiogenesis model. *Sci. Rep.* 11, 19903. doi:10.1038/s41598-021-99434-2

Coulibaly, L. M., Sacu, S., Fuchs, P., Bogunovic, H., Faustmann, G., Unterrainer, C., et al. (2023). Personalized treatment supported by automated quantitative fluid analysis in active neovascular age-related macular degeneration (nAMD)-a phase III, prospective, multicentre, randomized study: design and methods. *Eye (Lond.).* 37, 1464–1469. doi:10.1038/s41433-022-02154-8

Cui, L., Xue, W., Yao, W., Huang, X., Xue, W., Wang, Y., et al. (2022). Quantitative changes in iris and retinal blood flow after femtosecond laser-assisted *in situ* keratomileusis and small-incision lenticule extraction. *Front. Med. (Lausanne).* 9, 862195. doi:10.3389/fmed.2022.862195

Dumitrascu, O. M., and Koronyo-Hamaoui, M. (2020). Retinal vessel changes in cerebrovascular disease. *Curr. Opin. Neurol.* 33, 87–92. doi:10.1097/WCO.0000000000000779

Durbin, M. K., An, L., Shemonski, N. D., Soares, M., Santos, T., Lopes, M., et al. (2017). Quantification of retinal microvascular density in optical coherence tomographic angiography images in diabetic retinopathy. *JAMA Ophthalmol.* 135, 370–376. doi:10.1001/jamaophthalmol.2017.0080

Farinha, C., Santos, T., Santos, A. R., Lopes, M., Alves, D., Silva, R., et al. (2020). OPTICAL COHERENCE TOMOGRAPHY LEAKAGE IN neovascular AGE-RELATED MACULAR DEGENERATION: identification of choroidal neovascularization activity by location and quantification of abnormal fluid under anti-vascular endothelial growth factor therapy. *Retina* 40, 881–890. doi:10.1097/IAE.0000000000002470

Feng, Z., Wang, G., Xia, H., Li, M., Liang, G., Dong, T., et al. (2021). Macular vascular geometry changes with sex and age in healthy subjects: a fundus photography study. *Front. Med. (Lausanne).* 8, 778346. doi:10.3389/fmed.2021.778346

Frost, S., Bhuiyan, A., Offerman, D., Doecke, J. D., Macaulay, S. L., Sohrabi, H. R., et al. (2017). Modulation of retinal arteriolar central reflection by APOE genotype. *Curr. Alzheimer Res.* 14, 916–923. doi:10.2174/1567205014666170309115016

Fu, Y., Yusufu, M., Wang, Y., He, M., Shi, D., and Wang, R. (2023). Association of retinal microvascular density and complexity with incident coronary heart disease. *Atherosclerosis* 380, 117196. doi:10.1016/j.atherosclerosis.2023.117196

Fuchs, P., Coulibaly, L., Reiter, G. S., and Schmidt-Erfurth, U. (2022). Artificial intelligence in the management of anti-VEGF treatment: the Vienna fluid monitor in clinical practice. *Ophthalmol. Ophthalmol.* 119, 520–524. doi:10.1007/s00347-022-01618-2

Gadde, S. G., Anegondi, N., Bhanushali, D., Chidambara, L., Yadav, N. K., Khurana, A., et al. (2016). Quantification of vessel density in retinal optical coherence tomography angiography images using local fractal dimension. *Invest. Ophthalmol. Vis. Sci.* 57, 246–252. doi:10.1167/iovs.15-18287

Gao, L., Wang, J., You, Q., Guo, Y., Flaxel, C. J., Hwang, T. S., et al. (2022). Plexus-specific retinal capillary avascular area in exudative age-related macular degeneration with projection-resolved OCT angiography. *Br. J. Ophthalmol.* 106, 719–723. doi:10.1136/bjophthalmol-2020-317562

Hayreh, S. S. (2021). Photocoagulation for retinal vein occlusion. *Prog. Retin. Eye Res.* 85, 100964. doi:10.1016/j.preteyeres.2021.100964

Hu, Z., Su, Y., Xie, P., Chen, L., Ji, J., Feng, T., et al. (2019). OCT angiography-based monitoring of neovascular regression on fibrovascular membrane after preoperative intravitreal conbercept injection. *Graefes Arch. Clin. Exp. Ophthalmol.* 257, 1611–1619. doi:10.1007/s00417-019-04315-0

Huang, P. W., Lai, C. C., Hwang, Y. S., Wu, W. C., Wu, C. H., Huang, J. C., et al. (2022). Treatment responses for branch retinal vein occlusion predicted by semiautomated fluorescein angiography quantification. *BMC Ophthalmol.* 22, 50. doi:10.1186/s12886-022-02245-w

Hughes, A. D., Martinez-Perez, E., Jabbar, A. S., Hassan, A., Witt, N. W., Mistry, P. D., et al. (2006). Quantification of topological changes in retinal vascular architecture in essential and malignant hypertension. *J. Hypertens.* 24, 889–894. doi:10.1097/01.hjh.0000222759.61735.98

Ikram, M. K., de Jong, F. J., Vernooij, M. W., Hofman, A., Niessen, W. J., van der Lugt, A., et al. (2013). Retinal vascular calibers associate differentially with cerebral gray matter and white matter atrophy. *Alzheimer Dis. Assoc. Disord.* 27, 351–355. doi:10.1097/WAD.0b013e31829344ed

Jiang, A., Srivastava, S., Figueiredo, N., Babiuch, A., Hu, M., Reese, J., et al. (2020). Repeatability of automated leakage quantification and microaneurysm identification utilising an analysis platform for ultra-widefield fluorescein angiography. *Br. J. Ophthalmol.* 104, 500–503. doi:10.1136/bjophthalmol-2019-314416

Jinnouchi, H., Kitamura, A., Yamagishi, K., Kiyama, M., Imano, H., Okada, T., et al. (2017). Retinal vascular changes and prospective risk of disabling dementia: the circulatory risk in communities study (CIRCS). *J. Atheroscler. Thromb.* 24, 687–695. doi:10.5551/jat.37291

Kadomoto, S., Muraoka, Y., Uji, A., Tamiya, R., Oritani, Y., Kawai, K., et al. (2021). Nonperfusion AREA QUANTIFICATION IN BRANCH RETINAL VEIN OCCLUSION: a widefield optical coherence tomography angiography study. *Retina* 41, 1210–1218. doi:10.1097/IAE.0000000000002999

Keskinbora, K., and Güven, F. (2020). Artificial intelligence and ophthalmology. *Turk. J. Ophthalmol.* 50, 37–43. doi:10.4274/tjo.galenos.2020.78989

Kim, J. T., Chun, Y. S., Lee, J. K., Moon, N. J., and Yi, D. Y. (2020). Comparison of vessel density reduction in the deep and superficial capillary plexuses in branch retinal vein occlusion. *Ophthalmologica* 243, 66–74. doi:10.1159/000502385

Kim, K., You, J., Park, J. R., Kim, E. S., Oh, W. Y., and Yu, S. Y. (2021). Quantification of retinal microvascular parameters by severity of diabetic retinopathy using wide-field swept-source optical coherence tomography angiography. *Graefes Arch. Clin. Exp. Ophthalmol.* 259, 2103–2111. doi:10.1007/s00417-021-05099-y

Koutsiaris, A. G., Batis, V., Liakopoulou, G., Tachmitzi, S. V., Detorakis, E. T., and Tsironi, E. E. (2023). Optical coherence tomography angiography (OCTA) of the eye: a review on basic principles, advantages, disadvantages and device specifications. *Clin. Hemorheol. Microcirc.* 83, 247–271. doi:10.3233/CH-221634

Kromer, R., Glusa, P., Framme, C., Pielen, A., and Junker, B. (2019). Optical coherence tomography angiography analysis of macular flow density in glaucoma. *Acta Ophthalmol.* 97, e199–e206. doi:10.1111/aos.13914

Leal-González, M., Pessanha, F., Azevedo González-Oliva, M., Pérez-Fernández, E., and Gili, P. (2020). Study of peripapillary vascular flow using optical coherence tomography angiography in optic nerve head drusen. *Clin. Exp. Ophthalmol.* 48, 775–782. doi:10.1111/ceo.13783

Leclaire, M. D., Eter, N., and Alnawaiseh, M. (2021). Optical coherence tomography angiography and cardiovascular diseases. An overview of the current knowledge. *Ophthalmol. Ophthalmol.* 118, 1119–1127. doi:10.1007/s00347-021-01336-1

Lee, H., Lee, M., Chung, H., and Kim, H. C. (2018). Quantification of retinal vessel tortuosity in diabetic retinopathy using optical coherence tomography angiography. *Retina* 38, 976–985. doi:10.1097/IAE.0000000000001618

Li, Y. K., Fung, N. S., Chan, J. C. H., Choy, B. N. K., Chow, L. L. W., Shih, K. C., et al. (2023). Octa biomarkers in adults aged 50 and above: a prospective and cross-sectional community-based study. *BMC Ophthalmol.* 23, 71. doi:10.1186/s12886-023-02815-6

Liefers, B., Taylor, P., Alsaedi, A., Bailey, C., Balaskas, K., Dhingra, N., et al. (2021). Quantification of key retinal features in early and late age-related macular degeneration using deep learning. *Am. J. Ophthalmol.* 226, 1–12. doi:10.1016/j.ajo.2020.12.034

Lommatzsch, A. (2020). OCT angiography. *Monbl Augenheilkd* 237, 95–111. doi:10.1055/a-1023-9678

London, A., Benhar, I., and Schwartz, M. (2013). The retina as a window to the brain-from eye research to CNS disorders. *Nat. Rev. Neurol.* 9, 44–53. doi:10.1038/nrneurol.2012.227

Lu, B., Chao, G., and Xie, L. (2022). Optical coherence tomography angiography in retinitis pigmentosa: a narrative review. *Med. (Baltim.).* 101, e30068. doi:10.1097/MD.0000000000030068

Meng, Y., Lan, H., Hu, Y., Chen, Z., Ouyang, P., and Luo, J. (2022). Application of improved U-net convolutional neural network for automatic quantification of the foveal avascular zone in diabetic macular ischemia. *J. Diabetes Res.* 2022, 1–8. doi:10.1155/2022/4612554

Michl, M., Fabianska, M., Seeböck, P., Sadeghipour, A., Haj Najeeb, B., Bogunovic, H., et al. (2022). Automated quantification of macular fluid in retinal diseases and their response to anti-VEGF therapy. *Br. J. Ophthalmol.* 106, 113–120. doi:10.1136/bjophthalmol-2020-317416

Middel, C. S., Dietrich, N., Hammes, H. P., and Kroll, J. (2023). Analysis of the morphology of retinal vascular cells in zebrafish (*Danio rerio*). *Front. Cell Dev. Biol.* 11, 1267232. doi:10.3389/fcell.2023.1267232

Miguel, A., Legeai, J., and Silva, B. (2021). A software for quantification of vessel density in glaucoma: an OCT-Angiography study. *J. Fr. Ophtalmol.* 44, 376–381. doi:10.1016/j.jfo.2020.06.038

Mihailovic, N., Eter, N., and Alnawaiseh, M. (2019). Foveal avascular zone and OCT angiography. An overview of current knowledge. *Ophthalmol. Ophthalmol.* 116, 610–616. doi:10.1007/s00347-018-0838-2

Muste, J. C., Iyer, A. I., Kalur, A., Talcott, K. E., and Singh, R. P. (2022). The quantification and impact of persistent retinal fluid compartments on best-corrected visual acuity of patients with retinal vein occlusion. *Ophthalmic Surg. Lasers Imaging Retina* 53, 139–147. doi:10.3928/23258160-20220215-03

Nardini, J. T., Pugh, C. W. J., and Byrne, H. M. (2023). Statistical and topological summaries aid disease detection for segmented retinal vascular images. *Microcirculation* 30, e12799. doi:10.1111/micc.12799

Nascimento E Silva, R., Chiou, C. A., Wang, M., Devlin, J., Li, D., Lovelace, S., et al. (2021). Quantification of the peripapillary microvasculature in eyes with glaucomatous paracentral visual field loss. *Ophthalmol. Glaucoma* 4, 286–294. doi:10.1016/j.ogla.2020.10.009

Nguyen, U. T., Bhuiyan, A., Park, L. A., Kawasaki, R., Wong, T. Y., Wang, J. J., et al. (2013). Automated quantification of retinal arteriovenous nicking from colour fundus images. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2013, 5865–5868. doi:10.1109/embc.2013.6610886

Nunez do Rio, J. M., Sen, P., Rasheed, R., Bagchi, A., Nicholson, L., Dubis, A. M., et al. (2020). Deep learning-based segmentation and quantification of retinal capillary non-perfusion on ultra-wide-field retinal fluorescein angiography. *J. Clin. Med.* 9, 2537. doi:10.3390/jcm9082537

Olcay, K., Cakir, A., Sagdic, S. K., Duzgun, E., and Yildirim, Y. (2015). Bilateral iris atrophy after the femtosecond assisted laser *in situ* keratomileusis surgery. *Case Rep. Ophthalmol. Med.* 2015, 1–3. doi:10.1155/2015/127806

Ouederni, M., Sassi, H., Nefaa, F., Kharroubi, A., Kellil, N., and Cheour, M. (2019). Anatomo-functional study in branch retinal vein occlusion using swept source optical coherence tomography angiography. *J. Fr. Ophtalmol.* 42, 255–261. doi:10.1016/j.jfo.2018.09.010

Pauleikhoff, D., Gunnemann, F., Book, M., and Rothaus, K. (2019). Progression of vascular changes in macular telangiectasia type 2: comparison between SD-OCT and OCT angiography. *Graefes Arch. Clin. Exp. Ophthalmol.* 257, 1381–1392. doi:10.1007/s00417-019-04323-0

Pauleikhoff, D., Pauleikhoff, L., and Chew, E. Y. (2022). Imaging endpoints for clinical trials in MacTel type 2. *Eye (Lond.).* 36, 284–293. doi:10.1038/s41433-021-01723-7

Pournaras, C. J., and Riva, C. E. (2013). Retinal blood flow evaluation. *Ophthalmologica* 229, 61–74. doi:10.1159/000338186

Ramos, L., Novo, J., Rouco, J., Romeo, S., Álvarez, M. D., and Ortega, M. (2019). Computational assessment of the retinal vascular tortuosity integrating domain-related information. *Sci. Rep.* 9, 19940. doi:10.1038/s41598-019-56507-7

Ratra, D., Dalan, D., Prakash, N., Kaviarasan, K., Thanikachalam, S., Das, U. N., et al. (2021). Quantitative analysis of retinal microvascular changes in prediabetic and diabetic patients. *Indian J. Ophthalmol.* 69, 3226–3234. doi:10.4103/ijo.IJO_1254_21

Reiter, G. S., and Schmidt-Erfurth, U. (2022). Quantitative assessment of retinal fluid in neovascular age-related macular degeneration under anti-VEGF therapy. *Ther. Adv. Ophthalmol.* 14, 251584142210833. doi:10.1177/25158414221083363

Saraf, S. S., Tyring, A. J., Chen, C. L., Le, T. P., Kalina, R. E., Wang, R. K., et al. (2019). Familial retinal arteriolar tortuosity and quantification of vascular tortuosity using swept-source optical coherence tomography angiography. *Am. J. Ophthalmol. Case Rep.* 14, 74–78. doi:10.1016/j.ajoc.2019.03.001

Sariyildiz, C., Çiloğlu, E., and Yetkin, E. (2023). Quantification of macular perfusion following panretinal photocoagulation for diabetic retinopathy: an optical coherence tomography angiography study. *Photodiagnosis Photodyn.Ther.* 41, 103233. doi:10.1016/j.pdpdt.2022.103233

Schmidt-Erfurth, U., Vogl, W. D., Jampol, L. M., and Bogunović, H. (2020). Application of automated quantification of fluid volumes to anti-VEGF therapy of neovascular age-related macular degeneration. *Ophthalmology* 127, 1211–1219. doi:10.1016/j.ophtha.2020.03.010

Shokr, H., Dias, I. H., and Gherghel, D. (2021). Oxysterols and retinal microvascular dysfunction as early risk markers for cardiovascular disease in normal, ageing individuals. *Ageing Individ. Antioxidants (Basel)* 10, 1756. doi:10.3390/antiox10111756

Stahl, A. (2020). The diagnosis and treatment of age-related macular degeneration. *Dtsch. Ärztebl. Int.* 117, 513–520. doi:10.3238/arztebl.2020.0513

Tan, W., Yao, X., Le, T. T., Tan, B., Schmetterer, L., and Chua, J. (2022). The New Era of retinal imaging in hypertensive patients. *Asia Pac. J. Ophthalmol. (Phila).* 11, 149–159. doi:10.1097/APO.0000000000000509

Tang, Z., Zhang, X., Yang, G., Zhang, G., Gong, Y., Zhao, K., et al. (2021). Automated segmentation of retinal nonperfusion area in fluorescein angiography in retinal vein occlusion using convolutional neural networks. *Med. Phys.* 48, 648–658. doi:10.1002/mp.14640

Ting, D. S. J., Foo, V. H., Yang, L. W. Y., Sia, J. T., Ang, M., Lin, H., et al. (2021). Artificial intelligence for anterior segment diseases: emerging applications in ophthalmology. *Br. J. Ophthalmol.* 105, 158–168. doi:10.1136/bjophthalmol-2019-315651

Ting, D. S. W., Pasquale, L. R., Peng, L., Campbell, J. P., Lee, A. Y., Raman, R., et al. (2019). Artificial intelligence and deep learning in ophthalmology. *Br. J. Ophthalmol.* 103, 167–175. doi:10.1136/bjophthalmol-2018-313173

Türkcü, F. M., Şahin, A., Karaalp, Ü., Çınar, Y., Şahin, M., Özkurt, Z. G., et al. (2020). Automated quantification of foveal avascular zone and vascular density in Behçet's disease. *Ir. J. Med. Sci.* 189, 349–354. doi:10.1007/s11845-019-02051-2

Van Melkebeke, L., Barbosa-Breda, J., Huygens, M., and Stalmans, I. (2018). Optical coherence tomography angiography in glaucoma: a review. *Ophthalmic Res.* 60, 139–151. doi:10.1159/000488495

Wang, X., Chen, J., Kong, X., and Sun, X. (2021). Quantification of retinal microvascular density using optic coherence tomography angiography in primary angle closure disease. *Curr. Eye Res.* 46, 1018–1024. doi:10.1080/02713683.2020.1849728

Wang, X. N., Zhao, Q., Li, D. J., Wang, Z. Y., Chen, W., Li, Y. F., et al. (2019). Quantitative evaluation of primary retinitis pigmentosa patients using colour Doppler flow imaging and optical coherence tomography angiography. *Acta Ophthalmol.* 97, e993–e997. doi:10.1111/aos.14047

Wardlaw, J. M., Smith, E. E., Biessels, G. J., Cordonnier, C., Fazekas, F., Frayne, R., et al. (2013). Neuroimaging standards for research into small vessel disease and its contribution to ageing and neurodegeneration. *Lancet Neurol.* 12, 822–838. doi:10.1016/S1474-4422(13)70124-8

Wigdahl, J., Guimarães, P., Leontidis, G., Triantafyllou, A., and Ruggeri, A. (2015). Automatic Gunn and Salus sign quantification in retinal images. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2015, 5251–5254. doi:10.1109/embc.2015.7319576

Wu, H., Zhang, X., Geng, X., Dong, J., and Zhou, G. (2014). Computer aided quantification for retinal lesions in patients with moderate and severe non-proliferative diabetic retinopathy: a retrospective cohort study. *BMC Ophthalmol.* 14, 126. doi:10.1186/1471-2415-14-126

Wu, S., Wu, S., Feng, H., Hu, Z., Xie, Y., Su, Y., et al. (2021). An optimized segmentation and quantification approach in microvascular imaging for octa-based neovascular regression monitoring. *BMC Med. Imaging.* 21, 13. doi:10.1186/s12880-021-00546-y

Xu, B., Chen, J., Zhang, S., Shen, S., Lan, X., Chen, Z., et al. (2021). Association between the severity of diabetic retinopathy and optical coherence tomography angiography metrics. *Front. Endocrinol. (Lausanne).* 12, 777552. doi:10.3389/fendo.2021.777552

Xu, X., Chen, C., Ding, W., Yang, P., Lu, H., Xu, F., et al. (2019). Automated quantification of superficial retinal capillaries and large vessels for diabetic retinopathy on optical coherence tomographic angiography. *J. Biophot.* 12, e201900103. doi:10.1002/jbio.201900103

Yao, T., He, Y., Huang, L., Chen, J., Zhang, Z., Yang, W., et al. (2022). Quantitative vessel density analysis of macular and peripapillary areas by optical coherence tomography angiography in adults with primary nephrotic syndrome. *Microvasc. Res.* 144, 104407. doi:10.1016/j.mvr.2022.104407

Yarmohammadi, A., Zangwill, L. M., Diniz-Filho, A., Saunders, L. J., Suh, M. H., Wu, Z., et al. (2017). Peripapillary and macular vessel density in patients with glaucoma and single-hemifield visual field defect. *Ophthalmology* 124, 709–719. doi:10.1016/j.ophtha.2017.01.004

Yeung, L., Lee, Y. C., Lin, Y. T., Lee, T. W., and Lai, C. C. (2021). Macular ischemia quantification using deep-learning denoised optical coherence tomography angiography in branch retinal vein occlusion. *Transl. Vis. Sci. Technol.* 10, 23. doi:10.1167/tvst.10.7.23

Yoon, S. P., Thompson, A. C., Polascik, B. W., Calixte, C., Burke, J. R., Petrella, J. R., et al. (2019). Correlation of octa and volumetric MRI in mild cognitive impairment and Alzheimer's disease. *Ophthalmic Surg. Lasers Imaging Retina* 50, 709–718. doi:10.3928/23258160-20191031-06

Yu, G., Aaberg, M. T., Patel, T. P., Iyengar, R. S., Powell, C., Tran, A., et al. (2020). Quantification of retinal nonperfusion and neovascularization with Ultrawidefield fluorescein angiography in patients with diabetes and associated characteristics of advanced disease. *JAMA Ophthalmol.* 138, 680–688. doi:10.1001/jamaophthalmol.2020.1257

Yucel Gencoglu, A., Ağın, A., Un, Y., and Ozturk, Y. (2023). Quantification of retinal vein and artery trajectories using second-order polynomial equation in eyes with vitreomacular traction. *Photodiagnosis Photodyn. Ther.* 42, 103616. doi:10.1016/j.pdpdt.2023.103616

Zhou, D. B., Castanos, M. V., Pinhas, A., Gillette, P., Migacz, J. V., Rosen, R. B., et al. (2021). Quantification of intermittent retinal capillary perfusion in sickle cell disease. *Biomed. Opt. Express.* 12, 2825–2840. doi:10.1364/BOE.418874

Zhou, Y., Wagner, S. K., Chia, M. A., Zhao, A., Woodward-Court, P., Xu, M., et al. (2022). AutoMorph: automated retinal vascular morphology quantification via a deep learning pipeline. *Transl. Vis. Sci. Technol.* 11, 12. doi:10.1167/tvst.11.7.12

Zhu, Q. J., Liang, J., Xu, C. L., Wu, X., Zhu, M. H., Ma, L., et al. (2019). Quantification of optical coherence tomographic angiography images in type 2 diabetic patients. *Zhonghua Yan Ke Za Zhi* 55, 273–279. doi:10.3760/cma.j.issn.0412-4081.2019.04.009

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 10 May 2024
DOI 10.3389/fbioe.2024.1315398

Check for updates

OPEN ACCESS

EDITED BY
Laura Cercenelli,
University of Bologna, Italy

REVIEWED BY
Sana Boudabbous,
HUG, Switzerland
Yue Feng,
Zhejiang Provincial People's Hospital, China

*CORRESPONDENCE
Junna Wang,
junna@zcmu.edu.cn
Zengru Xie,
xiezenru@126.com
Chuang Ma,
1516175919@qq.com

†These authors have contributed equally to this work

RECEIVED 10 October 2023
ACCEPTED 17 April 2024
PUBLISHED 10 May 2024

CITATION
Jia Q, Zheng H, Lin J, Guo J, Fan S, Alimujiang A, Wang X, Fu L, Xie Z, Ma C and Wang J (2024), Optimizing diagnosis and surgical decisions for chronic osteomyelitis through radiomics in the precision medicine era.
Front. Bioeng. Biotechnol. 12:1315398.
doi: 10.3389/fbioe.2024.1315398

COPYRIGHT
© 2024 Jia, Zheng, Lin, Guo, Fan, Alimujiang, Wang, Fu, Xie, Ma and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Optimizing diagnosis and surgical decisions for chronic osteomyelitis through radiomics in the precision medicine era

Qiyu Jia[1†], Hao Zheng[2†], Jie Lin[2†], Jian Guo[1], Sijia Fan[2], Abudusalamu Alimujiang[1], Xi Wang[1], Lanqi Fu[2], Zengru Xie[1*], Chuang Ma[1*] and Junna Wang[2*]

[1]The First Affiliated Hospital of Xinjiang Medical University, Urumqi, China, [2]The First Affiliated Hospital of Zhejiang Chinese Medical University (Zhejiang Provincial Hospital of Chinese Medicine), Hangzhou, China

Introduction: Chronic osteomyelitis is a complex clinical condition that is associated with a high recurrence rate. Traditional surgical interventions often face challenges in achieving a balance between thorough debridement and managing resultant bone defects. Radiomics is an emerging technique that extracts quantitative features from medical images to reveal pathological information imperceptible to the naked eye. This study aims to investigate the potential of radiomics in optimizing osteomyelitis diagnosis and surgical treatment.

Methods: Magnetic resonance imaging (MRI) scans of 93 suspected osteomyelitis patients were analyzed. Radiomics features were extracted from the original lesion region of interest (ROI) and an expanded ROI delineated by enlarging the original by 5 mm. Feature selection was performed and support vector machine (SVM) models were developed using the two ROI datasets. To assess the diagnostic efficacy of the established models, we conducted receiver operating characteristic (ROC) curve analysis, employing histopathological results as the reference standard. The model's performance was evaluated by calculating the area under the curve (AUC), sensitivity, specificity, and accuracy. Discrepancies in the ROC between the two models were evaluated using the DeLong method. All statistical analyses were carried out using Python, and a significance threshold of $p < 0.05$ was employed to determine statistical significance.

Results and Discussion: A total of 1,037 radiomics features were extracted from each ROI. The expanded ROI model achieved significantly higher accuracy (0.894 vs. 0.821), sensitivity (0.947 vs. 0.857), specificity (0.842 vs. 0.785) and AUC (0.920 vs. 0.859) than the original ROI model. Key discriminative features included shape metrics and wavelet-filtered texture features. Radiomics analysis of MRI exhibits promising clinical translational potential in enhancing the

Abbreviations: ROI, Region of Interest; STIR, short tau inversion recovery; PACS, Picture Archiving and Communication Systems; T1WI, T1-weighted images; T2WI, T2-weighted images; LoG, Laplacian of Gaussian; SVM, support vector machine; ROC, receiver operating characteristic; AUC, area under the curve; MCC, maximum correlation coefficient; GLCM, gray-level co-occurrence matrix.

diagnosis of chronic osteomyelitis by accurately delineating lesions and identifying surgical margins. The inclusion of an expanded ROI that encompasses perilesional tissue significantly improves diagnostic performance compared to solely focusing on the lesions. This study provides clinicians with a more precise and effective tool for diagnosis and surgical decision-making, ultimately leading to improved outcomes in this patient population.

KEYWORDS

chronic osteomyelitis, radiomics, MRI, diagnostic accuracy, surgical decisionmaking, shape features, region of interest

1 Introduction

Chronic osteomyelitis has long been recognized as one of the most challenging diseases in the medical field, often referred to as the "second cancer" (Zhang et al., 2019; Masters et al., 2022). Despite the refined antimicrobial activity of new-generation antibiotics and the efficacy of surgical intervention, the recurrence rate of chronic osteomyelitis remains as high as 20%–30% (Conterno and Turchi, 2013; Chastain and Davis, 2019; Zeitlinger, 2019; Wang X. et al., 2023).

Accurate diagnosis of chronic osteomyelitis is essential, as misdiagnosis can lead to the worst outcomes due to differences in treatment approaches. Magnetic Resonance imaging (MRI), with its excellent contrast between bone and soft tissue, is currently one of the most valuable tools for diagnosing chronic osteomyelitis. However, relying solely on MRI can be challenging for differentiating diseases with similar radiographic features, such as bone tuberculosis and osteosarcoma. These conditions often present with bone marrow edema, soft tissue masses, and inflammatory changes, which may overlap in imaging characteristics, thus complicating differential diagnosis and potentially affecting subsequent treatment strategies. Furthermore, once a definitive diagnosis is made, thorough debridement is necessary, as incomplete debridement can lead to recurrent infections (Bosse et al., 2002). Therefore, relying on intraoperative judgment based on experience, experienced orthopedic surgeons have increasingly adopted expanding debridement as the preferred approach. However, this practice can result in extensive bone defects, which pose a significant challenge during the postoperative period (Heng et al., 2023). On the other hand, narrowing the debridement area increases the risk of infection recurrence. Thus, it seems challenging to strike a balance between the two approaches. Moreover, even with the expansion of the debridement area, there are instances where chronic osteomyelitis still recurs, highlighting the limitations of current interventional approaches and the harsh reality faced in clinical practice (Hotchen et al., 2020).

Radiomics, an emerging diagnostic and adjunctive imaging technique, has witnessed rapid development in recent years, offering hope in addressing this issue (Huang et al., 2016; Witt et al., 2020; Bera et al., 2022; Wang W. et al., 2023). High throughput radiomics transforms traditional medical images into highly reliable, reproducible, and non-redundant data that can be mined for valuable information by extracting and analyzing a large volume of advanced and quantitative image feature data. These extracted features provide insights into pathological and physiological phenomena that are not readily discernible by the naked eye in chronic osteomyelitis, particularly in terms of structural damage and alterations in image texture (Gillies et al., 2015; Huang et al., 2016; Bera et al., 2022; Cuce et al., 2023). Currently, this technique is primarily applied in the classification and prediction of various cancers (Lambin et al., 2017; Hassani et al., 2019). Its high diagnostic accuracy can also be leveraged to reduce the misdiagnosis rate of chronic osteomyelitis by mining big data from radiographic images and facilitating early diagnosis and treatment.

For orthopedic surgeons, the future potential primary advantage of this technique lies in optimizing surgical decision-making. The infected area and its extent can be accurately determined by analyzing preoperative imaging data and establishing precise three-dimensional reconstruction images combined with machine learning and image segmentation techniques. The development of expanded detection technology provides the foothold for optimal debridement in cases of osteomyelitis (Wu et al., 2021), which involves improving the discrimination of the primary lesion area and defining a reliable and safe zone based on "artificial intelligence" judgments. This approach can assist in preoperative planning and optimizing debridement strategies, allowing surgeons to treat patients more accurately and efficiently while ensuring complete removal of infected tissue and minimizing bone damage. To achieve the previously stated objectives, we first need to determine whether radiomics technology is sufficiently effective in diagnosing chronic osteomyelitis by assessing the combined area of lesions visible to the naked eye and their surrounding expanded regions (potential lesions).

Herein, we sought to develop an imaging-based osteomyelitis diagnosis model using radiomics analysis by comprehensively analyzing and extracting features from patient MRI data and evaluating its accuracy and reliability in determining the nature and extent of lesions. This approach could facilitate diagnosis and surgical decision-making, achieving a breakthrough in treating chronic osteomyelitis.

2 Materials and methods

2.1 Data collection

A retrospective analysis was conducted on the clinical and imaging data of 93 patients with an initial diagnosis suspected to be chronic osteomyelitis of the long bones (Figure 1) who attended the First Affiliated Hospital of Xinjiang Medical University from January 2016 to May 2022. The study population comprised predominantly of males (n = 63/93, 67.7%), with a mean age of 35.5 (range: 19–67 years). Inclusion/exclusion criteria: Patients with a high suspicion of having

FIGURE 1
Macroscopic photo of a patient with chronic osteomyelitis.

chronic osteomyelitis of the long bones and requiring surgical intervention were included. These patients had well-established bone marrow infection persisting for more than 10 weeks, and the diagnosis was based on intraoperative histopathological examination or at least two sites with the same pathogen cultured or well-defined sinus tracts directly connected to the long bones, excluding cases with specific infections such as mycobacteria. The medical records of patients diagnosed with chronic osteomyelitis of the long bones included the following data: gender, age, anatomical site of infection, intraoperative microbiological culture results, treatment strategies, serum biomarkers, and MRI images after admission. Patients who were pregnant, breastfeeding, had metal implants, or were diagnosed with acute osteomyelitis, Charcot disease, diabetes, or chronic osteomyelitis in non-long bone locations were all excluded from the study. Moreover, if a patient had multiple medical records (multiple hospitalizations), only the most relevant record related to chronic osteomyelitis of the long bones was retained for analysis. Based on retrospective pathological analysis, the study ultimately included 48 patients with chronic osteomyelitis and 45 patients with non-chronic osteomyelitis. The Ethics Committee of The First Affiliated Hospital of Xinjiang Medical University approved the study with an informed consent exemption (K202308-11). Patients' personal information was anonymized and de-identified prior to analysis.

2.2 MR scanning method

The patient image was acquired by a 1.5T MR scanner (SIEMENS). The following scanning parameters were used: T1WI TR 600 ms TE 9.5 ms; T2WI TR 3000 ms TE 88 ms; FS T2WI TR3600 ms TE 83 ms, FOV320 mm, matrix: 256 × 256. All patients underwent routine scanning, including T1-weighted sequences, T2-weighted sequences, and a T2-based short tau inversion recovery (STIR) sequence. Scans were performed in the coronal, sagittal, and axial planes according to the location of lesions, with a slice thickness of 4 mm and an interslice gap of 0.4 mm.

2.3 Lesion segmentation and radiomics feature extraction

2.3.1 Image selection

MRI images of selected patients were extracted from the Picture Archiving and Communication Systems (PACS). The images were reviewed by a radiologist with over 10 years of musculoskeletal imaging diagnostic experience. T1-weighted images (T1WI) exhibited low signal intensity, whereas T2-weighted images (T2WI) and Short Tau Inversion Recovery images displayed high signal intensity for bone marrow inflammation. The presence of complete sequences without artifacts was confirmed before proceeding with image delineation.

2.3.2 Lesion and perilesional area delineation

1) Lesion delineation was performed using 3D Slicer (version 5.3.0) software, simultaneously outlining both the outer contour and bounding box of the lesion. 2) Using the STIR sequence based on the MRI T2 sequence as a reference, the region of interest (ROI) was delineated in the coronal plane. 3) Manual delineation was conducted on each lesion plane while avoiding areas of necrosis and hemorrhage, with an emphasis on comprehensive coverage of

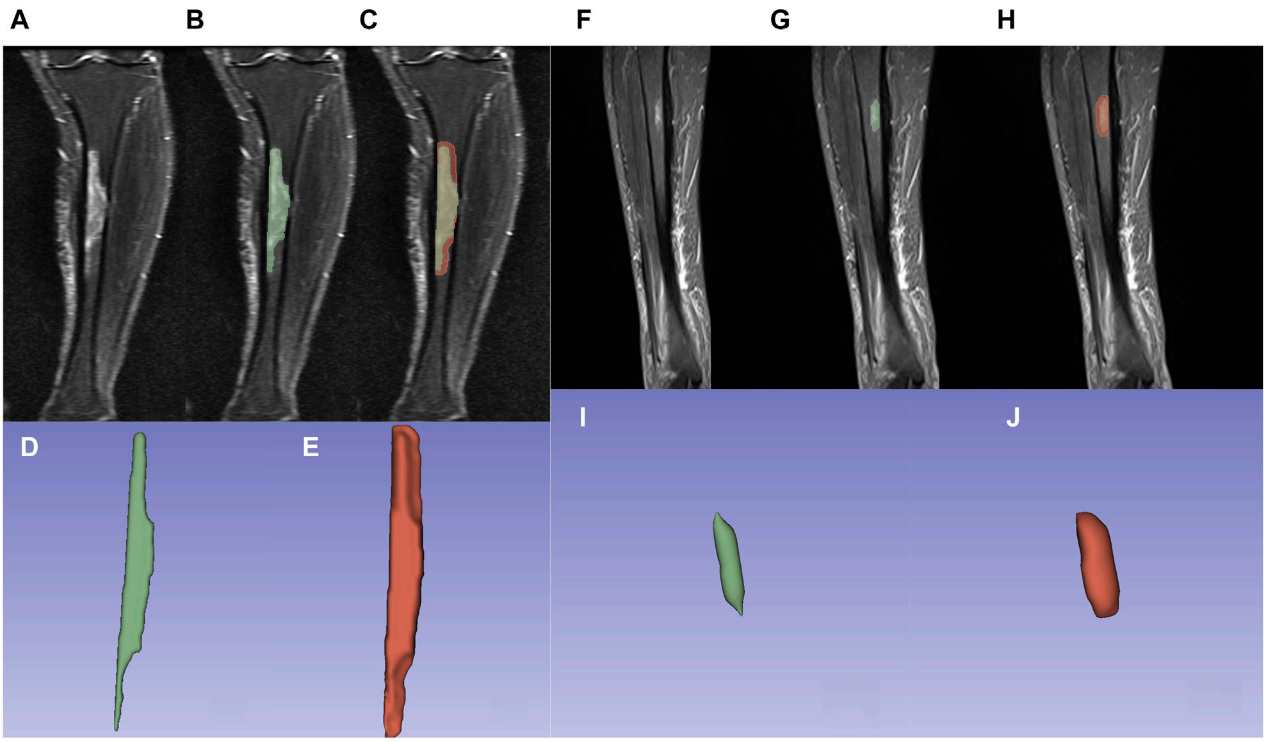


FIGURE 2
Osteomyelitis outlined using 3Dslicer. **(A,F)** Osteomyelitis in the proximal-medial tibia on the coronal view of the STIR sequence. **(B,G)** Osteomyelitis outlined within original ROI. **(C,H)** Osteomyelitis outlined within expanded ROI. **(D,I)** 3D model of osteomyelitis reconstructed from original ROI. **(E,J)** 3D model of osteomyelitis reconstructed from expanded ROI.

the lesion substance. 4) Semi-automatic delineation was applied to the original region of interest (original ROI), displaying the lesion on the image, as well as the lesion area expanded by 5 mm (expanded ROI) from the original ROI. Subsequent manual adjustments were made to confirm the delineation scope, preventing any extension of the delineation beyond the bone structure. 5) For lesions with unclear borders, distinct high-signal areas were delineated. 6) In the case of multiple lesions, only the largest lesion was delineated. The results of lesion and perilesional area delineation are illustrated in Figure 2. 7) The delineation and review of the Regions of Interest (ROIs) were conducted by two radiologists, each boasting over a decade of expertise in musculoskeletal imaging diagnostics.

2.3.3 Raiomics feature extraction

In this study, radiomics feature extraction was conducted using the pyradiomics module within the 3D Slicer software. A total of 1,037 features were automatically extracted for each ROI, encompassing various categories of features, including shape, firstorder, glcm, glszm, gldm, glrlm, and ngdtm. Furthermore, wavelet filtering and Laplacian of Gaussian (LoG) filtering techniques were applied. Wavelet filtering was applied to iteratively break down the initial image into various scales, thereby extracting valuable insights across diverse levels. In contrast, LoG filtering functioned as an edge enhancement filter, predominantly highlighting areas with significant variations in gray levels. By manipulating the sigma parameter in the LoG filtering process, we could regulate the prominence of texture characteristics. Smaller sigma values were employed to enhance intricate texture intricacies, while larger sigma values highlighted textural attributes on a broader scale (Lambin et al., 2012; van Griethuysen et al., 2017).

2.4 Statistical analysis

We used Python (ver 3.9.13) to process the two sets of radiomics features extracted from 3D Slicer, encompassing both the original and expanded ROI. Initially, data standardization was executed on the two datasets through a standardized method. The patient cohort was subsequently randomly partitioned into a training set and a testing set, maintaining a ratio of 7:3. Following this, dimensionality reduction was implemented on the extracted radiomics features via *t*-test, least absolute shrinkage and selection operator (LASSO) regression analysis, and the SelectKBest classifier.

The support vector machine (SVM) model was employed for modeling the kernel functions of both the original and expanded ROIs. The dimensionality-reduced features were employed for classification tasks. To assess the diagnostic efficacy of the established models, we conducted receiver operating characteristic (ROC) curve analysis, employing histopathological results as the reference standard. The model's performance was evaluated by calculating the area under the curve (AUC), sensitivity, specificity, and accuracy. Discrepancies in the ROC between the two models were evaluated using the DeLong method (DeLong et al., 1988). All statistical analyses were carried out using Python (ver 3.9.13), and a significance threshold of $p < 0.05$ was employed to determine statistical significance.

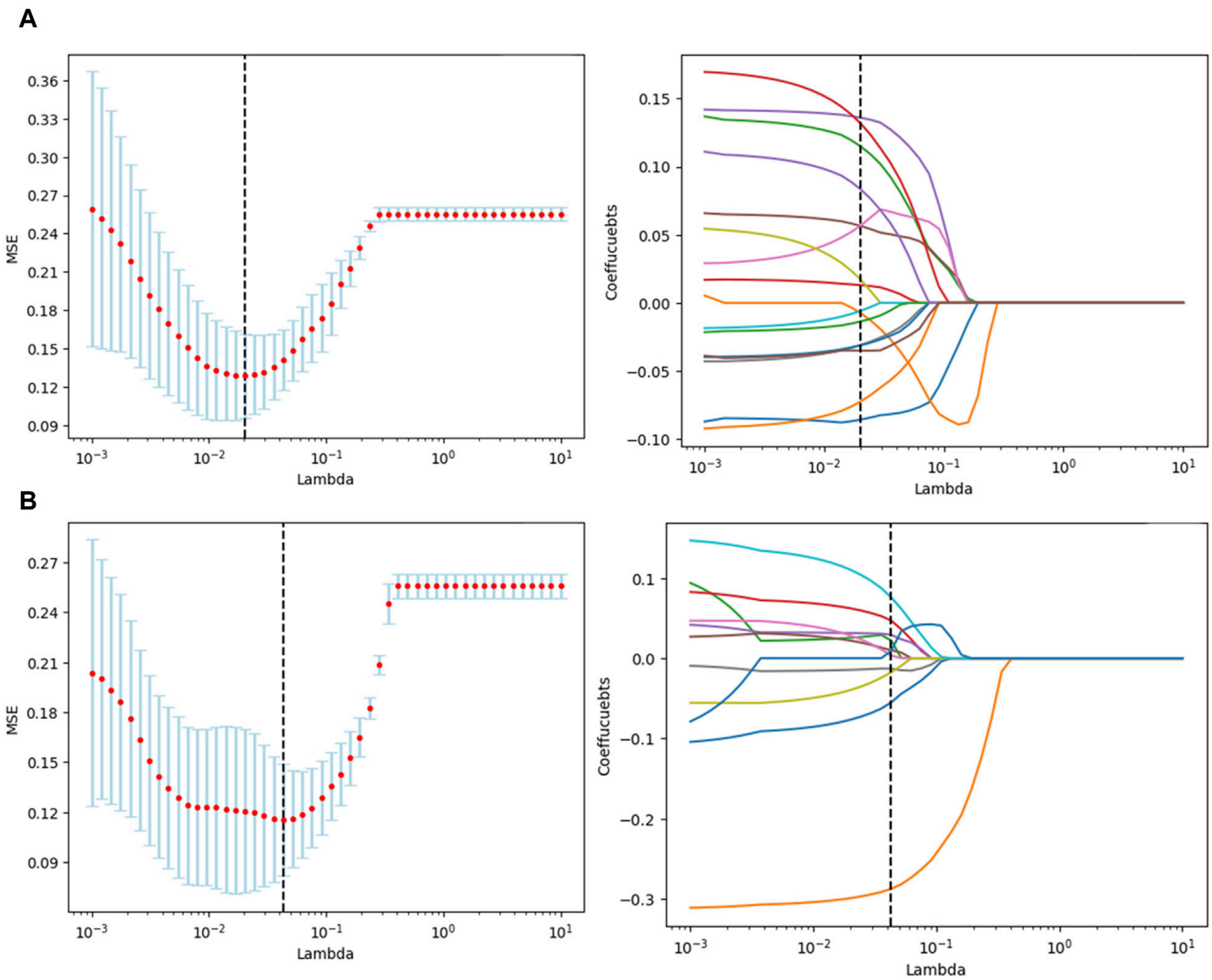


FIGURE 3
Texture feature selection using t-tests and selection operator (LASSO) for radiomics. **(A)** Original ROI texture feature selection. **(B)** Expanded ROI texture feature selection.

3 Results

3.1 Radiomics feature extraction

Utilizing the pyradiomics module within the 3D Slicer software, feature extraction was conducted separately on the original ROI and the expanded ROI, yielding a total of 1,037 features. Following feature selection through t-tests and the LASSO method, 16 and 11 features were retained for the original and expanded ROI, respectively.

Figures 3A,B illustrate the feature selection results, displaying the LASSO-driven selection of lesion texture features. To further reduce feature dimensionality and construct the SVM model, optimal polynomial degrees of 2 were determined during the hyperparameter grid search for SVM. The cost variables for the original ROI and the expanded ROI were set at 0.5 and 3.05, respectively, with scaling variables of 0.15 and 0.0625. Subsequent feature refinement was performed using the SelectKBest classifier, which employs statistical methods such as the chi-squared test, F-test, or mutual information to evaluate the relationship between each feature and the target variable. Features are ranked and selected based on the magnitude of the computed statistic (Bisong, 2019). By applying the SelectKBest classifier, we ultimately identified the top 10 features for the original and expanded ROI. Figures 4A,B depicts the feature weights selected for the original and expanded ROIs. The correlations between features within the original and expanded ROIs are visually represented through heatmap matrices in Figures 5A,B, respectively. All coefficients of total feature values were less than 0.7, implying the absence of collinearity among the features. The feature selection process underscores the significance of all gathered parameters as pivotal predictive elements for machine learning algorithms.

3.2 Comparison of model performance between original ROI and expanded ROI

The original ROI model demonstrated an excellent diagnostic performance with an accuracy of 0.821, sensitivity of 0.857, and

10.3389/fbioe.2024.1315398

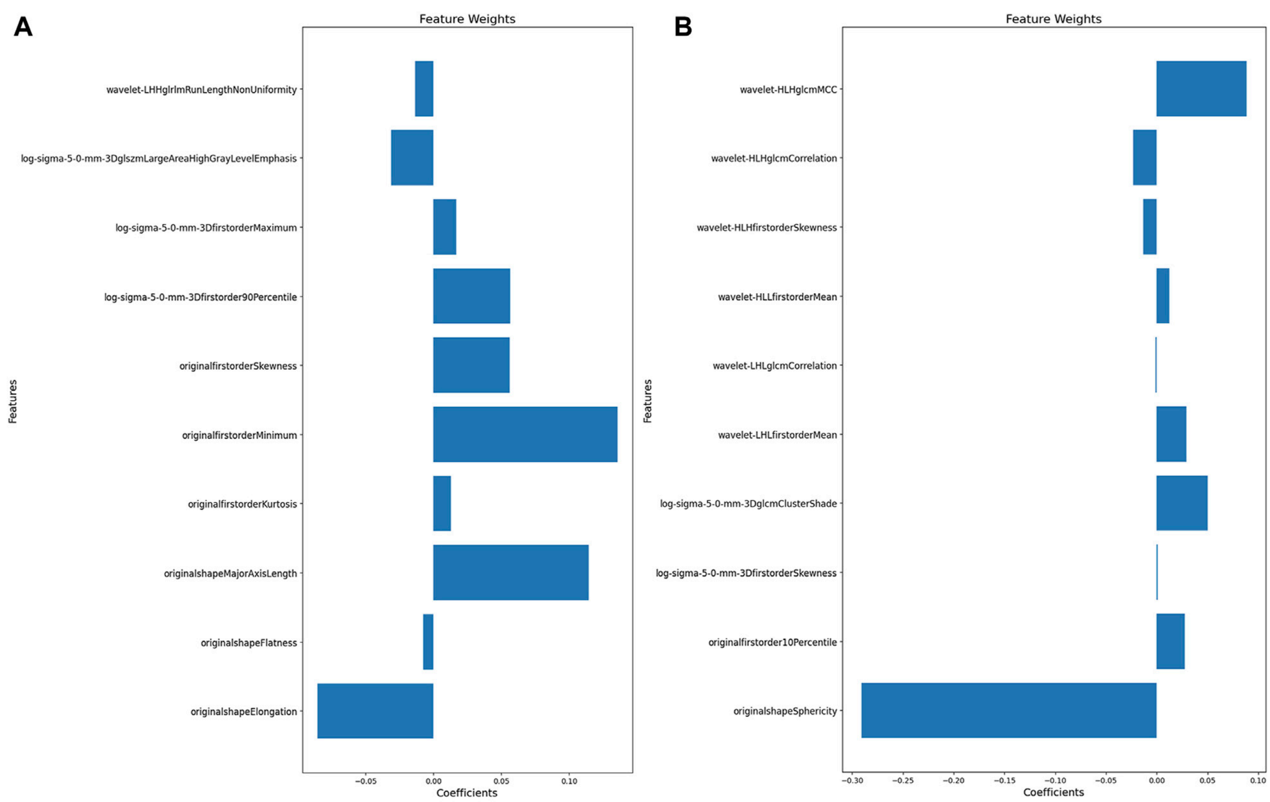


FIGURE 4
The features weight selected for original and expanded ROIs. **(A)** Features weight within the original ROI. **(B)** Features weight within the expanded ROI.

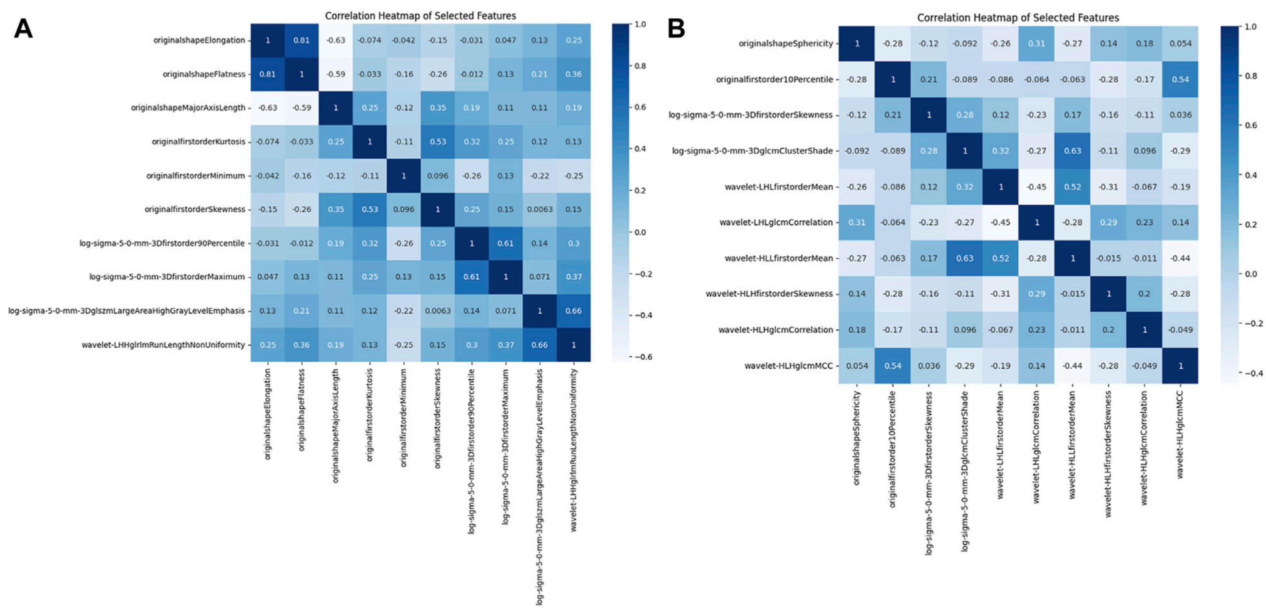


FIGURE 5
The correlation matrix heatmap. **(A)** Original ROI correlation matrix heat map. **(B)** Expanded ROI correlation matrix heat map.

TABLE 1 Diagnostic performance comparison between original ROI and expanded ROI models.

	Original ROI model	Expanded ROI model
Accuracy	0.821	0.894
Sensitivity	0.857	0.947
Specificity	0.785	0.842
AUC	0.859	0.920

specificity of 0.785. Nonetheless, the expanded ROI model exhibited higher accuracy (0.894), sensitivity (0.947), and specificity (0.842), significantly outperforming the SVM model based on the original ROI radiomics features (Table 1). As depicted in Figures 6A,B, the expanded ROI model's predictive performance was significantly superior to the original ROI (AUC value 0.920 vs. 0.859). DeLong's test confirmed a significant difference between the two approaches ($Z = 3.336$, $p < 0.001$), indicating the enhanced diagnostic efficacy of the expanded ROI model.

4 Discussion

MRI stands out for its absence of electromagnetic radiation, swift examination times, and exceptional contrast depiction between bone and soft tissues. Beyond its capacity for multi-directional imaging, MRI boasts superior sensitivity in pinpointing lesion locations and extents compared to X-rays and CT scans (Hatzenbuehler and Pulling, 2011; Acikgoz and Averill, 2014). Furthermore, contrast-enhanced MRI can effectively delineate abscesses and sinus tracts related to chronic osteomyelitis, thus improving diagnostic precision. However, the utilization of gadolinium-based contrast agents presents iatrogenic hazards to patients. While MRI currently plays a pivotal role in osteomyelitis diagnosis, significant challenges remain for clinical diagnostics (Lee et al., 2016; Wong et al., 2019). In cases of chronic osteomyelitis, MRI reveals distinct patterns with low signal intensity on T1WI and high signal intensity on T2WI and STIR sequences. Additionally, the surrounding soft tissues often exhibit edema, inflammatory alterations, and localized osteolysis. However, distinguishing chronic osteomyelitis from conditions like osteosarcoma and osteotuberculosis which present very similar imaging features on MRI, can be challenging on plain MR scans. In this context, radiomics becomes particularly important since it can extract countless quantitative features from MRI images, dynamically observe lesions and their microenvironments in a non-invasive manner, discover a large amount of information hidden in MR image layers to predict clinical endpoints or lesion properties and provide possibilities for a comprehensive assessment of lesion heterogeneity (van Griethuysen et al., 2017; Muraoka et al., 2022), as well as providing more possibilities for precise guidance of the surgical scope. Studies have shown that the microenvironment around the lesion has important biological significance in terms of lesion growth, cell migration, inflammation, and other aspects (Lambin et al., 2012). However, most radiomics studies of osteomyelitis have focused on the lesion itself, and the perilesional area has not been comprehensively explored (DeLong et al., 1988). Integrating radiomics features extracted from the perilesional area could improve lesion predictive diagnostic efficiency and offers extensive prospects for precise surgical scope guidance by incorporating microtexture intricacies and multifaceted data.

Although previous studies have evaluated osteomyelitis using MRI (Oda et al., 2018; Iwasaki and Muraoka, 2020; Massel et al., 2021), there is limited research on the application of MRI radiomics in assessing chronic osteomyelitis. Therefore, further research and exploration in this field are valuable. In this study, we constructed two different regions of interest for osteomyelitis: the original ROI based on the lesion area of osteomyelitis and an expanded ROI

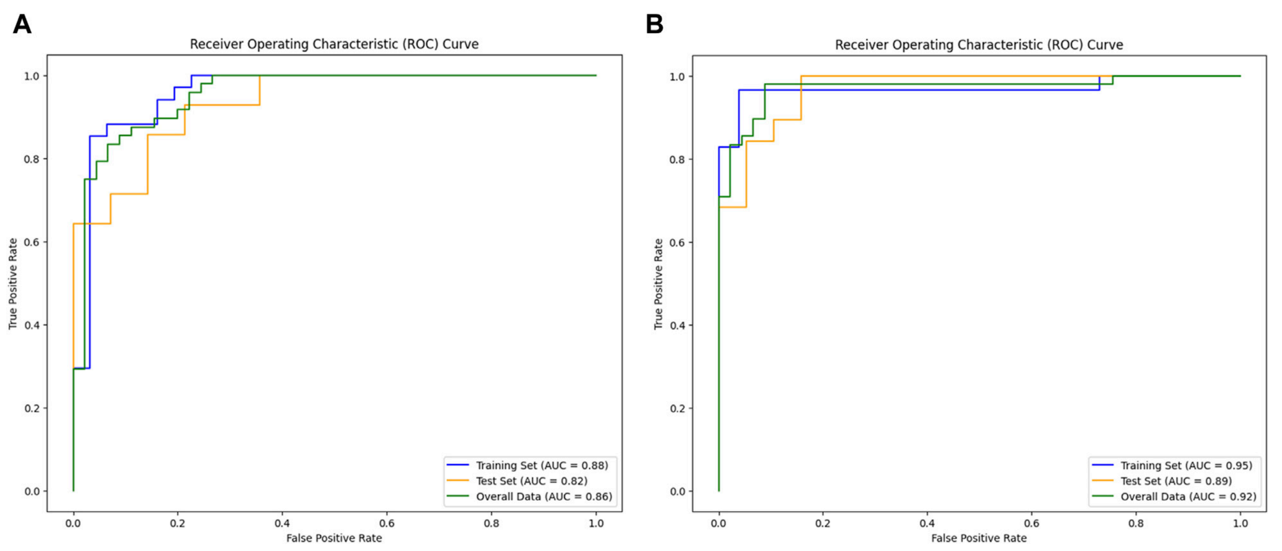


FIGURE 6
Comparison of model performance. **(A)** AUC of the original ROI. **(B)** AUC of the expanded ROI.

expanding 5 mm beyond the original. Subsequently, two SVM models were developed using these ROIs to assess the diagnostic performance of MRI radiomics in osteomyelitis diagnosis. SVM, a supervised learning algorithm, endeavors to identify a hyperplane within the feature space that maximizes the classification of distinct classes of data points, thereby facilitating effective classification. In our study, SVM demonstrated effective applicability; MRI radiomics-based osteomyelitis diagnosis achieved enhanced sensitivity compared to previous studies on MRI's diagnostic performance in osteomyelitis. For instance, compared to Hulsen et al.'s study (Hulsen et al., 2022), the sensitivity increased from 0.780 to 0.857, consistent with Hirotaka et al.'s findings regarding MRI radiomics' robust diagnostic performance in pyogenic osteomyelitis (Muraoka et al., 2022). Expanding the original ROI by 5 mm resulted in the SVM model yielded marked enhancements in accuracy (0.894 vs. 0.821), sensitivity (0.947 vs. 0.857), specificity (0.842 vs. 0.785), and AUC value (0.920 vs. 0.859), attributed to the fact that not all osteomyelitis lesions manifest as high signal areas in MRI images, and subtle lesions beyond the high signal region might remain imperceptible to the naked eye. Importantly, the radiomics-based SVM model can identify these subtle features accurately, improving diagnotic accuracy. Our findings are consistent with the widely accepted perspective that eradicating osteomyelitis necessitates expanding the surgical scope (Lazzarini et al., 2002; McNally et al., 2022; Wu et al., 2023). Our results indicate that expanding the delineation range from 0 to 5 mm substantially enhanced MRI radiomics' diagnostic performance in osteomyelitis, providing valuable insights for osteomyelitis surgical scope guidance.

Differentiating healthy tissue from non-viable tissue during early-stage surgery is inherently complex. Therefore, recognizing the significance of early, thorough debridement is widely acknowledged (Li et al., 2020). Eckardt et al. initially proposed aggressive debridement, akin to managing giant cell bone tumors, for chronic osteomyelitis treatment (Eckardt et al., 1994). Some scholars even advocate treating it like malignancy (Simpson et al., 2001), consistent with our findings. Thus, applying radiomics techniques to accurately determine lesion extent holds significant practical benefits for patients, impacting short-term surgical outcomes and long-term disease control and prognosis.

After applying t-tests and LASSO-based feature extraction, the original ROI model retained 16 features, while the expanded ROI model retained 11. The potential drawbacks of excessive features, including increased false discovery rates, overfitting, and diminished model generalization efficacy, have been highlighted (Kumar et al., 2012; Gillies et al., 2015). To mitigate these risks and enhance model accuracy, the SelectKBest classifier was employed, further reducing feature dimensionality to 10 for both models.

Based on our feature selection, the original ROI model highlighted Elongation, Major Axis Length in the shape features, and the Minimum feature in the first-order statistical features as optimal descriptors for characterizing the texture attributes within the osteomyelitis region. In contrast, within the expanded ROI model, the wavelet filter's maximum correlation coefficient (MCC) feature and sphericity from the shape-related feature set emerged as the most informative in delineating the textural characteristics of the expanded osteomyelitis suspicious area. In the original ROI model, Elongation and Major Axis Length encapsulate the primary directional characteristics of the region of interest's shape and the length of its primary axis within the enclosed ellipsoid. Elongation assesses the extent to which the ROI's shape appears elongated, while Major Axis Length quantifies the principal dimension of the ROI. During osteomyelitis imaging, a lesion can lead to anomalous enlargement and morphological alterations within the adjacent bone marrow architecture. This often manifests as an elongated lesion region with an irregularly expanded shape, collectively indicating the structural attributes of the bone and the extent of osteomyelitis infiltration (He et al., 2021). Utilizing the Minimum feature within the first-order statistical features primarily assesses the minimum grayscale intensity within the lesion region. In cases where the original ROI exclusively encompasses the osteomyelitis area displaying high signal intensity on the MRI image, the overall image's minimum grayscale value tends to be elevated. Conversely, the expanded ROI's delineation encompasses regions outside the lesion, often represented by darker areas on the MRI image. This inclusion leads to a comparatively lower overall minimum grayscale value, resulting in the reduced significance of the Minimum feature. Meanwhile, sphericity quantifies an object's resemblance to a perfect sphere, with lower values indicating deviation towards irregularity. In the context of the expanded ROI model, sphericity emerges as a pivotal discriminatory feature, reflecting the extent of roundness within the lesion area in relation to a spherical shape (Priya et al., 2021) Its prominently negative weight might stem from the indistinct boundaries and unevenness of the expanded ROI, resulting in a modified shape. Additionally, the MCC feature linked to the wavelet filter is a gray-level co-occurrence matrix (GLCM) component, signifying the intricacy of texture patterns (Su et al., 2023). Given the broader delineation of regions within the expanded ROI model, implying heightened textural complexity compared to the original ROI, the MCC feature was more prominent in the expanded ROI model.

Across the original and expanded ROI models, the shape feature retained a notably substantial weight ratio among all the screened features. Previous studies have effectively employed shape features in delineating tumor aggressiveness (Limkin et al., 2019). Similar to the invasive characteristics exhibited by tumors, osteomyelitis also demonstrates comparable aggressiveness. Shape features possess a unique capability in assessing bone erosion, which can effectively differentiate chronic osteomyelitis from other conditions in magnetic resonance imaging.

The application of radiomics ensures a relatively high level of accuracy in differentiating between residual lesions in chronic osteomyelitis, such as infected tissue, inflammatory tissue, and edema, despite their similar high signal intensity on imaging. Implementing an expanded lesion detection strategy can be likened to providing young doctors with a vantage point on the accomplishments of experts, reflecting the inevitable trajectory of rapid artificial intelligence advancement. However, it should be borne in mind that despite the immense potential of radiomics technology, its effective utilization still hinges on the expertise and experience of medical professionals for thorough analysis and interpretation. Radiomics technology functions as a supplementary tool, and doctors must still integrate elements like clinical history, physical examinations, and other supplementary test outcomes to arrive at the ultimate diagnosis and treatment decisions.

Our study has several limitations. Firstly, it is a retrospective study, potentially subject to information bias. Secondly, our discussion solely pertains to the impact of MRI radiomics on the diagnostic efficacy of osteomyelitis. Thirdly, comparing radiomics features with imaging morphological characteristics is necessary. In our future research, we plan to construct a fusion model based on both radiomics and morphological features to further explore and validate their combined diagnostic value. However, previous studies have shown that certain clinical factors, such as a history of previous wounds (Lavery et al., 2009) and microbial infections (Bury et al., 2021), may also influence the qualitative diagnosis of osteomyelitis. Therefore, in future research, integrating clinical features with radiomics may become an expanded focus to provide a more comprehensive disease diagnosis and treatment guidance. Additionally, further research is needed to investigate the extent of expanded delineation in conjunction with radiomics, aiming to achieve optimal accuracy, sensitivity, specificity, and other information. We also plan to explore additional radiomics models, such as random forest models and logistic models, to compare the diagnostic performance of different models and identify the most suitable radiomics model for osteomyelitis diagnosis.

5 Conclusion

MRI radiomics-based methods yielded promising results for diagnosing chronic osteomyelitis, especially when utilizing an expanded ROI model that enhances diagnostic accuracy. With further validation from larger-scale, high-quality studies in the future, this approach can potentially become a valuable tool for guiding surgical interventions in chronic osteomyelitis, providing accurate diagnosis and precise localization of the affected lesion areas, ultimately optimizing surgical decision-making and improving patient outcomes.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by the Ethics Committee of the First Affiliated Hospital of Xinjiang Medical University. The studies were conducted in accordance with the local legislation and institutional requirements. The ethics committee/institutional review board waived the requirement of written informed consent for participation from the participants or the participants' legal guardians/next of kin because 1. Retrospective Data: The data used in this study is derived from previously collected clinical records or biological specimens. These materials were obtained during routine clinical diagnosis and treatment, and no additional procedures or interventions were performed specifically for this research. Therefore, obtaining written informed consent from each individual patient is not feasible or necessary. 2. No Adverse Impact: The waiver of written informed consent does not pose any adverse effects on the rights, welfare, or health of the subjects involved. The study solely relies on de-identified and anonymized data, ensuring the privacy and confidentiality of the individuals' personal information. The analysis and interpretation of the data do not involve any interventions, risks, or potential harm to the participants. 3. Minimal Risk: The research poses minimal risk to the subjects, as it is based on retrospective data analysis. There are no interventions, treatments, or experimental procedures involved that could potentially harm the participants. The study strictly adheres to ethical guidelines and regulations to ensure the protection and wellbeing of the subjects. 4. Confidentiality: All data used in this study will be handled with strict confidentiality. Any personal identifiers will be removed or anonymized to ensure the privacy and confidentiality of the subjects. The research findings will be reported in an aggregated and de-identified manner to maintain the anonymity of the participants. Overall, given the retrospective nature of the study, the absence of interventions or risks, and the protection of subject confidentiality, the waiver of written informed consent is justified and will not compromise the rights or health of the individuals involved. Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article because The Ethics Committee of The First Affiliated Hospital of Xinjiang Medical University approved the study with an informed consent exemption (K202308-11).

Author contributions

QJ: Conceptualization, Data curation, Investigation, Software, Writing–original draft. HZ: Methodology, Project administration, Software, Writing–original draft. JL: Formal Analysis, Project administration, Software, Supervision, Writing–original draft. JG: Investigation, Methodology, Project administration, Writing–review and editing. SF: Data curation, Investigation, Writing–review and editing. AA: Methodology, Software, Writing–review and editing. XW: Formal Analysis, Investigation, Writing–review and editing. LF: Data curation, Writing–review and editing. ZX: Conceptualization, Resources, Supervision, Writing–review and editing. CM: Funding acquisition, Resources, Supervision, Writing–review and editing. JW: Conceptualization, Supervision, Validation, Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This research was funded by the National Natural Science Foundation of China (82260425 and 82260409) and the Natural Science Foundation of Xinjiang Uygur Autonomous Region (2021D01D19). A Project Supported by Scientific Research Fund of Zhejiang Provincial Education Department (No. Y202249555).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Acikgoz, G., and Averill, L. W. (2014). Chronic recurrent multifocal osteomyelitis: typical patterns of bone involvement in whole-body bone scintigraphy. *Nucl. Med. Commun.* 35, 797–807. doi:10.1097/mnm.0000000000000126

Bera, K., Braman, N., Gupta, A., Velcheti, V., and Madabhushi, A. (2022). Predicting cancer outcomes with radiomics and artificial intelligence in radiology *Nat. Rev. Clin. Oncol.* 19, 132–146. doi:10.1038/s41571-021-00560-7

Bisong, E. (2019). "Building machine learning and deep learning models on google cloud platform: a comprehensive guide for beginners," in *More supervised machine learning techniques with scikit-learn*. (Berkeley, CA: Apress), 287–308. doi:10.1007/978-1-4842-4470-8_24

Bosse, M. J., MacKenzie, E. J., Kellam, J. F., Burgess, A. R., Webb, L. X., Swiontkowski, M. F., et al. (2002). An analysis of outcomes of reconstruction or amputation after leg-threatening injuries. *N. Engl. J. Med.* 347, 1924–1931. doi:10.1056/nejmoa012604

Bury, D. C., Rogers, T. S., and Dickman, M. M. (2021). Osteomyelitis: diagnosis and treatment. *Am. Fam. Physician* 104, 395–402.

Chastain, D. B., and Davis, A. (2019). Treatment of chronic osteomyelitis with multidose oritavancin: a case series and literature review. *Int. J. Antimicrob. Agents* 53, 429–434. doi:10.1016/j.ijantimicag.2018.11.023

Conterno, L. O., and Turchi, M. D. (2013). Antibiotics for treating chronic osteomyelitis in adults. *Cochrane Database Syst. Rev.*, 52, CD004439. doi:10.1002/14651858.cd004439.pub3

Cuce, F., Tulum, G., Yilmaz, K. B., Osman, O., and Aralasmak, A. (2023). Radiomics method in the differential diagnosis of diabetic foot osteomyelitis and charcot neuroarthropathy. *Br. J. Radiol.* 96, 20220758. doi:10.1259/bjr.20220758

DeLong, E. R., DeLong, D. M., and Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 44, 837–845. doi:10.2307/2531595

Eckardt, J. J., Wirganowicz, P. Z., and Mar, T. (1994). An aggressive surgical approach to the management of chronic osteomyelitis. *Clin. Orthop. Relat. Res.* 298, 229–239. doi:10.1097/00003086-199401000-00030

Gillies, R., Kinahan, P., and Hricak, H. (2015). Radiomics: images are more than pictures, they are data. *Radiology* 278, 563–577. doi:10.1148/radiol.2015151169

Hassani, C., Varghese, B. A., Nieva, J., and Duddalwar, V. (2019). Radiomics in pulmonary lesion imaging. *AJR Am. J. Roentgenol.* 212, 497–504. doi:10.2214/ajr.18.20623

Hatzenbuehler, J., and Pulling, T. J. (2011). Diagnosis and management of osteomyelitis. *Am. Fam. Physician* 84, 1027–1033.

He, L., Liu, Z., Liu, C., Gao, Z., Ren, Q., Lei, L., et al. (2021). Radiomics based on lumbar spine magnetic resonance imaging to detect osteoporosis. *Acad. Radiol.* 28, e165–e171. doi:10.1016/j.acra.2020.03.046

Heng, B. C., Bai, Y., Li, X., Lim, L. W., Li, W., Ge, Z., et al. (2023). Electroactive biomaterials for facilitating bone defect repair under pathological conditions. *Adv. Sci. (Weinh)* 10, e2204502. doi:10.1002/advs.202204502

Hotchen, A. J., Dudareva, M., Corrigan, R. A., Ferguson, J. Y., and McNally, M. A. (2020). Can we predict outcome after treatment of long bone osteomyelitis? *Bone Jt. J.* 102, 1587–1596. doi:10.1302/0301-620x.102b11.bjj-2020-0284.r1

Huang, Y. Q., Liang, C. H., He, L., Tian, J., Liang, C. S., Chen, X., et al. (2016). Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J. Clin. Oncol.* 34, 2157–2164. doi:10.1200/jco.2015.65.9128

Hulsen, D. J. W., Mitea, C., Arts, J. J., Loeffen, D., and Geurts, J. (2022). Diagnostic value of hybrid FDG-PET/MR imaging of chronic osteomyelitis. *Eur. J. Hybrid. Imaging* 6, 15. doi:10.1186/s41824-022-00125-6

Iwasaki, T., and Muraoka, H. (2020). Assessment of acute osteomyelitis in the mandible using diffusion weighted MR imaging. *Int. J. Oral-Medical Sci.* 18, 172–176. doi:10.5466/ijoms.18.172

Kumar, V., Gu, Y., Basu, S., Berglund, A., Eschrich, S. A., Schabath, M. B., et al. (2012). Radiomics: the process and the challenges. *Magn. Reson Imaging* 30, 1234–1248. doi:10.1016/j.mri.2012.06.010

Lambin, P., Leijenaar, R., Deist, T., Peerlings, J., de Jong, E., Van Timmeren, J., et al. (2017). Radiomics: the bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* 14, 749–762. doi:10.1038/nrclinonc.2017.141

Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G., Granton, P., et al. (2012). Radiomics: extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer* 48, 441–446. doi:10.1016/j.ejca.2011.11.036

Lavery, L. A., Peters, E. J., Armstrong, D. G., Wendel, C. S., Murdoch, D. P., and Lipsky, B. A. (2009). Risk factors for developing osteomyelitis in patients with diabetic foot wounds. *Diabetes Res. Clin. Pract.* 83, 347–352. doi:10.1016/j.diabres.2008.11.030

Lazzarini, L., De Lalla, F., and Mader, J. T. (2002). Long bone osteomyelitis. *Curr. Infect. Dis. Rep.* 4, 439–445. doi:10.1007/s11908-002-0012-4

Lee, Y. J., Sadigh, S., Mankad, K., Kapse, N., and Rajeswaran, G. (2016). The imaging of osteomyelitis. *Quant. Imaging Med. Surg.* 6, 184–198. doi:10.21037/qims.2016.04.01

Li, R., Zhu, G., Chen, C., Chen, Y., and Ren, G. (2020). Bone transport for treatment of traumatic composite tibial bone and soft tissue defects: any specific needs besides the ilizarov technique? *Biomed. Res. Int.* 2020, 1–13. doi:10.1155/2020/2716547

Limkin, E. J., Reuze, S., Carre, A., Sun, R., Schernberg, A., Alexis, A., et al. (2019). The complexity of tumor shape, spiculatedness, correlates with tumor radiomic shape features. *Sci. Rep.* 9, 4329. doi:10.1038/s41598-019-40437-5

Massel, D. H., Jenkins, N. W., Rush, A. J., 3rd, Trapana, J. E., Foremny, G. B., Donnally, C. J., 3rd, et al. (2021). MRI and clinical risk indicators for osteomyelitis. *Foot Ankle Spec.* 14, 415–426. doi:10.1177/1938640020921572

Masters, E. A., Ricciardi, B. F., Bentley, K. L. M., Moriarty, T. F., Schwarz, E. M., and Muthukrishnan, G. (2022). Skeletal infections: microbial pathogenesis, immunity and clinical management. *Nat. Rev. Microbiol.* 20, 385–400. doi:10.1038/s41579-022-00686-0

McNally, M. A., Ferguson, J. Y., Scarborough, M., Ramsden, A., Stubbs, D. A., and Atkins, B. L. (2022). Mid-to long-term results of single-stage surgery for patients with chronic osteomyelitis using a bioabsorbable gentamicin-loaded ceramic carrier. *Bone Jt. J.* 104, 1095–1100. doi:10.1302/0301-620x.104b9.bjj-2022-0396.r1

Muraoka, H., Ito, K., Hirahara, N., Ichiki, S., Kondo, T., and Kaneda, T. (2022). Magnetic resonance imaging texture analysis in the quantitative evaluation of acute osteomyelitis of the mandibular bone. *Dentomaxillofac Radiol.* 51, 20210321. doi:10.1259/dmfr.20210321

Oda, T., Sue, M., Sasaki, Y., and Ogura, I. (2018). Diffusion-weighted magnetic resonance imaging in oral and maxillofacial lesions: preliminary study on diagnostic ability of apparent diffusion coefficient maps. *Oral Radiol.* 34, 224–228. doi:10.1007/s11282-017-0303-y

Priya, S., Liu, Y., Ward, C., Le, N. H., Soni, N., Pillenahalli Maheshwarappa, R., et al. (2021). Machine learning based differentiation of glioblastoma from brain metastasis using MRI derived radiomics. *Sci. Rep.* 11, 10478. doi:10.1038/s41598-021-90032-w

Simpson, A. H., Deakin, M., Latham, J. M., and Chronic osteomyelitis (2001). Chronic osteomyelitis: THE EFFECT OF THE EXTENT OF SURGICAL RESECTION ON INFECTION-FREE SURVIVAL. *J. Bone Jt. Surg. Br.* 83, 403–407. doi:10.1302/0301-620x.83b3.0830403

Su, Y., Lu, C., Zheng, S., Zou, H., Shen, L., Yu, J., et al. (2023). Precise prediction of the sensitivity of platinum chemotherapy in SCLC: establishing and verifying the feasibility of a CT-based radiomics nomogram. *Front. Oncol.* 13, 1006172. doi:10.3389/fonc.2023.1006172

van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., et al. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 77, e104–e107. doi:10.1158/0008-5472.can-17-0339

Wang, W., Fan, Z., and Zhen, J. (2023b). MRI radiomics-based evaluation of tuberculous and brucella spondylitis. *J. Int. Med. Res.* 51, doi:10.1177/03000605231195156

Wang, X., Zhang, M., Zhu, T., Wei, Q., Liu, G., and Ding, J. (2023a). Flourishing antibacterial strategies for osteomyelitis therapy. *Adv. Sci. (Weinh)* 10, e2206154. doi:10.1002/advs.202206154

Witt, J. S., Rosenberg, S. A., and Bassetti, M. F. (2020). MRI-guided adaptive radiotherapy for liver tumours: visualising the future. *Lancet Oncol.* 21, e74–e82. doi:10.1016/s1470-2045(20)30034-6

Wong, D., Holtom, P., and Spellberg, B. (2019). Osteomyelitis complicating sacral pressure ulcers: whether or not to treat with antibiotic therapy. *Clin. Infect. Dis.* 68, 338–342. doi:10.1093/cid/ciy559

Wu, H., Sun, D., Wang, S., Jia, C., Shen, J., Wang, X., et al. (2023). Incidence and risk factors of recurrence in limb osteomyelitis patients after antibiotic-loaded cement spacer for definitive bone defect treatment. *Bone Jt. Res.* 12, 467–475. doi:10.1302/2046-3758.128.bjr-2022-0413.r2

Wu, L., Gao, C., Ye, J., Tao, J., Wang, N., Pang, P., et al. (2021). The value of various peritumoral radiomic features in differentiating the invasiveness of adenocarcinoma manifesting as ground-glass nodules. *Eur. Radiol.* 31, 9030–9037. doi:10.1007/s00330-021-07948-0

Zeitlinger, M. (2019). A pragmatic trial in bone and joint infection. *Lancet Infect. Dis.* 19, 804–805. doi:10.1016/s1473-3099(19)30354-8

Zhang, X., Lu, Q., Liu, T., Li, Z., and Cai, W. (2019). Bacterial resistance trends among intraoperative bone culture of chronic osteomyelitis in an affiliated hospital of South China for twelve years. *BMC Infect. Dis.* 19, 823. doi:10.1186/s12879-019-4460-y

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 16 May 2024
DOI 10.3389/fbioe.2024.1398237

Check for updates

OPEN ACCESS

EDITED BY
Adrian Elmi-Terander,
Stockholm Spine Center, Sweden

REVIEWED BY
Shireen Y. Elhabian,
The University of Utah, United States
Nguyen Quoc Khanh Le,
Taipei Medical University, Taiwan

*CORRESPONDENCE
Guanqun Sun,
sun.guanqun@hmc.edu.cn
Le-Minh Nguyen,
nguyenml@jaist.ac.jp
Junyi Xin,
xinjunyi@hmc.edu.cn

RECEIVED 09 March 2024
ACCEPTED 18 April 2024
PUBLISHED 16 May 2024

CITATION
Sun G, Pan Y, Kong W, Xu Z, Ma J, Racharak T, Nguyen L-M and Xin J (2024), DA-TransUNet: integrating spatial and channel dual attention with transformer U-net for medical image segmentation.
Front. Bioeng. Biotechnol. 12:1398237.
doi: 10.3389/fbioe.2024.1398237

COPYRIGHT
© 2024 Sun, Pan, Kong, Xu, Ma, Racharak, Nguyen and Xin. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

DA-TransUNet: integrating spatial and channel dual attention with transformer U-net for medical image segmentation

Guanqun Sun[1,2]*, Yizhi Pan[1], Weikun Kong[3], Zichang Xu[4], Jianhua Ma[5], Teeradaj Racharak[2], Le-Minh Nguyen[2]* and Junyi Xin[1,6,7]*

[1]School of Information Engineering, Hangzhou Medical College, Hangzhou, China, [2]School of Information Science, Japan Advanced Institute of Science and Technology, Nomi, Japan, [3]Department of Electronic Engineering, Tsinghua University, Beijing, China, [4]Department of Systems Immunology, Immunology Frontier Research Institute (IFReC), Osaka University, Suita, Japan, [5]Faculty of Computer and Information Sciences, Hosei University, Tokyo, Japan, [6]Zhejiang Engineering Research Center for Brain Cognition and Brain Diseases Digital Medical Instruments, Hangzhou Medical College, Hangzhou, China, [7]Academy for Advanced Interdisciplinary Studies of Future Health, Hangzhou Medical College, Hangzhou, China

Accurate medical image segmentation is critical for disease quantification and treatment evaluation. While traditional U-Net architectures and their transformer-integrated variants excel in automated segmentation tasks. Existing models also struggle with parameter efficiency and computational complexity, often due to the extensive use of Transformers. However, they lack the ability to harness the image's intrinsic position and channel features. Research employing Dual Attention mechanisms of position and channel have not been specifically optimized for the high-detail demands of medical images. To address these issues, this study proposes a novel deep medical image segmentation framework, called DA-TransUNet, aiming to integrate the Transformer and dual attention block (DA-Block) into the traditional U-shaped architecture. Also, DA-TransUNet tailored for the high-detail requirements of medical images, optimizes the intermittent channels of Dual Attention (DA) and employs DA in each skip-connection to effectively filter out irrelevant information. This integration significantly enhances the model's capability to extract features, thereby improving the performance of medical image segmentation. DA-TransUNet is validated in medical image segmentation tasks, consistently outperforming state-of-the-art techniques across 5 datasets. In summary, DA-TransUNet has made significant strides in medical image segmentation, offering new insights into existing techniques. It strengthens model performance from the perspective of image features, thereby advancing the development of high-precision automated medical image diagnosis. The codes and parameters of our model will be publicly available at https://github.com/SUN-1024/DA-TransUnet.

KEYWORDS

U-net, medical image segmentation, dual attention, transformer, deep learning

1 Introduction

Machine learning and deep learning techniques have emerged as powerful tools in biomedical research, revolutionizing disease diagnosis, treatment planning, and personalized medicine (Le, 2024; Tran and Le, 2024). Medical image segmentation is the process of delineating regions of interest within medical images for diagnosis and treatment planning. It serves as a cornerstone in medical image analysis. Manual segmentation is both accurate and affordable for pathology diagnosis but vital in standardized clinical settings. Conversely, automated segmentation ensures a reliable and consistent process, boosting efficiency, cutting down on labor and costs, and preserving accuracy. Consequently, there is a substantial demand for exceptionally accurate automated medical image segmentation technology within the realm of clinical diagnostics. However, medical image segmentation faces unique challenges, such as the need for precise delineation of complex anatomical structures, variability across patients, and the presence of noise and artifacts in the images (Tran et al., 2023). These challenges necessitate the development of advanced segmentation techniques that can capture fine-grained details while maintaining robustness and efficiency.

In the last decade, the traditional U-net structure has been widely employed in numerous segmentation tasks, yielding commendable outcomes. Notably, the U-Net model (Ronneberger et al., 2015), along with its various enhanced iterations, has achieved substantial success. ResUnet (Diakogiannis et al., 2020) emerged during this period, influenced by the residual concept. Similarly, UNet++ (Zhou et al., 2018) emphasizes enhancements in skip connections. Moving beyond these CNN-based approaches, the Transformer architecture introduces a completely new perspective. The transformer (Vaswani et al., 2017), originally developed for sequence-to-sequence modeling in Natural Language Processing (NLP), has also found utility in the field of Computer Vision (CV). ViTs segment images into patches and input their embeddings into a transformer network for strong performance. (Dosovitskiy et al., 2020). This signifies a trend of shifting from traditional CNN models to more flexible Transformer models. While the above-mentioned U-Net structures have enhanced the capabilities of models in segmentation tasks (Ronneberger et al., 2015; Zhou et al., 2018; Diakogiannis et al., 2020), they do not integrate the more powerful feature extraction abilities inherent in the Transformer and attention mechanisms, which limits their potential for further improvement. On the one hand, several studies have made progress in image segmentation by leveraging Dual Attention (DA) mechanisms for both channels and positions. The Dual Attention Network (DANet) utilizes a Position Attention Block (PAM) and Channel Attention Block (CAM) from the DA Network for natural scene image segmentation (Fu et al., 2019). This research primarily focuses on scene segmentation and does not explore the unique characteristics of medical imagery. Also, DAResUnet (Shi et al., 2020) introduces a dual attention block combined with a residual block (Res-Block) in a U-net architecture for medical image segmentation, demonstrating significant improvements in this domain. However, in the realm of medical image segmentation, existing models, including those employing Dual Attention mechanisms, have not yet extensively explored the optimal integration of Dual Attention with Transformer models for enhanced feature extraction; this oversight represents a significant research opportunity in the task of medical image segmentation. Therefore, addressing this gap and optimizing the integration of Transformers and Dual Attention mechanisms in the context of medical image segmentation poses a significant challenge for future research in the field.

To overcome the above drawbacks, recent studies have explored the application of Transformer models in medical image segmentation. Inspired by ViTs, TransUNet (Chen et al., 2021) further combines the functionality of ViTs with the advantages of U-net in the field of medical image segmentation. Specifically, it employs a transformer's encoder to process the image and employs CNN and hopping connections for accurate up-sampling feature recovery, yet it neglects image-specific features like position and channel. These aspects are crucial for capturing the nuanced variations and complex structures often present in medical images, which are essential for accurate diagnosis and analysis. Swin-Unet (Cao et al., 2022) combines the Swin-transform block with the U-net structure and achieves good results. Yet, adding extensive Transformer blocks inflates the parameter count without significantly improving results. This study merely stacked multiple Transformers to enhance models, resulting in inflated parameters and computational complexity with marginal gains in performance. Moreover, some studies have specifically focused on incorporating position and channel attention mechanisms in medical image segmentation. For instance, DA-DSUnet has been applied to head-and-neck tumor segmentation, but it doesn't combine Position Attention Module (PAM) and Channel Attention Module (CAM), nor does it discuss the potential filtering role of DA blocks in skip connections (Tang et al., 2021). Additionally, it doesn't leverage ViT for feature extraction. Another example is research on brain tumor segmentation, which, while applying DA blocks, limits its scope to brain tumors without validating other types of medical images (Sahayam et al., 2022). These studies integrate DA blocks with other blocks but do not thoroughly explore the role of DA in skip connections or optimize DA blocks for the unique intricacies of medical imaging.

However, Despite the progress made by these transformer-based approaches, they often overlook the importance of integrating image-specific features, such as position and channel information, which are crucial for capturing the nuanced variations and complex structures in medical images. Moreover, the existing methods that incorporate dual attention mechanisms have not been optimized for the unique characteristics of medical imagery, leaving room for further improvement. To address these limitations, we propose DA-TransUNet, which strategically integrates the Dual Attention Block (DA-Block) into the transformer-based U-Net architecture, specifically tailored for medical image segmentation.

In this research, our proposed model DA-TransUNet is an innovative approach for medical image segmentation that integrates the Transformer mechanism, specifically the Vision Transformer (ViT) and a Dual Attention (DA) mechanism within a U-Net architecture. First, the Transformer ViT is combined with DA in the encoder of the U-Net structure, enhancing feature extraction capabilities by leveraging the detailed characteristics of medical images. This integration allows the model to capture both local and global contextual information, which is essential for accurate segmentation of complex anatomical structures. Then, to further refine feature extraction tailored to medical images, DA is optimized for specific channels and incorporated into every module of the skip connections, enabling the model to effectively filter out irrelevant information and focus on the most discriminative features. The skip connections pass the shallow positional information from the encoder, while the DA module refines the crucial detailed features. This targeted optimization is substantiated by extensive ablation studies, demonstrating its significance in improving the model's performance. Lastly, this architecture has been rigorously tested across five medical image segmentation datasets and extensive ablation studies, demonstrating its effectiveness and superiority (Candemir et al., 2013; Jaeger et al., 2013; Bernal et al., 2015; Landman et al., 2015; Tschandl et al., 2018; Codella et al., 2019; Jha et al., 2020; Jha et al., 2021).

The main contributions of this article are summarized as follows:

1) The model of DA-TransUnet is proposed by integrating Transformer ViT and Dual Attention in U-net architecture's encoder and skip connections. This design enhances feature extraction capabilities in better extracting detailed features of medical images.
2) We propose an optimized Dual Attention (DA) Block that is designed for medical image segmentation with two key enhancements: the optimization of intermediate channel configurations within the DA block, and its integration into each skip-connection layer for effectively filtering irrelevant information. These are validated through comprehensive ablation experiments.
3) The segmentation performance and generalization ability of DA-TransUnet are validated on five medical datasets. In comparison to recent related studies, DA-TransUnet exhibits superior results in medical image segmentation, demonstrating its effectiveness in this field.

The rest of this article is organized as follows. Section 2 reviews the related works of automatic medical image segmentation, and the description of our proposed DA-TransUNet is given in Section 3. Next, the comprehensive experiments and visualization analyses are conducted in Section 4. Finally, Section 5 makes a conclusion of the whole work.

2 Related work

2.1 U-net model

Recently, attention mechanisms have gained popularity in U-net architectures (Ronneberger et al., 2015). For example, Attention U-net incorporates attention mechanisms to enhance pancreas localization and segmentation performance (Oktay et al., 2018); DAResUnet integrates both double attention and residual mechanisms into U-net (Shi et al., 2020); Attention Res-UNet explores the substitution of hard-attention with soft-attention (Maji et al., 2022); Sa-unet incorporates a spatial attention mechanism in U-net (Guo et al., 2021). Following this, TransUNet innovatively combines Transformer and U-net structure (Chen et al., 2021). Building on TransUNet, TransU-Net++ incorporates attention mechanisms into both skip connections and feature extraction (Jamali et al., 2023). Swin-Unet (Cao et al., 2022) improves by replacing every convolution block in U-net with Swin-Transformer (Liu et al., 2021). DS-TransUNet proposes to incorporate the tif module (which is a multi-scale module using Transformer) to the skip connection to improve the model (Lin et al., 2022). AA-transunet leverages Block Attention Model (CBAM) and Deep Separable Convolution to further optimize TransUNet (Yang and Mehrkanoon, 2022). TransFuse uses dual attention Bifusion blocks and AG to fuse features of two different parts of CNN and Transformer (Zhang et al., 2021). Numerous attention mechanisms have been added to U-net and TransUNet models, yet further exploration is warranted. Diverging from prior approaches, our experiment introduces a dual attention mechanism and Transformer module into the traditional U-shaped encoder-decoder and skip connections, yielding promising results.

2.2 Application of skip connections in medical image segmentation modeling

Skip connections in U-net aim to bridge the semantic gap between the encoder and decoder, effectively recovering fine-grained object details (Drozdzal et al., 2016; He et al., 2016; Huang et al., 2017). There are three primary modifications to skip connections: firstly, increasing their complexity (Azad et al., 2022a). U-Net++ redesigned the skip connection to include a Dense-like structure in the skip connection (Zhou et al., 2018), and U-Net3++(Huang et al., 2020) changed the skip connection to a full-scale skip connection. Secondly, RA-UNet introduces a 3D hybrid residual attention-aware method for precise feature extraction in skipped connections (Jin et al., 2020). The third is a combination of encoder and decoder feature maps: An alternative extension to the classical skip connection was introduced in BCDU-Net with a bidirectional convolutional long-term-short-term memory (LSTM) module was added to the skip connection (Azad et al., 2019). Aligning with the second approach, we integrate Dual Attention Blocks into each skip connection layer, enhancing decoder feature extraction and thereby improving image segmentation accuracy.

2.3 The use of attentional mechanisms in medical images

Attention mechanisms are essential for directing model focus towards relevant features, thereby enhancing performance. In recent years, dual attention mechanisms have seen diverse applications across multiple fields. In scene segmentation, the Dual Attention Network (DANet) employs position and channel attention mechanisms to improve performance (Fu et al., 2019). A modularized DANs framework is presented that adeptly merges visual and textual attention mechanisms (Nam et al., 2017). This cohesive approach enables selective focus on pivotal features in both types of data, thereby improving task-specific performance. Additionally, the introduction of the Dual Attention Module

10.3389/fbioe.2024.1398237

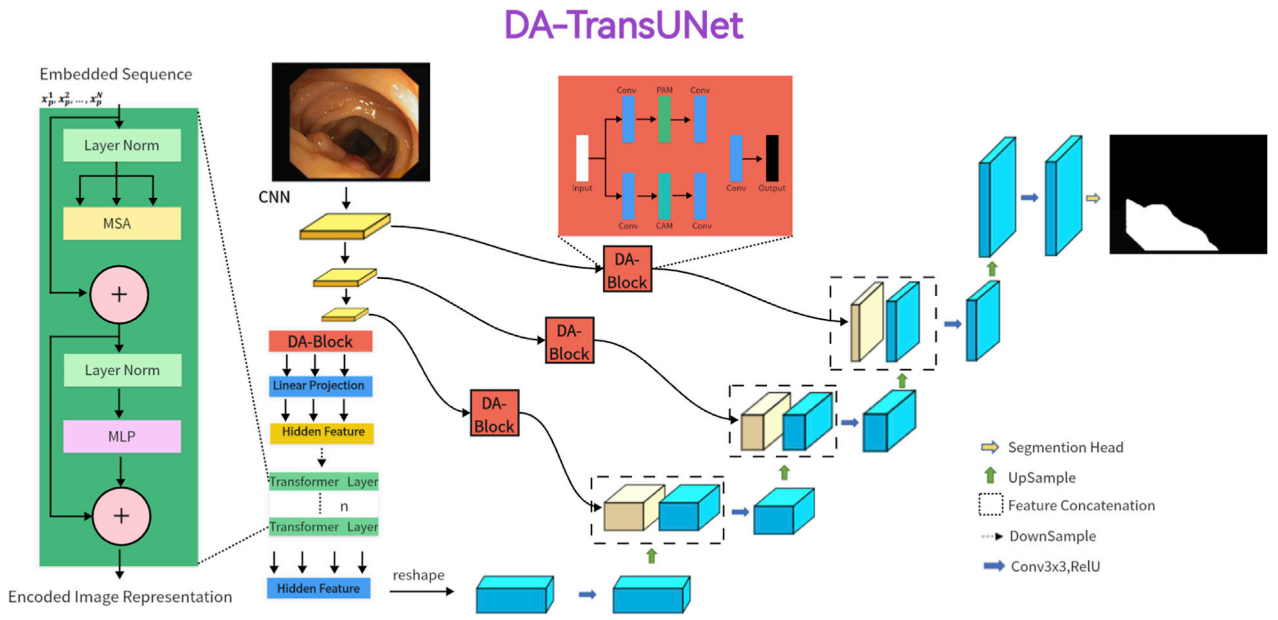


FIGURE 1
Illustration of the proposed dual attention transformer U-Net(DA-TransUNet). For the input medical images, we feed them into an encoder with transformer and Dual Attention Block (DA-Block). Then, the features of each of the three different scales are purified by DA-Block. Finally, the purified skip connections are fused with the decoder, which subsequently undergoes CNN-based up-sampling to restore the channel to the same resolution as the input image. In this way, the final image prediction result is obtained.

(DuATM) has been groundbreaking in the field of audio-visual event localization. This model excels at learning context-aware feature sequences and performing attention sequence comparisons in tandem, effectively incorporating auditory-oriented visual attention mechanisms (Si et al., 2018). Moreover, dual attention mechanisms have been applied to medical segmentation, yielding promising results (Shi et al., 2020). The Multilevel Dual Attention U-net for Polyp Segment combines dual attention and U-net in medical image segmentation (Cai et al., 2022). While significant progress has been made in medical image segmentation, there is still ample room for further research to explore the potential of position and channel attention mechanism in the field of medical image segmentation.

3 Methods

In the subsequent section, we propose the DA-TransUNet architecture, illustrated in Figure 1. We start with a comprehensive overview of the architecture. Next, we detailed the architecture's key components in the following order: the dual attention blocks (DA-Block), the encoder, the skip connections, and the decoder.

3.1 Overview of DA-TransUNet

In Figure 1, the architecture of DA-TransUNet is presented. The model comprises three core components: the encoder, the decoder, and the skip connections. In particular, the encoder fuses a conventional convolutional neural network (CNN) with a Transformer layer and is further enriched by the DA-Block, which are exclusively introduced in this model architecture. In contrast, the decoder primarily employs conventional convolutional mechanisms. For the optimization of skip connections, DA-Blocks serve as pivotal components within the DA-TransUNet architecture. DA-Blocks filter irrelevant information in skip connections, enhancing image reconstruction accuracy. In summary, in contrast to traditional convolutional approaches and the extensive use of Transformers, DA-TransUNet uniquely leverages DA-Blocks for the extraction and utilization of image-specific features of position and channel. This strategic incorporation significantly elevates the overall performance of the model.

Compared to traditional U-Net architectures, DA-TransUNet integrates the Transformer layer in the encoder to capture global dependencies, while the U-Net relies solely on convolutional layers for local feature extraction. Moreover, the inclusion of DA-Blocks in the encoder and skip connections sets DA-TransUNet apart from both U-Net and Transformer-based models. These DA-Blocks enable the extraction and utilization of image-specific position and channel features, enhancing the model's ability to capture fine-grained details crucial for medical image segmentation.

To elucidate the rationale behind our proposed DA-TransUNet model's design, it's imperative to consider the limitations and strengths of both U-Net architectures and Transformers in the context of feature extraction. While Transformers excel in global feature extraction through their self-attention mechanisms, they are inherently limited to unidirectional focus on positional attributes, thus neglecting multi-faceted feature perspectives. On the other hand, traditional U-Net architectures are proficient in local feature extraction but lack the capability for comprehensive global

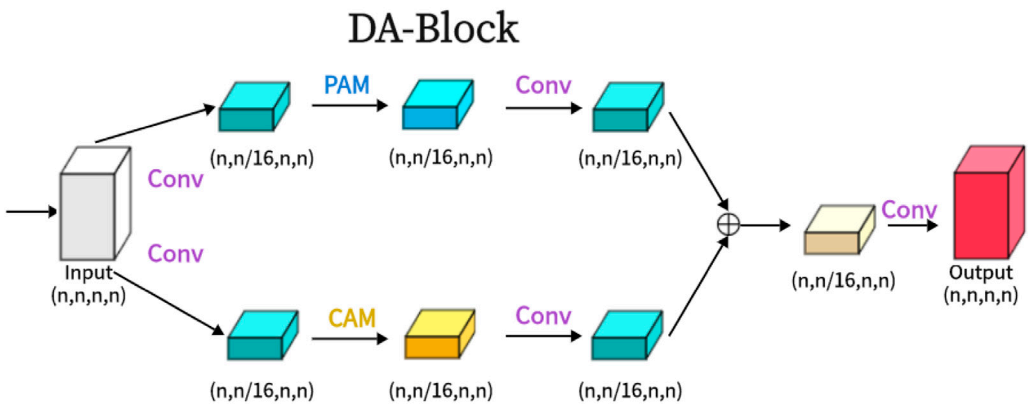


FIGURE 2
The proposed Dual Attention Block (DA-Block) is shown in the Figure. The same input feature map is input into two feature extraction layers, one is the position feature extraction block and the other is the channel feature extraction block, and finally, the two different features are fused to obtain the final DA-Block output.

contextualization. To address these constraints, we integrate DA-Blocks both preceding the Transformer layers and within the encoder-decoder skip connections. This achieves two goals: firstly, it refines the feature map input to the Transformer, enabling more nuanced and precise global feature extraction; secondly, the DA-Block in the skip connections optimize the transmitted features from the encoder, facilitating the decoder in reconstructing a more accurate feature map. Thus, our proposed architecture amalgamates the strengths and mitigates the weaknesses of both foundational technologies, resulting in a robust system capable of image-specific feature extraction.

3.2 Dual attention block (DA-Block)

As shown in the attached Figure 2, the Dual Attention Block (DA-Block) serves as a feature extraction module that integrates image-specific features of position and channel. This enables feature extraction tailored to the unique attributes of the image. Particularly in the context U-Net shaped architectures, the specialized feature extraction capabilities of the DA-Block are crucial. While Transformers are adept at using attention mechanisms to extract global features, they are not specifically tailored for image-specific attributes. In contrast, the DA-Block excels in both position-based and channel-based feature extraction, enabling a more detailed and accurate set of features to be obtained. Therefore, we incorporate it into the encoder and skip connections to enhance the model's segmentation performance. The DA-Block consists of two primary components: one featuring a Position Attention Module (PAM), and the other incorporating a Channel Attention Module (CAM), both borrowed from the Dual Attention Network for scene segmentation (Fu et al., 2019).

3.2.1 PAM (position attention module)

As shown in Figure 3, PAM captures spatial dependencies between any two positions of feature maps, updating specific features through a weighted sum of all position features. The weights are determined by the feature similarity between two positions. Therefore, PAM is effective at extracting meaningful spatial features.

PAM initially takes a local feature, denoted as A $\in R^{C\times H\times W}$ (C represents Channel, H represents, and W represents Width). We then feed A into a convolutional layer, resulting in three new feature maps, namely, B, C, and D, each of size $R^{C\times H\times W}$. Next, we reshape B and C to $R^{C\times N}$, where N = H× W denotes the number of pixels. We perform a matrix multiplication between the transpose of C and B and subsequently use a softmax layer to compute the spatial attention map S $\in R^{N\times N}$:

$$S_{ji} = \frac{\exp\left(B_i \cdot C_j\right)}{\sum_{i=1}^{N} \exp\left(B_i \cdot C_j\right)} \tag{1}$$

Here, S_{ji} measures the impact of the i-th position on the j-th position. We then reshape matrix D to $R^{C\times N}$. A matrix multiplication is performed between D and the transpose of S, followed by reshaping the result to $R^{C\times H\times W}$. Finally, we multiply it by a parameter α and perform an element-wise sum operation with the features A to obtain the final output E $\in R^{C\times H\times W}$:

$$E_j = \alpha \sum_{i=1}^{N} \left(S_{ji} D_i\right) + A_j \tag{2}$$

The weight α is initialized as 0 and is learned progressively. PAM has a strong capability to extract spatial features. It can be inferred from Eq. 2 that the resulting feature E at each position is a weighted sum of the features across all positions and original features, it possesses global contextual features and aggregates context based on the spatial attention map. This ensures effective extraction of position features while maintaining global contextual information.

3.2.2 CAM (channel attention module)

As shown in Figure 4, this is CAM, which excels in extracting channel features. Unlike PAM, we directly reshape the original feature A $\in R^{C\times H\times W}$ to $R^{C\times N}$, and then perform a matrix

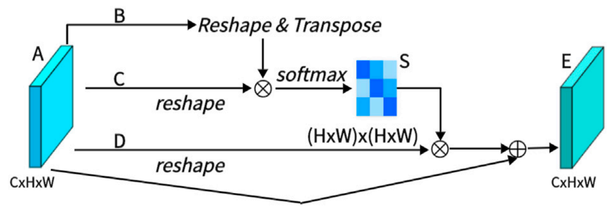


FIGURE 3
Architecture of position attention Mechanism (PAM).

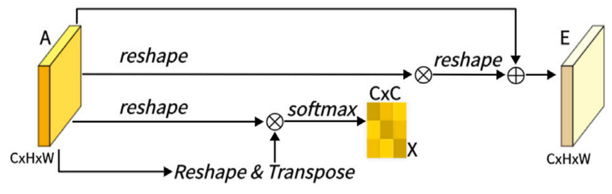


FIGURE 4
Architecture of channel attention Mechanism (CAM).

multiplication between A and its transpose. Subsequently, we apply a softmax layer to obtain the channel attention map X $\in R^{C\times C}$:

$$X_{ji} = \frac{\exp\left(A_i \cdot A_j\right)}{\sum_{i=1}^{C} \exp\left(A_i \cdot A_j\right)} \tag{3}$$

Here, x_{ji} measures the impact of the i-th channel on the j-th channel. Next, we perform a matrix multiplication between the transpose of X and A, reshaping the result to $R^{C\times H\times W}$. We then multiply the result by a scale parameter β and perform an element-wise sum operation with A to obtain the final output E $\in R^{C\times H\times W}$:

$$E_j = \beta \sum_{i=1}^{N} \left(X_{ji} A_i\right) + Aj \tag{4}$$

Like α, β is learned through training. Similar to PAM, during the extraction of channel features in CAM, the final feature for each channel is generated as a weighted sum of all channels and original features, thus endowing CAM with powerful channel feature extraction capabilities.

3.2.3 DA (dual attention module)

As shown in the Figure 2, we present the architecture of the Dual Attention Block (DA-Block). This architecture merges the robust position feature extraction capabilities of the Position Attention Module (PAM) with the channel feature extraction strengths of the Channel Attention Module (CAM). Furthermore, when coupled with the nuances of traditional convolutional methodologies, the DA-Block emerges with superior feature extraction capabilities. DA-Block consists of two components, the first one is dominated by PAM and the second one is dominated by CAM. The first component takes the input features and performs one convolution to scale the number of channels by one-sixteenth to get α^1. This convolution operation not only simplifies feature extraction by PAM but also helps to adjust the scale and dimension of features, making them more suitable for the subsequent attention mechanism computations. Following a PAM feature extraction and another convolution, $\widehat{\alpha^1}$ is obtained, which further refines the extracted features.

$$\alpha^1 = Conv(input) \tag{5}$$

$$\widehat{\alpha^1} = Conv\left(PAM\left(\alpha^1\right)\right) \tag{6}$$

The other component is the same, with the only difference being that the PAM block is replaced with a CAM with the following formula:

$$\alpha^2 = Conv(input) \tag{7}$$

$$\widehat{\alpha^2} = Conv\left(CAM\left(\alpha^2\right)\right) \tag{8}$$

After extracting $\widehat{\alpha^1}$ and $\widehat{\alpha^2}$ from the two layers of attention, the output is obtained by aggregating and summing the two layers of attention and recovering the number of channels in one convolution.

$$output = Conv\left(\widehat{\alpha^1} + \widehat{\alpha^2}\right) \tag{9}$$

To optimize the DA-Block for medical image segmentation, we fine-tuned the number of intermediate channels. This optimization allows the model to focus on the most critical features, enhancing its sensitivity to key information in the medical images. By adapting the DA-Block to the specific characteristics of medical images, we enable the model to better capture the fine-grained details necessary for accurate segmentation. This targeted optimization sets our approach apart from previous works, which often overlook the importance of tailoring attention mechanisms to the unique demands of medical image segmentation.

3.3 Encoder with transformer and dual attention

As illustrated in Figure 1, the encoder architecture consists of four key components: convolution blocks, DA-Block, embedding layers, and transformer layers. Of particular significance is the inclusion of the DA block before the Transformer layer. This design is aimed at performing specialized image processing on the post-convolution features, enhancing the Transformer's feature extraction for image content. While the Transformer architecture plays a crucial role in preserving global context, the DA block strengthens the Transformer's capability to capture image-specific features, enhancing its ability to capture global contextual information in the image. This approach effectively combines global features with image-specific spatial and channel characteristics.

The first component comprises the three convolutional blocks of the architecture of the U-Net and its diverse iterations, seamlessly integrating convolutional operations with downsampling processes. Each convolutional layer halves the size of the input feature map and doubles its dimension, a configuration empirically found to maximize feature expressiveness while maintaining computational efficiency. The second component uses DA-Block extract features at both positional and channel levels, enhancing the depth of feature representation while preserving the intrinsic characteristics of the input map. The third component is the embedding layer serves as a critical intermediary, enabling the requisite dimensional adaptation, a prelude to the subsequent Transformer strata. The fourth component integrates Transformer layers for enhanced global feature extraction, beyond the reach of traditional CNNs. Putting the above parts together, it works as follows: the input image traverses three consecutive convolutional blocks, systematically expanding the receptive field to encompass vital features. Subsequently, the DA-Block refines features through the application of both position-based and channel-based attention mechanisms. Following this, the remodeled features undergo a dimensionality transformation courtesy of the embedding stratum before they are channeled into the Transformer framework for the extraction of all-encompassing global features. This orchestrated progression safeguards the comprehensive retention of information across the continuum of successive convolutional layers. Ultimately, the Transformer-generated feature map is restructured and navigated through skip connection layers to feed into the decoder.

By combining convolutional neural networks, transformer architectures, and dual-attention mechanisms, the encoder configuration culminates in a robust capability for feature extraction, resulting in a symbiotic powerhouse of capabilities.

3.4 Skip-connections with dual attention

Similar to other U-structured models, we have also incorporated skip connections between the encoder and decoder to bridge the semantic gap that exists between them. To further minimize this semantic gap, we introduced dual-attention blocks (DA-Blocks), as depicted in Figure 1, in each of the three skip connection layers. This decision was based on our observation that traditional skip connections often transmit redundant features, which DA-Blocks effectively filter. Integrating DA-Blocks into the skip connections allows them to refine the sparsely encoded features from both positional and channel perspectives, extracting more valuable information while reducing redundancy. By doing so, DA-Blocks assist the decoder in more accurate feature map reconstruction. Moreover, the inclusion of DA-Blocks not only enhances the model's robustness but also effectively mitigates sensitivity to overfitting, contributing to the overall performance and generalization capability of the model.

3.5 Decoder

As depicted in Figure 1, the right half of the diagram corresponds to the decoder. The primary role of the decoder is to reconstruct the original feature map by utilizing features acquired from the encoder and those received through skip connections, employing operations like upsampling.

The decoder's components include feature fusion, a segmentation head, and three upsampling convolution blocks. The first component: feature fusion entails the integration of feature maps transmitted through skip connections with the existing feature maps, thereby assisting the decoder in faithfully reconstructing the original feature map. The second component: the segmentation head is responsible for restoring the final output feature map to its original dimensions. The third component: the three upsampling convolution blocks incrementally double the size of the input feature map in each step, effectively restoring the image's resolution.

Putting the above parts together, the workflow begins by passing the input image through convolution blocks and subsequently performing upsampling to augment the size of the feature maps. These feature maps undergo a twofold size increase while their dimensions are reduced by half. The features received through the skip connections are then fused, followed by continued upsampling and convolution. After three iterations of this process, the generated feature map undergoes one final round of upsampling and is accurately restored to its original size by the segmentation head.

Thanks to this architecture, the decoder demonstrates robust decoding capabilities, effectively revitalizing the original feature map using features from both the encoder and skip connections.

Furthermore, compared to other transformer-based approaches that extensively utilize transformer blocks throughout the architecture, such as Swin-Unet, DA-TransUNet achieves a more favorable balance between performance and computational efficiency. The judicious integration of DA-Blocks in the encoder and skip connections allows DA-TransUNet to enhance feature representation while maintaining a manageable computational footprint.

4 Experiments

To evaluate the proposed method, we performed experiments on Synapse (Landman et al., 2015), CVC-ClinicDB dataset (Bernal et al., 2015), Chest X-ray mask and label dataset (Candemir et al., 2013; Jaeger et al., 2013) Analysis, Kvasir SEG dataset (Jha et al., 2020), Kvasir-Instrument dataset (Jha et al., 2021), 2018ISIC-Task (Tschandl et al., 2018; Codella et al., 2019). The experimental results demonstrate that DA-TransUNet outperforms existing methods across all six datasets. In the following subsections, we first introduce the dataset and implementation details. Then show the results on each of the six datasets.

4.1 Datasets

4.1.1 Synapse

The Synapse dataset consists of 30 scans of eight abdominal organs. These eight organs include the left kidney, right kidney, aorta, spleen, gallbladder, liver, stomach and pancreas. There are a total of 3779 axially enhanced abdominal clinical CT images.

4.1.2 CVC—ClinicDB

CVC-ClinicDB is a database of frames extracted from colonoscopy videos, which is part of the Endoscopic Vision Challenge. This is a dataset of endoscopic colonoscopy frames for the detection of polyps. CVC-ClinicDB contains 612 still images from 29 different sequences. Each image has its associated manually annotated ground truth covering the polyp.

4.1.3 Chest Xray

Chest Xray Masks and Labels X-ray images and corresponding masks are provided. The X-rays were obtained from the

Montgomery County Department of Health and Human Services Tuberculosis Control Program, Montgomery County, Maryland, United States. The set of images contains 80 anterior and posterior X-rays, of which 58 X-rays are normal and 1702 X-rays are abnormal with evidence of tuberculosis. All images have been de-identified and presented in DICOM format. The set contains a variety of abnormalities, including exudates and corneal morphology. It contains 138 posterior-anterior radiographs, of which 80 radiographs were normal and 58 radiographs showed abnormal manifestations of tuberculosis.

4.1.4 Kvasir SEG

Kvasir SEG is an open-access dataset of gastrointestinal polyp images and corresponding segmentation masks, manually annotated and verified by an experienced gastroenterologist. It contains 1000 polyp images and their corresponding groudtruth, the resolution of the images contained in Kvasir-SEG varies from 332 × 487 to 1920 × 1072 pixels, and the file format is jpg.

4.1.5 Kvasir-instrument

Kvasir-Instrument a gastrointestinal instrument Dataset. It contains 590 endoscopic tool images and their groud truth mask, the resolution of the image in the dataset varies from 720 × 576 to 1280 × 1024, which consists of 590 annotated frames comprising of GI procedure tools such as snares, balloons, biopsy forceps, etc. The file format is jpg.

4.1.6 2018ISIC-task

The dataset used in the 2018 ISIC Challenge addresses the challenges of skin diseases. It comprises a total of 2512 images, with a file format of JPG. The images of lesions were obtained using various dermatoscopic techniques from different anatomical sites (excluding mucous membranes and nails). These images are sourced from historical samples of patients undergoing skin cancer screening at multiple institutions. Each lesion image contains only a primary lesion.

4.2 Implementation settings

4.2.1 Baselines

In our endeavor to innovate in the field of medical image segmentation, we benchmark our proposed model against an array of highly-regarded baselines, including the U-net, UNet++, DA-Unet, Attention U-net, and TransUNet. The U-net has been a foundational model in biomedical image segmentation (Ronneberger et al., 2015). Unet++ brings added sophistication with its implementation of intermediate layers (Zhou et al., 2018). The DA-Unet goes a step further by integrating dual attention blocks, amplifying the richness of features extracted (Cai et al., 2022). The Attention U-net employs an attention mechanism for improved feature map weighting (Oktay et al., 2018), and finally, the TransUNet deploys a transformer architecture, setting a new bar in segmentation precision (Chen et al., 2021). Through this comprehensive comparison with these eminent baselines, we aim to highlight the unique strengths and expansive potential applications of our proposed model. Additionally, we benchmarked our model against advanced state-of-the-art algorithms. UCTansNet allocates skip connections through the attention module in the traditional U-net model (Wang et al., 2022a). TransNorm integrates the Transformer module into the encoder and skip connections of standard U-Net (Azad et al., 2022b). A novel Transformer module was designed and a model named MIM was built with it (Wang et al., 2022b). By extensively comparing our model with current state-of-the-art solutions, we intend to showcase its superior segmentation performance.

4.2.2 Implementation details

We implemented DA-TransUNet using the PyTorch framework and trained it on a single NVIDIA RTX 3090 GPU (Paszke et al., 2019). The model was trained with an image resolution of 256 × 256 and a patch size of 16. We employed the Adam optimizer, configured with a learning rate of 1e-3, momentum of 0.9, and weight decay of 1e-4. All models were trained for 500 epochs unless stated otherwise. In order to ensure the convergence of the indicators, but due to different data set sizes, we used 50 epochs for training on the two data sets, Chest Xray Masks and Labels and ISIC 2018-Task.

During the training phase on five datasets, including CVC-ClinicDB, the proposed DA-TransUNet model is trained in an end-to-end manner. Its objective function consists of a weighted binary cross-entropy loss function (BCE) and a Dice coefficient loss function. To facilitate training, the final loss function, termed "Loss," is formulated as follows:

$$\text{Loss} = \frac{1}{2} \times \text{BCE} + \frac{1}{2} \times \text{DiceLoss} \tag{10}$$

To ensure a fair evaluation of the Synapse dataset, we utilized the pre-trained model "R50-ViT" with input resolution and patch size set to 224 × 224 and 16, respectively. We trained the model using the SGD optimizer, setting the learning rate to 0.01, momentum of 0.9, and weight decay of 1e-4. The default batch size was set to 24. The loss function employed for the Synapse dataset is defined as follows:

$$\text{Loss} = \frac{1}{2} \times \text{Cross} - \text{Entropy Loss} + \frac{1}{2} \times \text{DiceLoss} \tag{11}$$

This loss function balances the contributions of cross-entropy and Dice losses, ensuring impartial evaluation during testing on the Synapse dataset.

When using the datasets, we use a 3 to 1 ratio, where 75% is the training set and 25% is the test set, to ensure adequacy of training.

4.2.3 Model evaluation

In evaluating the performance of DA-TransUNet, we utilize a comprehensive set of metrics including Intersection over Union (IoU), Dice Coefficient (DSC), and Hausdorff Distance (HD). These metrics are industry standards in computer vision and medical image segmentation, providing a multifaceted assessment of the model's accuracy, precision, and robustness.

The choice of these metrics is based on their complementary nature and ability to capture different aspects of segmentation quality. IoU and DSC measure the overlap between the predicted and ground truth segmentation masks, providing a global assessment of the model's ability to accurately identify and delineate target structures. HD, on the other hand, captures the maximum distance between the predicted and ground truth segmentation boundaries, ensuring that the predicted segmentation closely adheres to the true boundaries of the target structures, even in the presence of small segmentation errors or irregularities.

IOU (Intersection over Union) is one of the commonly used metrics to evaluate the performance of computer vision tasks such as object detection, image segmentation and instance segmentation. It measures the degree of overlap between the predicted region of the model and the actual target region, which helps us to understand the accuracy and precision of the model. In target detection tasks, IOU is usually used to determine the degree of overlap between the predicted bounding box (Bounding Box) and the real bounding box. In image segmentation and instance segmentation tasks, IOU is used to evaluate the degree of overlap between the predicted region and the ground truth segmentation region.

$$IOU = \frac{TP}{FP + TP + FN} \tag{12}$$

The Dice coefficient (also known as the Sørensen-Dice coefficient, F1-score, DSC) is a measure of model performance in image segmentation tasks, and is particularly useful for dealing with class imbalance problems. It measures the degree of overlap between the predicted results and the ground truth segmentation results, and is particularly effective when dealing with segmentation of objects with unclear boundaries. The Dice coefficient is commonly used as a measure of the model's accuracy on the target region in image segmentation tasks, and is particularly suitable for dealing with relatively small or uneven target regions.

$$\mathrm{Dice}(P,T) = \frac{|P_1 \cap T_1|}{|P_1| + |T_1|} \Leftrightarrow \mathrm{Dice} = \frac{2|T \cap P|}{|F| + |P|} \tag{13}$$

Hausdorff Distance (HD) is a distance measure for measuring the similarity between two sets and is commonly used to evaluate the performance of models in image segmentation tasks. It is particularly useful in the field of medical image segmentation to quantify the difference between predicted and true segmentations. The computation of Hausdorff distance captures the maximum difference between the true segmentation result and the predicted segmentation result, and is particularly suitable for evaluating the performance of segmentation models in boundary regions.

$$H(A,B) = \max\{\max_{a\in A}\min_{b\in B}\|a-b\|, \max_{b\in B}\min_{a\in A}\|b-a\|\} \tag{14}$$

We evaluate using both Dice and HD in the Synapse dataset and both Dice and IOU in other datasets.

4.3 Comparison to the state-of-the-art methods

4.3.1 Segmentation performance and comparison

We have chosen U-net (Ronneberger et al., 2015), Res-Unet (Diakogiannis et al., 2020), TransUNet (Chen et al., 2021), U-Net++(Zhou et al., 2018), Att-Unet (Oktay et al., 2018), TransNorm (Azad et al., 2022b), UCTransNet (Wang et al., 2022a), MultiResUNet (Ibtehaz and Rahman, 2020), swin-unet (Cao et al., 2022) and MIM (Wang et al., 2022b) to compare with our DA-TransUNet, and the experimental data are tabulated below.

In order to demonstrate the superiority of the DA-TransUNet model proposed in this paper, we conducted the main experiments using the Synapse dataset and compared it with its 11 state-of-the-art models (SOTA) (see Table1).

As shown in the Figure 5, we can see that the average DSC and average HD evaluation criteria are 79.80% and 23.48 mm, respectively, which are improved by 2.32% and 8.21 mm, respectively, compared with TransUNet, which indicates that our DA-TransUNet has better segmentation ability than TransUNer in terms of overall segmentation results and organ edge prediction. As shown in the Figure 6, on the other hand, we can see that DSC has the highest value of our model. Although HD is higher than Swin-Unet, it is still an improvement compared to several newer models and TransUNet. The segmentation time for an image is 35.98 ms for our DA-TransUNet and 33.58 ms for TransUNet, which indicates that there is not much difference in the segmentation speed between the two models, but our DA-TransUNet has better segmentation results. In the segmentation results of 8 organs, DA-TransUNet outperforms TransUNet by 2.14%, 3.43%, 0.48%, 3.45%, and 4.11% for the five datasets of Gallbladder, right kidney, liver, spleen, and stomach, respectively. The segmentation rate for the pancreas is notably higher at 5.73%. In a comparative evaluation across six distinct organs, DA-TransUNet demonstrates superior segmentation capabilities relative to TransUNet. Nevertheless, it exhibits a marginal decrement in the segmentation accuracy for the aorta and left kidney by 0.69% and 0.17%, respectively. The model achieves the best segmentation rates for the right kidney, liver, pancreas, and stomach, indicating superior feature learning capabilities on these organs.

To further confirm the better segmentation of our model compared to TransUNet, we visualized the segmentation plots of TransUNet and DA-TransUNet (see Figure 5). From the yellow and purple parts in the first column, we can see that our segmentation effect is obviously better than that of TransUNet; from the second column, the extension of purple is better than that of TransUNet, and there is no vacancy in the blue part; from the third column, there is a semicircle in the yellow part, and the vacancy in red is smaller than that of TransUNet, etc. It is evident that DA-TransUNet outperforms TransUNet in segmentation quality. In summary, DA-TransUNet significantly surpasses TransUNet in segmenting the left kidney, right kidney, spleen, stomach, and pancreas. It also offers superior visualization performance in image segmentation.

We simultaneously took DA-TransUNet in five datasets, CVC-ClinicDB, Chest Xray Masks and Labels, ISIC2018-Task, kvasir-instrument, and kvasir-seg, and compared it with some classical models (see Table 2). In the table, the values of IOU and Dice of DA-TransUNet are higher than TransUNet in all five datasets, CVC-ClinicDB, Chest Xray Masks and Labels, ISIC2018-Task, kvasir-instrument, and kvasir-seg. Also DA-TransUNet has the best dataset segmentation in four of the five datasets. As seen in the table, our DA-TransUNet has more excellent feature learning and image segmentation capabilities.

We also show the results of image segmentation visualization of DA-TransUNet in these five datasets, and we also show the results of the comparison models for the comparison. The visualization results for Chest X-ray Masks and Labels, Kvasir-Seg, Kvasir-Instrument, ISIC2018-Task, and CVC-ClinicDB datasets are presented in

TABLE 1 Experimental results on the Synapse dataset.

Model	Year	DSC ↑ (%)	HD ↓	Aorta	Gallbladder	Kidney(L)	Kidney(R)	Liver	Pancreas	Spleen	Stomach
U-net (Ronneberger et al., 2015)	2015	76.85	39.70	89.07	**69.72**	77.77	68.6	93.43	53.98	86.67	75.58
U-Net++(Zhou et al., 2018)	2018	76.91	36.93	88.19	68.89	81.76	75.27	93.01	58.20	83.44	70.52
Residual U-Net (Diakogiannis et al., 2020)	2018	76.95	38.44	87.06	66.05	**83.43**	76.83	93.99	51.86	85.25	70.13
Att-Unet (Oktay et al., 2018)	2018	77.77	36.02	**89.55**	68.88	77.98	71.11	93.57	58.04	87.30	75.75
MultiResUNet (Ibtehaz and Rahman, 2020)	2020	77.42	36.84	87.73	65.67	82.08	70.43	93.49	60.09	85.23	75.66
TransUNet (Chen et al., 2021)	2021	77.48	31.69	87.23	63.13	81.87	77.02	94.08	55.86	85.08	75.62
UCTransNet (Wang et al., 2022a)	2022	78.23	26.75	84.25	64.65	82.35	77.65	94.36	58.18	84.74	79.66
TransNorm (Azad et al., 2022b)	2022	78.40	30.25	86.23	65.1	82.18	78.63	94.22	55.34	89.50	76.01
MIM(Wang et al., 2022b)	2022	78.59	26.59	87.92	64.99	81.47	77.29	93.06	59.46	87.75	76.81
swin-unet (Cao et al., 2022)	2022	79.13	**21.55**	85.47	66.53	83.28	79.61	94.29	56.58	**90.66**	76.60
DA-TransUNet(Ours)	2023	**79.80**	23.48	86.54	65.27	81.70	**80.45**	**94.57**	**61.62**	88.53	**79.73**
Average Relative Improvement	-	**2.03**	**−9.00**	**−0.73%**	**−1.09%**	**0.28%**	**5.21%**	**0.82%**	**4.86%**	**1.97%**	**4.5%**

The bold values indicate the best performance among all the methods compared in each respective evaluation metric. Specifically, for each row in a table, the bold number represents the method that achieves the highest score or lowest error on that particular metric, demonstrating its superior performance relative to the other approaches.

Figure 7, Figure 8, Figure 9, Figure 10, and Figure 11, respectively. In the Figure, it can be seen that the segmentation effect of DA-TransUNet has a good performance. Firstly, DA-TransUNet has better segmentation results than TransUNet. In addition, compared with the four classical models of U-net, Unet++, Attn-Unet, and Res-Unet, DA-TransUNet has a certain improvement. It can be seen that the effectiveness of DA-TransUNet for model segmentation is not only confirmed in the Synapse dataset, but also in the five datasets (CVC-ClinicDB, Chest Xray Masks and Labels, ISIC2018-Task, kvasir-instrument, kvasir-seg). We further establish that DA-TransUNet excels in both 3D and 2D medical image segmentation.

4.3.2 Computational complexity and efficiency

The integration of DA-Blocks in the encoder and skip connections introduces additional computational overhead compared to the standard TransUNet architecture. Let the input feature map have a spatial resolution of $H \times W$ and C channels. The computational complexity of the Position Attention Module (PAM) is $\mathcal{O}(H^2W^2C)$, while the Channel Attention Module (CAM) has a complexity of $\mathcal{O}(C^2HW)$. As the DA-Block consists of both PAM and CAM, its overall computational complexity is $\mathcal{O}(H^2W^2C + C^2HW)$. However, it is worth noting that the DA-Block itself is not computationally intensive, as it only involves simple matrix multiplications and element-wise operations.

Table 3 compares the number of parameters, Dice Similarity Coefficient (DSC), and Hausdorff Distance (HD) between DA-TransUNet and TransUNet. The incorporation of DA-Blocks leads to a modest increase of 2.54% in the number of parameters compared to TransUNet. This incremental increase in parameters is justifiable considering the substantial performance gains achieved by DA-TransUNet, as demonstrated in our experimental results (Section 4). DA-TransUNet achieves an average improvement of 2.99% in DSC and 25.9% in HD compared to TransUNet. The strategic placement of DA-Blocks allows for efficient feature refinement while maintaining a reasonable model size.

4.4 Ablation study

We conducted ablation experiments on the DA-TransUNet model using the Synapse dataset to discuss the effects of different factors on model performance. Specifically, it includes: 1) DA-Block in Encoder. 2) DA-Block in Skip Connection.

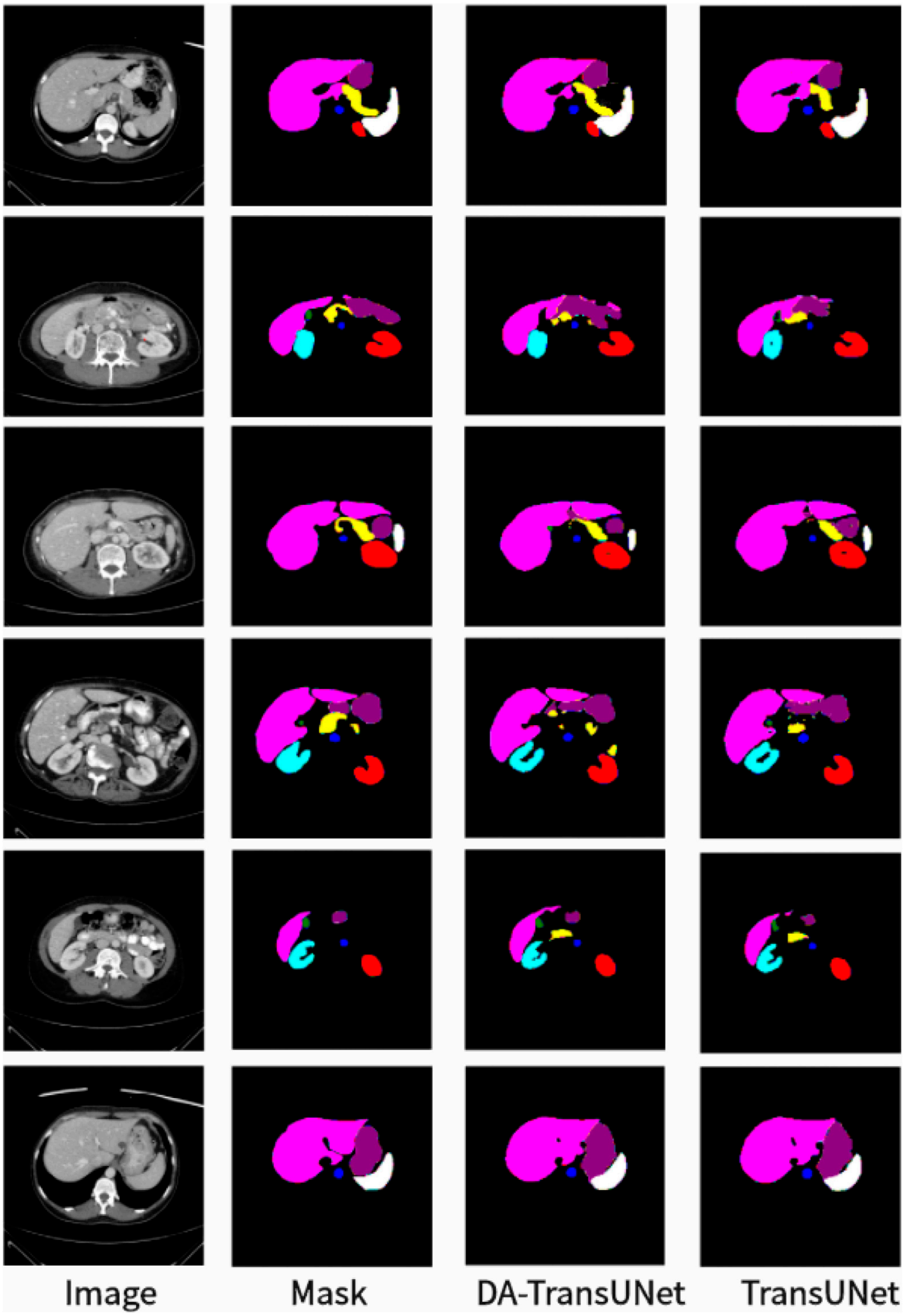


FIGURE 5
Segmentation results of TransUNet and DA-TransUNet on the Synapse dataset.

4.4.1 Effect of the DA-Block in encoder and skip connection

In this research (see Table 4), we conducted experiments to assess the impact of integrating DA-Blocks into the encoder and skip connections on the model's segmentation performance. To be specific, we introduced DA-Blocks into each layer of the skip connections. The results demonstrated an improvement: the DSC baseline saw an increase from 77.48% to 78.28%, HD index dropped from 31.69 mm to 29.09 mm. This indicates that the addition of DA-Blocks at each skip connection layer provided the decoder with more refined features, mitigating feature loss during the upsampling process, thereby reducing the risk of overfitting and enhancing model stability. Furthermore, incorporating DA-Blocks into the encoder before the Transformer yielded an enhancement, with

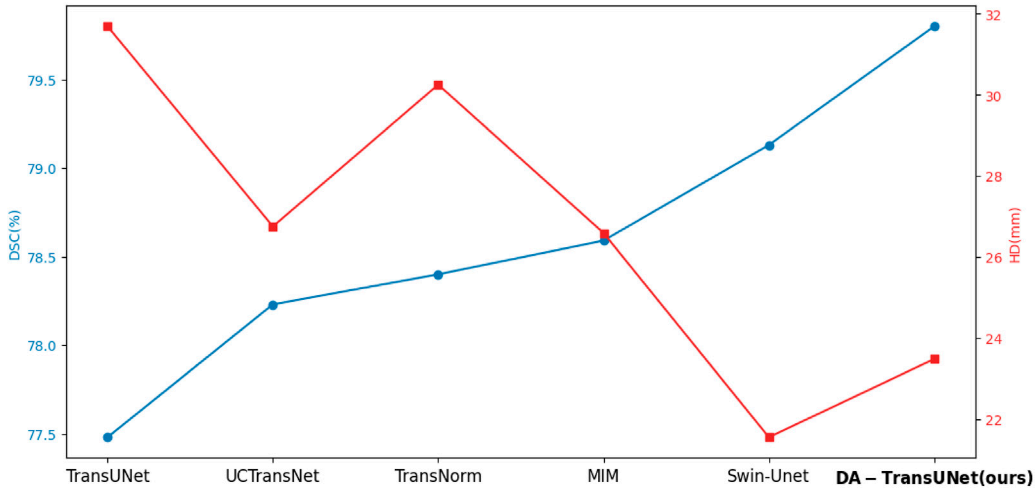


FIGURE 6
Line chart of DSC and HD values of several advanced models in the Synapse dataset.

TABLE 2 Experimental results of datasets (CVC-ClinicDB, Chest Xray Masks and Labels, ISIC2018-Task, kvasir-instrument, kvasir-seg).

	CVC-ClinicDB		Chest xray masks and labels		ISIC2018-task		Kvasir-instrument		Kvasir-seg	
	Iou ↑	Dice ↑	Iou ↑	Dice ↑	Iou ↑	Dice ↑	Iou ↑	Dice ↑	Iou ↑	Dice ↑
U-net (Ronneberger et al., 2015)	0.7821	0.8693	0.9303	0.9511	0.8114	0.8722	0.8957	0.9358	0.8012	0.8822
Attn-Unet (Oktay et al., 2018)	0.7935	0.8741	0.9274	0.9503	0.8151	0.876	0.8949	0.9359	0.7801	0.8661
Unet++(Zhou et al., 2018)	0.7847	0.8714	0.9289	0.9505	0.8133	0.873	**0.8995**	**0.9389**	0.7767	0.8657
ResUNet (Diakogiannis et al., 2020)	0.5902	0.7422	0.9262	0.9505	0.7651	0.8332	0.8572	0.9141	0.6604	0.7785
TransUNet (Chen et al., 2021)	0.8163	0.8901	0.9301	0.9535	0.8263	0.8878	0.8926	0.9363	0.8003	0.8791
DA-TransUNet(Ours)	**0.8251**	**0.8947**	**0.9317**	**0.9538**	**0.8278**	**0.8888**	0.8973	0.9381	**0.8102**	**0.8847**

The bold values indicate the best performance among all the methods compared in each respective evaluation metric. Specifically, for each row in a table, the bold number represents the method that achieves the highest score or lowest error on that particular metric, demonstrating its superior performance relative to the other approaches.

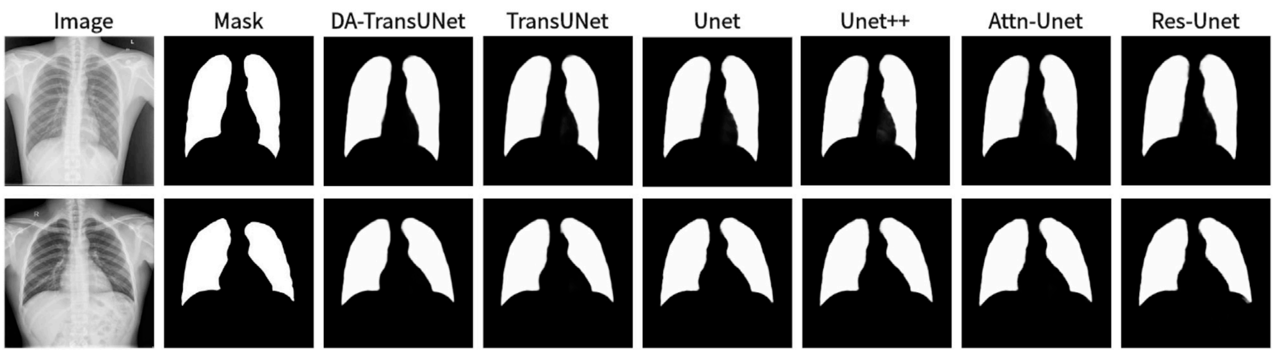


FIGURE 7
Comparison of qualitative results between DA-TransUNet and existing models on the task of segmenting Chest X-ray Masks and Labels X-ray datasets.

the DSC baseline increasing from 77.48% to 78.87%, even though the HD metric decreased from 31.69 mm to 27.71 mm. In conclusion, based on the findings presented in Table 4, we can assert that the inclusion of DA-Blocks both before the Transformer layer and within the skip connections effectively boosts medical image segmentation capabilities.

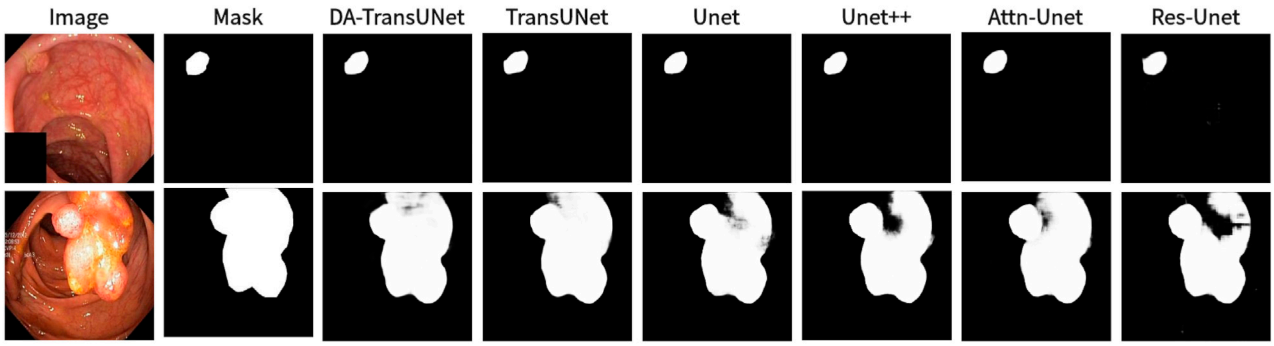


FIGURE 8
Comparison of qualitative results between DA-TransUNet and existing models on the task of segmenting Kvasir-Seg datasets.

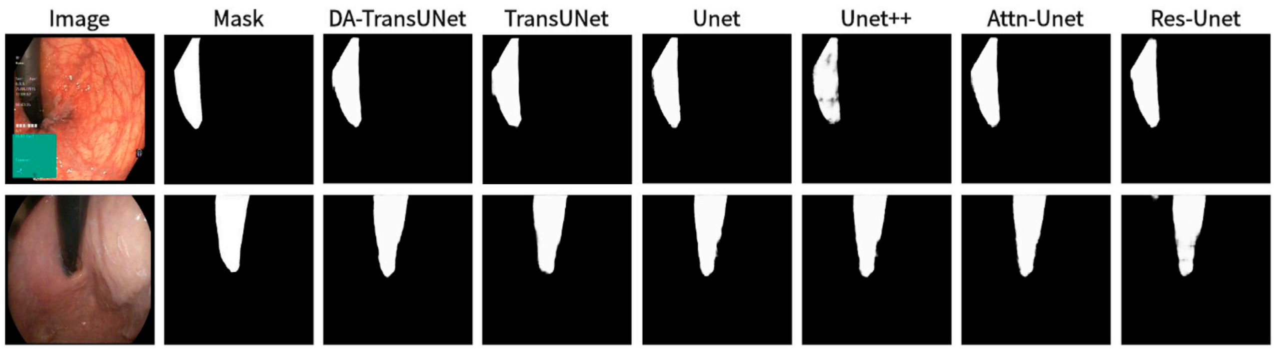


FIGURE 9
Comparison of qualitative results between DA-TransUNet and existing models on the task of segmenting Kavsir-Instrument datasets.

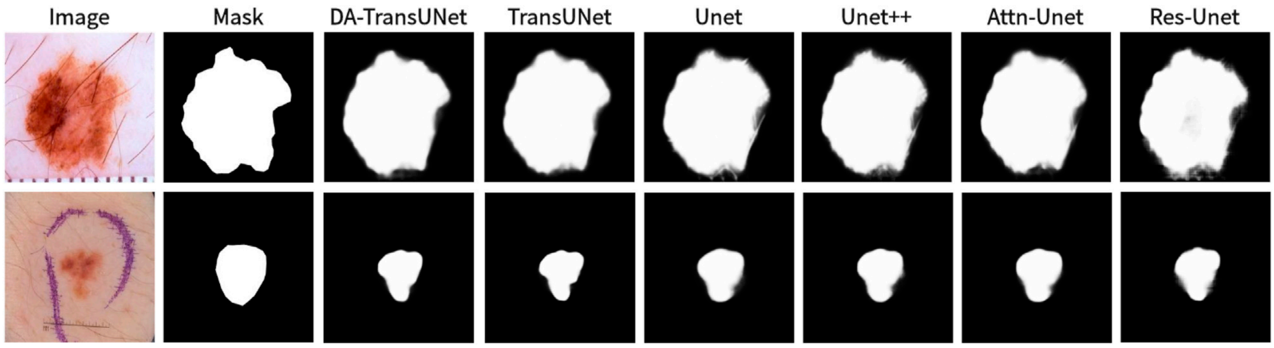


FIGURE 10
Comparison of qualitative results between DA-TransUNet and existing models on the task of segmenting 2018ISIC-Task datasets.

4.4.2 Effect of adding DA-Blocks to skip connections in different layers

Building on the quantitative results from Table 5, we experimented with various configurations of DA-Block placement across three different layers of skip connections to identify the optimal architectural layout for enhancing the model's performance. Specifically, when DA-Blocks were added to just the first layer, the DSC metric improved to 79.36% from a baseline of 78.87%, and the HD metric decreased to 25.80 mm from 27.71 mm. Adding DA-Blocks to the second and third layers resulted in some progress. When DA-Blocks were integrated across all layers, there was an improvement, reflected by a DSC of 79.80% and a HD of 23.48 mm. In contrast to traditional architectures where skip connections indiscriminately pass features from the encoder to the decoder, our approach with DA-Blocks selectively improves feature quality at each layer. The results, as corroborated by Table 5, reveal that introducing DA-Blocks to even

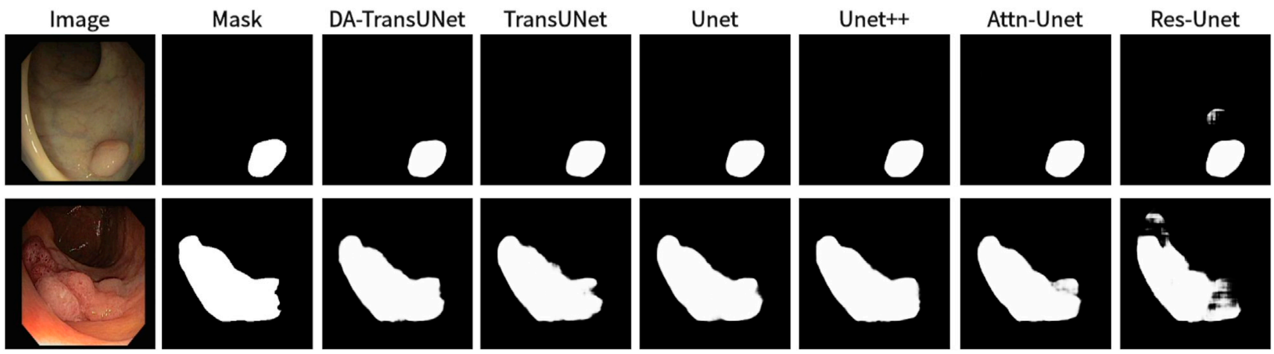


FIGURE 11
Comparison of qualitative results between DA-TransUNet and existing models on the task of segmenting CVC-ClinicDB datasets.

TABLE 3 Comparison of model parameters and performance between DA-TransUNet and TransUNet.

Model	Params	Params increase	DSC improvement	HD improvement
TransUNet	105,276,066	-	-	-
DA-TransUNet	**107,950,840**	**2.54%**	**2.99%**	**25.9%**

The bold values indicate the best performance among all the methods compared in each respective evaluation metric. Specifically, for each row in a table, the bold number represents the method that achieves the highest score or lowest error on that particular metric, demonstrating its superior performance relative to the other approaches.

TABLE 4 Effects of combinatorial placement of DA-Blocks in the encoder and through skip connections on performance metrics.

	Encoder with DA	Skip with DA	DSC ↑	HD ↓
DA-TransUNet			77.48	31.69
DA-TransUNet		√	78.28	29.09
DA-TransUNet	√		78.87	27.71
DA-TransUNet	√	√	**79.80**	**23.48**

The bold values indicate the best performance among all the methods compared in each respective evaluation metric. Specifically, for each row in a table, the bold number represents the method that achieves the highest score or lowest error on that particular metric, demonstrating its superior performance relative to the other approaches.

TABLE 5 Effects of incorporating DA-Block in the encoder and skip connections at different layers on performance metrics.

	1st layer	2nd layer	3rd layer	DSC ↑	HD ↓
DA-TransUNet				78.87	27.71
DA-TransUNet	√			79.36	25.80
DA-TransUNet		√		78.65	23.43
DA-TransUNet			√	79.49	30.71
DA-TransUNet	√	√	√	**79.80**	**23.48**

The bold values indicate the best performance among all the methods compared in each respective evaluation metric. Specifically, for each row in a table, the bold number represents the method that achieves the highest score or lowest error on that particular metric, demonstrating its superior performance relative to the other approaches.

a single layer enhances performance, and the greatest gains are observed when applied across all layers. This indicates the effectiveness of integrating DA-Blocks within skip connections for enhancing both feature extraction and medical image segmentation. Therefore, the table clearly supports the idea that layer-wise inclusion of DA-Blocks in skip connections is an effective strategy for enhancing medical image segmentation.

4.4.3 Effect of the number of intermediate channels in DA-Block

Based on the results shown in the Table 6, we conducted a discussion regarding the size of the intermediate layer in the DA-Block, which demonstrates the effectiveness of convolutional layers from an experimental perspective. The original DA-Block had an intermediate layer size that is one-fourth of the input layer size. However, since its

TABLE 6 Effect of the number of intermediate channels in DA-Block.

	1	2	4	8	16	32	DSC ↑	HD ↓
DA-TransUNet	√						78.55	28.22
DA-TransUNet		√					79.35	23.77
DA-TransUNet			√				79.71	25.90
DA-TransUNet				√			79.35	25.66
DA-TransUNet					√		**79.80**	**23.48**
DA-TransUNet						√	79.71	24.45

The bold values indicate the best performance among all the methods compared in each respective evaluation metric. Specifically, for each row in a table, the bold number represents the method that achieves the highest score or lowest error on that particular metric, demonstrating its superior performance relative to the other approaches.

TABLE 7 Statistical analysis of DSC improvements and model performance.

Model	Mean DSC ± SD	95% CI for DSC
DA-TransUNet	79.80 ± 5.01	[74.79, 84.81]
TransUNet	75.84 ± 6.77	[69.06, 82.61]

Comparison of DSC improvements achieved by DA-TransUNet and TransUNet relative to U-net			
Metric	Mean difference	95% CI for difference	t-Test *p*-value
Improvement	3.96	[0.40, 7.53]	0.032

intended application is for road scene segmentation and not specifically tailored for medical image segmentation, we deemed that setting the intermediate layer size to one-fourth of the input layer size might not be suitable for the medical image segmentation domain. As seen in the graph, when we set the intermediate layer size to be the same as the input size, the evaluation results show a DSC of 78.55% and HD of 28.22 mm. In the related research DANet (Fu et al., 2019), where the intermediate layer was set to one-fourth of the input layer, the DSC result was 79.71%, and HD was 25.90 mm. However, when we further reduced the size of the intermediate layer to one-sixteenth of the input layer size, we observed an improvement in DSC to 79.80%, and HD decreased further to 23.48 mm. It is evident that setting the intermediate layer to one-sixteenth of the input layer size is more suitable for medical image segmentation tasks. The reduction in the intermediate layer size can help the model mitigate the risk of overfitting, optimize computational resources, and, given the precision requirements of medical image segmentation tasks, enable the model to focus more on selecting the most crucial features, thereby enhancing sensitivity to critical information for the task.

5 Discussion

In this present study, we have discovered promising outcomes from the integration of DA-Blocks with the Transformer and their combination with skip-connections. Encouraging results were consistently achieved across all six experimental datasets.

5.1 Statistical validation of the improvements by DA-TransUNet

To enhance the credibility of our results and further validate the superiority of DA-TransUNet, We evaluated the performance of the models discussed in the Experiment Section 4 (U-Net, TransUNet, and DA-TransUNet) on 12 subsets of the Synapse dataset, constituting 40% of the total data, and obtained their Dice Similarity Coefficients (DSC). It is important to note that both DA-TransUNet and TransUNet are based on the U-Net architecture, which serves as the baseline model. Therefore, using U-Net as the benchmark to assess whether the improvements of DA-TransUNet over TransUNet are significant is a valid approach.

We first assessed the normality of the DSC improvement values for both DA-TransUNet and TransUNet relative to U-Net using the Shapiro-Wilk test. The results showed *p*-values of 0.36 and 0.82 for the improvements of DA-TransUNet and TransUNet, respectively. Since both *p*-values are greater than 0.05, we cannot reject the null hypothesis of normality. This indicates that the DSC improvement values for both DA-TransUNet and TransUNet relative to U-Net can be considered approximately normally distributed. We then performed a paired *t*-test to compare the significance of the improvements. As shown in Table 7, the test yielded a t-statistic of 2.45 and a *p*-value of 0.032, demonstrating a significant difference between the improvements achieved by DA-TransUNet and TransUNet.

Moreover, to further quantify the superiority of DA-TransUNet over TransUNet, we calculated the 95% confidence interval for the difference in improvements between DA-TransUNet and TransUNet. The results showed that the mean difference was 3.96, with a standard deviation of 5.61, and the confidence interval was [0.40, 7.53]. This means that, at a 95% confidence level, the magnitude of the difference in DSC improvements between DA-TransUNet and TransUNet lies between 0.40 and 7.53.

To provide a comprehensive overview of the models' performance, we calculated the 95% confidence intervals for their DSC scores. DA-TransUNet achieved a mean DSC of 79.80 ± 5.01, with a confidence interval of [74.79, 84.81], while TransUNet achieved a mean DSC of 75.84 ± 6.77, with a confidence interval of [69.06, 82.61]. These results, summarized in Table 7, suggest that DA-TransUNet not only achieves higher average performance but also exhibits more consistent results compared to TransUNet.

The statistical analysis, confidence intervals, and the quantification of the relative improvement provide strong evidence for the superiority of DA-TransUNet over TransUNet in the task of medical image segmentation. These results highlight the effectiveness of our proposed approach and its potential to advance the field of medical image analysis.

5.2 Enhancing feature extraction and segmentation with DA-Blocks

To start with, drawing from empirical results in Table 4, it is demonstrated that the integration of DA-Block within the encoder significantly enhances the feature extraction capabilities as well as its segmentation performance. In the landscape of computer vision, Vision Transformer (ViT) has been lauded for its robust global feature extraction capabilities (Dosovitskiy et al., 2020). However, its falls short in specialized tasks like medical image segmentation, where attention to image-specific features is crucial. To remedy this, in DA-TransUNet we strategically place DA-Blocks ahead of the Transformer module. These DA-Blocks are tailored to first extract and filter image-specific features, such as spatial positioning and channel attributes. Following this initial feature refinement, the processed data is then fed into the Transformer for enhanced global feature extraction. This approach results in significantly improved feature learning and segmentation performance. In summary, the strategic placement of DA-Blocks prior to the Transformer layer constitutes a pioneering approach that significantly elevates both feature extraction efficacy and medical image segmentation precision.

Morever, building on empirical data in Table 5, our integration of DA-Blocks with skip connections significantly improves semantic continuity and the decoder's ability to reconstruct accurate feature maps. While traditional U-Net architectures (Ronneberger et al., 2015) utilize skip connections to bridge the semantic gap between encoder and decoder, our novel incorporation of Dual Attention Blocks within the skip-connection layers yields promising results. By incorporating DA-Blocks across skip-connection layers, we focus on relevant features and filter out extraneous information, making the image reconstruction process more efficient and accurate. In summary, the strategic inclusion of DA-Blocks in skip connections represents a groundbreaking approach that not only enhances feature extraction but also improves the model's performance in medical image segmentation.

Lastly, our extensive evaluation across six diverse medical image segmentation datasets demonstrates the effectiveness and generalizability of the DA-TransUNet. The consistent improvements over state-of-the-art methods (Table 1) highlight the impact of our targeted integration of the DA-Block. Moreover, the ablation studies (4.4) provide valuable insights into the individual contributions of the DA-Block in different components of the architecture. These findings not only underscore the novelty of our approach but also shed light on the importance of strategically integrating attention mechanisms for enhanced medical image segmentation. The DA-TransUNet represents a significant step forward in leveraging the power of attention mechanisms and transformers for accurate and robust segmentation across a wide range of medical imaging modalities. Our work paves the way for further exploration of targeted attention mechanisms in medical image analysis and has the potential to impact clinical decision-making and patient care.

5.3 Limitations and future directions

Despite the advantages, our model also has some limitations. Firstly, the introduction of the DA-Blocks contributes to an increase in computational complexity. This added cost could potentially be a hindrance in real-time or resource-constrained applications. Although this increase in parameters is relatively modest considering the performance gains achieved, it could still be a concern in resource-constrained scenarios or when dealing with very large-scale datasets. Secondly, the decoder part of our model retains the original U-Net architecture. While this design choice preserves some of the advantages of U-Net, it also means that the decoder has not been specifically optimized for our application. This leaves room for further research and improvements, particularly in the decoder section of the architecture. Thirdly, one potential limitation of our DA-TransUNet architecture is the risk of losing fine-grained details during the tokenization process, which occurs after the convolution and pooling operations in the encoder. This is particularly concerning for medical images with thin and complex structures, where preserving intricate details is crucial for accurate segmentation. Although our proposed integration of the Dual Attention (DA) module before the Transformer in the encoder and within the skip connections helps mitigate this issue to some extent, as evidenced by the improved segmentation performance, we acknowledge that there may still be room for further enhancement in capturing and retaining fine-grained information.

6 Conclusion

In this paper, we innovatively proposed a novel approach to image segmentation by integrating DA-Blocks with the Transformer in the architecture of TransUNet. The DA-Blocks, focusing on image-specific position and channel features, were further integrated into the skip connections to enhance the model's performance. Our experimental results, validated by an extensive ablation study, showed significant improvements in the model's performance across various datasets, particularly the Synapse dataset.

Our research revealed the potential of image-special features position and channel (DA-Block) in enhancing the feature extraction capability and global information retention of the Transformer. The integration of DA-Block and Transformer substantially improved the model's performance without creating redundancy. Furthermore, the introduction of DA-Blocks into skip connections not only effectively bridges the semantic gap between the encoder and decoder, but also refines the feature maps, leading to an enhanced image segmentation performance.

Our model also has some limitations. Firstly, the introduction of DA blocks increases computational complexity. This added cost may pose obstacles for real-time or resource-constrained applications. Secondly, the decoder part of our model retains the original U-Net architecture. Lastly, the utilization of image feature positions and channels is only superficial, with deeper exploration possible.

This study has paved the way for the further use of image-special features position and channel (DA-Block) in the field of medical image segmentation. At the same time, it provides the idea of leveraging image characteristics to achieve high-precision medical image segmentation. Future work may focus on optimizing the decoder part of our architecture and exploring methods to reduce the computational complexity introduced by DA blocks without compromising the model's performance. We believe our approach can inspire future research in the domain of medical image segmentation and beyond.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: B. Landman, Z. Xu, J. E. Igelsias, M. Styner, T. Langerak, and A. Klein, "Segmentation outside the cranial vault challenge," in MICCAI: Multi Atlas Labeling Beyond Cranial Vault-Workshop Challenge, 2015. J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño, "Wm-dova maps for accurate polyp

highlighting in colonoscopy: Validation vs. saliency maps from physicians," Computerized medical imaging and graphics, vol. 43, pp. 99–111, 2015. N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti et al., "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic)," arXiv preprint arXiv:1902.03368, 2019. P. Tschandl, C. Rosendahl, and H. Kittler, "The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions," Scientific data, vol. 5, no. 1, pp. 1–9, 2018. D. Jha, P. H. Smedsrud, M. A. Riegler, P. Halvorsen, T. de Lange, D. Johansen, and H. D. Johansen, "Kvasir-seg: A segmented polyp dataset," in MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020, Proceedings, Part II 26. Springer, 2020, pp. 451–462. D. Jha, S. Ali, K. Emanuelsen, S. A. Hicks, V. Thambawita, E. GarciaCeja, M. A. Riegler, T. de Lange, P. T. Schmidt, H. D. Johansen et al., "Kvasir-instrument: Diagnostic and therapeutic tool segmentation dataset in gastrointestinal endoscopy," in MultiMedia Modeling: 27th International Conference, MMM 2021, Prague, Czech Republic, June 22–24, 2021, Proceedings, Part II 27. Springer, 2021, pp. 218–229. S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. K. Singh, S. Antani et al., "Automatic tuberculosis screening using chest radiographs," IEEE transactions on medical imaging, vol. 33, no. 2, pp. 233–245, 2013 S. Candemir, S. Jaeger, K. Palaniappan, J. P. Musco, R. K. Singh, Z. Xue, A. Karargyris, S. Antani, G. Thoma, and C. J. McDonald, "Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration," IEEE transactions on medical imaging, vol. 33, no. 2, pp. 577–590, 2013.

Author contributions

GS: Writing–original draft, Writing–review and editing. YP: Software, Visualization, Writing–original draft, Writing–review and editing. WK: Investigation, Validation, Writing–review and editing. ZX: Conceptualization, Writing–review and editing. JM: Conceptualization, Writing–review and editing. TR: Methodology, Writing–review and editing. L-MN: Conceptualization, Supervision, Writing–review and editing. JX: Investigation, Project administration, Resources, Supervision, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by The Soft Science Research Planning Project of Zhejiang Province under Grant 2024C35064 for the project "Study on Performance Evaluation and Optimization Path of Digital Aging Transformation Driven by User Experience."

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Azad, R., Aghdam, E. K., Rauland, A., Jia, Y., Avval, A. H., Bozorgpour, A., et al. (2022a) *Medical image segmentation review: the success of u-net*. arXiv preprint arXiv:2211.14830.

Azad, R., Al-Antary, M. T., Heidari, M., and Merhof, D. (2022b). Transnorm: transformer provides a strong spatial normalization mechanism for a deep segmentation model. *IEEE Access* 10, 108205–108215. doi:10.1109/access.2022.3211501

Azad, R., Asadi-Aghbolaghi, M., Fathy, M., and Escalera, S. (2019). "Bi-directional convlstm u-net with densley connected convolutions," in Proceedings of the IEEE/CVF international conference on computer vision workshops.

Bernal, J., Sánchez, F. J., Fernández-Esparrach, G., Gil, D., Rodríguez, C., and Vilariño, F. (2015). Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Comput. Med. imaging Graph.* 43, 99–111. doi:10.1016/j.compmedimag.2015.02.007

Cai, Y., Li, H., Xin, J., and Sun, G. (2022). "Mlda-unet: multi level dual attention unet for polyp segmentation," in 2022 16th ICME International Conference on Complex Medical Engineering (CME) (IEEE), 372–376.

Candemir, S., Jaeger, S., Palaniappan, K., Musco, J. P., Singh, R. K., Xue, Z., et al. (2013). Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. *IEEE Trans. Med. imaging* 33, 577–590. doi:10.1109/tmi.2013.2290491

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., et al. (2022). "Swin-unet: unet-like pure transformer for medical image segmentation," in European conference on computer vision (Springer), 205–218.

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., et al. (2021) *Transunet: transformers make strong encoders for medical image segmentation*. arXiv preprint arXiv:2102.04306.

Codella, N., Rotemberg, V., Tschandl, P., Celebi, M. E., Dusza, S., Gutman, D., et al. (2019) *Skin lesion analysis toward melanoma detection 2018: a challenge hosted by the international skin imaging collaboration (isic)*. arXiv preprint arXiv:1902.03368.

Diakogiannis, F. I., Waldner, F., Caccetta, P., and Wu, C. (2020). Resunet-a: a deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J. Photogrammetry Remote Sens.* 162, 94–114. doi:10.1016/j.isprsjprs.2020.01.013

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2020) *An image is worth 16x16 words: transformers for image recognition at scale*. arXiv preprint arXiv:2010.11929.

Drozdzal, M., Vorontsov, E., Chartrand, G., Kadoury, S., and Pal, C. (2016). "The importance of skip connections in biomedical image segmentation," in International Workshop on Deep Learning in Medical Image Analysis, International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis (Springer), 179–187.

Fu, J., Liu, J., Tian, H., Li, Y., Bao, Y., Fang, Z., et al. (2019). "Dual attention network for scene segmentation," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 3146–3154.

Guo, C., Szemenyei, M., Yi, Y., Wang, W., Chen, B., and Fan, C. (2021). "Sa-unet: spatial attention u-net for retinal vessel segmentation," in 2020 25th international conference on pattern recognition (ICPR) (IEEE), 1236–1242.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 770–778.

Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). "Densely connected convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, 4700–4708.

Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., et al. (2020). "Unet 3+: a full-scale connected unet for medical image segmentation," in ICASSP 2020-2020 IEEE international conference on acoustics, speech and signal processing (ICASSP) (IEEE), 1055–1059.

Ibtehaz, N., and Rahman, M. S. (2020). Multiresunet: rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural Netw.* 121, 74–87. doi:10.1016/j.neunet.2019.08.025

Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., et al. (2013). Automatic tuberculosis screening using chest radiographs. *IEEE Trans. Med. imaging* 33, 233–245. doi:10.1109/tmi.2013.2284099

Jamali, A., Roy, S. K., Li, J., and Ghamisi, P. (2023). Transu-net++: rethinking attention gated transu-net for deforestation mapping. *Int. J. Appl. Earth Observation Geoinformation* 120, 103332. doi:10.1016/j.jag.2023.103332

Jha, D., Ali, S., Emanuelsen, K., Hicks, S. A., Thambawita, V., Garcia-Ceja, E., et al. (2021). "Kvasir-instrument: Diagnostic and therapeutic tool segmentation dataset in gastrointestinal endoscopy," in Proceedings, Part II 27 MultiMedia Modeling: 27th International Conference, MMM 2021, Prague, Czech Republic, June 22–24, 2021 (Springer), 218–229.

Jha, D., Smedsrud, P. H., Riegler, M. A., Halvorsen, P., de Lange, T., Johansen, D., et al. (2020). "Kvasir-seg: a segmented polyp dataset," in Proceedings, Part II 26 MultiMedia Modeling: 26th International Conference, MMM 2020, Daejeon, South Korea, January 5–8, 2020 (Springer), 451–462.

Jin, Q., Meng, Z., Sun, C., Cui, H., and Su, R. (2020). Ra-unet: a hybrid deep attention-aware network to extract liver and tumor in ct scans. *Front. Bioeng. Biotechnol.* 8, 605132. doi:10.3389/fbioe.2020.605132

Landman, B., Xu, Z., Igelsias, J. E., Styner, M., Langerak, T., and Klein, A. (2015). "Segmentation outside the cranial vault challenge," in *MICCAI: multi Atlas labeling beyond cranial vault-workshop challenge*.

Le, N. Q. K. (2024). Hematoma expansion prediction: still navigating the intersection of deep learning and radiomics. *Eur. Radiol.*, 1–3. doi:10.1007/s00330-024-10586-x

Lin, A., Chen, B., Xu, J., Zhang, Z., Lu, G., and Zhang, D. (2022). Ds-transunet: dual swin transformer u-net for medical image segmentation. *IEEE Trans. Instrum. Meas.* 71, 1–15. doi:10.1109/tim.2022.3178991

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., et al. (2021). "Swin transformer: hierarchical vision transformer using shifted windows," in Proceedings of the IEEE/CVF international conference on computer vision, 10012–10022.

Maji, D., Sigedar, P., and Singh, M. (2022). Attention res-unet with guided decoder for semantic segmentation of brain tumors. *Biomed. Signal Process. Control* 71, 103077. doi:10.1016/j.bspc.2021.103077

Nam, H., Ha, J.-W., and Kim, J. (2017). "Dual attention networks for multimodal reasoning and matching," in Proceedings of the IEEE conference on computer vision and pattern recognition, 299–307.

Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., et al. (2018) *Attention u-net: learning where to look for the pancreas.* arXiv preprint arXiv:1804.03999.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). Pytorch: an imperative style, high-performance deep learning library. *Adv. neural Inf. Process. Syst.* 32. doi:10.48550/arXiv.1912.01703

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: convolutional networks for biomedical image segmentation," in Proceedings, Part III 18 Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015 (Springer), 234–241.

Sahayam, S., Nenavath, R., Jayaraman, U., and Prakash, S. (2022). Brain tumor segmentation using a hybrid multi resolution u-net with residual dual attention and deep supervision on mr images. *Biomed. Signal Process. Control* 78, 103939. doi:10.1016/j.bspc.2022.103939

Shi, Z., Miao, C., Schoepf, U. J., Savage, R. H., Dargis, D. M., Pan, C., et al. (2020). A clinically applicable deep-learning model for detecting intracranial aneurysm in computed tomography angiography images. *Nat. Commun.* 11, 6090. doi:10.1038/s41467-020-19527-w

Si, J., Zhang, H., Li, C.-G., Kuen, J., Kong, X., Kot, A. C., et al. (2018). "Dual attention matching network for context-aware feature sequence based person re-identification," in Proceedings of the IEEE conference on computer vision and pattern recognition, 5363–5372.

Tang, P., Zu, C., Hong, M., Yan, R., Peng, X., Xiao, J., et al. (2021). Da-dsunet: dual attention-based dense su-net for automatic head-and-neck tumor segmentation in mri images. *Neurocomputing* 435, 103–113. doi:10.1016/j.neucom.2020.12.085

Tran, T.-O., and Le, N. Q. K. (2024). Sa-ttca: an svm-based approach for tumor t-cell antigen classification using features extracted from biological sequencing and natural language processing. *Comput. Biol. Med.* 174, 108408. doi:10.1016/j.compbiomed.2024.108408

Tran, T.-O., Vo, T. H., and Le, N. Q. K. (2023). Omics-based deep learning approaches for lung cancer decision-making and therapeutics development. *Briefings Funct. Genomics*, elad031. doi:10.1093/bfgp/elad031

Tschandl, P., Rosendahl, C., and Kittler, H. (2018). The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Sci. data* 5, 180161–180169. doi:10.1038/sdata.2018.161

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Adv. neural Inf. Process. Syst.* 30. doi:10.48550/arXiv.1706.03762

Wang, H., Cao, P., Wang, J., and Zaiane, O. R. (2022a). Uctransnet: rethinking the skip connections in u-net from a channel-wise perspective with transformer. *Proc. AAAI Conf. Artif. Intell.* 36, 2441–2449. doi:10.1609/aaai.v36i3.20144

Wang, H., Xie, S., Lin, L., Iwamoto, Y., Han, X.-H., Chen, Y.-W., et al. (2022b). "Mixed transformer u-net for medical image segmentation," in ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (IEEE), 2390–2394.

Yang, Y., and Mehrkanoon, S. (2022). "Aa-transunet: attention augmented transunet for nowcasting tasks," in 2022 International Joint Conference on Neural Networks (IJCNN) (IEEE), 01–08.

Zhang, Y., Liu, H., and Hu, Q. (2021). "Transfuse: fusing transformers and cnns for medical image segmentation," in Proceedings, Part I 24 Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021 (Springer), 14–24.

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., and Liang, J. (2018). "Unet++: a nested u-net architecture for medical image segmentation," in Proceedings 4 Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018 (Springer), 3–11.

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Original Research
PUBLISHED 27 June 2024
DOI 10.3389/fbioe.2024.1414605

Check for updates

OPEN ACCESS

EDITED BY
Adrian Elmi-Terander,
Stockholm Spine Center, Sweden

REVIEWED BY
Rizwan Ali Naqvi,
Sejong University, Republic of Korea
Shireen Y. Elhabian,
The University of Utah, United States

*CORRESPONDENCE
Zhiping Song,
2452671073@qq.com
ZeKui Liu,
704603595@qq.com

†These authors have contributed equally to this work

RECEIVED 09 April 2024
ACCEPTED 30 May 2024
PUBLISHED 27 June 2024

CITATION
Mao K, Li R, Cheng J, Huang D, Song Z and Liu Z (2024), PL-Net: progressive learning network for medical image segmentation.
Front. Bioeng. Biotechnol. 12:1414605.
doi: 10.3389/fbioe.2024.1414605

COPYRIGHT
© 2024 Mao, Li, Cheng, Huang, Song and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

PL-Net: progressive learning network for medical image segmentation

Kunpeng Mao[1†], Ruoyu Li[2†], Junlong Cheng[2], Danmei Huang[1], Zhiping Song[3*] and ZeKui Liu[3*]

[1]Chongqing City Management College, Chongqing, China, [2]College of Computer Science, Sichuan University, Chengdu, China, [3]Chongqing University of Engineering, Chongqing, China

In recent years, deep convolutional neural network-based segmentation methods have achieved state-of-the-art performance for many medical analysis tasks. However, most of these approaches rely on optimizing the U-Net structure or adding new functional modules, which overlooks the complementation and fusion of coarse-grained and fine-grained semantic information. To address these issues, we propose a 2D medical image segmentation framework called Progressive Learning Network (PL-Net), which comprises Internal Progressive Learning (IPL) and External Progressive Learning (EPL). PL-Net offers the following advantages: 1) IPL divides feature extraction into two steps, allowing for the mixing of different size receptive fields and capturing semantic information from coarse to fine granularity without introducing additional parameters; 2) EPL divides the training process into two stages to optimize parameters and facilitate the fusion of coarse-grained information in the first stage and fine-grained information in the second stage. We conducted comprehensive evaluations of our proposed method on five medical image segmentation datasets, and the experimental results demonstrate that PL-Net achieves competitive segmentation performance. It is worth noting that PL-Net does not introduce any additional learnable parameters compared to other U-Net variants.

KEYWORDS

progressive learning, coarse-grained to fine-grained semantic information, complementation and fusion, medical image segmentation, computer version

1 Introduction

Medical image segmentation is a technique used to extract regions of interest for quantitative and qualitative analysis. For example, it can be used for cell segmentation in electron microscopy (EM) recordings (Ronneberger et al., 2015), melanoma segmentation in dermoscopy images (Berseth, 2017; Cheng et al., 2020), thyroid nodule segmentation in ultrasound images, and heart segmentation in MRI images (Bernard et al., 2018). Traditionally, medical image segmentation methods relied on manually designed features to generate segmentation results (Xu et al., 2007; Tong et al., 2015). However, this approach requires distinct feature designs for various applications. Furthermore, the large variety of medical image modalities makes it difficult or impossible to transfer a specific type of feature design method to different image types. Therefore, the development of a universal feature extraction technique is crucial in the field of medical image analysis.

The emergence of deep learning technology has revolutionized medical image segmentation by overcoming the limitations of traditional manual feature extraction methods. Convolutional neural networks (CNN) based automatic feature learning algorithms, such as the fully convolutional network (FCN) proposed by Shelhamer et al. (Long et al., 2015) and the U-Net framework for biomedical image segmentation proposed by Ronneberger et al., 2015, have shown promising results. The FCN model structure is designed to be end-to-end, which eliminates the need for manual feature extraction and image post-processing steps. On the other hand, the U-Net framework's encoder-decoder-skip connection network structure has shown good results on medical image segmentation datasets with small amounts of data.

To further enhance the adaptability of U-Net for different medical image segmentation tasks, researchers have continuously explored and innovated, proposing numerous variant models of U-Net. These variant models aim to achieve better performance in medical image segmentation by adding new functional modules or optimizing the network structure. For instance, Vanilla U-Net introduces channel/spatial attention mechanisms or self-attention mechanisms to capture crucial information in medical images, significantly improving its performance in various segmentation tasks. Additionally, researchers have optimized the encoder-decoder structure of Vanilla U-Net or adjusted the skip connections to generate more refined and abundant feature representations.

However, the introduction of these methods has also brought new challenges. Although the addition of new parameters and functional modules enhances model performance, it also increases model complexity and the risk of overfitting. More importantly, these methods often overlook the complementarity and fusion of coarse-grained and fine-grained semantic information. Most existing semantic segmentation methods assume that the entire segmentation process can be completed through a single feedforward process, resulting in homogeneous feature representations that struggle to excel in extracting fine-grained feature representations. Therefore, for medical environments with limited computational resources, it is highly beneficial to ensure model simplicity while fully integrating and utilizing semantic information at different scales while maintaining a small number of parameters. Such a design can not only enhance the generalization and robustness of the model but also ensure its efficiency and practicality in real-world applications.

In this paper, we propose a new medical image segmentation method called progressive learning networks (PL-Net). PL-Net divide the feature learning process within the U-Net architecture into two distinct depth "steps" to achieve the combination of different receptive field sizes, enabling the network to learn semantic information at varying granularities. The entire segmentation process is performed through two feedforward processes (referred to as "stages"). At the end of each stage, the features obtained from that stage are transferred to the next stage for fusion. This transfer operation allows the model to leverage the knowledge learned in the previous training stage to extract finer-grained information, thereby refining the coarse segmentation output. Unlike previous works, our proposed method does not add any additional parameters or functional modules to the U-Net. Instead, our method fully explores the complementary relationships between features through a progressive learning strategy. The main contributions can be summarized as follows:

1) We propose a progressive learning network (PL-Net) designed specifically for medical image segmentation tasks. Through its unique design, this network deeply explores the potential of feature complementarity and fusion in medical image segmentation. By adjusting the scale of output channels, we designed both a standard PL-Net (15.03 M) and a smaller version, PL-Net† (Ocs = 0.5, 3.77 M), to accommodate medical scenarios with different computational resources.
2) We introduce internal progressive learning (IPL) and external progressive learning (EPL) strategies. The IPL strategy effectively captures different receptive field sizes, thereby learning and integrating multi-granularity semantic information. The EPL strategy allows the model to extract finer information based on the knowledge from the previous stage, thus optimizing the segmentation results.
3) We applied the proposed method to tasks such as skin lesion segmentation and cell nucleus segmentation. Experimental results indicate that PL-Net outperforms other state-of-the-art methods such as U-NeXt and BiO-Net. Moreover, despite its smaller parameter size, the smaller version of PL-Net† still demonstrates superior segmentation performance.

2 Releat work

Currently, most semantic segmentation methods assume that the entire segmentation process can be executed through a single feedforward pass of the input image, which often overlooks global information. To address this, researchers have added new functional modules or optimized the U-Net structure to achieve performance improvements. These methods can be classified into: 1) U-Net variants focused on functional optimization; 2) U-Net variants focused on structural optimization.

U-Net variants focused on functional optimization. Due to the large number of irrelevant features in medical images, it is crucial to focus on target features and suppress irrelevant features during the segmentation process. Recent works have extended U-Net by adding different novel functional modules, demonstrating its potential in various visual analysis tasks. Squeeze-and-Excitation (SE) (Hu et al., 2018) has facilitated the development of U-Net by automatically learning the importance of each feature channel through an attention mechanism. Additionally, ScSE (Roy et al., 2018) and FCANet (Cheng et al., 2020) integrate concurrent spatial and channel attention modules into U-Net to improve segmentation performance. Oktay et al., 2018 proposed an attention gate for medical imaging to focus on target structures of different shapes and sizes and suppress irrelevant areas of the input image. In addition to plug-and-play attention modules, researchers have designed specific functional modules for different medical image segmentation tasks. For example, Zhou et al., 2019 proposed a contour-aware information aggregation network with a multi-level information aggregation module between two task-specific decoders. The SAUNet (Sun et al., 2020) uses both a secondary shape stream and a regular texture stream in parallel to capture rich shape-related information, enabling multi-level interpretation of the external network and reducing the need for additional computations. The CE-Net (Gu et al., 2019) uses a dense atrous convolution (DAC) block to extract a rich feature representation and residual multi-kernel pooling (RMP) operation to further encode the multi-scale context features extracted from the DAC block without additional learning weights.

The emergence of the Vision Transformer (ViT) (Dosovitskiy et al., 2020) has had a significant impact on the progress of medical image analysis. Compared to CNN methods, ViT has less inductive bias. The U-Transformer (Petit et al., 2021) takes inspiration from ViT and incorporates multi-head self-attention modules into U-Net, which helps to obtain global contextual information. The UNeXt (Valanarasu and Patel, 2022) is the first fast medical image segmentation network that uses both convolution and MLP. It reduces the number of parameters and computational complexity by using tokenized MLP. In contrast to the aforementioned U-Net variants, our work explores the effectiveness of progressive learning techniques in capturing both coarse-grained and fine-grained semantic information. The PL-Net enhances the performance of different stage U-Nets by reusing learned features.

U-Net variants focused on structural optimization. Unlike U-Net variants focused on functional optimization, optimizing its structure allows it to extract feature information at different levels, which is feasible and effective for many computer vision problems. One of the simplest and most effective ways to optimize the encoder-decoder structure is to replace the basic building blocks with more advanced ones, such as (Jégou et al., 2017; Diakogiannis et al., 2020; Hasan et al., 2020), which benefit from residual or dense connections in deeper network structures. In addition to replacing the building blocks, performance can also be improved for different tasks by increasing the number of U-shaped network structures, as demonstrated in (Jégou et al., 2017; Isensee et al., 2021). One of the most famous networks in this category is nnU-Net (Isensee et al., 2021), which proposes three networks based on the original U-Net structure: 2D U-Net, 3D U-Net, and U-Net cascade. The first stage performs coarse segmentation of downsampled low-resolution images, and the second stage combines the results of the first stage for fine-tuning. ResGANet (Cheng et al., 2022) achieved segmentation performance improvement by replacing the encoder in U-Net with a lightweight and efficient backbone. TransUNet (Chen J. et al., 2021) and FATNet (Wu et al., 2022) replaced the encoder structure of U-Net with CNN and Transformer branches in a parallel or serial manner.

Skip connections are considered a key component of U-Net's success. U-Net++ (Zhou et al., 2018) has redesigned skip connections through a series of nested and dense connections, reducing the semantic gap between the subnet feature map encoders and decoders. R2U-Net (Alom et al., 2019) effectively increases the network depth by utilizing residual networks and RCNN and obtaining more expressive features through feature summation with different time steps. Xiang et al. designed BiO-Net (Xiang et al., 2020) with backward skip connections based on R2UNet, which can reuse the features of each decoding level to achieve more intermediate information aggregation. The emergence of BiO-Net allows building blocks to be reused by U-Net in a circular manner without introducing any additional parameters.

3 Progressive learning network

We now describe PL-Net, a progressive learning framework for medical image segmentation. As is shown in Figure 1, PL-Net is a multi-level U-Net network architecture that does not rely on additional functional modules but has paired bidirectional connections. The core of our framework is to enhance the feature representation required for image segmentation through two progressive learning approaches (internal and external) and to fuse coarse-grained as well as fine-grained semantic information.

Two U-Nets with different depths form the "stages" of external progressive learning. In each stage, as the "step" of internal progressive learning increases, the shallow network is expanded to a deeper network, learning stable multi-granularity information from it. In brief, the number of parameters is not increased through internal progressive learning, but it can learn feature maps with different sizes of receptive fields. External progressive learning is defined as the coarse segmentation stage (Stage 1) and the fine segmentation stage (Stage 2). The input image will be examined at multiple scales to achieve the fusion of coarse-grained and fine-grained information.

3.1 Internal progressive learning

Bidirectional skip connections are used in internal progressive learning to reuse building blocks. In order to enable the network at each stage to learn distinctive feature representations, we use two "steps" to gradually mine the features from shallow to deep.

Forward Skip Connections (FSC) are used to assist up-sampling learning, restore the contour of the segmentation target, and retain the low-level visual features of encoding. The feature f_s^{FSC} after FSC can be expressed as:

$$f_s^{FSC} = [Conv_s(x), \hat{x}] \tag{1}$$

Backward Skip Connections (BSC) are used for flexible aggregation of low-level visual features and high-level semantic features. In order to realize the complementation and fusion of semantic information at different stages, multi-granularity information of different "steps" and "stages" is combined in feature f^{BSC}:

$$f^{BSC} = \begin{cases} [x, x] & s = 1, stage1 \\ \left[Conv_s\left(f_s^{FSC}\right), x\right] & s = 1, stage2 \\ [Conv_s(\hat{x}), x] & s = 2, stage1, 2 \end{cases} \tag{2}$$

Among them, $[\cdot]$ is the concatenation layer, $Conv_s$ means that the convolution operation of the sth "steps" ($s \in \{1, 2\}$) is applied to the input feature map, x and $\hat{x}$ are feature maps of the same size in the down-sampling and up-sampling path respectively.

It is worth noting that the reasoning path of internal progressive learning can be extended to multiple recursions to obtain instant performance gains. More importantly, a larger receptive field will be got in each output of this learning strategy than the previous "steps." We use K_i to represent the ith complete encoding-decoding process, and x_{out}^i is used to represent the output. Therefore, x_{out}^i can be written as:

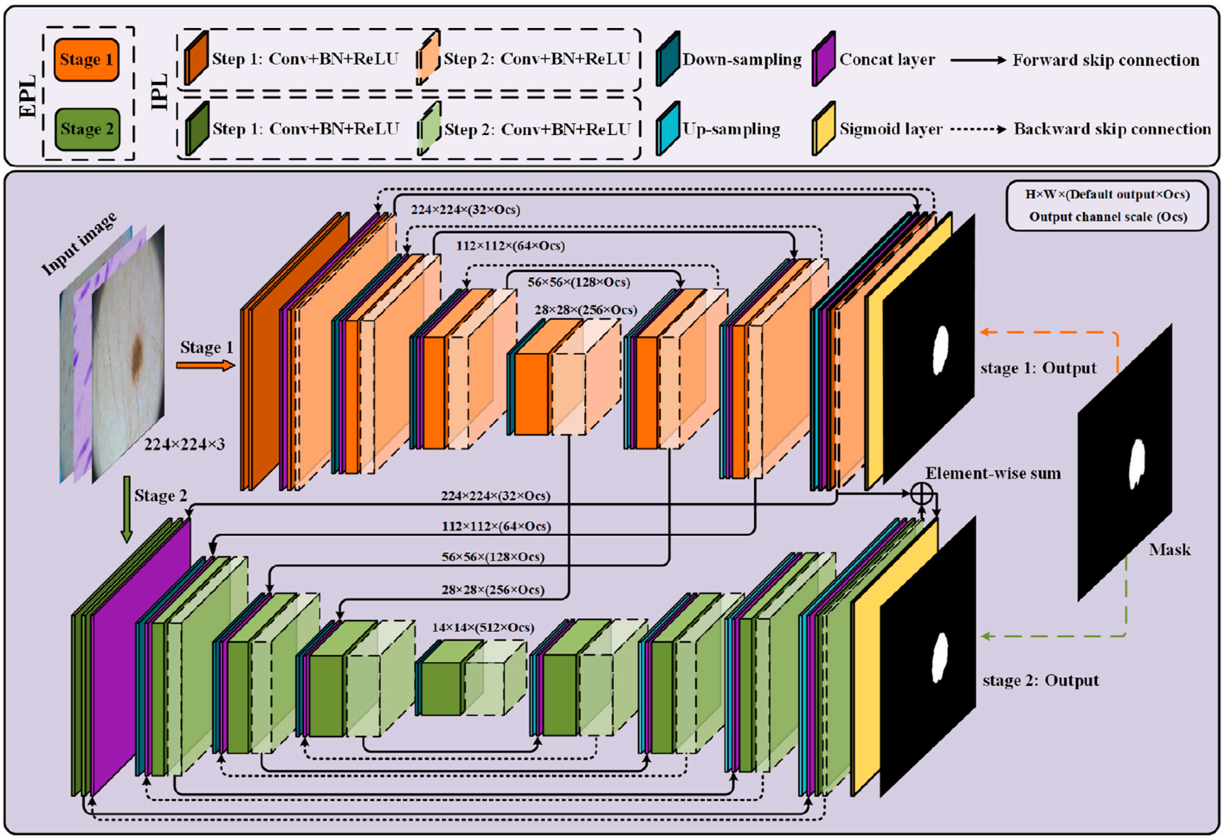


FIGURE 1
Overview of the progressive learning network (PL-Net). PL-Net consists of two parts: internal progressive learning (IPL) and external progressive learning (EPL).

$$x_{out}^{i} = \begin{cases} x_{in} & i = 0 \\ K_i(x_{in}) & i = 1 \\ K_i\left(x_{out}^{i-1}\right) & i \geq 2 \end{cases} \tag{3}$$

In this study, we define $i = 2$, and through such a setting the parameters equivalent to BiO-Net can be maintained. In future research, the setting of $i > 2$ can be used to further improve the segmentation accuracy, but the exploration of the optimal hyperparameter setting is beyond the scope of this paper.

3.2 External progressive learning

The external progressive learning strategy first trains the low stage (stage 1), and then gradually trains toward the high stage (stage 2). Since "stage1″ is relatively shallow in depth and limited by the perceptual field and performance ability, it will focus on local information extraction, while "stage 2″ incorporates the local texture information learned from "stage 1." Compared with directly training the entire network in series, in the model, it is allowed by this incremental nature to pay attention to global information as the features gradually enter a higher stage.

For each stage of training, we calculate the loss based on the Dice coefficient ($\mathcal{L}_{Dice}$) (Eelbode et al., 2020) between the ground truth (y_{true}) and the predicted probability (y_{pred}^{n}) distribution of different stages:

$$\mathcal{L}_{Dice}\left(y_{pred}^{n}, y_{true}\right) = 1 - \frac{2 \times \left|y_{pred}^{n} \cap y_{true}\right|}{\left|y_{pred}^{n}\right| + \left|y_{true}\right|} \tag{4}$$

Here $|\cdot|$ is an operator through which the number of pixels is found in the qualified area. In each iteration, the input data will be used in each learning stage (where $n \in \{1, 2\}$). What needs to be clear is that when the latter stage is predicted, all the parameters of the previous stage are optimized and updated, which helps each stage in the model to work together.

Since the low stage is mainly to assist the feature expression and knowledge reasoning of the high-stage network, we can delete the low-stage prediction layer (Sigmoid layer) when predicting, thereby reducing the reasoning time. In addition, the predictions at different stages are unique, but they can form complementary information among different granularities. When we combine all outputs with an equal weight, it will result in a better performance. In other words, the final output of the model is determined by all stages:

$$y = \frac{1}{1 + e^{-\sum_{n=1}^{N-1} y^{n}}} \tag{5}$$

3.3 PL-Net architecture

Our framework has a trade-off between performance and parameters. Like U-Net, the down-sampling and up-sampling

TABLE 1 Detailed configuration of U-Net, BiO-Net, and our PL-Net architecture. We use " [kernel, kernel, channel]" to represent the convolution configuration.

Input	Encoder			Output	Decoder		
	U-Net	BiO-Net	PL-Net		U-Net	BiO-Net	PL-Net
224^2	[3, 3, 64]× 2	[3, 3, 32]× 2	$\left\{\begin{matrix}[3,3,32],\ \mathrm{step1}\\ [3,3,32],\ \mathrm{step2}\\ \mathrm{stage1, stage2}\end{matrix}\right\}$	7^2	—	[3, 3, 256]× 2	—
112^2	[3, 3, 128]× 2	[3, 3, 32]× 2	$\left\{\begin{matrix}[3,3,64],\ \mathrm{step1}\\ [3,3,64],\ \mathrm{step2}\\ \mathrm{stage1, stage2}\end{matrix}\right\}$	28^2	[3, 3, 512]× 2	[3, 3, 128]× 2	$\left\{\begin{matrix}[3,3,256],\ \mathrm{step1}\\ [3,3,256],\ \mathrm{step2}\\ \mathrm{stage2}\end{matrix}\right\}$
56^2	[3, 3, 256]× 2	[3, 3, 64]× 2	$\left\{\begin{matrix}[3,3,128],\ \mathrm{step1}\\ [3,3,128],\ \mathrm{step2}\\ \mathrm{stage1, stage2}\end{matrix}\right\}$	56^2	[3, 3, 256]× 2	[3, 3, 64]× 2	$\left\{\begin{matrix}[3,3,128],\ \mathrm{step1}\\ [3,3,128],\ \mathrm{step2}\\ \mathrm{stage1, stage2}\end{matrix}\right\}$
28^2	[3, 3, 512]× 2	[3, 3, 128]× 2	$\left\{\begin{matrix}[3,3,256],\ \mathrm{step1}\\ [3,3,256],\ \mathrm{step2}\\ \mathrm{stage1, stage2}\end{matrix}\right\}$	112^2	[3, 3, 128]× 2	[3, 3, 32]× 2	$\left\{\begin{matrix}[3,3,64],\ \mathrm{step1}\\ [3,3,64],\ \mathrm{step2}\\ \mathrm{stage1, stage2}\end{matrix}\right\}$
14^2	[3, 3, 1024]× 2	[3, 3, 256]× 2	$\left\{\begin{matrix}[3,3,512],\ \mathrm{step1}\\ [3,3,512],\ \mathrm{step2}\\ \mathrm{stage2}\end{matrix}\right\}$	224^2	[3, 3, 64]× 2	[3, 3, 32]× 2	$\left\{\begin{matrix}[3,3,32],\ \mathrm{step1}\\ [3,3,32],\ \mathrm{step2}\\ \mathrm{stage1, stage2}\end{matrix}\right\}$
7^2	—	[3, 3, 512]× 2	—	224^2	[1, 1, 1], *Sigmoid*		
Parameters					25.59 M	14.30 M	15.03 M
Model size					118 MB	57.7 MB	60.7 MB

stages of PL-Net only use standard convolutional layers, batch normalization layers and ReLU layers without introducing any additional functional modules. Table 1 is the detailed configuration of U-Net, BiO-Net and our PL-Net.

As shown in Table 1, BiO-Net has a maximum coding depth of 4, using BSC from the deepest coding level, and inputting the decoded features in each iteration as a whole into the last-stage block. BSC is also used in PL-Net. Unlike the previous methods, the convolutional layer is allowed to be used in the model to mine features from coarse-grained to fine-grained ones in a progressive manner. It should be noted that a smaller version of PL-Net† can be obtained only by adjusting the Ocs, whose depth and connection method will not change.

4 Experiments

4.1 Datasets

ISIC 2017 (Berseth, 2017) is a dataset consisting of 2000 training images, 150 validation images, and 600 test images. The images in the original dataset provided by ISIC have different resolutions. To address this, we first use the gray world color consistency algorithm to normalize the colors of the images and then adjust the size of all images to 2242 pixels. All experimental results reported in this paper for ISIC 2017 are from the official test set results.

PH2 (Mendonça et al., 2013) is a dataset containing 200 dermoscopic images, with a fixed size of 768 × 560 pixels. The dataset contains 80% benign mole cases and 20% melanoma cases, with ground truth annotated by dermatologists. Due to the small scale of the dataset, we use the preprocessing method of the ISIC 2017 dataset and the trained model to directly predict all images in the dataset to evaluate the performance of different models.

Kaggle 2018 Data Science Bowl (referred to as Nuclei) (Caicedo et al., 2019) is a dataset provided by the Booz Allen Foundation, containing 670 cell feature maps with ground truth for each image. To prepare the dataset for training and testing, we adjust all images and corresponding ground truth to 2242 pixels and use 80% of the images for training and the remaining 20% for testing.

The TN-SCUI (Pedraza et al., 2015) dataset is a collection of 3644 nodular thyroid images, each annotated by experienced physicians. The dataset was originally part of the TN-SCUI 2020 challenge and was processed to remove personal labels to protect patient privacy. In this study, we randomly divided the dataset into a training set (60%), validation set (20%), and test set (20%). To ensure consistency, we uniformly adjusted the resolution of all images to 2242 pixels.

ACDC (Bernard et al., 2018) is a dataset that includes cardiac MRI images of 150 patients, from which we collected 1489 slices for 3D images. For training and testing purposes, we used 951 and 538 slices, respectively. Notably, in contrast to the four other datasets mentioned earlier, ACDC comprises three different categories: left ventricle, right ventricle, and myocardium. Hence, we employed this dataset to investigate how various models perform on multi-class segmentation. Figure 2 displays sample images from these datasets and their corresponding ground truth.

4.2 Implementation details

We conducted all experiments on Tesla V100 GPUs using Keras and expanded the training data for all datasets by applying random rotations (±25°), random horizontal and vertical shifts (15%), and random flips (horizontal and vertical). For all models, we trained for more than 200 epochs with a batch size of 16, a fixed learning rate of 1e-4, and an Adam optimizer with a momentum of 0.9 to minimize

10.3389/fbioe.2024.1414605

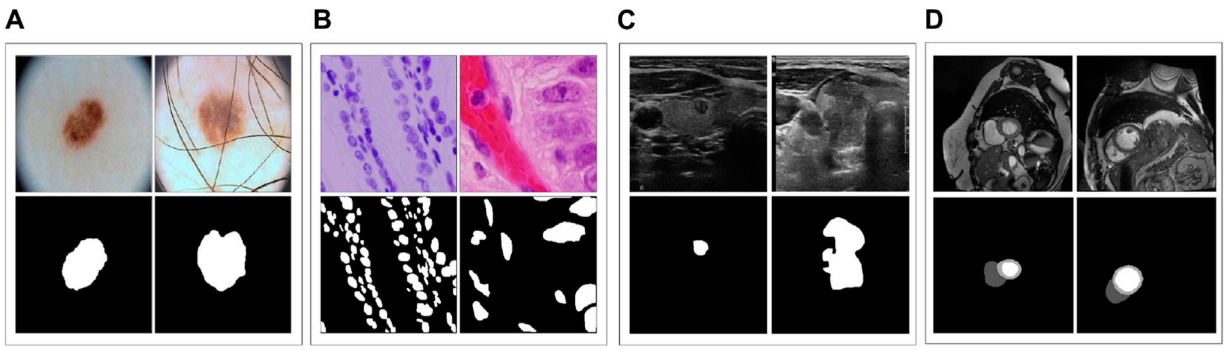


FIGURE 2
(A–D) represent samples from the five datasets.

TABLE 2 Ablative results. IoU (Dice), number of parameters, and model size are reported.

Dataset	EPL	Without IPL		With IPL	
		n = 1	n = 2	n = 2	n = 3
ISIC 2017	×	76.05 (84.43)	77.09 (85.23)	77.15 (85.37)	77.07 (85.44)
	✓	76.69 (84.94)	77.04 (85.27)	**77.92** (**85.94**)	77.49 (85.56)
Nuclei	×	85.54 (91.89)	85.13 (91.53)	85.93 (**92.14**)	84.78 (91.28)
	✓	85.60 (91.84)	85.80 (92.00)	**86.14** (92.13)	85.37 (91.70)
PH2	×	83.90 (90.74)	85.88 (91.61)	86.69 (92.47)	86.48 (92.45)
	✓	84.97 (91.00)	86.63 (92.44)	**87.27** (**92.86**)	87.03 (92.77)
TN-SCUI	×	72.32 (81.38)	73.72 (82.61)	75.95 (85.36)	76.67 (84.60)
	✓	74.20 (83.33)	75.63 (84.33)	76.66 (85.10)	**77.05** (**85.55**)
ACDC	×	74.44 (81.56)	74.66 (82.03)	77.78 (83.84)	77.49 (83.60)
	✓	75.30 (82.19)	76.80 (83.42)	**78.06** (**84.36**)	77.96 (83.91)
Parameters	—	**10.33M**	15.03 M	15.03 M	19.73 M
Model size	—	**41.60 MB**	60.70 MB	60.70 MB	79.70 MB

The bold values indicates the best performance.

Dice loss. We used an early stop mechanism and stopped training when the validation loss reached a stable level with no significant change for 20 epochs. Unless explicitly specified, PL-Net had two "steps" and "stages," and BSC was established at each stage of the network. When testing, all prediction layers are deleted before the last "stage," and other configurations remain unchanged.

4.3 Ablation study

To understand the effectiveness of IPL and EPL strategies, we conducted ablation studies. When there is no IPL strategy, features are extracted by naturally stacking benchmark blocks, and we conducted experiments on stacking 1-layer and 2-layer benchmark blocks, respectively. Adopting an IPL strategy means that the encoder-decoder must be iterated for n times in different stages, and we set n = 2 and n = 3. When external progressive learning is not performed, different "stages" are connected in series through PL-Net to transfer the feature information learned in each stage. Only the parameters in the last stage are optimized, and the segmentation results are output through the model. That is to say, the feature information obtained in the current "stage" of training is transferred to the next training "stage" and fused through the EPL method, allowing fine-grained information to be mined through the model based on learning in the previous training "stage".

Table 2 presents our IoU (Dice) scores without/with a progressive learning strategy on five different medical image segmentation datasets. We provide the parameters and model sizes for different scenarios to comprehensively analyze segmentation performance. In most cases, the best segmentation performance is achieved through PL-Net when both internal (n = 2) and external progressive learning strategies are used simultaneously. Compared to the model with the same parameter settings without IPL, the segmentation performance is significantly improved. These results demonstrate the effectiveness of EPL and IPL. Moreover, we observed that the progressive learning strategy has a significant

 10.3389/fbioe.2024.1414605

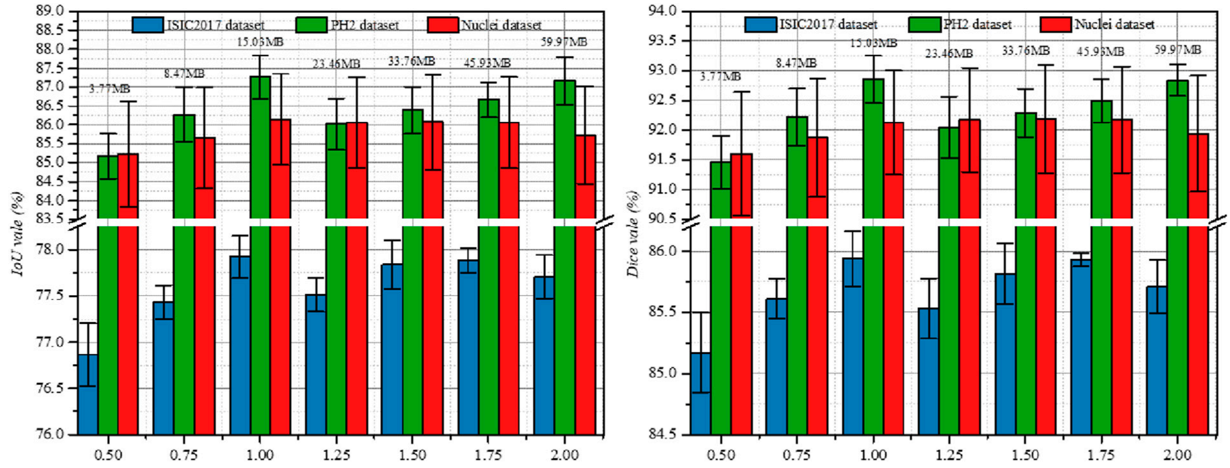


FIGURE 3
Study the impact of Ocs on three public datasets. Results are calculated over 5 runs and are shown with standard errors. We label the parameters of the model at the top of the bar chart.

impact on datasets with complex boundaries or multi-category datasets. On the TN-SCUI dataset, for instance, the IoU improvement is as high as 2.94% with the same parameter setting (n = 2). To balance factors such as performance and parameters, we used the setting of n = 2 in the following experiments. However, we believe that the setting of n = 3 may be more effective as the size of the dataset increases.

In addition to the above ablation studies, we also investigated the impact of the output channel scale (Ocs) on the segmentation performance of different datasets. Figure 3 shows the experimental results on three datasets, where we set $Ocs \in [0.5, 2.0]$ and take values at an interval of 0.25. Note that Ocs = 0.5 represents a smaller version of PL-Net†.We found that when Ocs = 1.0, the best segmentation result can be obtained, and the parameter amount (15.03 MB) is well-balanced. When Ocs > 1.0, the segmentation performance improves as the number of channels increases, but it does not exceed that of the standard PL-Net. We attribute this to the limitation of the data size and the complexity of the segmentation content. While PL-Net† has slightly lower segmentation performance than other networks, it has very few parameters. Thus, it is recommended for use on small datasets. Additionally, it can be configured to run on servers or mobile devices with lower hardware requirements.

4.4 Comparison with state-of-the-arts

4.4.1 Quantitative comparison

For the ISIC 2017 and PH2 datasets, we compared our PL-Net to the baseline U-Net and other state-of-the-art methods (Ronneberger et al., 2015; Badrinarayanan et al., 2017; Al-Masni et al., 2018; Oktay et al., 2018; Zhou et al., 2018; Alom et al., 2019; Kaul et al., 2019; Cheng et al., 2020; Hasan et al., 2020; Jha et al., 2020; Lei et al., 2020; Xiang et al., 2020; Cao et al., 2021; Isensee et al., 2021; Cheng et al., 2022; Valanarasu and Patel, 2022; Wu et al., 2022). The functional optimization-oriented variants of U-Net include (Badrinarayanan et al., 2017; Al-Masni et al., 2018; Oktay et al., 2018; Alom et al., 2019; Kaul et al., 2019; Cheng et al., 2020; Valanarasu and Patel, 2022) while the structural optimization-oriented variants of U-Net include (Zhou et al., 2018; Alom et al., 2019; Hasan et al., 2020; Jha et al., 2020; Lei et al., 2020; Xiang et al., 2020; Cao et al., 2021; Isensee et al., 2021; Cheng et al., 2022; Wu et al., 2022). To ensure fairness, we either used the experimental results provided by the authors on the same test set or ran their models published in the same environment.

Table 3 presents the accuracy (Acc), intersection over union (IoU), Dice coefficient (Dice), sensitivity (Sens), and specificity (Spec) scores of different segmentation methods on the ISIC2017 and PH2 datasets. Our PL-Net outperforms other methods in terms of both IoU and Dice metrics on the ISIC2017 dataset. Specifically, the IoU and Dice scores of PL-Net are 0.6% and 0.3% higher than those of BiO-Net (t = 3, INT), respectively. The smaller sized PL-Net† (3.77 M) achieves the same Dice score as BiO-Net (t = 3) (14.30 M). Although nnU-Net (Petit et al., 2021) achieves the best sensitivity on the ISIC2017 test set, its model size is 3.76 times larger than that of the standard PL-Net. The PH2 dataset also involves the dermoscopic image segmentation task. While the number of parameters in UNeXt-L (Dosovitskiy et al., 2020) is similar to that of our smaller version of PL-Net†, UNeXt-L completes the entire segmentation process through a single feed-forward pass of the input image, resulting in low parameter utilization and insufficient learning. When compared with other state-of-the-art methods, PL-Net demonstrates superior performance on the PH2 dataset. Furthermore, PL-Net† hasmuch fewer parameters than other methods, yet it still achieves competitive segmentation performance.

Nuclei dataset. The datasets used for nucleus segmentation have non-uniform feature distributions, and the shapes of positive and negative samples vary greatly. Table 4 presents the quantitative comparison results of our method against 14 other methods. Compared to the state-of-the-art TransAttUnet-R (Chen B. et al., 2021), our PL-Net achieves better overall segmentation performance, with improvements ranging from 0.27% to 1.16% for different evaluation metrics. The segmentation performance

TABLE 3 Performance comparison with SOTA methods on ISIC 2017 and PH2 datasets. Red, Green, and Blue indicate the best, second best and third best performance.

Network	ISIC 2017 dataset					PH2 dataset					#Params	Model size
	Acc	IoU	Dice	Sens	Spec	Acc	IoU	Dice	Sens	Spec		
FrCN Al-Masni et al. (2018)	0.940	0.771	0.871	0.854	0.967	0.951	0.848	0.918	0.937	0.957	—	—
FocusNet Tong et al. (2015)	0.921	0.756	0.832	0.767	0.990	—	—	—	—	—	—	—
SegNet Kaul et al. (2019)	0.918	0.696	0.821	0.801	0.954	0.934	0.808	0.894	0.865	0.966	28.09M	112 MB
DSNet Gu et al. (2019)	—	0.775	—	0.875	0.967	—	0.870	—	0.929	0.969	10.00 M	—
DAGAN Valanarasu and Patel (2022)	0.935	0.771	0.859	0.835	0.976	—	—	—	—	—	—	—
FATNet Wu et al. (2022)	0.933	0.765	0.850	0.839	0.973	—	—	—	—	—	27.43 M	109 MB
ResGANet Cheng et al. (2022)	0.936	0.764	0.862	0.842	0.950	—	—	—	—	—	39.21M	—
U-Net Ronneberger et al. (2015)	0.926	0.736	0.825	0.828	0.964	0.943	0.851	0.915	0.946	0.957	29.59 M	118 MB
U-Net++ Berseth (2017)	0.929	0.753	0.840	0.848	0.965	0.948	0.853	0.917	0.973	0.937	34.48 M	138 MB
Double U-Net Isensee et al. (2021)	0.936	0.765	0.847	0.830	0.970	0.942	0.860	0.915	0.934	0.953	27.94 M	112 MB
FCANet Hu et al. (2018)	0.935	0.776	0.856	0.869	0.962	0.952	0.868	0.926	0.968	0.926	59.97 M	241 MB
Att U-Net Cheng et al. (2020)	0.939	0.757	0.840	0.859	0.957	0.951	0.868	0.926	0.953	0.955	30.42 M	121 MB
R2U-Net Cheng et al. (2022)	0.938	0.776	0.858	0.859	0.969	0.952	0.871	0.927	0.954	0.960	91.61 M	366 MB
Att R2U-Net Cheng et al. (2022)	0.939	0.775	0.857	0.857	0.961	0.954	0.873	0.928	0.949	0.962	92.11 M	368 MB
BiO-Net (t = 3) Zhou et al. (2018)	0.937	0.772	0.852	0.845	0.973	0.944	0.851	0.915	0.963	0.931	14.30 M	57.7 MB
BiO-Net (t = 3, INT) Zhou et al. (2018)	0.934	0.754	0.840	0.821	0.976	0.945	0.851	0.909	0.968	0.944	3.99 M	15.2 MB
UNeXt-L Valanarasu and Patel (2022)	0.935	0.773	0.856	0.864	0.966	0.949	0.859	0.919	0.941	0.955	14.30 M	57.7 MB
PL-Net† (Our)	0.932	0.769	0.852	0.871	0.953	0.945	0.852	0.915	0.955	0.957	3.77 M	15.6 MB
PL-Net (Our)	0.940	0.779	0.859	0.848	0.975	0.957	0.873	0.929	0.965	0.966	15.03 M	60.7 MB

The bold values indicates the best performance.

of U-Net++ falls between our PL-Net† and PL-Net, with an IoU of 85.56% and a Dice of 91.59%. Across five cross-validation experiments, standard PL-Net showed higher stability than PL-Net†, with a 14% reduction in standard deviation. Although the Dice score of Att R2U-Net is higher than that of PL-Net, its overall performance and stability are slightly inferior. Notably, both PL-Net and BiO-Net use BSC, but our method shows better overall performance. With a smaller PL-Net† size, almost the same IoU and Dice scores as BiO-Net (t = 3, INT) can be achieved.

TN-SCUI and ACDC datasets. The boundary of the TN-SCUI dataset is blurred compared to other datasets, and we found that methods including CNN may obtain better experimental results in this case. As shown in Table 5, even lightweight approaches like PL-Net† can achieve performance similar to Swin-UNet. UNeXt-L, a hybrid network based on CNN and MLP, has the smallest model size, but its segmentation performance is inferior to baseline methods. Our analysis shows that this is because the method has fewer learnable parameters and cannot make good use of the learned features. In the ACDC dataset (Table 6), we demonstrate the segmentation performance of different methods on different classes. The target area of the myocardium (Myo) is ringed between the left atrium (LV) and right atrium (RV) and is relatively small overall. The segmentation accuracy of different methods on this category tends to be lower than that of the other two categories. Our PL-Net achieves the highest IoU and Dice scores. Although TransUNet and EANet can achieve better average segmentation performance, their model size is increased by 6 times, making them more complex and requiring more computing resources than our proposed method. Additionally, the experimental results of PL-Net on the ACDC dataset show that our method is also suitable for multi-category segmentation tasks.

The above quantitative comparison demonstrates that our proposed network can be applied to different segmentation tasks, which can include different modalities and categories. Even for images with blurred boundaries, PL-Net can produce good segmentation results. Although the overall segmentation performance of PL-Net† is not as good as that of standard PL-Net, its smaller parameters and model size will promote its application in memory-constrained environments. Additionally, other U-Net variants, which are oriented towards functional optimization or structural optimization, can improve the segmentation performance of the original U-Net to some extent, but the increased computational cost is a difficult problem to avoid. As PL-Net is a progressive learning framework, it achieves a good trade-off between segmentation performance and parameters.

TABLE 4 Performance comparison with SOTA methods on Nuclei dataset. **Red**, **Green**, and **Blue** indicate the best, second-best, and third-best performance. For the original implementation methods, we report mean ± standard deviation.

Network	Nuclei dataset				#Params	Model size
	Acc	IoU	Dice	Sens		
PraNet Fan et al. (2020)	95.59	71.08	81.03	80.62	—	—
Channel-UNet Chen et al. (2019)	96.27	79.75	87.55	90.70	—	—
ResUNet Diakogiannis et al. (2020)	97.05	82.44	89.91	90.00	—	—
Double U-Net Jha et al. (2020)	—	84.07	91.33	64.07	27.94 M	112 MB
TransAttUnet D Chen et al. (2021a)	97.37	84.62	91.34	91.86	—	—
TransAttUnet R Chen et al. (2021a)	97.46	84.98	91.62	91.85	—	—
TransUNet Chen et al. (2021b)	97.84	85.21	91.69	91.62	100.4 M	401 MB
FATNet Wu et al. (2022)	98.11	85.24	91.69	91.73	27.43 M	109 MB
U-Net Ronneberger et al. (2015)	97.84 ± 0.24	85.68 ± 1.40	91.90 ± 1.00	92.61 ± 0.52	29.59 M	118 MB
U-Net++ Zhou et al. (2018)	97.87 ± 0.22	85.91 ± 1.35	92.06 ± 1.00	92.48 ± 1.08	34.48 M	138 MB
FCANet Cheng et al. (2020)	97.68 ± 0.31	84.87 ± 1.39	91.33 ± 1.09	91.70 ± 1.50	59.97 M	241 MB
Att U-Net Oktay et al. (2018)	97.84 ± 0.18	85.46 ± 1.20	91.77 ± 0.88	91.93 ± 0.66	30.42 M	121 MB
R2U-Net Alom et al. (2019)	97.93 ± 0.18	85.68 ± 1.26	91.89 ± 0.92	92.28 ± 1.20	91.61 M	366 MB
Att R2U-Net Alom et al. (2019)	97.76 ± 0.34	85.86 ± 1.04	92.15 ± 0.92	92.51 ± 1.46	92.11 M	368 MB
BiO-Net (t = 3) Xiang et al. (2020)	97.81 ± 0.22	85.09 ± 1.42	91.53 ± 1.04	91.99 ± 0.72	14.30 M	57.7 MB
BiO-Net (t = 3, INT) Xiang et al. (2020)	97.84 ± 0.20	85.31 ± 1.27	91.68 ± 0.93	91.94 ± 0.76	14.30 M	57.7 MB
UNeXt-L Valanarasu and Patel (2022)	97.43 ± 0.15	81.26 ± 1.46	88.75 ± 1.31	88.71 ± 1.65	3.99 M	15.2 MB
PL-Net† (Our)	97.79 ± 0.22	85.23 ± 1.39	91.60 ± 1.03	91.79 ± 0.59	3.77 M	15.6 MB
PL-Net (Our)	97.96 ± 0.16	86.14 ± 1.20	92.13 ± 0.88	92.12 ± 1.11	15.03 M	60.7 MB

The bold values indicates the best performance.

TABLE 5 Performance comparison with SOTA methods on TN-SCUI datasets. **Red**, **Green**, and **Blue** indicate the best, second-best, and third-best performance.

Network	TN-SCUI datasets		#Params (M)	Model size (MB)
	IoU	Dice		
U-Net Ronneberger et al. (2015)	0.718	0.806	29.59	118
SegNet Badrinarayanan et al. (2017)	0.726	0.819	17.94	71.8
FATNet Wu et al. (2022)	0.751	0.842	27.43	109
Swin-UNet Cao et al. (2021)	0.744	0.835	25.86	105
TransUNet Chen et al. (2021b)	0.746	0.837	88.87	401
EANet Wang et al. (2022)	0.751	0.839	47.07	118
UNeXt-L Valanarasu and Patel (2022)	0.693	0.794	3.99	15.2
PL-Net† (Our)	0.742	0.830	3.77	15.6
PL-Net (Our)	0.767	0.851	15.03	60.7

The bold values indicates the best performance.

4.4.2 Qualitative comparison

To better understand the excellent performance of our method, we present example results of PL-Net and several other methods in Figure 4 and Figure 5. As shown, our PL-Net and PL-Net† can handle different types of targets and produce accurate segmentation results.

TABLE 6 Performance comparison with SOTA methods on ACDC datasets. Red, Green, and Blue indicate the best, second-best, and third-best performance.

Network	ACDC datasets				#Params (M)	Model size (MB)
	RV	Myo	LV	Average		
U-Net Ronneberger et al. (2015)	0.743 (0.792)	0.717 (0.812)	0.861 (0.897)	0.774 (0.834)	29.59	118
SegNet Badrinarayanan et al. (2017)	0.738 (0.790)	0.720 (0.817)	0.864 (0.902)	0.774 (0.836)	17.94	71.8
FATNet Wu et al. (2022)	0.743 (0.799)	0.702 (0.805)	0.859 (0.899)	0.768 (0.834)	27.43	109
Swin-UNet Cao et al. (2021)	0.754 (0.805)	0.722 (0.820)	0.865 (0.903)	0.780 (0.843)	25.86	105
TransUNet Chen et al. (2021b)	0.750 (0.800)	0.715 (0.812)	0.872 (0.905)	0.779 (0.839)	88.87	401
EANet Wang et al. (2022)	0.742 (0.791)	0.732 (0.825)	0.864 (0.902)	0.779 (0.839)	47.07	118
UNeXt-L Valanarasu and Patel, (2022)	0.719 (0.779)	0.675 (0.810)	0.840 (0.882)	0.745 (0.824)	3.99	15.2
PL-Net† (Our)	0.723 (0.778)	0.692 (0.796)	0.845 (0.887)	0.753 (0.820)	3.77	15.6
PL-Net (Our)	0.761 (0.807)	0.738 (0.828)	0.872 (0.907)	0.790 (0.847)	15.03	60.7

The bold values indicates the best performance.

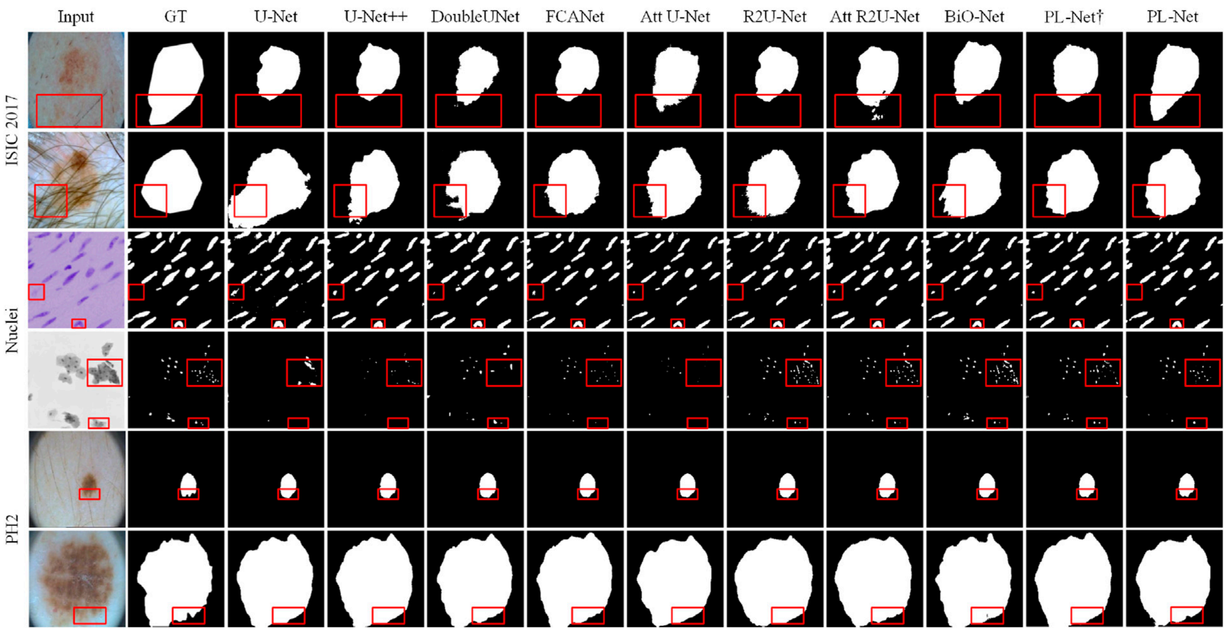


FIGURE 4
Qualitative segmentation results of ISIC 2017, Nuclei, and PH2 datasets using different methods.

The first and second rows of Figure 4 respectively show the segmentation results of an ambiguous target area and a small amount of occlusion (hair). As observed, although the results produced by PL-Net are not as accurate, our method is still effective for areas with ambiguous targets compared to other methods. When segmenting occluded images, other models either tend to divide boundaries incorrectly or mistake masked areas as target areas. The segmentation target of the image in the third row is clear, and relatively accurate segmentation results can be produced through other methods. However, for the content marked in the red box, most methods mistake interfering pixels for target pixels for segmentation, and better results are produced through our method compared to other methods. The fourth row shows the performance of different models for targets consisting of tiny targets and dispersed structures. As observed, U-Net and Att U-Net either recognize the saliency area as the target area or lose the target area, resulting in poor segmentation results. The fifth and sixth rows show the segmentation results of different methods for smaller and larger targets. As seen, our model makes a good decision on the boundary of the small target, while the area marked in the red box cannot be segmented well by other models. Compared to the fifth row, the lesion area shown in the sixth row covers almost the entire image. Although more accurate segmentation results can be produced through other methods, our PL-Net produces more perfect results as far as the area marked in the red box is concerned.

Figure 5 presents qualitative comparison results of different methods on the TN-SCUI and ACDC datasets. From the

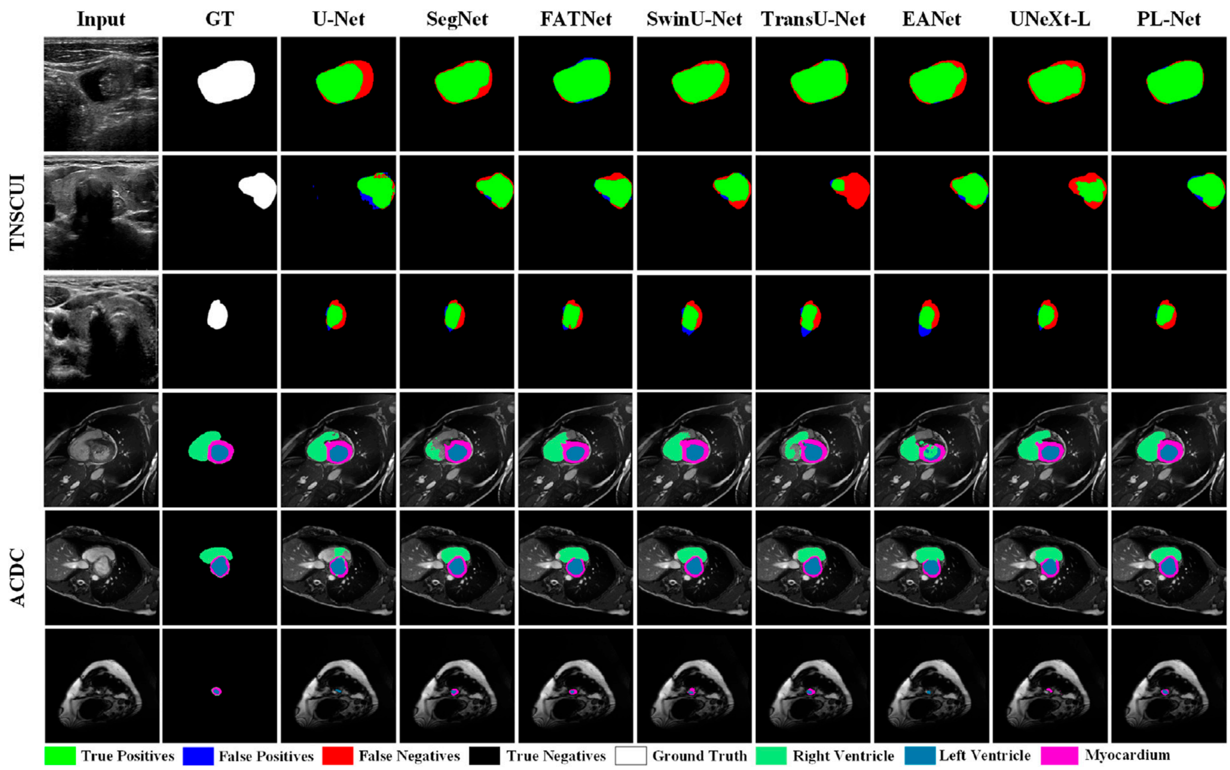


FIGURE 5
Qualitative segmentation results of TN-SCUI and ACDC datasets using different methods.

experimental results in the first two rows of the TN-SCUI dataset, PL-Net has a larger true positive area compared to other methods and is more accurate in lesion boundary segmentation. The third row shows an example where different methods perform poorly. Although there is a certain difference between our segmentation results and the ground truth, the false positive area is significantly lower than that of other methods, which is particularly important in medical image analysis. We highlighted different targets in the ACDC dataset using different colors, and the experimental results in the fourth-row show that SegNet, TransUNet, and EANet have poor segmentation results and incomplete segmentation of the RV area. In the example image in the fifth row, the Myocardium area accounts for a relatively small proportion, and FATNet, EANet, and UNeXt do not correctly segment the ring area, while PL-Net clearly segments the Myocardium area. The experimental results in the last row demonstrate the advantage of PL-Net in segmenting small targets. Although U-Net, EANet, and UNeXt segment the target area, their category definitions are inaccurate. These experimental results cover different situations in medical images, including large, medium, and small lesions, as well as targets of different categories. These results indicate that PL-Net has good generalization ability and can handle different types of medical image semantic segmentation problems.

In addition to the visualization results mentioned above, we present the features learned by different "stages" and "steps" of PL-Net in the form of a heat map, as shown in Figure 6. During internal progressive learning (i.e., "Step1″ and "Step2″), the shallower "Step1″ tends to focus on coarse-grained semantic information first, such as the outline of hair or lesions. As the network depth increases, "Step2″ gives less weight to texture features and focuses more on fine-grained semantics. PL-Net captures semantic information from coarse to fine granularity at different "stages" using internal progressive learning and does not introduce additional parameters compared to other approaches that replace deeper encoders. Through the visualization results of different "stages," we observe that the heat value of "Stage2″ is higher than that of "Stage1″ at the same position (i.e., the corresponding weight value is larger), which benefits from the fusion of coarse-grained and fine-grained information of the two stages. In addition, "Fusion" represents the feature map of the last convolutional layer after the two-stage fusion, with a distribution of thermal values similar to that of the ground truth and masked from irrelevant background regions.

4.4.3 Expanding to 3D medical image segmentation

In this section, we will detail how to effectively apply the proposed progressive strategy to 3D medical image segmentation tasks. To validate the effectiveness of this strategy, we chose 3D U-Net as the baseline network and conducted preliminary experimental validation on the standard prostate MRI segmentation dataset, PROMISE 2012 Litjens et al. (2014). This dataset contains 50 MRI volumes, which we split into a training set and a test set in a 3:2 ratio.

In the 3D U-Net Çiçek et al. (2016) structure, we employed basic blocks consisting of two 3 × 3 × 3 convolutions (excluding batch

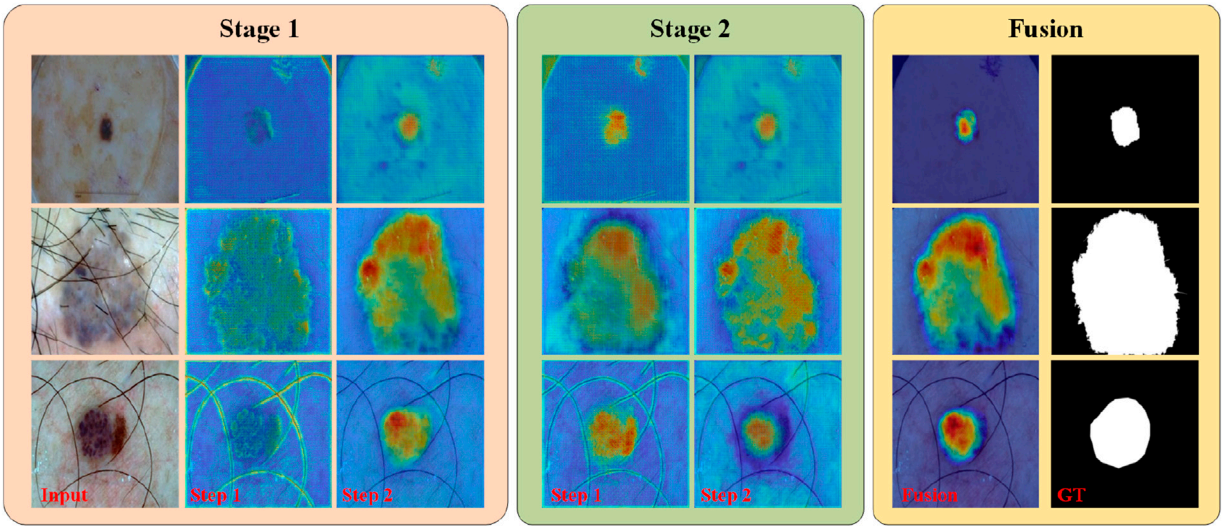


FIGURE 6
PL-Net's heat map of different "stages" and "steps" on the ISIC2017 dataset.

normalization and activation functions here). The encoder part includes four such basic blocks, each followed by a downsampling operation to progressively reduce the spatial dimensions of the feature maps. The decoder part restores the feature space through four upsampling operations, each also followed by a basic block.

To integrate the progressive learning strategy into the 3D U-Net, we converted the two $3 \times 3 \times 3$ convolutions in the basic block into an internal progressive learning process, where each convolution layer is considered a "step." In the second "step," we introduced backward skip connections to fuse features of different levels at the same scale. This design allows us to effectively incorporate the internal progressive learning strategy without altering the original 3D U-Net's basic structure.

Next, we regarded the above network as the first "stage" of external progressive learning. To construct the second "stage," we added a downsampling layer and a basic block at the end of the encoder, and an upsampling layer and a basic block at the beginning of the decoder. Similar to the design concept of PL-Net, we built the second "stage" by reusing the network structure from the first stage along with the newly added basic blocks. In the second stage, we used skip connections to effectively fuse the coarse-grained information from the first stage with the fine-grained features of the second stage. Through these steps, we extended the original 3D U-Net into a progressive learning network with two "steps" and two "stages."

As shown in Table 7, we compared the experimental results of the original 3D U-Net with those after introducing the progressive learning strategy. The data indicates that introducing only the internal progressive learning strategy improved the Dice score by 0.6% compared to the baseline 3D U-Net. When both progressive learning strategies were applied, the Dice score improvement was even more significant, reaching 1.36%. These preliminary experimental results fully demonstrate the effectiveness and practicality of our proposed method. Encouraged by these positive findings, we plan to further explore the potential performance of progressive learning strategies in a broader range of 3D medical image segmentation tasks in the future.

TABLE 7 Extended experiments on the application of progressive learning strategies to 3D U-Net.

Method	PROMISE 2012 dataset	
	IoU (%)	Dice (%)
3D U-Net Çiçek et al. (2016)	56.84	72.48
3D U-Net + IPL	57.58	73.08
3D U-Net + IPL + EPL	**58.53**	**73.84**

The bold values indicates the best performance.

5 Discussion

U-Net has been widely used as a benchmark model for medical image segmentation due to its simple and easily modifiable structure. Most of its variant approaches enhance segmentation performance by adding functional modules (e.g., attention module) or modifying its original structure (e.g., residual, and densely connected structures) in the feed-forward process. In this paper, we adopt an alternative approach by recognizing that coarse-grained and fine-grained discriminative information naturally exists at different stages of the network, which can be learned incrementally, similar to how humans learn through shallow and deep network structures. Based on this intuition, we design a framework with internal and external progressive learning strategies, called PL-Net. Internal progressive learning strategies are used to mine semantic information at different granularities, while external progressive learning strategies further refine segmentation details based on the features learned in the previous training phase.

Researchers have proposed numerous network architectures based on U-Net to address various medical image semantic segmentation problems. However, some approaches that add functional modules (such as FCANet and Att U-Net) do not consistently improve performance across different datasets. Our experimental results demonstrate that while FCANet improves IoU by 4% over vanilla U-Net on the ISIC2017 dataset, it degrades performance by 0.81% on the Nuclei dataset, indicating that performance variation is related to the type, size, and complexity of the dataset. Our proposed PL-Net achieves consistent performance improvements over vanilla U-Net on five datasets without adding new functional modules or structural modifications and remains competitive with state-of-the-art network frameworks (EANet and ResGANet). Moreover, PL-Net has lower computational overhead and fewer parameters, resulting in a model size reduction of 3.8 times and 6.6 times compared to widely used nnUNet and TransUNet, respectively. We also provide PL-Net† with a smaller number of parameters, which can offer options for different medical imaging scenarios, although the decrease in the number of parameters results in reduced segmentation accuracy. Our method can run on a GPU with limited memory, reducing the complex configuration and tedious preprocessing steps of nnUNet. In other words, designing such a network is crucial to translate medical imaging from the laboratory to clinical practice.

On the other hand, similar to most existing state-of-the-art methods, our proposed segmentation network still has limitations in handling cases with complex boundaries and small targets. As shown in the first row of Figure 4, when the boundary between the skin lesion and the background region is difficult to distinguish, our method and other approaches fail to accurately delineate the boundary. As shown in the third row of Figure 5, PL-Net's segmentation performance is lower when the target region is very small. However, in these cases, our method is closest to the ground truth, and the segmentation results are still better than those of other competitors. From the experimental results in Table 1, we found that the best results were obtained by performing three internal progressive learning experiments on the large-scale TN-SCUI dataset, indicating the necessity of setting different internal progressive learning strategies. Finally, we believe that introducing robust functional modules may further improve the segmentation performance of PL-Net, and we will explore this in future work. The ideas proposed in this paper mainly provide inspiration for researchers who are committed to designing feature representations to improve convolutional neural networks.

6 Conclusion

In this study, we propose a new variant of U-Net called PL-Net for 2D medical image segmentation, which mainly consists of internal and external progressive learning strategies. Compared to U-Net methods that optimize functional or structural aspects, our PL-Net achieves consistent performance improvements without additional trainable parameters. We provide both a standard PL-Net (15.03 M) and a smaller version, PL-Net† (3.77 M), to address different medical image segmentation scenarios in real-world situations. We conduct comprehensive experiments on five public medical image datasets, and the results show that PL-Net can improve the segmentation IoU of the baseline network by 0.46%–4.9%, demonstrating high competitiveness with other state-of-the-art methods.

Although our proposed method has shown promising results, it still has some limitations that need to be further addressed in future research: 1) Impact of data size: Exploring the parameter settings of internal and external progressive learning under different data sizes will help researchers understand the potential of the model under different scales of data. In the future, we will further explore the performance of PL-Net on larger datasets. 2) Due to the limitations of computing power and data, our method mainly focuses on 2D medical image segmentation. This article has initially demonstrated the feasibility of the progressive learning strategy in 3D medical image segmentation. In the future, we will extend PL-Net to more advanced 3D medical image segmentation frameworks to further enhance its capabilities in 3D medical image segmentation. 3) Design of functional modules: How to design functional modules suitable for PL-Net to improve segmentation performance while maintaining a concise framework is also a topic for further research in the future.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

KM: Conceptualization, Methodology, Project administration, Writing–review and editing. RL: Writing–review and editing, Conceptualization, Investigation, Methodology. JC: Writing–original draft, Conceptualization, Software. DH: Writing–review and editing, Formal Analysis. ZS: Writing–review and editing, Validation. ZL: Writing–review and editing, Visualization.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was partially funded by Chongqing Municipal Education Commission Science and Technology Research Project (KJQN202203302, KJZD-K202303301, and KJQN202203309).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Al-Masni, M. A., Al-Antari, M. A., Choi, M.-T., Han, S.-M., and Kim, T.-S. (2018). Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Comput. methods programs Biomed.* 162, 221–231. doi:10.1016/j.cmpb.2018.05.027

Alom, M. Z., Yakopcic, C., Hasan, M., Taha, T. M., and Asari, V. K. (2019). Recurrent residual u-net for medical image segmentation. *J. Med. imaging* 6, 1–014006. doi:10.1117/1.jmi.6.1.014006

Badrinarayanan, V., Kendall, A., and Cipolla, R. (2017). Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. pattern analysis Mach. Intell.* 39, 2481–2495. doi:10.1109/tpami.2016.2644615

Bernard, O., Lalande, A., Zotti, C., Cervenansky, F., Yang, X., Heng, P.-A., et al. (2018). Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE Trans. Med. imaging* 37, 2514–2525. doi:10.1109/tmi.2018.2837502

Berseth, M. (2017) *Isic 2017 - skin lesion analysis towards melanoma detection. ArXiv* abs/1703.00523.

Caicedo, J. C., Goodman, A., Karhohs, K. W., Cimini, B. A., Ackerman, J., Haghighi, M., et al. (2019). Nucleus segmentation across imaging experiments: the 2018 data science bowl. *Nat. methods* 16, 1247–1253. doi:10.1038/s41592-019-0612-7

Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., et al. (2021). "Swin-unet: unet-like pure transformer for medical image segmentation," in *ECCV workshops*.

Chen, B., Liu, Y., Zhang, Z., Lu, G., and Kong, A. W.-K. (2021a). Transattunet: multi-level attention-guided u-net with transformer for medical image segmentation. *IEEE Trans. Emerg. Top. Comput. Intell.* 8, 55–68. doi:10.1109/tetci.2023.3309626

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., et al. (2021b) *Transunet: transformers make strong encoders for medical image segmentation*. arXiv preprint arXiv:2102.04306.

Chen, Y., Wang, K., Liao, X., Qian, Y., Wang, Q., Yuan, Z., et al. (2019). Channel-unet: a spatial channel-wise convolutional neural network for liver and tumors segmentation. *Front. Genet.* 10, 1110. doi:10.3389/fgene.2019.01110

Cheng, J., Tian, S., Yu, L., Gao, C., Kang, X., Ma, X., et al. (2022). Resganet: residual group attention network for medical image classification and segmentation. *Med. Image Anal.* 76, 102313. doi:10.1016/j.media.2021.102313

Cheng, J., Tian, S., Yu, L., Lu, H., and Lv, X. (2020). Fully convolutional attention network for biomedical image segmentation. *Artif. Intell. Med.* 107, 101899. doi:10.1016/j.artmed.2020.101899

Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). "3d u-net: learning dense volumetric segmentation from sparse annotation," in Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, 424–432. doi:10.1007/978-3-319-46723-8_49

Diakogiannis, F. I., Waldner, F., Caccetta, P., and Wu, C. (2020). Resunet-a: a deep learning framework for semantic segmentation of remotely sensed data. *ISPRS J. Photogrammetry Remote Sens.* 162, 94–114. doi:10.1016/j.isprsjprs.2020.01.013

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2020) *An image is worth 16x16 words: transformers for image recognition at scale. arXiv preprint arXiv:2010.11929*.

Eelbode, T., Bertels, J., Berman, M., Vandermeulen, D., Maes, F., Bisschops, R., et al. (2020). Optimization for medical image segmentation: theory and practice when evaluating with dice score or jaccard index. *IEEE Trans. Med. imaging* 39, 3679–3690. doi:10.1109/tmi.2020.3002417

Fan, D.-P., Ji, G.-P., Zhou, T., Chen, G., Fu, H., Shen, J., et al. (2020). "Pranet: parallel reverse attention network for polyp segmentation," in *International conference on medical image computing and computer-assisted intervention* (Berlin, Germany: Springer), 263–273.

Gu, Z., Cheng, J., Fu, H., Zhou, K., Hao, H., Zhao, Y., et al. (2019). Ce-net: context encoder network for 2d medical image segmentation. *IEEE Trans. Med. imaging* 38, 2281–2292. doi:10.1109/tmi.2019.2903562

Hasan, M. K., Dahal, L., Samarakoon, P. N., Tushar, F. I., and Martí, R. (2020). Dsnet: automatic dermoscopic skin lesion segmentation. *Comput. Biol. Med.* 120, 103738. doi:10.1016/j.compbiomed.2020.103738

Hu, J., Shen, L., and Sun, G. (2018). "Squeeze-and-excitation networks," in Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu, Hawaii, USA, 21-26 July 2017, 7132–7141. doi:10.1109/cvpr.2018.00745

Isensee, F., Jaeger, P. F., Kohl, S. A., Petersen, J., and Maier-Hein, K. H. (2021). nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. methods* 18, 203–211. doi:10.1038/s41592-020-01008-z

Jégou, S., Drozdzal, M., Vazquez, D., Romero, A., and Bengio, Y. (2017). "The one hundred layers tiramisu: fully convolutional densenets for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition workshops, Long Beach, CA, USA, June 16 2019 to June 17 2019, 11–19. doi:10.1109/cvprw.2017.156

Jha, D., Riegler, M. A., Johansen, D., Halvorsen, P., and Johansen, H. D. (2020). "Doubleu-net: a deep convolutional neural network for medical image segmentation," in 2020 IEEE 33rd International symposium on computer-based medical systems (CBMS) (IEEE), Rochester, MN, USA, July 28 2020 to July 30 2020, 558–564. doi:10.1109/cbms49503.2020.00111

Kaul, C., Manandhar, S., and Pears, N. (2019). "Focusnet: an attention-based fully convolutional network for medical image segmentation," in 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019) (IEEE), Venice, Italy, 8-11 April 2019, 455–458. doi:10.1109/isbi.2019.8759477

Lei, B., Xia, Z., Jiang, F., Jiang, X., Ge, Z., Xu, Y., et al. (2020). Skin lesion segmentation via generative adversarial networks with dual discriminators. *Med. Image Anal.* 64, 101716. doi:10.1016/j.media.2020.101716

Litjens, G., Toth, R., Van De Ven, W., Hoeks, C., Kerkstra, S., Van Ginneken, B., et al. (2014). Evaluation of prostate segmentation algorithms for mri: the promise12 challenge. *Med. image Anal.* 18, 359–373. doi:10.1016/j.media.2013.12.002

Long, J., Shelhamer, E., and Darrell, T. (2015). "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, Boston, MA, USA, June 7 2015 to June 12 2015, 3431–3440. doi:10.1109/cvpr.2015.7298965

Mendonça, T., Ferreira, P. M., Marques, J. S., Marcal, A. R., and Rozeira, J. (2013). "PH^2 - a dermoscopic image database for research and benchmarking," in 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC) (IEEE), Osaka, Japan, 3-7 July 2013, 5437–5440. doi:10.1109/embc.2013.6610779

Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M. J., Heinrich, M. P., Misawa, K., et al. (2018) *Attention u-net: learning where to look for the pancreas. ArXiv abs/1804.03999*.

Pedraza, L., Vargas, C., Narváez, F., Durán, O., Muñoz, E., and Romero, E. (2015). An open access thyroid ultrasound image database. *10th Int. symposium Med. Inf. Process. analysis (SPIE)* 9287, 188–193. doi:10.1117/12.2073532

Petit, O., Thome, N., Rambour, C., Themyr, L., Collins, T., and Soler, L. (2021). "U-net transformer: self and cross attention for medical image segmentation," in Machine Learning in Medical Imaging: 12th International Workshop, MLMI 2021, Held in Conjunction with MICCAI 2021, Strasbourg, France, September 27, 2021, 267–276. doi:10.1007/978-3-030-87589-3_28

Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: convolutional networks for biomedical image segmentation. *Fronrs Neurosci. abs/1505*, 04597. doi:10.1007/978-3-319-24574-4_28

Roy, A. G., Navab, N., and Wachinger, C. (2018). "Concurrent spatial and channel 'squeeze & excitation'in fully convolutional networks," in Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, 421–429. doi:10.1007/978-3-030-00928-1_48

Sun, J., Darbehani, F., Zaidi, M., and Wang, B. (2020). "Saunet: shape attentive u-net for interpretable medical image segmentation," in Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, 797–806. doi:10.1007/978-3-030-59719-1_77

Tong, T., Wolz, R., Wang, Z., Gao, Q., Misawa, K., Fujiwara, M., et al. (2015). Discriminative dictionary learning for abdominal multi-organ segmentation. *Med. image Anal.* 23, 92–104. doi:10.1016/j.media.2015.04.015

Valanarasu, J. M. J., and Patel, V. M. (2022). "Unext: mlp-based rapid medical image segmentation network," in International conference on medical image computing and computer-assisted intervention, Vancouver, BC, Canada, 8-12 October 2022, 23–33. doi:10.1007/978-3-031-16443-9_3

Wang, K., Zhang, X., Zhang, X., Lu, Y., Huang, S., and Yang, D. (2022). Eanet: iterative edge attention network for medical image segmentation. *Pattern Recognit.* 127, 108636. doi:10.1016/j.patcog.2022.108636

Wu, H., Chen, S., Chen, G., Wang, W., Lei, B., and Wen, Z. (2022). Fat-net: feature adaptive transformers for automated skin lesion segmentation. *Med. image Anal.* 76, 102327. doi:10.1016/j.media.2021.102327

Xiang, T., Zhang, C., Liu, D., Song, Y., Huang, H., and Cai, W. (2020). "Bio-net: learning recurrent bi-directional connections for encoder-decoder architecture," in Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, 74–84. doi:10.1007/978-3-030-59710-8_8

Xu, J., Chutatape, O., Sung, E., Zheng, C., and Kuan, P. C. T. (2007). Optic disk feature extraction via modified deformable model technique for glaucoma analysis. *Pattern Recognit.* 40, 2063–2076. doi:10.1016/j.patcog.2006.10.015

Zhou, Y., Onder, O. F., Dou, Q., Tsougenis, E., Chen, H., and Heng, P.-A. (2019). "Cia-net: robust nuclei instance segmentation with contour-aware information aggregation," in Information Processing in Medical Imaging: 26th International Conference, IPMI 2019, Hong Kong, China, June 2–7, 2019, 682–693. doi:10.1007/978-3-030-20351-1_53

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., and Liang, J. (2018). "Unet++: a nested u-net architecture for medical image segmentation," in Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, 3–11. doi:10.1007/978-3-030-00889-5_1

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Review
PUBLISHED 22 July 2024
DOI 10.3389/fbioe.2024.1392807

Check for updates

OPEN ACCESS

EDITED BY
Peter Kokol,
University of Maribor, Slovenia

REVIEWED BY
Parvez Ahmad,
Huazhong University of Science and Technology, China
Bojan Žlahtič,
University of Maribor, Slovenia

*CORRESPONDENCE
Daesik Jeong,
jungsoft97@smu.ac.kr
Seung Won Lee,
swleemd@g.skku.edu

†These authors have contributed equally to this work

RECEIVED 28 February 2024
ACCEPTED 14 June 2024
PUBLISHED 22 July 2024

CITATION
Abidin ZU, Naqvi RA, Haider A, Kim HS, Jeong D and Lee SW (2024), Recent deep learning-based brain tumor segmentation models using multi-modality magnetic resonance imaging: a prospective survey.
Front. Bioeng. Biotechnol. 12:1392807.
doi: 10.3389/fbioe.2024.1392807

COPYRIGHT
© 2024 Abidin, Naqvi, Haider, Kim, Jeong and Lee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Recent deep learning-based brain tumor segmentation models using multi-modality magnetic resonance imaging: a prospective survey

Zain Ul Abidin[1†], Rizwan Ali Naqvi[1†], Amir Haider[1], Hyung Seok Kim[1], Daesik Jeong[2*] and Seung Won Lee[3*]

[1]Department of Intelligent Mechatronics Engineering, Sejong University, Seoul, Republic of Korea, [2]College of Convergence Engineering, Sangmyung University, Seoul, Republic of Korea, [3]School of Medicine, Sungkyunkwan University, Suwon, Republic of Korea

Radiologists encounter significant challenges when segmenting and determining brain tumors in patients because this information assists in treatment planning. The utilization of artificial intelligence (AI), especially deep learning (DL), has emerged as a useful tool in healthcare, aiding radiologists in their diagnostic processes. This empowers radiologists to understand the biology of tumors better and provide personalized care to patients with brain tumors. The segmentation of brain tumors using multi-modal magnetic resonance imaging (MRI) images has received considerable attention. In this survey, we first discuss multi-modal and available magnetic resonance imaging modalities and their properties. Subsequently, we discuss the most recent DL-based models for brain tumor segmentation using multi-modal MRI. We divide this section into three parts based on the architecture: the first is for models that use the backbone of convolutional neural networks (CNN), the second is for vision transformer-based models, and the third is for hybrid models that use both convolutional neural networks and transformer in the architecture. In addition, in-depth statistical analysis is performed of the recent publication, frequently used datasets, and evaluation metrics for segmentation tasks. Finally, open research challenges are identified and suggested promising future directions for brain tumor segmentation to improve diagnostic accuracy and treatment outcomes for patients with brain tumors. This aligns with public health goals to use health technologies for better healthcare delivery and population health management.

KEYWORDS

deep learning, brain tumor segmentation, medical images, multi-modality analysis, vision transformers, convolutional neural network

1 Introduction

1.1 Background

The brain contains around one hundred billion neurons and is an essential organ in the human body (Stiles and Jernigan, 2010). Brain and other nervous system tumors are a significant cause of mortality in developed nations, ranking 10th among the leading causes of death (Siegel et al., 2023). This condition impacts individuals throughout various age

TABLE 1 Overview of MRI modalities.

Modality	Properties
T1	• T1 images provide good anatomical detail • Highlights differences in tissue composition • Sensitive to variations in proton density and T1 relaxation times • Brighter tissue denotes shorter relaxation time
T1ce	• T1ce images acquired after administering a contrast agent (e.g., gadolinium) • Contrast agent enhances regions with the disrupted blood-brain barrier • Helps identify areas of increased vascularity
T2	• T2 images are sensitive to variations in proton density and T2 relaxation times • Emphasizes differences in tissue water content • Good for highlighting edema and lesions • Longer relaxation time is associated with brighter tissues
FLAIR	• Designed to suppress the signal from cerebrospinal fluid (CSF) • Particularly useful for highlighting abnormalities in white matter and gray matter • Minimizes the signal from the fluid

groups, including adults and children. According to estimates, the United States witnessed approximately 18,280 fatalities in 2022 due to primary brain tumors (Tahir et al., 2022). The brain comprises various cell types with individual characteristics, rendering generalizations concerning malignancies in other organs irrelevant (Charles et al., 2011). Common symptoms of brain tumors frequently encompass feelings of high blood pressure, severe fatigue, nausea attacks, physical discomfort with fever, skin eruptions, and increased cardiac pulsations. Although professionals attempt to establish a correlation between symptoms and a definitive diagnosis, it is essential to note that brain tumors do not consistently exhibit observable symptoms (Desjardins et al., 2019; Kotia et al., 2020).

Over the last few decades, researchers have conducted comprehensive fundamental research on brain tumors (Rao and Karunakara, 2021; Dhole and Dixit, 2022; Jyothi and Singh, 2023). The primary objective of this research is to understand biological properties and their transformation into malignant tumors. Over time, there has been significant progress in comprehending the genetic and molecular changes associated with brain tumors. This has significantly contributed to advancing novel methods for diagnosing and treating brain tumors. Additionally, researchers have explored the use of several imaging modalities, such as magnetic resonance imaging (MRI), to aid in identifying brain tumors and tracking their subsequent development (Guo et al., 2020; Yang et al., 2021). Due to its exceptional accuracy and clarity, MRI has emerged as the primary method for examining brain tumors. Consequently, this technological advancement has paved the way for innovative surgical techniques, including minimally invasive procedures (Privitera et al., 2022). These technological breakthroughs facilitate the accurate removal of brain tumors while minimizing damage to surrounding tissues. To be more specific, the primary objective of segmenting brain tumors is to accurately delineate different areas of tumors by modifying the representations obtained from MRI. The segmentation outcomes are subsequently applied to the prognosis and prediction of survival for brain malignancies.

The broad use of multi-modal MRI images in the segmentation of brain tumors has been facilitated by advancements in MRI technology. This method provides a detailed interpretation of the tumors and the neighboring tissues. MRI includes four unique modalities: T1-weighted (T1), T2-weighted (T2), T1-weighted with contrast enhancement (T1ce), and fluid attenuation inversion recovery (FLAIR). These modalities provide extra information for diagnosing and monitoring brain tumors (Menze et al., 2014). Table 1 provides a detailed overview of these modalities along with their properties, and Figure 1 shows the MRI modalities of brain tumors. The T1 is frequently utilized to generate high-resolution brain images. On the other hand, T2 is useful for evaluating the fluid content in tissues, which serves as a key differentiator between tumors and healthy brain tissue. Additionally, the T1ce provides relevant details on the vascular structures and the enhancing characteristics of tumors, hence facilitating the classification of tumor types (Wang et al., 2023a).

Integrating several MRI modalities provides a comprehensive and accurate depiction of tumors and adjacent brain tissue, which is essential for successful segmentation (Salvador et al., 2019). By employing multi-modal MRI images, researchers can assess the efficacy of various segmentation algorithms and make comparisons of their outcomes. This comparative study aims to stimulate the development of novel methodologies and improve the precision of brain tumor segmentation. The Brain Tumor Segmentation (BraTS) Challenge dataset is generally recognized as the principal resource for assessing brain tumor segmentation (Menze et al., 2014). The dataset consists of a wide array of MRI modalities, such as T1, T2, T1ce, and FLAIR, accompanied by precisely annotated tumor segmentation masks. The BraTS dataset is a significant resource for academics and clinicians involved in segmenting gliomas and diagnosing brain tumors.

Recent advancements in deep learning (DL) have significantly enhanced the capabilities of computer-aided analysis in various domains. In particular, the segmentation of multi-modal brain tumors has witnessed substantial progress, offering a plethora of techniques with varying degrees of accuracy and effectiveness (Rao and Karunakara, 2021; Dhole and Dixit, 2022; Jyothi and Singh, 2023). Initially, brain tumor segmentation relied on manual tracking, where skilled practitioners manually delineated tumor boundaries on medical images. However, this method is time-consuming and prone to inter-observer variability. As computer vision gains the limelight, automatic ways to separate brain tumors have become more popular. There are two main groups of these methods: traditional and DL methods. Some examples of traditional methods are atlas-based segmentation and region-growing level-set methods (Hamamci et al., 2011; Hamamci et al., 2012). Such approaches employ things about the image, like sharpness, color, etc., to segment the tumor from the surrounding tissue.

DL methods, especially convolutional neural networks (CNNs), have gotten much attention lately for how well they work at segmenting brain tumors than traditional methods. Traditional models estimate the tumor borders and locations

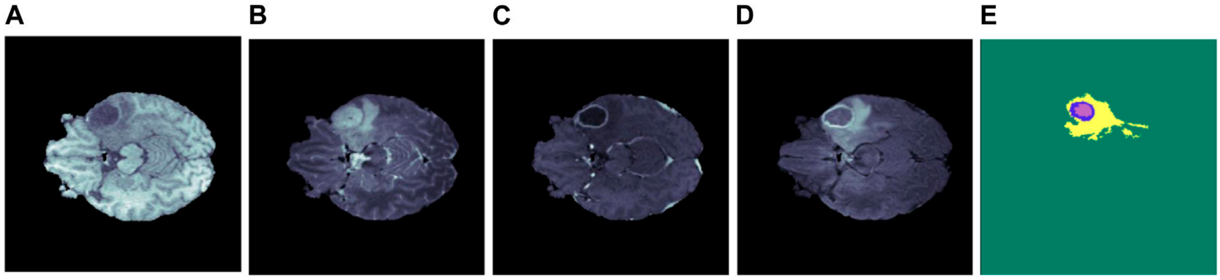


FIGURE 1
Illustration of several brain MRI modalities **(A)** T1, **(B)** T2, **(C)** T1ce, **(D)** FLAIR, and **(E)** ground truth. The yellow, blue, and purple colors in the ground truth represent edema, enhancing, and necrosis, respectively.

using statistical learning techniques (Ilhan and Ilhan, 2017; Biratu et al., 2021; Khilkhal et al., 2022; Nyo et al., 2022). These models rely on preprocessing techniques to improve the quality and clarity of the tumor images before the tumor lesions are delineated. Traditional models use these techniques to help with the later investigation, characterization of the tumor, and precise estimations of tumor boundaries. On the other hand, CNN leverages DL techniques to autonomously learn hierarchical representations of features directly from the data (Pereira et al., 2016). This enables CNNs to adapt and optimize their performance based on the specific characteristics of brain tumor images, ultimately leading to more precise and reliable segmentation results compared to traditional approaches. Recently, vision transformers have made amazing progress and are now better at separating brain tumors into their different parts (Liu et al., 2022a; He et al., 2022). Some researchers have employed transformer layers that integrate a self-attention mechanism featuring multiple heads, aiming to capture additional distinctive global characteristics. Meanwhile, other researchers have devised transformer-based modules for modal fusion, facilitating the alignment of multi-modal inputs and enhancing the integration of diverse data types. This approach simplifies the process of segregating multi-modal MRI data. The objective of these studies is to discover more effective methods for visualizing multi-modal brain tumors and leveraging appropriate data to enhance tumor segmentation outcomes.

In this section, we highlight the significance of brain tumor segmentation through statistical insights. The discussion then shifts toward the importance of research conducted in the last decade and the use of MRI. Following this, we explored various modalities of MRI, such as T1, T2, T1ce, and FLAIR, along with their properties and utilization of these modalities in brain tumor segmentation. In the end, we explored the various techniques used in brain tumor segmentation and the superiority of DL-based methods. Acknowledging the complexities associated with segmenting multi-modal brain MRI due to inherent challenges, the ultimate objective of the study becomes clearer: to provide a comprehensive overview of recent DL-based models from 2021 to 2023 designed to segment brain tumor lesions in multi-modal MRI autonomously. The list of all used abbreviations is summarized in Table 2.

1.2 Related work

AI has demonstrated notable advancements in medical imaging, specifically in image processing and computer vision. AI models have emerged as a powerful tool for automating tasks like classifying, detecting, and segmenting tumor lesions. Thus, utilizing the capabilities of AI, these models improve the accuracy of segmentation results, consequently improving the quality of patient care. Current research contains comprehensive surveys that dive into cutting-edge advancements, particularly multi-modality MRI segmentation. Table 3 concisely summarizes previous studies conducted on the segmentation of brain tumors using multi-modality MRI with their essential features and weaknesses while briefly describing our proposed survey. The research conducted in (Wang et al., 2023b) provides a comprehensive overview of state-of-the-art (SOTA) vision transformers (ViT) employed in the segmentation of multi-modal brain MRI, along with the associated challenges and their potential future directions. However, it focuses more on statistical analysis of brain tumor segmentation.

In (Liu et al., 2023), authors analyzed various DL methods for brain tumor segmentation. In (Mohammed et al., 2023), authors analyzed machine learning, DL, and hybrid techniques for brain tumor segmentation using multi-modality MRI. However, the discussion on challenges and their future direction was neglected. In (Ranjbarzadeh et al., 2023), the author analyzed the supervised and unsupervised DL models used in multi-modal brain tumor segmentation. However, they did not cover the main limitations and possible ways forward. In (Ali et al., 2022), the authors covered BraTS challenges from 2012 to 2020 but did not discuss the problems with BraTS challenges. and the survey is more specific to the BraTS challenges than the architectural and performance improvements. Our proposed review aims to conduct a comprehensive analysis and comparison of DL-based methods and their architectures for multi-modal MRI. Additionally, we will include statistical analyses of recent research articles, widely utilized datasets, evaluation measures, and a thorough comparison of segmentation performance. To address existing knowledge gaps and improve the reliability and efficiency of DL-based models for multi-modal brain tumor lesion segmentation using MRI, we will emphasize open research challenges and suggest potential future directions.

TABLE 2 List of abbreviations.

Abbreviation	Full form	Abbreviation	Full form	Abbreviation	Full form
AI	Artificial Intelligence	DL	Deep Learning	MRI	Magnetic Resonance Imaging
FLAIR	Fluid Attenuation Inversion Recovery	T1	T1-weighted	T1ce	T1-weighted with contrast enhancement
T2	T2-weighted	BraTS	Brain Tumor Segmentation	CNN	Convolutional Neural Network
ViT	Vision Transformer	MAAB	Multiple Atrous Convolutions Attention Block	MM-BiFPN	Multimodal Fusion Bi-directional Feature Pyramid Network
CMFT	Cross-Modality Feature Transition	CMFF	Cross-Modality Feature Fusion	FeG	Feature-enhanced Generator
CC	Correlation Constraints	RFNet	Region-aware Fusion Network	RFM	Region-aware Fusion Module
AABTS-Net	Axial Attention Brain Tumor Segmentation Network	MCC	Modality-level Cross Connection	AFFM	Attentional Feature Fusion Module
MSFF	Multi-Scale Spatial Feature Fusion	DCFF	Dual Path Channel Feature Fusion	MAF-Net	Modality-Level Attention Fusion Network
DP	Dual Path	MAF	Multi-scale Attention Fusion	IDCM	Iterative Dilated Convolution Merging
mPMRI	Multi Parametric MRI	FFCM	Fast Fuzzy C Means	WT	Whole Tumor
TC	Tumor Core	ET	Enhance Tumor	MFD-Net	Modality Fusion Diffractive Network
GAM-Net	Gradient Assisted Multi-category Network	MSFR-Net	Multi-modality and Single-modality Feature Recalibration Network	DRM	Dual Recalibration Module
ViTBIS	Vision Transformer for Biomedical Image Segmentation	NMaFA	Nested Modalityaware Feature Aggregation	EMSViT	Efficient Multi-Scale Vision Transformer
MMCFormer	Missing Modality Compensation Transformer	TC-inception	Transformer-Convolution Inception mechanism	CAFGL	Cross-Attention Fusion with a Global and Local feature
SCCAF	Skip Connection with Cross-Attention Fusion	GAN	Generative Adversarial Network	AST	Axial Spatial Transformer
MLP	Multilayer Perceptron	VAE	Variational Autoencoder	CBAM	Convolution Block Attention Module
ESAB	Edge Spatial Attention Block	MFIB	Multi-Feature Inference Block	F2 Net	Flexible Fusion Network
CFM	Cross-modal Feature-enhanced Module	MCM	Multi-modal Collaboration Module	GSP	Ghost Spatial Pyramid
GSA	Ghost Self Attention	DRG	Dense Residual Ghost	TP	True Positive
FP	False Positive	FN	False Negative	TN	True Negative
DSC	Dice Similarity Coefficient	IoU	Intersection over Union	HD	Hausdorff distance
NLP	Natural Language Processing	MICCAI	Medical Imaging Computing and Computer-Aided Intervention Association	CE	Cross Entropy
BCE	Binary Cross Entropy	WCE	Weighted Cross Entropy	PWCE	Pixel Wise Cross Entropy
SHAP	SHapley Additive exPlanations	Lime	Local Interpretable Model-agnostic Explanations	XAI	Explainable AI

1.3 Contributions

This review predominantly focuses on using DL in multi-modal brain tumor MRI segmentation. Presently, DL demonstrates exceptional proficiency in this, exhibiting SOTA performance. In addition, we endeavor to examine the existing challenges and provide potential direction for future research from diverse perspectives. In summary, this review presents the subsequent significant contributions.

- We investigate several aspects of current DL-based methodologies employed for brain tumor segmentation.

TABLE 3 Summary of related survey articles.

Reference	Essential features	Weaknesses	Year
Wang et al. (2023b)	• Provide SOTA analysis of ViT for brain tumor segmentation • Include brain tumor databases	• It focuses more on the statistical analysis of the model	2023
Liu et al. (2023)	• Analysed DL-based multi-modality MRI brain tumor segmentation • Future trends were discussed	• Do not include a discussion on multi-modality brain MRI.	2023
Mohammed et al. (2023)	• Analysed machine learning, DL, and hybrid techniques for brain tumor segmentation using multi-modality MRI.	• Focus more on a general discussion of various techniques than architecture • Discussion on existing challenges and their possible future direction is neglected	2023
Ranjbarzadeh et al. (2023)	• Described machine learning and DL models • Overview of performance measures used in the segmentation	• Do not include the state-of-the-art vision transformers • Discussion on existing challenges and their possible future direction is neglected	2022
Ali et al. (2022)	• Discussed BraTS challenges from 2012 to 2018	• Do not discuss the problems in the BraTS challenges • The survey is more specific to the BraTS challenges than the architectural and performance improvements	2019
Proposed Survey	• Examine the prior research, which is based on CNN, transformer, and hybrid models, and cover their architecture in depth • Perform a thorough comparison of the multi-modal brain tumor segmentation model's performance • Provide statistical analysis of recent research articles, widely utilized datasets, and evaluation metrics • Highlight open research challenges for brain tumor segmentation using multi-modal MRI images and suggest a possible future direction		

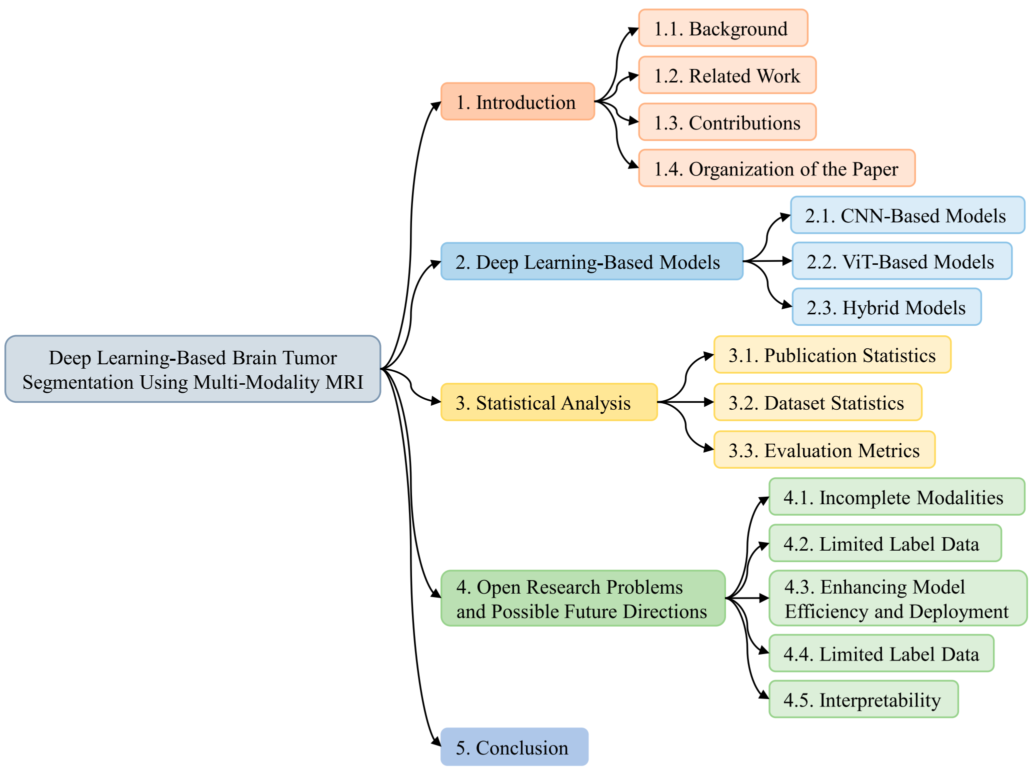


FIGURE 2
Organization of the DL-based brain tumor segmentation using multi-modality MRI survey.

These aspects encompass the background, datasets utilized, models employed, and current progress trends in this field.

- We summarize the most recent CNN-based, transformer-based, and hybrid models to segment brain tumors. These models specifically focus on utilizing multi-modal MRI data and thoroughly comparing segmentation performance.
- Our study offers an extensive statistical analysis of recent research articles, widely utilized datasets, and evaluation metrics used in the multi-modal brain tumor segmentation.

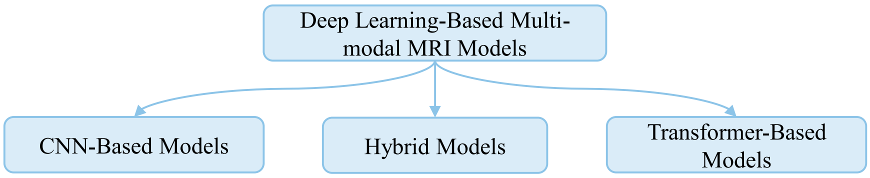


FIGURE 3
Classification of DL-based multi-modal MRI brain tumor segmentation models.

- We highlight open research challenges for DL-based brain tumor segmentation using multi-modal MRI images and suggest a possible future direction, emphasizing extending the ability to enhance segmentation.

1.4 Organization of the paper

This study is structured to provide a comprehensive understanding of multi-modal brain tumor segmentation. Each section highlights the various aspects involved in segmenting and evaluating brain tumors using multi-modal MRI, as shown in Figure 2. Section 1 provides an overview of this study. This section is divided into four subsections: background, related work, contributions, and organization of the paper. In Section 2, recent SOTA studies focusing on DL-based brain tumor segmentation using MRI are described. This section is divided into three subsections based on the model architecture: CNN, vision transformer, and hybrid models. Section 3 comprises a comprehensive statistical analysis and is divided into three subsections: publication statistics, datasets statistics, and evaluation metrics. Section 4 highlights some open research challenges in DL-based multi-modal MRI brain tumor segmentation and proposes possible future directions. Finally, Section 5 concludes the survey.

2 Deep learning-based multi-modality MRI brain tumor segmentation models

Medical image analysis has experienced an enormous revolution in recent years with the advent of powerful DL models. This paradigm change is most visible and important in multi-modal MRI brain tumor segmentation. The precise and efficient delineation of brain tumors is critical for clinical diagnosis, therapy planning, and ongoing patient monitoring (Philip et al., 2022). In response to this necessity, this analysis thoroughly examines improvements in DL-based models specifically designed for the challenging task of segmentation of brain tumors using MRI data. These models have demonstrated unparalleled success in extracting meaningful features from the diverse information encapsulated in MRI modalities by incorporating cutting-edge technologies such as CNNs, ViT models, and innovative hybrid architectures, combining both strengths. Figure 3 shows the classification of DL-based models according to the architecture and organization of this section.

The research highlights the technical complexities inherent in the models and underscores their novel influence on transforming the domain of medical imaging analysis. The ongoing purpose for improved accuracy, generalization, and interpretability urges scholars to continuously analyze and develop new systems and methodologies. The core purpose of this endeavor is to enhance the practicality of brain tumor segmentation models using MRI data, hence introducing an era of effectiveness in patient treatment. A detailed assessment of the DL-based models for multi-modal MRI brain tumor segmentation indicates a dynamic interaction of evolving architectures. Researchers are exploring this challenging domain with a desire for innovation, from the impressive power of CNNs to the transformational promise of vision transformer models. The use of CNNs, with their inherent capacity to acquire hierarchical features automatically, has prepared the path for ground-breaking advances.

The section on CNN-based models delves into the different architectural details, training methodologies, and data that explain their effectiveness. Concurrently, the introduction of vision transformer models marked a new era in image processing. Vision transformers provide a new viewpoint for feature extraction and fusion in multi-modal MRI data by relying on self-attention processes and the capacity to perceive global contextual information. The investigation of vision transformer-based models helps to reveal the distinct characteristics and exciting possibilities they bring ahead. Recognizing the combined effect and complementarities of both CNNs and vision transformers, the section on hybrid models delves into how these integrated architectures strive to strike an optimal balance between local and global information, aiming to push the boundaries of accuracy and robustness in brain tumor segmentation. The ongoing interaction of these diverse techniques and the dynamic growth of model architectures highlight the vibrant field of DL in multi-modal MRI brain tumor segmentation, offering a future where precision and clinical relevance merge to improve patient outcomes.

2.1 CNN-based models

CNNs have emerged as important tools in the fast-developing field of medical image processing, showing exceptional proficiency across a wide range of imaging applications (Singha et al., 2021). Within the domain of multi-modal MRI brain tumor segmentation, CNN-based models distinguish themselves through their exceptional performance and ongoing evolution. This section embarks on a detailed exploration, delving into the complexities of the various architectures and tactics used by CNN-based models. The exploration intends to highlight the complexity inherent in multi-modal MRI data, where the fusion of many imaging modalities brings distinct challenges. The present study provides nuanced insights into the CNNs for brain tumor segmentation, ranging from the typical effectiveness of 2D CNNs to the more advanced and volumetric capabilities of 3D variants.

One of the most important CNN-based architectures is UNet, designed for semantic segmentation tasks, particularly in medical image analysis (Ronneberger et al., 2015). It was introduced by Ronneberger et al., in 2015 and has since become widely used due to its effectiveness in producing accurate segmentation masks while efficiently handling limited training data. The UNet architecture consists of a contracting path, which captures context and reduces resolution, followed by an expanding path, which enables precise localization. Its unique feature is the skip connections that concatenate feature maps from the contracting path to the corresponding layers in the expanding path. These skip connections help preserve spatial information, allowing the model to produce detailed segmentations even for small structures in the input images. Despite its success, the original UNet architecture has certain limitations, such as struggles with handling class imbalance (Oktay et al., 2018) and difficulties in segmenting objects of varying sizes effectively (Zhou et al., 2018).

The authors in (Akbar et al., 2021) enhanced the UNet model for brain tumor segmentation by including attention, multiple atrous convolutions, and a residual route. This modified version is referred to as the Multiple Atrous Convolutions Attention Block (MAAB). The expansion part is included by extracting pyramid characteristics from each level and using them to generate the ultimate segmentation result. In (Syazwany et al., 2021), the authors proposed a multi-modal fusion network that incorporates a bi-directional feature pyramid network (MM-BiFPN). This network performs feature extraction from each modality using a separate encoder. The main objective is to use the intricate interactions across these modalities effectively. Furthermore, via the use of the bi-directional feature pyramid network (Bi-FPN) layer, they specifically concentrate on the combination of various modalities to examine the interrelationship between different modalities and the features at numerous scales.

The work in (Zhang et al., 2021) presented an innovative approach for segmenting brain tumors using a cross-modality deep feature learning framework. The fundamental concept is to extract valuable patterns to compensate for the limited amount of data available. The proposed framework for deep feature learning across different modalities comprises two distinct learning processes: the cross-modality feature transition (CMFT) process and the cross-modality feature fusion (CMFF) process. The CMFT process focuses on transferring knowledge between different modalities to learn comprehensive feature representations. On the other hand, the CMFF process aims to merge knowledge from various modalities to enhance the feature representations.

In (Fang et al., 2021), the proposed framework utilizes a hybrid fusion technique to combine data from different modalities. The authors also include a self-supervised learning method in this approach, and it relies on a fully CNN. Initially, they provide an architecture with multiple inputs that acquire distinct characteristics from multi-modal data. The model outperforms single-modal multi-channel networks by offering an improved feature extractor for segmentation tasks. This feature extractor effectively captures cross-modal information from multi-modal input. Furthermore, they provide a novel method for combining features, which they refer to as hybrid attentional fusion. This technique allows the acquisition of the hybrid representation of various characteristics and the collection of correlation information via an attention mechanism. Contrary to commonly used techniques like feature map concatenation, this approach has a complementary nature of multi-modal data, resulting in remarkable progress in the segmentation outcomes of certain areas.

The study in (Zhou et al., 2021) introduced an innovative neural network for segmenting brain tumors when one or more modalities are absent. The network has three sub-networks: a feature-enhanced generator (FeG), a correlation constraint (CC), and segmentation. The FeG employs the existing modalities to create a three-dimensional image that enhances the features and represents the missing modality. The CC block leverages the multi-source correlation and restricts the generator to produce a modality enriched with features consistent with the existing modalities. The segmentation network utilizes a U-Net architecture with multiple encoders to perform brain tumor segmentation accurately. In (Wang et al., 2021b), the authors developed an innovative end-to-end modality-pairing learning approach for segmenting brain tumors. The goal of paralleled branches is to use distinct modality traits, while a network of layer connections is employed to collect intricate interactions and ample information across modalities. In addition, they use consistency loss to reduce the variability in predictions across two branches. Finally, they use an average ensemble of different models together with various post-processing approaches to obtain the ultimate outcomes.

The authors in (Ding et al., 2021a) introduced a Region-aware Fusion Network (RFNet) that can intelligently and efficiently use various combinations of multi-modal data for tumor segmentation. The researchers have developed a Region-aware Fusion Module (RFM) in RFNet to combine features from multiple image modalities based on specific brain tumor locations since different modalities are sensitive to different regions. RFNet utilizes RFM to intelligently segment tumor areas from a limited collection of multi-modal images by efficiently combining modal data. In addition, they also create a segmentation-based regularizer to address the issue of inadequate and imbalanced training in RFNet due to missing multi-modal data. More precisely, in addition to acquiring segmentation outcomes from combined modal features, they also segment each imaging modality separately using the associated encoded features. By using this approach, every modal encoder is compelled to acquire distinguishing characteristics, hence enhancing the capacity of the combined features to represent information.

The CNN model in (Tong and Wang, 2023) has a distinctive architecture with two prominent characteristics. The feature extraction block has three pathways to extract full feature information from the multi-modality input. Each path is responsible for extracting features from mono-modality, paired-modality, and cross-modality data. Furthermore, it possesses a distinct tri-sectioned categorization system to differentiate pixels belonging to three intra-tumoral groups from the surroundings. The branches are trained individually to ensure that the updating process is applied to the parameters precisely using the matching annotations of the target tumor locations. In (Zhao L. et al., 2022), a multi-modality feature fusion network called MM-UNet was developed. This network utilizes a structure with several encoders and a single decoder to perform brain tumor segmentation. Within the proposed network, individual encoders autonomously extract low-level characteristics from their respective imaging modalities, while the hybrid attention block enhances the features. The decoder utilizes skip connections to include high-level semantic information and provide accurate pixel-level segmentation results.

The researchers in (Tian et al., 2022) devised an axial attention brain tumor segmentation network (AABTS-Net) to automatically delineate tumor sub-regions using multi-modality MRIs. The axial attention mechanism aids in the acquisition of deeper semantic information, facilitating models by offering both local and global information while reducing computing complexity. The use of the deep supervision mechanism serves the purpose of preventing the occurrence of vanishing gradients and providing guidance to the AABTS-Net to provide enhanced feature representations. The authors in (Zhou et al., 2023) introduced a modality-level cross-connection (MCC) network, which is a 3D UNet based on several encoders designed for brain tumor segmentation. The MCC network leverages beneficial information between the different modalities. Additionally, to improve its ability to learn features, the researchers introduced the attentional feature fusion module (AFFM). This module combines many modalities and extracts valuable feature representations for segmentation. The AFFM comprises two main elements: the multi-scale spatial feature fusion (MSFF) block and the dual-path channel feature fusion (DCFF) block. Their objective is to acquire multi-scale spatial and channel-wise feature information to enhance the accuracy of segmentation.

The authors in (Liu et al., 2022) proposed a multi-modal image fusion approach that combines pixel- and feature-level fusion to improve the effectiveness and precision of brain tumor segmentation. The goal is to enhance the exploitation of multi-modal information. They introduced a convolutional network called PIF-Net for 3D MR image fusion at the pixel level, enhancing the segmentation model's input modalities. The integration of numerous source modalities might increase the correlation between various forms of disease information, resulting in an amplification of modality effects. At the feature level, attention-based modality selection feature fusion is designed to improve multi-modal features by addressing the variations among different modalities for a certain segmentation objective. In the (Huang et al., 2022), the authors introduced a modality-level attention fusion network (MAF-Net), which uses patchwise contrastive learning to extract latent features from several modalities. Additionally, attention weights are dynamically assigned to fuse the distinct modalities uniquely.

The work in (Chang et al., 2023) introduced a 3D segmentation model called DPAFNet. This model is based on integrating a dual-path (DP) module and a multi-scale attention fusion (MAF) module. The DPAFNet utilizes DP convolution to expand capacity and incorporates residual connections to prevent deterioration. An attention fusion module combines global and local information at the channel level. This module fuses feature maps of various sizes to provide enriched features with enhanced semantic information. This prioritizes the comprehensive examination of tiny cancers. In addition, the 3D iterative dilated convolution merging (IDCM) module also enhances the receptive field and contextual awareness.

A novel approach is presented in (Sahoo et al., 2023), which combines the Inception V2 network with 16 newly developed layered segmentation nets to create a hybrid deep neural network. The network undergoes testing using the BraTs 2020 and BraTs 2017 multi-parametric MRI (mPMRI) datasets to identify the whole tumor. To recognize the tumor core (TC) and the edema, the fast fuzzy C-means (FFCM) algorithm is used. In (Hou et al., 2023), the authors proposed a modality fusion diffractive network (MFD-Net) for accurately and automatically segmenting brain tumors. The MFD-Net consists of diffractive blocks and modality feature extractors. The diffractive block, constructed using Fraunhofer's single-slit diffraction principle, highlights nearby feature points with high confidence while reducing the prominence of low-quality or isolated feature points. This improves the interconnectedness of the features. Adopting a global passive reception mode resolves the problem of fixed receptive fields. The self-supervised technique efficiently exploits the inherent generalization information of each modality to extract modality features. This allows the main segmentation branch to prioritize the fusion of multi-modal feature information.

The work in (Çetiner and Metlek, 2023) introduced DenseUNet+, a novel DL method for achieving precise segmentation of multi-modal images. The DenseUNet + model included data from four distinct modalities in dense block structures. Subsequently, the data underwent linear operations followed by the concatenate operation. The findings acquired using this method were transmitted to the decoder layer. In (Wang et al., 2023), the authors introduced a novel segmentation network called a gradient-assisted multi-category network (GAM-Net). GAM-net consists of three components: a double convolution encoder, a gradient extraction branch, and a gradient-driven decoder. A double convolution encoder extracts detailed features from MRI images; a gradient extraction branch generates gradient features to aid in area segmentation, and a gradient-driven decoder effectively combines contour information and encoding features.

The researchers in (Li et al., 2023a) introduced multi-modality and single-modality feature recalibration network (MSFR-Net). Distinct pathways handle the flow of multi-modality and single-modality information. The multi-modality network captures the correlations relating to different modalities and various tumor sub-components. A single-modality network is trained to understand the connection between a single modality and its closely related tumor subcomponents. Subsequently, a dual recalibration module (DRM) is devised to establish a connection between the parallel single-

TABLE 4 CNN-based models for multi-modal MRI brain tumor segmentation.

Segmentation models	Dataset	Experimental parameters		Segmentation performance			Ref.
		Optimizer	Loss function	WT	TC	ET	
UNet with Multiple Atrous convolutions Attention Block (MAAB)	BraTS 2021	Adam	dice	DSC = 0.884 HD = 10.70	DSC = 0.829 HD = 23.01	DSC = 0.817 HD = 19.70	Akbar et al. (2021)
Multi-Modality Bi-directional Feature Pyramid Network (MM-BiFPN)	BraTS 2018	Adam	CE	DSC = 0.811	DSC = 0.777	DSC = 0.735	Syazwany et al. (2021)
	BraTS 2020			DSC = 0.836	DSC = 0.815	DSC = 0.779	
Cross Modality Deep Feature Learning	BraTS 2017	Adam	-	DSC = 0.898 HD = 5.155	DSC = 0.823 HD = 6.999	DSC = 0.762 HD = 3.170	Zhang et al. (2021)
	BraTS 2018			DSC = 0.903 HD = 4.998	DSC = 0.836 HD = 6.639	DSC = 0.791 HD = 3.992	
Self-Supervised Learning Model	BraTS 2019	Adam	CE and dice	DSC = 0.927 HD = 2.446	DSC = 0.895 HD = 1.783	DSC = 0.835 HD = 1.623	Fang et al. (2021)
Feature Enhance Generation and Multi-Modality Fusion Based Network	BraTS 2018	Nadam	dice	DSC = 0.866	DSC = 0.858	DSC = 0.769	Zhou et al. (2021)
Madality-Paring Learning	BraTS 2020	SGD	CE and dice	DSC = 0.891 HD = 6.24	DSC = 0.842 HD = 19.54	DSC = 0.816 HD = 17.79	Wang et al., 2021b
Region-aware Fusion Network (RFNet))	BraTS 2015	Adam	WCE and dice	DSC = 0.861	DSC = 0.719	DSC = 0.589	Ding et al. (2021a)
	BraTS 2018			DSC = 0.857	DSC = 0.765	DSC = 0.571	
	BraTS 2020			DSC = 0.869	DSC = 0.782	DSC = 0.615	
CNN Model with Feature Extraction (FE) Block	BraTS 2018	-	-	DSC = 0.886	DSC = 0.801	DSC = 0.787	Tong and Wang (2023)
	BraTS 2019			DSC = 0.885	DSC = 0.776	DSC = 0.751	
Multi-Modality Fusion network (MM-UNet)	BraTS 2020	Adam	dice and focal	DSC = 0.850 HD = 8.243	DSC = 0.765 HD = 10.76	DSC = 0.762 HD = 6.389	Zhao et al. (2022a)
Axial Attention CNN for Brain Tumor Segmentation (AABTS-Net)	BraTS 2019	Adam	BCE and dice	DSC = 0.911 HD = 3.988	DSC = 0.838 HD = 6.028	DSC = 0.777 HD = 3.246	Tian et al. (2022)
	BraTS 2021			DSC = 0.922 HD = 3.996	DSC = 0.861 HD = 11.18	DSC = 0.830 HD = 17.73	
Modality-Level Cross Connection (MCC)	BraTS 2018	Nadam	dice	DSC = 0.865 HD = 4.60	DSC = 0.870 HD = 3.60	DSC = 0.794 HD = 2.50	Zhou et al. (2023)

(Continued on following page)

TABLE 4 (*Continued*) CNN-based models for multi-modal MRI brain tumor segmentation.

Segmentation models	Dataset	Experimental parameters		Segmentation performance			Ref.
		Optimizer	Loss function	WT	TC	ET	
Pixel level and Feature level Image Fusion Network	BraTS 2019	Adam	BCE and dice	DSC = 0.894 HD = 5.349	DSC = 0.814 HD = 10.89	DSC = 0.771 HD = 5.855	Liu et al. (2022b)
	BraTS 2020			DSC = 0.895 HD = 5.312	DSC = 0.817 HD = 9.429	DSC = 0.775 HD = 4.472	
Modality-level Attention Fusion Network (MAF-Net)	BraTS 2020	Adam	CE	DSC = 0.880	DSC = 0.679	DSC = 0.418	Huang et al. (2022a)
Pixel-level and Feature-level Image Fusion Network for Brain Tumor Segmentation	BraTS 2018	Adam	CE	DSC = 0.900 HD = 6.51	DSC = 0.839 HD = 5.71	DSC = 0.795 HD = 2.92	Chang et al. (2023)
	BraTS 2019			DSC = 0.890 HD = 8.53	DSC = 0.812 HD = 7.43	DSC = 0.782 HD = 3.82	
	BraTS 2020			DSC = 0.894	DSC = 0.832	DSC = 0.781	
Improve DNN with fast Fuzzy C-Means (FFCM)	BraTS 2017	-	-	DSC = 0.891	DSC = 0.847	DSC = 0.865	Sahoo et al. (2023)
	BraTS 2020			DSC = 0.904	DSC = 0.858	DSC = 0.865	
Modality Fusion Diffractive Network (MFD-Net)	BraTS 2018		BCE and dice	DSC = 0.908 HD = 5.986	DSC = 0.856 HD = 6.995	DSC = 0.767 HD = 3.409	Hou et al. (2023)
	BraTS 2019			SGD	DSC = 0.857 HD = 5.83	DSC = 0.767 HD = 3.41	
	BraTS 2021			DSC = 0.927 HD = 3.51	DSC = 0.887 HD = 5.77	DSC = 0.854 HD = 13.98	
DenseUNet + Model	FeTS 2021	Adam	dice	DSC = 0.883	DSC = 0.862	DSC = 0.865	Çetiner and Metlek (2023)
	BraTS 2021			DSC = 0.958	DSC = 0.955	DSC = 0.937	
Gradient Assisted Multi-Category (GAM-Net) Network	BraTS 2020	Adam	dice	DSC = 0.899 HD = 5.076	DSC = 0.840 HD = 5.096	DSC = 0.758 HD = 5.296	Wang et al. (2023c)
Multi-Modality and Single-Modality Feature Recalibration Network (MSFR-Net)	BraTS 2015	Adam	CE and dice	DSC = 0.860	DSC = 0.740	DSC = 0.650	Li et al. (2023a)
	BraTS 2018			DSC = 0.909 HD = 4.24	DSC = 0.858 HD = 6.72	DSC = 0.807 HD = 2.73	

10.3389/fbioe.2024.1392807

modality network and the multi-modality network at various phases. The purpose of the DRM is to integrate the two kinds of features into a single feature space.

In this subsection, the advancement of CNN-based brain tumor segmentation models using multi-modal MRI signifies notable progress in medical image analysis. Since its first implementation to the current advanced 3D versions, CNNs have been crucial in improving the precision of segmentation. However, their inherent limitation in capturing global characteristics has facilitated the development of later advancements. As we recognize the accomplishments and continued difficulties in this field, the persistent effort to improve CNN designs and methodologies highlights their ongoing importance in accurately and therapeutically useful brain tumor segmentation. Finally, the reviewed studies that used the CNN-based model are summarized in Table 4.

2.2 Vision transformer-based models

The introduction of ViT models represents an architectural change in image analysis, demonstrating effectiveness across several domains. Vision transformers provide a unique viewpoint for feature extraction and fusion by depending on self-attention mechanisms and acquiring global contextual information. This section examines the use of vision transformer models in brain tumor segmentation using multi-modal MRI, shedding light on their potential to improve segmentation accuracy and resilience in the context of multi-modal MRI data. In (Sagar et al., 2021a), the authors proposed a ViT for biomedical image segmentation (ViTBIS) model. The model divides input feature maps into three parts using 1×1, 3×3, and 5×5 convolutions in the encoder and decoder. The concatenation operator merges features before feeding them to three transformer blocks with attention mechanisms. Skip connections link encoder and decoder transformer blocks. Before linearly projecting the output segmentation map, decoders employ transformer blocks and a multi-scale architecture.

In (Pinaya et al., 2022), authors use vector quantized variational autoencoders' latent representation and an ensemble of autoregressive transformers to identify and segment unsupervised anomalies based on brain imaging data deviation at a low computing cost. They achieve improved image- and pixel-wise anomaly detection without post-processing. These findings highlight transformers' potential in this most difficult imaging job. The work in (Peiris et al., 2022a) presents a novel Transformer architecture designed specifically for volumetric segmentation. This is challenging as it effectively captures and incorporates local and global spatial inputs while conserving information across volume axes. The proposed design's encoder leverages a self-attention mechanism to simultaneously encode local and global cues. Meanwhile, the decoder utilizes a parallel formulation of self and cross-attention to effectively capture intricate features for boundary refinement. The proposed model is computationally efficient and exhibits competitive and promising outcomes when applied to the BraTS Task.

The authors of (Peiris et al., 2022b) introduced a model that constructs a U-shaped Volumetric Transformer (CR-Swin2-VT) using two well-known window-based attention mechanisms: the Cross-shaped window attention-based Swin Transformer block and the Shifted window attention-based Swin Transformer block. The CR-Swin2-VT model employs a parallel configuration of Swin Transformer blocks and CSWin Transformer blocks to capture voxel information on the encoder side. However, on the decoder side, only Swin Transformer blocks are utilized. In (Xing et al., 2022), authors presented a Nested Modality Aware Transformer (NestedFormer) that investigates the inter- and intra-modality relationships. They implemented modality-sensitive gating (MSG) at lower scales to facilitate more efficient skip connections and conduct nested multi-modal fusion for high-level representations of distinct modalities, utilizing a transformer-based multi-encoder and single-decoder architecture. Their proposed Nested Modality-aware Feature Aggregation (NMaFA) module provides the basis for performing multi-modal fusion. This module utilizes a cross-modality attention transformer to supplement critical contextual information among modalities and a tri-orientated spatial attention transformer to enhance long-term dependencies within individual modalities.

The authors in (Sagar et al., 2021b) proposed an efficient multi-scale ViT (EMSViT) that divides the input into three parts with various convolution sizes. Feature maps are merged before being fed into the three transformer blocks. In the decoder, transformer blocks and a multi-scale architecture are used to facilitate the linear projection of the input, resulting in the generation of the output segmentation map. In (Liu et al., 2022c), authors introduced a self-attention-based fusion block (SFusion). The proposed block automatically fuses available modalities without zero-padding missing ones. To produce latent multi-modal correlations, project feature representations from the upstream processing model as tokens and feed them into the self-attention module. The self-attention module generates latent multi-modal correlations from upstream processing model feature representations projected as tokens. A modal attention technique builds a common representation for the downstream decision model. The proposed SFusion integrates readily into multi-modal analytic networks, and they use SFusion on several backbone networks to segment brain tumors.

The authors in (Karimijafarbigloo et al., 2023) proposed the missing modality compensation transformer (MMCFormer) to handle missing information. They utilized 3D-efficient transformer blocks and co-training to efficiently train a missing modality network. MMCFormer uses global contextual agreement modules in each encoder scale to maintain feature consistency across many scales. Further, they used auxiliary tokens at the bottleneck stage to depict the interaction between full and missing-modality channels to transmit modality-specific concepts. Moreover, they included feature consistency losses to minimize network prediction domain gaps and enhance reliability for missing modality paths.

In this subsection, we investigate the ViT for brain tumor segmentation using multi-modal MRI. ViT overcomes a significant drawback of CNNs by using self-attention processes to gather global contextual information effectively. As we acknowledge the fundamental change introduced by transformers, their incorporation into the rapidly evolving field of medical image processing has significant potential to improve accuracy and reliability in brain tumor segmentation tasks. Finally, the reviewed studies that used the transformer model are summarized in Table 5.

2.3 Hybrid models

Hybrid models combine the benefits of both CNN and transformer. Many studies prefer to combine these two to improve the model's performance. CNNs struggle to capture global feature relations, affecting segmentation accuracy (Li et al., 2024). Thus, a Transformer network is developed, which can capture global information but not local details and requires pre-training on big datasets (Zhang et al., 2023). Therefore, the hybrid model overcomes the limitations by combining their strengths and aims to strike a superior balance between local and global information. The authors in (Wang et al., 2021) utilize a Transformer in 3D CNN for the first time and propose a TransBTS. To obtain the local 3D context information, the encoder initially extracts the volumetric spatial feature maps using 3D CNN. In the meantime, the tokens from the feature maps are precisely transformed and input into a Transformer to model global features. To predict the detailed segmentation map, the decoder employs progressive upsampling and utilizes the features embedded by the Transformer.

In (Li et al., 2021), the authors introduced Segtran, a transformer-based segmentation technique with infinite effective receptive fields at high feature resolutions. Segtran uses a unique squeeze-and-expansion transformer to regularize self-attention and learn diverse representations. Additionally, they introduced a transformer positional encoding method with a continuous inductive bias for images. The authors in (Jun et al., 2021) introduced a medical transformer, a transfer learning architecture that models 3D volumetric images as 2D image slices. For improved 3D-form representation of spatial relations, they utilized a multi-view technique that integrates information from the three planes of 3D volume and offers parameter-efficient training. They use a large-scale normal, healthy brain MRI dataset to pre-train a source model for masked encoding vector prediction, which may be used for numerous purposes.

The work in (Liang et al., 2022a) introduced a TransConver, a U-shaped segmentation network that utilizes convolution and transformer to provide automated and precise brain tumor segmentation in MRI images. In contrast to the transformer and convolution models that have been previously proposed, they have introduced a parallel module called transformer-convolution inception (TC-inception). This module utilizes convolution blocks to extract local information and transformer blocks to extract global information. These two types of information are integrated through a cross-attention fusion with a global and local feature (CAFGL) mechanism. The skip connection with cross-attention fusion (SCCAF) method is an enhanced structure that may mitigate the semantic disparities between encoder and decoder features, resulting in improved feature fusion.

In (Zhang et al., 2022a), the authors introduced a new multi-modal medical transformer (mmFormer) for incomplete multi-modal learning. It consists of three main parts: a hybrid modality-specific encoder that models both local and global contexts in every modality; An inter-modal transformer is designed to construct and synchronize long-range correlations among modalities to identify modality-invariant features that correspond to the global semantics of the tumor region; and a decoder that generates robust segmentation by progressive up-sampling and fusion with the modality-invariant features. Additionally, to make the model even more resistant to incomplete modalities, auxiliary regularizers are included in the encoder and decoder.

The authors in (Chen and Wang, 2022) introduced TSEUnet, a 3D nnUNet-based network. This network uses a parallel interactive transformer module in the encoder to extract local features and global contexts effectively. The decoder additionally uses SE-Attention to increase brain tumor segmentation and provide useful information. The authors in (Wang et al., 2022) designed a hybrid encoder-decoder that included lightweight convolution modules as well as an axial-spatial transformer (AST) module in the encoder. They intergrade axial and spatial attention in the AST module to capture better multi-view and multi-scale characteristics to learn long-range relationships, while convolution operations extract local dependencies and rich local characteristics.

To simplify the process of segmentation, the authors of (Liu et al., 2022) take advantage of a 2D backbone for segmenting a 3D brain tumor (Transition Net). To segment 3D brain tumor images, they make use of the Swin transformer as the encoder, in conjunction with a decoder that is produced by the process of 3D convolution. To address the issue of cross-domain variation, they developed the components known as the transition head to turn the input data into feature maps that are acceptable for Swin Transformer and the transition decoder to convert the multi-scale feature maps that were recovered by the backbone. After a series of stages, these maps are fused with the features sampled on CNN to obtain the final segmentation results.

In (Li et al., 2022), the authors aimed to use the Transformer model in a 3D CNN to segment 3D medical image volumes. They introduced a new model called TransBTSV2, built upon an encoder-decoder architecture. The proposed TransBTSV2 is not just restricted to brain tumors but emphasizes the broader medical image segmentation domain. It offers a more robust and efficient 3D foundation for the volumetric segmentation of medical images. TransBTSV2 is a hybrid CNN-transformer architecture that can accurately segment medical images without the need for pre-training. It integrates the strong permanent bias of CNNs with the excellent global context modeling capacity of transformers. By proposing a new approach to restructure the internal structure of the transformer block and introducing the deformable bottleneck module to capture shape-aware local information, they have produced a highly efficient architecture with higher performance.

The work in (Huang et al., 2022) introduced a generative adversarial network (GAN) based on transformers. To optimize the segmentation process, the network integrates the "generative adversarial" and "transformer" concepts. The generator network segments multi-modal MRI brain tumors using a transformer with the Resnet module in 3D CNN. The transformer and Resnet block efficiently capture local and global features, thereby facilitating the progressive upsampling of embedded features to generate full-resolution predicted maps. In (Liang et al., 2022b), the authors introduce an effective transformer-based model that incorporates a 3D parallel shifted window-based transformer module (3D PSwinBTS) to capture long-range contextual information.

Additionally, to achieve efficient semantic modeling, they make use of semantic supervision to incorporate eight semantic priors into the encoder of the 3D PSwinBTS model.

In (Jia et al., 2021), the authors proposed a combined CNN-transformer model called BiTr-UNet. It contains the main characteristics and backbone of TransBTS. They validated their model on the BraTS 2021 datasets and achieved good performance. The authors in (Dobko et al., 2021) modified the original TransBTS by adding more CNN layers, squeeze-and-excitation (SE) blocks, and trainable multilayer perceptron (MLP) embeddings instead of positional encoding in the transformer block. This modification enables the transformer to be adjusted to accommodate inputs of any size while performing inference. In addition, they chose to integrate our improved TransBTS into the nnU-Net framework by making architectural modifications to the nnUNet model according to our custom model.

The authors in (Pham et al., 2022) introduced a novel model called SegTransVAE, which utilizes an encoder-decoder design, including a transformer and a variational autoencoder (VAE) in the model. SegTransVAE is a multitask learning model that can simultaneously achieve brain tumor segmentation and image reconstruction. In (Yang et al., 2021), the authors proposed a convolution-and-transformer network (COTRNet) to accurately gather global information, along with the implementation of a topology-aware (TA) loss to restrict the learning process to topological information. In addition, they use transfer learning by using pre-trained parameters from ImageNet and implement deep supervision by including multi-level predictions to enhance segmentation performance.

In (Futrega et al., 2021), the authors introduced a segmentation model called Swin UNEt TRansformer (Swin UNETR). The objective of 3D brain tumor semantic segmentation is transformed into a prediction problem where multi-modal input data is converted into a one-dimensional sequence of embeddings. This series is then fed into a hierarchical Swin transformer, which serves as the encoder. The Swin transformer encoder employs shifted windows to compute self-attention and extract features at five distinct resolutions. The authors in (Wang et al., 2022) introduced a Trans-NUNet model, they used a convolution block attention module (CBAM) in the model to improve the performance of each model while dealing with images of varying sizes throughout the stage. The CBAM models provide rapid identification of the region of interest within the feature map by the whole network, followed by a thorough analysis of that specific area.

The authors in (Hu et al., 2023) proposed a novel combination of R-Transformer and U-Net, an efficient R-Transformer with dual encoders (ERTN). To capture global information and complicated semantic characteristics, ERTN builds a feature branch and a patch branch. To achieve accurate localization, the decoder augments low- and high-resolution CNN data with up-sampled features produced by the feature branch and patch branch. Finally, ERTN uses the Transformer's ranking attention mechanism (RTransformer), assisting the model in focusing on relevant data for enhanced training efficiency and decreased computing cost.

In (Zhu et al., 2023a), the authors proposed a model that fuses deep semantics with edge information. Semantic segmentation, edge detection, and feature fusion are the primary components of the proposed model. This module's semantic segmentation makes use of the Swin Transformer for feature extraction and introduces a shifting patch tokenization technique for enhanced training. A CNN-based edge detection module is introduced, together with an edge spatial attention block (ESAB) for feature improvement. They developed a graph convolution-based multi-feature inference block (MFIB) to conduct feature reasoning and information dissemination to achieve successful feature fusion in the feature fusion module, which is responsible for merging the derived semantic and edge features.

The study in (Gao et al., 2023) incorporates transformer layers into a U-shaped design's encoder and decoder using a deep mutual learning method. Due to the inherent complementarity between shallow features and deep features in a layer, where shallow features encompass plentiful spatial details but lack semantic information, conversely, the feature map of the shallowest layer is employed to guide the feature map of the deeper layers. This approach ensures that the deeper layers, which retain more edge information, guide the accuracy of sub-region segmentation. Employing the most profound classification logits to oversee the less profound logits to preserve a greater amount of semantic information for the differentiation of tumor sub-regions. Moreover, the shallow feature map and the deep logit mutually supervise each other, leading to an improvement in the overall accuracy of tumor segmentation.

The researchers in (Yang et al., 2023) introduced a flexible fusion network (F2 Net) for the segmentation of brain tumors. The F2 Net is built around an encoder-decoder structure, including two Transformer-based streams for feature learning and a cross-modal shared learning network to extract distinct and common feature representations. To efficiently incorporate information from multiple types of data, they suggested the use of a cross-modal feature-enhanced module (CFM) and a multi-modal collaboration module (MCM). The CFM is designed to combine features from different modalities in a shared learning network, while the MCM integrates features from encoders into a shared decoder.

The authors in (Lu et al., 2023) introduced a new 3D multi-scale Ghost CNN with an additional MetaFormer decoding path (GMetaNet). Efficient semantic information extraction was carried out Through the integration of CNN's localized modeling and the Transformer's capability for long-range representation. Three new modules are introduced, notably the lightweight Ghost spatial pyramid (GSP) module, the Ghost self-attention (GSA) module, and the dense residual Ghost (DRG) module, which are built upon the existing Ghost module. Furthermore, the GSP module efficiently acquires knowledge about various receptive fields to enhance the multiscale representation while reducing computational expenses. The GSA module allows the model to capture long-range relationships effectively. The DRG module, functioning as a local decoder, enhances information and prevents deterioration. Furthermore, a comprehensive decoder

TABLE 5 Vision transformer-based models for multi-modal brain tumor segmentation.

Segmentation models	Dataset	Experimental parameters		Segmentation performance			Year	Ref.
		Optimizer	Loss function	WT	TC	ET		
ViT for biomedical image segmentation (ViTBIS)	BraTS 2019	Adam	BCE and dice	DSC = 0.903 HD = 5.621	DSC = 0.822 HD = 7.129	DSC = 0.792 HD = 3.71	2021	Sagar et al. (2021a)
Vector Quantised Variational Autoencoder with an Ensemble of Autoregressive Transformer	BraTS 2018	-	-	Avg. DSC = 0.537 Avg. AUPRC = 0.555			2022	Pinaya et al. (2022)
Volumetric transformer UNet (VT UNet)	MSD	AdamW	-	DSC = 0.919 HD = 3.51	DSC = 0.872 HD = 4.10	DSC = 0.822 HD = 2.68	2022	Peiris et al. (2022a)
CR-Swin2-VT	FeTS	Adam	CE, dice and VAT	DSC = 0.914 HD = 3.93	DSC = 0.854 HD = 11.19	DSC = 0.817 HD = 14.81	2022	Peiris et al. (2022b)
Nested Modality-Aware Transformer (NestedFormer)	BraTS 2020	AdamW	CE and soft dice	DSC = 0.920 HD = 4.567	DSC = 0.864 HD = 5.316	DSC = 0.800 HD = 5.269	2022	Xing et al. (2022)
Efficient multi-scale ViT (EMSViT)	BraTS 2019	Adam	BCE and dice	DSC = 0.903 HD = 5.621	DSC = 0.822 HD = 7.129	DSC = 0.792 HD = 3.71	2022	Sagar et al. (2021b)
Self-attention based N-to-One multi-modal fusion (SFusion)	BraTS 2020	Adam	CE	DSC = 0.889	DSC = 0.822	DSC = 0.738	2023	Liu et al. (2022c)
Missing modality compensation transformer (MMCFormer)	BraTS 2018	Adam	dice	DSC = 0.890	DSC = 0.874	DSC = 0.801	2023	Karimijafarbigloo et al. (2023)

incorporating MetaFormer has been developed to combine local and global information successfully. Ultimately, the technique of deep supervision combines three outputs and enhances the rate at which the system reaches convergence.

To summarize, this section examines the latest research in brain tumor segmentation techniques, specifically focusing on the use of multi-modal MRI data. The field has seen a significant movement in segmentation methodologies, moving from the original use of CNNs to the introduction of transformers, and finally to the development of hybrid models. This transition has resulted in more comprehensive and effective segmentation techniques. The use of transformers, which excel at collecting global characteristics, complements the localized capabilities of CNNs in a mutually beneficial way. The ongoing development of multi-modal MRI brain tumor segmentation is driven by the junction of CNNs, transformers, and hybrid architectures, as we seek to achieve the most effective solutions. Finally, the reviewed studies that used the hybrid model are summarized in Table 6.

3 Statistical analysis

In this section, we will delve into DL-based brain tumor segmentation models with an emphasis on statistical insights. To commence, we look at the data, particularly focusing on the number of papers published in the preceding 3 years, spanning from 2021 to 2023. This analysis provides valuable insights into current trends and achievements, offering a glimpse into the pace of evolution within the field. Subsequently, we explore the datasets commonly utilized by researchers in modern brain tumor segmentation studies. Understanding these datasets is essential since they give actual data for testing DL models. It is similar to inspecting the tools in a toolbox: the more we understand them, the more efficiently we can utilize them. Finally, we outline the assessment criteria commonly employed by researchers to evaluate the performance of DL models in the task of multimodal brain tumor segmentation. These metrics serve as benchmarks, enabling us to gauge the efficacy of these models accurately.

3.1 Publication statistics

The field of brain tumor segmentation using DL models has seen tremendous advancements in recent years, with notable contributions from several architectures. In 2021, Dosovitskiy et al. (Dosovitskiy et al., 2020) introduced the vision transformer, which successfully applied the transformer architecture from natural language processing (NLP) to computer vision, marking a significant advancement. This pioneering research marked the beginning of the effective use of transformers in areas outside natural language processing (NLP), expanding into other computer vision tasks such as image classification, segmentation, and detection. Since the introduction of the vision transformer, the field of DL models has seen a significant increase in innovation, with the emergence of models that use transformers, CNNs, and hybrid architectures. This survey presents a thorough overview of brain tumor segmentation methodologies based on CNN, transformer, and hybrid models between 2021 and 2023. Figure 4 graphically represents the patterns in publications over this time, demonstrating the continuous shifts and diverse contributions from all these models.

3.2 Dataset statistics

The presence of multi-modal MRI datasets is indispensable for the effective evaluation of DL-based brain tumor segmentation models. Beginning in 2012, the Medical Imaging Computing and Computer-Aided Intervention Association (MICCAI) initiated the annual BraTS challenge. This longstanding challenge serves a pivotal role in fostering research and establishing a benchmark for evaluating brain tumor segmentation methods in the field. The BraTS challenge provides a standardized multi-modal MRI dataset consisting of four distinct scans - T1, T1, T2, and FLAIR. These modalities collectively offer a comprehensive view of brain anatomy and pathology, enabling researchers to develop and assess DL-based brain tumor segmentation methods. The influence of the BraTS challenge on research methodologies is profound, with a majority of studies opting to utilize BraTS datasets for training and testing their segmentation approaches.

Figure 5 provides a quantitative analysis of the utilization of multi-modal MRI datasets in DL-based models over the past 3 years. Notably, over 97% of studies have leveraged BraTS datasets, with BraTS 2018, 2019, and 2020 emerging as the most commonly employed versions. While a few studies incorporate private datasets for segmentation performance comparisons, the prevailing trend emphasizes the use of publicly accessible BraTS datasets. The widespread availability and standardized nature of BraTS datasets make them the preferred choice, despite challenges posed by private datasets, such as the labor-intensive pixel-level annotations. As we anticipate future studies, the overarching trajectory is expected to continue toward the refinement and advancement of brain tumor segmentation methods utilizing the established and publicly accessible BraTS datasets. Table 7 provides the top BraTS databases. mostly used in the evaluation of brain tumor segmentation.

3.3 Evaluation metrics

Evaluation metrics are quantitative measures used to assess the performance of a segmentation model, as they provide objective insights into how well a particular model performs compared to the ground truth. The segmentation model used a binary classification method in which each pixel belongs to either the tumorous or non-tumorous regions, usually represented as 1 and 0, respectively. From an input image, we obtained the segmentation results produced by the segmentation model and compared them with the ground truth created by experts. Numerous quantitative segmentation assessment metrics can be produced using the true positive (TP), false negative (FN), false positive (FP), and true negative (TN) metrics. A TP is an outcome in which the model correctly predicts a positive class, whereas an FP is an outcome in which the model incorrectly predicts a positive class. Similarly, TN is the outcome in which the model correctly predicts the negative class, whereas FN is the outcome in which the model incorrectly predicts the negative class. The most widely used evaluation metrics for segmentation tasks are the dice similarity coefficient (DSC), intersection over union (IoU), accuracy, precision, recall, and Hausdorff distance (HD).

Firstly, DSC represents the ratio of the overlapping region of predicted and ground truth over the total region. Mathematically, DSC is expressed as shown in Eq. (1):

$$DSC = \frac{2 \cdot Y_{\text{pre}} \cap Y_{\text{GT}}}{Y_{\text{pre}} + Y_{\text{GT}}} \tag{1}$$

where Y_{pre} represents the segmentation or predicted pixel and Y_{GT} represents the groundtruth pixels.

Regarding the segmentation task, DSC is equal to the F1 score, and expressed in Eq. (2):

$$DSC = \frac{2 \cdot TP}{2 \cdot TP + FP + FN} \tag{2}$$

Then, IoU is a metric for quantifying the overlap between the segmentation prediction and the expert annotation (ground truth). This metric is defined as the proportion of the overlap between the segmentation outcome and the actual ground truth about their unions. The mathematical expression of IoU is formulated Eq. (3):

$$IoU = \frac{Y_{\text{pre}} \cap Y_{\text{GT}}}{Y_{\text{pre}} \cup Y_{\text{GT}}} \tag{3}$$

It is important to highlight that the Jaccard similarity coefficient and the IoU are equivalent. As a result, we may use TP, FP, and FN to rewrite the IoU expression as shown in Eq. (4):

$$IoU = \frac{TP}{TP + FP + FN} \tag{4}$$

Precision assesses the accuracy of positive predictions and is computed as the proportion of correctly identified positive outcomes relative to the combined total of true positives and false positives. It provides insight into the accuracy of positive predictions made by the model by indicating how many were correct. The mathematical expression for precision is shown in Eq. (5):

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

Recall assesses the model's capability to accurately identify all relevant positive instances by determining how many actual positive instances the model correctly identifies. and is computed as the proportion of correctly identified positive instances relative to the combined total of accurately identified positives and incorrectly identified negatives. The mathematical expression for recall is shown in Eq. (6):

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

Accuracy is a comprehensive metric measuring the overall correctness of the model's predictions, encompassing both positive and negative predictions. It is calculated as the sum of true positive and true negative results divided by the total number of predictions. Accuracy assesses how many predictions, both positive and negative, the model got correct out of all predictions made. The mathematical expression for accuracy is shown in Eq. (7):

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{7}$$

TABLE 6 Hybrid transformer models for multi-modal MRI brain tumor segmentation.

Segmentation models	Dataset	Experimental parameters		Segmentation performance			Year	Ref.
		Optimizer	Loss function	WT	TC	ET		
Multi-modal brain tumor segmentation using transformer (TransBTS)	BraTS 2019	Adam	dice	DSC = 0.900 HD = 5.644	DSC = 0.819 HD = 6.049	DSC = 0.789 HD = 3.736	2021	Wang et al., 2021a
	BraTS 2020			DSC = 0.901 HD = 4.964	DSC = 0.817 HD = 9.769	DSC = 0.787 HD = 17.947		
Segmentation model based on squeeze and expansion transformer (SegTran)	BraTS 2019	AdamW	PWCE and dice	DSC = 0.895	DSC = 0.817	DSC = 0.740	2021	Li et al. (2021)
Medical transformer	BraTS 2019	Adam	triplet	DSC = 0.873	DSC = 0.697	DSC = 0.588	2021	Jun et al. (2021)
Convolution and transformer-based segmentation model (TransConver)	BraTS 2018	Adam	CE and dice	DSC = 0.859 HD = 2.587	DSC = 0.838 HD = 1.607	DSC = 0.789 HD = 2.692	2022	Liang et al. (2022a)
	BraTS 2019			DSC = 0.859 HD = 2.587	DSC = 0.838 HD = 1.607	DSC = 0.789 HD = 2.692		
Multi-modal medical transformer (mmFormer)	BraTS 2018	Adam	dice	DSC = 0.896	DSC = 0.858	DSC = 0.776	2022	Zhang et al. (2022a)
Transformer and SE-Attention (TSEUNet)	BraTS 2018	SGD	CE and dice	DSC = 0.911	DSC = 0.873	DSC = 0.824	2022	Chen and Wang (2022)
Axial-spatial transformer network (AST-Net)	BraTS 2018	Adam	dice	DSC = 0.905 HD = 5.950	DSC = 0.850 HD = 9.200	DSC = 0.795 HD = 2.980	2022	Wang et al., 2022b
	BraTS 2019			DSC = 0.899 HD = 5.49	DSC = 0.843 HD = 6.32	DSC = 0.786 HD = 2.90		
	BraTS 2020			DSC = 0.904 HD = 6.05	DSC = 0.842 HD = 6.12	DSC = 0.778 HD = 30.83		
2D backbone to segment 3D brain tumor (Transition Net)	BraTS 2019	AdamW	weighted region	DSC = 0.913 HD = 20.15	DSC = 0.845 HD = 12.21	DSC = 0.749 HD = 10.09	2022	Liu et al. (2022d)
TransBTSV2	BraTS 2019	Adam	softmax dice	DSC = 0.904 HD = 5.432	DSC = 0.849 HD = 5.473	DSC = 0.802 HD = 3.696	2022	Li et al. (2022)
	BraTS 2020			DSC = 0.904 HD = 5.432	DSC = 0.849 HD = 5.473	DSC = 0.802 HD = 3.696		
Generative adversarial network (GAN) based on transformers	BraTS 2018	Adam	dice	DSC = 0.902 HD = 5.418	DSC = 0.809 HD = 9.405	DSC = 0.769 HD = 5.712	2022	Huang et al. (2022b)
	BraTS 2020			DSC = 0.903 HD = 4.909	DSC = 0.815 HD = 7.494	DSC = 0.708 HD = 37.579		

(Continued on following page)

10.3389/fbioe.2024.1392807

TABLE 6 (*Continued*) **Hybrid transformer models for multi-modal MRI brain tumor segmentation.**

Segmentation models	Dataset	Experimental parameters		Segmentation performance			Year	Ref.
		Optimizer	Loss function	WT	TC	ET		
3D parallel shifted windows for brain tumor segmentation (3D PSwinBTS)	BraTS 2020	Adam	CE and dice	DSC = 0.908 HD = 5.573	DSC = 0.842 HD = 7.252	DSC = 0.795 HD = 19.437	2022	Liang et al. (2022b)
	BraTS 2021			DSC = 0.926 HD = 3.738	DSC = 0.867 HD = 11.084	DSC = 0.826 HD = 17.531		
CNN-transformer combined model (BiTr-UNet)	BraTS 2021	Adam	-	DSC = 0.926 HD = 9.165	DSC = 0.935 HD = 8.200	DSC = 0.951 HD = 3.742	2022	Jia et al. (2021)
Ensemble modified TransBTS and nnUNet	BraTS 2021	-	CE and dice	DSC = 0.928 HD = 4.930	DSC = 0.876 HD = 17.203	DSC = 0.879 HD = 10.426	2022	Dobko et al. (2021)
Hybrid CNN-transformer model with regularization (SegTransVAE)	BraTS 2021	-	VAE and dice	DSC = 0.905 HD = 3.570	DSC = 0.926 HD = 5.840	DSC = 0.855 HD = 2.890	2022	Pham et al. (2022)
Convolution-and-transformer network (COTRNet)	BraTS 2021	Adam	WCE and dice	DSC = 0.951 HD = 9.772	DSC = 0.961 HD = 15.560	DSC = 0.935 HD = 3.255	2022	Yang et al., 2021a
Swin UNEt TRansformer (Swin UNETR)	BraTS 2021	-	soft dice	DSC = 0.926 HD = 5.831	DSC = 0.885 HD = 3.770	DSC = 0.858 HD = 6.016	2022	Futrega et al. (2021)
Trans-NUNet	Kaggle	-	CE and dice	Avg. DSC = 0.864			2022	Wang et al., 2022a
Efficient R-transformer network (ERTN)	BraTS 2017	AdamW	focal and dice	DSC = 0.832 HD = 5.300	DSC = 0.779 HD = 4.600	DSC = 0.726 HD = 5.500	2023	Hu et al. (2023)
Deep semantics and edge information for brain tumor segmentation	BraTS 2018	Adam	BCE and dice	DSC = 0.909 HD = 3.923	DSC = 0.879 HD = 5.217	DSC = 0.819 HD = 3.440	2023	Zhu et al. (2023a)
	BraTS 2019			DSC = 0.916 HD = 3.866	DSC = 0.892 HD = 5.118	DSC = 0.838 HD = 3.080		
	BraTS 2020			DSC = 0.910 HD = 4.719	DSC = 0.882 HD = 5.985	DSC = 0.846 HD = 3.051		
Deep mutual learning with fusion network for brain tumor segmentation	BraTS 2019	Adam	focal and active contour	DSC = 0.901 HD = 4.800	DSC = 0.840 HD = 6.112	DSC = 0.801 HD = 3.282	2023	Gao et al. (2023a)
Flexible Fusion Network (F2 Net)	BraTS 2019	SGD	CE and dice	DSC = 0.950 HD = 2.21	DSC = 0.943 HD = 1.63	DSC = 0.902 HD = 1.33	2023	Yang et al. (2023)
	BraTS 2020			DSC = 0.953 HD = 2.20	DSC = 0.945 HD = 1.59	DSC = 0.905 HD = 1.32		
Multi-scale ghost CNN with auxiliary MetaFormer decoding path (GMetaNet)	BraTS 2018	Adam	generalized dice	DSC = 0.901 HD = 5.16	DSC = 0.840 HD = 5.26	DSC = 0.820 HD = 2.62	2023	Lu et al. (2023)
	BraTS 2019			DSC = 0.902 HD = 4.530	DSC = 0.825 HD = 6.400	DSC = 0.785 HD = 3.590		

HD serves as a distance-based evaluation metric. Within the HD, the predictions and the expert annotations are regarded as two distinct subsets in the measurement space. The mathematical expression is articulated in Eq. (8):

$$HD = \max\left(\sup_{Y_{\text{pre}}} \inf_{Y_{\text{GT}}} d\left(Y_{\text{pre}}, Y_{\text{GT}}\right), \sup_{Y_{\text{GT}}} \inf_{Y_{\text{pre}}} d\left(Y_{\text{pre}}, Y_{\text{GT}}\right)\right) \quad (8)$$

Here, "Sup" denotes the supremum, which is the least upper bound of a set, while "inf" signifies the infimum, which is the greatest lower bound of a set. So, $sup_{Y_{\text{pre}}}$ denotes the supremum over all possible subsets of Y_{pre} and $inf_{Y_{\text{GT}}}$ denotes the infimum over all possible subsets of Y_{GT}.

In summary, this section has extensively examined deep learning (DL)-based models for brain tumor segmentation, focusing on statistical insights. We began by scrutinizing the publication landscape from 2021 to 2023, revealing dynamic trends and the rapid evolution of DL-based approaches. Moving to dataset statistics, the indispensability of multi-modal MRI datasets, particularly through the MICCAI BraTS challenge, was emphasized. Figure 5 visually portrays the predominant use of BraTS datasets, underlining their widespread adoption. Additionally, we presented a comprehensive overview of common evaluation metrics, including DSC, IoU, accuracy, precision, recall, and HD, providing quantitative benchmarks for assessing model performance. Anticipating future studies, the trajectory is poised to continue refining brain tumor segmentation methods, leveraging established datasets and standardized metrics for ongoing advancements in this critical medical imaging domain.

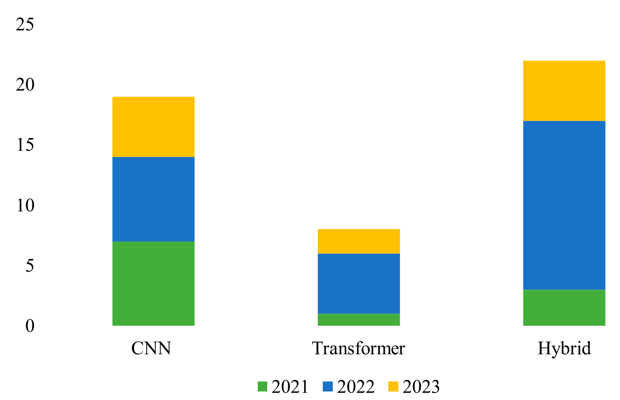


FIGURE 4
DL-based multi-modal MRI brain tumor segmentation model publication statistics from 2021 to 2023.

4 Open research problem and possible future directions

DL-based segmentation of brain tumors using MRI images is a prominent area of research in medical imaging and has achieved good results. The diagnosis, therapy planning, and ongoing observation of people with tumors depend on precise segmentation. The development of the DL model for brain tumor segmentation is complex. In this section, we examine major research challenges that must be resolved.

4.1 Incomplete modalities

Incomplete modalities pose a significant challenge in medical image analysis. While numerous studies demonstrate impressive results when equipped with complete modalities, their efficacy diminishes when utilizing incomplete modalities as input sources (Azad et al., 2022). In practice, acquiring all modalities is often impractical, leading medical institutions to possess only partial modalities. Leveraging established methods for the segmentation of brain tumors across multiple modalities. Becomes challenging in such scenarios, hampering accurate diagnoses. In clinical practices, medical institutions frequently encounter incomplete MRI modalities due to limitations in collection devices.

Previous works in brain tumor segmentation typically assume complete input MRI data, resulting in a notable decline in performance when confronted with incomplete modality inputs. For instance, RFNet (Ding et al., 2021) shows the effect of missing modality using the BraTS 2020 dataset. They achieved a maximum DSC of 87.32% on WT using only the FLAIR modality, but with the combination of three modalities, i.e., FLAIR, T1, and T1ce, RFNet achieved a DSC of 90.69%. On the other hand, RFNet achieved 91.11% DSC using all four modalities. Similarly, in (Zhang et al., 2022b) mmformer achieved a DSC of 86.10%, 88.14%, and 89.64% using only FLAIR, three modalities (FlAIR, T1ce, and T2), and all four modalities using BraTS 2018.

Moreover, various recent models (Zhou et al., 2020; Wang et al., 2021c; Yang et al., 2022; Diao et al., 2023; Ting and Liu, 2023) were developed to handle the missing modalities effectively, but there remains some degradation in the performance compared to all modalities. To address this limitation, it is imperative to devise robust segmentation methods capable of handling incomplete modalities. Recently, some works have been proposed to tackle this issue (Zhao et al., 2022; Shi et al., 2023), however, most are tailored to specific cases of incomplete modalities and lack adaptability to diverse scenarios. Future research efforts should prioritize the development of a unified framework capable of robustly handling all cases, both complete and incomplete modalities alike.

4.2 Limited label data

A primary challenge in training the transformer model for segmentation is insufficiently labeled data. Medical images, particularly MRI datasets, possess an inadequate number of sample images compared to non-medical datasets. For example, the BraTS 2012 data from MICCAI challenges contain fewer images, as shown in Table 7. In contrast, non-medical datasets, such as ImageNet (Deng et al., 2009), contain over 1.2 million images, and the MNIST dataset (Deng et al., 2012) comprises 70,000 images. These limitations make transformer-based segmentation models less robust and generalizable for medical tasks because they require extensive and diverse datasets to understand the complicated and high-level properties of the tumor and its surrounding tissues.

The problem of limited data, specifically for brain tumors, can be tackled by using different augmentation techniques. These techniques

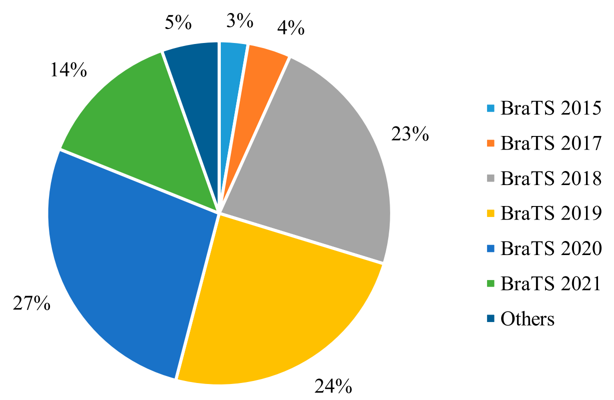


FIGURE 5
Statistical analysis of multi-modal brain tumor segmentation datasets used in the DL = based models from 2021 to 2023.

TABLE 7 Brain tumor segmentation (BraTS) datasets.

Dataset	Number of images	Available modalities
BraTS 2012	50	T1, T2, T1ce, FLAIR
BraTS 2013	60	T1, T2, T1ce, FLAIR
BraTS 2014	238	T1, T2, T1ce, FLAIR
BraTS 2015	253	T1, T2, T1ce, FLAIR
BraTS 2016	391	T1, T2, T1ce, FLAIR
BraTS 2017	477	T1, T2, T1ce, FLAIR
BraTS 2018	542	T1, T2, T1ce, FLAIR
BraTS 2019	651	T1, T2, T1ce, FLAIR
BraTS 2020	660	T1, T2, T1ce, FLAIR
BraTS 2021	2000	T1, T2, T1ce, FLAIR

generate training data and improve the model's performance. Generally, there are two types of data augmentation, i.e., conventional and GAN-based. In conventional augmentation approaches, different transformations, such as geometric and photometric transformations, are used to increase the data quantity. However, these techniques are not effective in dealing with diverse data. On the other hand, GAN-based augmentation has gained popularity owing to its ability to produce synthetic and diverse data that closely resemble input data. A GAN is composed of a generator and discriminator neural networks. The generator network learns to create synthetic samples, whereas the discriminator network determines the differences between the actual and created samples.

Furthermore, researchers have employed various GAN-based data-augmentation techniques such as conditional GAN (cGAN) (Isola et al., 2017), cycle GAN (Sandfort et al., 2019), and parasitic GAN (Sun et al., 2019) for this purpose. Moreover, test-time augmentation methods can also be explored, as (Amiri et al., 2022) suggest that test-time augmentation (TTA) is one of the influential factors in improving model performance. The list of open-source packages and frameworks for DL-based medical image data augmentation methods are as follows: Augmentor (Bloice et al., 2019), Albumentations (Buslaev et al., 2020), Batchgenerators (Isensee et al., 2020), CutBlue (Yoo et al., 2020), CLoDSA (Casado-García et al., 2019), Gryds (Eppenhof and Pluim, 2018), ImgAug (Gu et al., 2019), Keras ImagedataGenerator (Chollet et al., 2015), MONAI (Cardoso et al., 2022), Pymia (Jungo et al., 2021), PyTorch Transformer (Paszke et al., 2019), and TorchIO (Pérez-García et al., 2021).

4.3 Enhancing model efficiency and deployment

In real-world scenarios, adeptly trained deep models find applications on terminal devices characterized by constrained resource availability. These settings' requirements necessitate deploying efficient and lightweight deep models. During the training phase, emphasis is placed on ensuring the efficiency and compactness of deep models. Model compression strategies, including weight pruning, quantization, distillation of large models, and the incorporation of low-rank approximations, are employed to diminish the model's size. These techniques effectively minimize the memory and computational demands of the deep model. Additionally, optimizing network architectures and implementing tailored training regimens contribute to alleviating the demand for an excessive number of parameters while maintaining optimal performance. Regrettably, there exists a dearth of research addressing the crucial aspects of model efficiency and deployment in brain tumor segmentation. This represents a pivotal gap in understanding that needs attention to ensure the successful utilization of algorithms in upcoming clinical practices.

4.4 Class imbalance

Addressing class imbalance is an essential challenge in the task of multi-modal brain tumor segmentation due to the tendency of these tumors to occupy a very limited area of the brain, which in turn complicates the processing of MRI data. The disparity might lead to a skewed division that favors the more significant class (healthy tissue), affecting the precise segmentation of brain tumors with smaller areas (Akil et al., 2020; Deepak and Ameer, 2023). Conventional methods include using class reweighting strategies throughout the training process to tackle this problem. These strategies provide more importance to the minority class (small tumor areas) and lesser significance to the majority class, allowing the model to prioritize the smaller class during training.

Recently, some work has been done to overcome the issue of class imbalance for multimodal MRI brain tumor segmentation. Most of the work in the literature is based on the use of different loss functions, for instance, in (Li et al., 2023b) combined loss function is used to optimize the network. Here dice and cross entropy losses are used to overcome class imbalance and stable training process, respectively. In (Zhu et al., 2023b), the edge detection module is for the class imbalance problem and they introduce L_{edge} which is the combination of edge loss and dice. In (Yeung et al., 2022) introduce the unified focal loss for the class imbalance problem and also present the detailed discussion and effect of other loss functions. Various other loss functions are used in the literature to tackle the class imbalance problems (Jiao et al., 2023; Rehman et al., 2023; Li et al., 2024). Another way to deal with the class imbalance problem is to remove focus-free images in the dataset during training that are present in an excessive amount (Gao et al., 2023; Wu et al., 2024). Future research should focus on expanding current class balancing approaches and devise adaptive balancing approaches to improve the efficacy of tiny tumor areas.

4.5 Interpretability

Interpretability poses a difficulty in DL since these methods are often regarded as completely opaque models with limited insight into the reasoning behind predictions. The absence of interpretability is particularly critical in practical scenarios, notably in the field of clinical treatment, where understanding the functioning of deep models and the reasoning behind their choices is essential. One possible method to tackle this problem is using visual representations of feature maps, emphasizing prominent areas that influence the model's results. Researchers have developed many techniques to display intermediate layers in deep learning models, such as activation maximization, class activation maps (Muhammad et al., 2020), and conditional t-distributed stochastic neighbor embedding (ct-SNE) (Kang et al., 2021). Recent endeavors have also used feature attribution techniques to identify the most relevant characteristics for a certain prediction generated by a DL model. The techniques involved in this process include gradient-based attribution (Ancona et al., 2017), perturbation-based attribution (Ivanovs et al., 2021) etc.

To further enhance the interpretability of multimodal DL models, eXplainable AI (XAI) offers a suite of techniques aimed at providing transparency and insights into model decisions. Some notable XAI methodologies include SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME). SHAP assesses the outcome for any DL model by determining the relative contributions of every feature to the resulting estimation and prediction, making it particularly useful for multimodal models (Lundberg et al., 2017). LIME explains individual predictions by approximating the model locally with an interpretable model. It achieves interpretability by training these models on subsets of the dataset, enabling users to understand how different features influence the decision (Ribeiro et al., 2016). Various other techniques have been developed for XAI, recently one method was developed for contrasting the decision-making patterns of the black box and white box models (Žlahtič et al., 2024).

Future research in multimodal brain tumor segmentation, utilizing Vision Transformers and other advanced architectures, provides the potential to customize interpretability approaches for individual purposes. By integrating these XAI techniques, we can improve the understanding of model functioning, increasing its transparency and offering useful insights for therapeutic applications. Emphasizing interpretability in multimodal DL models not only aids in clinical decision-making but also builds trust among medical professionals and patients, facilitating the adoption of AI technologies in healthcare.

5 Conclusion

This study highlights the significance of DL in brain tumor segmentation using multi-modal MRI, offering critical insights into treatment planning and personalized care. Beginning with exploring MRI modalities and the advantages of DL-based segmentation models. DL models have significantly improved brain tumor segmentation using multi-modal MRI and offered numerous advantages for tumor segmentation tasks, such as saving time, eliminating human bias, and minimizing errors. We thoroughly investigated DL-based models for brain tumor segmentation using multi-modal MRI and evaluated the recent existing model. Our study categorizes current research into three main groups based on the model's architecture: CNN, transformer, and hybrid models. We have thoroughly investigated these models, considering their architectural design, dataset utilized, and experimental parameters. In addition, we perform a comprehensive statistical analysis of recent publications, brain tumor datasets, and evaluation metrics. Finally, open research challenges are identified and suggested promising future directions for multi-modal MRI brain tumor segmentation.

Author contributions

ZA: Conceptualization, Data curation, Writing–original draft. RN: Conceptualization, Formal Analysis, Methodology, Supervision, Writing–review and editing. AH: Software, Validation, Visualization, Writing–review and editing. HK: Investigation, Resources, Supervision, Writing–review and editing. DJ: Formal Analysis, Funding acquisition, Resources, Writing–review and editing. SL: Funding acquisition, Methodology, Validation, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was also supported by the MISP (Ministry of Science, ICT and Future Planning), Korea, under the National Program for Excellence in SW) (2019-0-01880) supervised by the IITP (Institute of Information and communications Technology Planning and Evaluation) (2019-0-01880).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Akbar, A. S., Fatichah, C., and Suciati, N. (2021). "Unet3d with multiple atrous convolutions attention block for brain tumor segmentation," in *International MICCAI brainlesion w orkshop* (Springer), 182–193.

Akil, M., Saouli, R., and Kachouri, R. (2020). Fully automatic brain tumor segmentation with deep learning-based selective attention using overlapping patches and multi-class weighted cross-entropy. *Med. image Anal.* 63, 101692. doi:10.1016/j.media.2020.101692

Ali, S., Li, J., Pei, Y., Khurram, R., Rehman, K. u., and Mahmood, T. (2022). A comprehensive survey on brain tumor diagnosis using deep learning and emerging hybrid techniques with multi-modal mr image. *Archives Comput. methods Eng.* 29, 4871–4896. doi:10.1007/s11831-022-09758-z

Amiri, S., and Ibragimov, B. (2022). Improved automated lesion segmentation in whole-body fdg/pet-ct via test-time augmentation. *arXiv Prepr. arXiv:2210.07761*. doi:10.48550/arXiv.2210.07761

Ancona, M., Ceolini, E., Öztireli, C., and Gross, M. (2017). *Towards a better understanding of gradient-based attribution methods for deep neural networks*. arXiv preprint arXiv:1711.06104.

Azad, R., Khosravi, N., Dehghanmanshadi, M., Cohen-Adad, J., and Merhof, D. (2022). *Medical image segmentation on mri images with missing modalities: a review*. arXiv preprint arXiv:2203.06217.

Biratu, E. S., Schwenker, F., Debelee, T. G., Kebede, S. R., Negera, W. G., and Molla, H. T. (2021). Enhanced region growing for brain tumor mr image segmentation. *J. Imaging* 7, 22. doi:10.3390/jimaging7020022

Bloice, M. D., Roth, P. M., and Holzinger, A. (2019). Biomedical image augmentation using augmentor. *Bioinformatics* 35, 4522–4524. doi:10.1093/bioinformatics/btz259

Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., and Kalinin, A. A. (2020). Albumentations: fast and flexible image augmentations. *Information* 11, 125. doi:10.3390/info11020125

Cardoso, M. J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., et al. (2022). *Monai: an open-source framework for deep learning in healthcare*. arXiv preprint arXiv: 2211.02701.

Casado-García, Áa, Domínguez, C., García-Domínguez, M., Heras, J., Inés, A., Mata, E., et al. (2019). Clodsa: a tool for augmentation in classification, localization, detection, semantic segmentation and instance segmentation tasks. *BMC Bioinforma.* 20, 323. doi:10.1186/s12859-019-2931-1

Çetiner, H., and Metlek, S. (2023). Denseunet+: a novel hybrid segmentation approach based on multi-modality images for brain tumor segmentation. *J. King Saud University-Computer Inf. Sci.* 35, 101663. doi:10.1016/j.jksuci.2023.101663

Chang, Y., Zheng, Z., Sun, Y., Zhao, M., Lu, Y., and Zhang, Y. (2023). Dpafnet: a residual dual-path attention-fusion convolutional neural network for multimodal brain tumor segmentation. *Biomed. Signal Process. Control* 79, 104037. doi:10.1016/j.bspc.2022.104037

Charles, N. A., Holland, E. C., Gilbertson, R., Glass, R., and Kettenmann, H. (2011). The brain tumor microenvironment. *Glia* 59, 1169–1180. doi:10.1002/glia.21136

Chen, Y., and Wang, J. (2022). "Tseunet: a 3d neural network with fused transformer and se-attention for brain tumor segmentation," in *2022 IEEE 35th international symposium on computer-based medical systems (CBMS) (IEEE)*, 131–136.

Chollet, F. (2015). *Keras: deep learning library for theano and tensorflow*, T1. Available at: https://keras.io/k7.

Deepak, S., and Ameer, P. (2023). Brain tumor categorization from imbalanced mri dataset using weighted loss and deep feature fusion. *Neurocomputing* 520, 94–102. doi:10.1016/j.neucom.2022.11.039

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L. (2009). "Imagenet: a large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition (Ieee)*, 248–255.

Deng, L. (2012). The mnist database of handwritten digit images for machine learning research [best of the web]. *IEEE signal Process. Mag.* 29, 141–142. doi:10.1109/msp.2012.2211477

Desjardins, L., Barrera, M., Schulte, F., Chung, J., Cataudella, D., Janzen, L., et al. (2019). Predicting social withdrawal, anxiety and depression symptoms in pediatric brain tumor survivors. *J. Psychosoc. Oncol.* 37, 22–36. doi:10.1080/07347332.2018.1535531

Dhole, N. V., and Dixit, V. V. (2022). Review of brain tumor detection from mri images with hybrid approaches. *Multimedia tools Appl.* 81, 10189–10220. doi:10.1007/s11042-022-12162-1

Diao, Y., Li, F., and Li, Z. (2023). Joint learning-based feature reconstruction and enhanced network for incomplete multi-modal brain tumor segmentation. *Comput. Biol. Med.* 163, 107234. doi:10.1016/j.compbiomed.2023.107234

Ding, Y., Yu, X., and Yang, Y. (2021). "Rfnet: region-aware fusion network for incomplete multi-modal brain tumor segmentation," in *Proceedings of the IEEE/CVF international conference on computer vision*, 3975–3984.

Dobko, M., Kolinko, D. I., Viniavskyi, O., and Yelisieiev, Y. (2021). "Combining cnns with transformer for multimodal 3d mri brain tumor segmentation," in *International MICCAI brainlesion workshop* (Springer), 232–241.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., et al. (2020). *An image is worth 16x16 words: transformers for image recognition at scale. arXiv preprint arXiv:2010.11929.*

Eppenhof, K. A., and Pluim, J. P. W. (2018). Pulmonary ct registration through supervised learning with convolutional neural networks. *IEEE Trans. Med. imaging* 38, 1097–1105. doi:10.1109/tmi.2018.2878316

Fang, F., Yao, Y., Zhou, T., Xie, G., and Lu, J. (2021). Self-supervised multi-modal hybrid fusion network for brain tumor segmentation. *IEEE J. Biomed. Health Inf.* 26, 5310–5320. doi:10.1109/jbhi.2021.3109301

Futrega, M., Milesi, A., Marcinkiewicz, M., and Ribalta, P. (2021). "Optimized u-net for brain tumor segmentation," in *International MICCAI brainlesion workshop* (Springer), 15–29.

Gao, H., Miao, Q., Ma, D., and Liu, R. (2023a). Deep mutual learning for brain tumor segmentation with the fusion network. *Neurocomputing* 521, 213–220. doi:10.1016/j.neucom.2022.11.038

Gao, L., Li, J., Zhang, R., Bekele, H. H., Wang, J., Cheng, Y., et al. (2023b). Mmgan: a multimodal mr brain tumor image segmentation method. *Front. Hum. Neurosci.* 17, 1275795. doi:10.3389/fnhum.2023.1275795

Gu, S., Bao, J., Yang, H., Chen, D., Wen, F., and Yuan, L. (2019). "Mask-guided portrait editing with conditional gans," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 3436–3445.

Guo, X., Yang, C., Lam, P. L., Woo, P. Y., and Yuan, Y. (2020). "Domain knowledge-based brain tumor segmentation and overall survival prediction," in *Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries: 5th international workshop, BrainLes 2019, held in conjunction with MICCAI 2019, shenzhen, China, october 17, 2019* (Springer), 285–295.

Hamamci, A., and Unal, G. (2012). Multimodal brain tumor segmentation using the tumor-cut method on the brats dataset. *Proc. MICCAI-BraTS*, 19–23.

Hamamci, A., Kucuk, N., Karaman, K., Engin, K., and Unal, G. (2011). Tumor-cut: segmentation of brain tumors on contrast-enhanced mr images for radiosurgery applications. *IEEE Trans. Med. imaging* 31, 790–804. doi:10.1109/tmi.2011.2181857

He, W., Zhou, X., Mao, Y., Wu, Y., Tang, X., Yan, S., et al. (2022). Circcrim1 promotes nasopharyngeal carcinoma progression via the mir-34c-5p/fosl1 axis. *Eur. J. Med. Res.* 27, 59. doi:10.1186/s40001-022-00667-2

Hou, Q., Peng, Y., Wang, Z., Wang, J., and Jiang, J. (2023). Mfd-net: modality fusion diffractive network for segmentation of multimodal brain tumor image. *IEEE J. Biomed. Health Inf.* 27, 5958–5969. doi:10.1109/jbhi.2023.3318640

Hu, Z., Li, L., Sui, A., Wu, G., Wang, Y., and Yu, J. (2023). An efficient r-transformer network with dual encoders for brain glioma segmentation in mr images. *Biomed. Signal Process. Control* 79, 104034. doi:10.1016/j.bspc.2022.104034

Huang, L., Zhu, E., Chen, L., Wang, Z., Chai, S., and Zhang, B. (2022b). A transformer-based generative adversarial network for brain tumor segmentation. *Front. Neurosci.* 16, 1054948. doi:10.3389/fnins.2022.1054948

Huang, Z., Lin, L., Cheng, P., Peng, L., and Tang, X. (2022a). *Multi-modal brain tumor segmentation via missing modality synthesis and modality-level attention fusion*. arXiv preprint arXiv:2203.04586.

Ilhan, U., and Ilhan, A. (2017). Brain tumor segmentation based on a new threshold approach. *Procedia Comput. Sci.* 120, 580–587. doi:10.1016/j.procs.2017.11.282

Isensee, F., Jäger, P., Wasserthal, J., Zimmerer, D., Petersen, J., Kohl, S., et al. (2020). *batchgenerators—a python framework for data augmentation*. Zenodo 3632567.

Isola, P., Zhu, J. Y., Zhou, T., and Efros, A. A. (2017). "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1125–1134.

Ivanovs, M., Kadikis, R., and Ozols, K. (2021). Perturbation-based methods for explaining deep neural networks: a survey. *Pattern Recognit. Lett.* 150, 228–234. doi:10.1016/j.patrec.2021.06.030

Jia, Q., and Shu, H. (2021). "Bitr-unet: a cnn-transformer combined network for mri brain tumor segmentation," in *International MICCAI brainlesion workshop* (Springer), 3–14.

Jiao, C., Yang, T., Yan, Y., and Yang, A. (2023). Rftnet: region–attention fusion network combined with dual-branch vision transformer for multimodal brain tumor image segmentation. *Electronics* 13, 77. doi:10.3390/electronics13010077

Jun, E., Jeong, S., Heo, D. W., and Suk, H. I. (2021). *Medical transformer: universal brain encoder for 3d mri analysis*. arXiv preprint arXiv:2104.13633.

Jungo, A., Scheidegger, O., Reyes, M., and Balsiger, F. (2021). pymia: a python package for data handling and evaluation in deep learning-based medical image analysis. *Comput. methods programs Biomed.* 198, 105796. doi:10.1016/j.cmpb.2020.105796

Jyothi, P., and Singh, A. R. (2023). Deep learning models and traditional automated techniques for brain tumor segmentation in mri: a review. *Artif. Intell. Rev.* 56, 2923–2969. doi:10.1007/s10462-022-10245-x

Kang, B., García García, D., Lijffijt, J., Santos-Rodríguez, R., and De Bie, T. (2021). Conditional t-sne: more informative t-sne embeddings. *Mach. Learn.* 110, 2905–2940. doi:10.1007/s10994-020-05917-0

Karimijafarbigloo, S., Azad, R., Kazerouni, A., Ebadollahi, S., and Merhof, D. (2023). "Mmcformer: missing modality compensation transformer for brain tumor segmentation," in *Medical imaging with deep learning*.

Khilkhal, R., and Ismael, M. (2022). "Brain tumor segmentation utilizing thresholding and k-means clustering," in *2022 muthanna international conference on engineering science and Technology (MICEST)* (IEEE), 43–48.

Kotia, J., Kotwal, A., and Bharti, R. (2020). "Risk susceptibility of brain tumor classification to adversarial attacks," in *Man-machine interactions 6: 6th international conference on man-machine interactions, ICMMI 2019, cracow, Poland, october 2-3, 2019*. Springer, 181–187.

Li, D., Chen, G., Wu, X., Yu, Z., and Tan, M. (2024a). Face anti-spoofing with cross-stage relation enhancement and spoof material perception. *Neural Netw.* 175, 106275. doi:10.1016/j.neunet.2024.106275

Li, J., Wang, W., Chen, C., Zhang, T., Zha, S., Wang, J., et al. (2022). *Transbtsv2: towards better and more efficient volumetric segmentation of medical images*. arXiv preprint arXiv:2201.12785.

Li, P., Li, Z., Wang, Z., Li, C., and Wang, M. (2024b). mresu-net: multi-scale residual u-net-based brain tumor segmentation from multimodal mri. *Med. Biol. Eng. Comput.* 62, 641–651. doi:10.1007/s11517-023-02965-1

Li, S., Sui, X., Luo, X., Xu, X., Liu, Y., and Goh, R. (2021). Medical image segmentation using squeeze-and-expansion transformers. *arXiv Prepr. arXiv:2105.09511*. doi:10.48550/arXiv.2105.09511

Li, X., Jiang, Y., Li, M., Zhang, J., Yin, S., and Luo, H. (2023a). Msfr-net: multi-modality and single-modality feature recalibration network for brain tumor segmentation. *Med. Phys.* 50, 2249–2262. doi:10.1002/mp.15933

Li, X., Jiang, Y., Li, M., Zhang, J., Yin, S., and Luo, H. (2023b). Msfr-net: multi-modality and single-modality feature recalibration network for brain tumor segmentation. *Med. Phys.* 50, 2249–2262. doi:10.1002/mp.15933

Liang, J., Yang, C., and Zeng, L. (2022b). 3d pswinbts: an efficient transformer-based unet using 3d parallel shifted windows for brain tumor segmentation. *Digit. Signal Process.* 131, 103784. doi:10.1016/j.dsp.2022.103784

Liang, J., Yang, C., Zeng, M., and Wang, X. (2022a). Transconver: transformer and convolution parallel network for developing automatic brain tumor segmentation in mri images. *Quantitative Imaging Med. Surg.* 12, 2397–2415. doi:10.21037/qims-21-919

Liu, J., Zheng, J., and Jiao, G. (2022d). Transition net: 2d backbone to segment 3d brain tumor. *Biomed. Signal Process. Control* 75, 103622. doi:10.1016/j.bspc.2022.103622

Liu, Y., Mu, F., Shi, Y., Cheng, J., Li, C., and Chen, X. (2022b). Brain tumor segmentation in multimodal mri via pixel-level and feature-level image fusion. *Front. Neurosci.* 16, 1000587. doi:10.3389/fnins.2022.1000587

Liu, Z., Lv, Q., Yang, Z., Li, Y., Lee, C. H., and Shen, L. (2022a). *Medical image analysis based on transformer: a review*. arXiv preprint arXiv:2208.06643.

Liu, Z., Wei, J., Li, R., and Zhou, J. (2022c). *Tfusion: transformer based n-to-one multimodal fusion block*. arXiv preprint arXiv:2208.12776.

Liu, Z., Tong, L., Chen, L., Jiang, Z., Zhou, F., Zhang, Q., et al. (2023). Deep learning-based brain tumor segmentation: a survey. *Complex and intelligent Syst.* 9, 1001–1026. doi:10.1007/s40747-022-00815-5

Lu, Y., Chang, Y., Zheng, Z., Sun, Y., Zhao, M., Yu, B., et al. (2023). Gmetanet: multi-scale ghost convolutional neural network with auxiliary metaformer decoding path for brain tumor segmentation. *Biomed. Signal Process. Control* 83, 104694. doi:10.1016/j.bspc.2023.104694

Lundberg, S. M., and Lee, S. I. (2017). A unified approach to interpreting model predictions. *Adv. neural Inf. Process. Syst.* 30.

Menze, B. H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., et al. (2014). The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans. Med. imaging* 34, 1993–2024. doi:10.1109/tmi.2014.2377694

Mohammed, Y. M., El Garouani, S., and Jellouli, I. (2023). A survey of methods for brain tumor segmentation-based mri images. *J. Comput. Des. Eng.* 10, 266–293. doi:10.1093/jcde/qwac141

Muhammad, M. B., and Yeasin, M. (2020). "Eigen-cam: class activation map using principal components," in *2020 international joint conference on neural networks (IJCNN) (IEEE)*, 1–7.

Nyo, M. T., Mebarek-Oudina, F., Hlaing, S. S., and Khan, N. A. (2022). Otsu's thresholding technique for mri image brain tumor segmentation. *Multimedia tools Appl.* 81, 43837–43849. doi:10.1007/s11042-022-13215-1

Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., et al. (2018). *Attention u-net: learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.*

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). Pytorch: an imperative style, high-performance deep learning library. *Adv. neural Inf. Process. Syst.* 32.

Peiris, H., Hayat, M., Chen, Z., Egan, G., and Harandi, M. (2022a). "A robust volumetric transformer for accurate 3d tumor segmentation," in *International conference on medical image computing and computer-assisted intervention* (Springer), 162–172.

Peiris, H., Hayat, M., Chen, Z., Egan, G., and Harandi, M. (2022b). *Hybrid window attention based transformer architecture for brain tumor segmentation*. arXiv preprint arXiv:2209.07704.

Pereira, S., Pinto, A., Alves, V., and Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in mri images. *IEEE Trans. Med. imaging* 35, 1240–1251. doi:10.1109/tmi.2016.2538465

Pérez-García, F., Sparks, R., and Ourselin, S. (2021). Torchio: a python library for efficient loading, preprocessing, augmentation and patch-based sampling of medical images in deep learning. *Comput. Methods Programs Biomed.* 208, 106236. doi:10.1016/j.cmpb.2021.106236

Pham, Q. D., Nguyen-Truong, H., Phuong, N. N., Nguyen, K. N., Nguyen, C. D., Bui, T., et al. (2022). "Segtransvae: hybrid cnn-transformer with regularization for medical image segmentation," in *2022 IEEE 19th international symposium on biomedical imaging (ISBI) (IEEE)*, 1–5.

Philip, A. K., Samuel, B., Bhatia, S., Khalifa, S., and El-Seedi, H. (2022). Artificial intelligence and precision medicine: a new frontier for the treatment of brain tumors. *Life* 13, 24. doi:10.3390/life13010024

Pinaya, W. H., Tudosiu, P. D., Gray, R., Rees, G., Nachev, P., Ourselin, S., et al. (2022). Unsupervised brain imaging 3d anomaly detection and segmentation with transformers. *Med. Image Anal.* 79, 102475. doi:10.1016/j.media.2022.102475

Privitera, L., Paraboschi, I., Dixit, D., Arthurs, O. J., and Giuliani, S. (2022). Image-guided surgery and novel intraoperative devices for enhanced visualisation in general and paediatric surgery: a review. *Innov. Surg. Sci.* 6, 161–172. doi:10.1515/iss-2021-0028

Ranjbarzadeh, R., Caputo, A., Tirkolaee, E. B., Jafarzadeh Ghoushchi, S., and Bendechache, M. (2023). Brain tumor segmentation of mri images: a comprehensive review on the application of artificial intelligence tools. *Comput. Biol. Med.* 152, 106405. doi:10.1016/j.compbiomed.2022.106405

Rao, C. S., and Karunakara, K. (2021). A comprehensive review on brain tumor segmentation and classification of mri images. *Multimedia Tools Appl.* 80, 17611–17643. doi:10.1007/s11042-020-10443-1

Rehman, M. U., Ryu, J., Nizami, I. F., and Chong, K. T. (2023). Raagr2-net: a brain tumor segmentation network using parallel processing of multiple spatial frames. *Comput. Biol. Med.* 152, 106426. doi:10.1016/j.compbiomed.2022.106426

Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). "Why should i trust you? explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 1135–1144.

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, october 5-9, 2015, proceedings, Part III 18* (Springer), 234–241.

Sagar, A. (2021a). "Vitbis: vision transformer for biomedical image segmentation," in *MICCAI workshop on distributed and collaborative learning* (Springer), 34–45.

Sagar, A. (2021b). "Emsvit: efficient multi scale vision transformer for biomedical image segmentation," in *International MICCAI brainlesion workshop* (Springer), 39–51.

Sahoo, A. K., Parida, P., Muralibabu, K., and Dash, S. (2023). "An improved dnn with ffcm method for multimodal brain tumor segmentation," in *Intelligent systems with applications*.

Salvador, R., Canales-Rodríguez, E., Guerrero-Pedraza, A., Sarró, S., Tordesillas-Gutiérrez, D., Maristany, T., et al. (2019). Multimodal integration of brain images for mri-based diagnosis in schizophrenia. *Front. Neurosci.* 13, 1203. doi:10.3389/fnins.2019.01203

Sandfort, V., Yan, K., Pickhardt, P. J., and Summers, R. M. (2019). Data augmentation using generative adversarial networks (cyclegan) to improve generalizability in ct segmentation tasks. *Sci. Rep.* 9, 16884. doi:10.1038/s41598-019-52737-x

Shi, J., Yu, L., Cheng, Q., Yang, X., Cheng, K. T., and Yan, Z. (2023). “M $\{2\}$ ftrans: modality-masked fusion transformer for incomplete multi-modality brain tumor segmentation,” in *IEEE journal of biomedical and health informatics*.

Siegel, R. L., Miller, K. D., Wagle, N. S., and Jemal, A. (2023). Cancer statistics, 2023. *Ca Cancer J. Clin.* 73, 17–48. doi:10.3322/caac.21763

Singha, A., Thakur, R. S., and Patel, T. (2021). “Deep learning applications in medical image analysis,” in *Techniques and applications biomedical data mining for information retrieval: methodologies*, 293–350. doi:10.1002/9781119711278.ch11

Stiles, J., and Jernigan, T. L. (2010). The basics of brain development. *Neuropsychol. Rev.* 20, 327–348. doi:10.1007/s11065-010-9148-4

Sun, Y., Zhou, Fu, Y., and Xue, X. (2019). “Parasitic gan for semi-supervised brain tumor segmentation,” in *2019 IEEE international conference on image processing (ICIP) (IEEE)*, 1535–1539.

Syazwany, N. S., Nam, J. H., and Lee, S. C. (2021). Mm-bifpn: multi-modality fusion network with bi-fpn for mri brain tumor segmentation. *IEEE Access* 9, 160708–160720. doi:10.1109/access.2021.3132050

Tahir, A., Asif, M., Ahmad, M. B., Mahmood, T., Khan, M. A., and Ali, M. (2022). Brain tumor detection using decision-based fusion empowered with fuzzy logic. *Math. Problems Eng.* 2022, 1–13. doi:10.1155/2022/2710285

Tian, W., Lv, M., and Huang, P. (2022). Axial attention convolutional neural network for brain tumor segmentation with multi-modality mri scans. *Brain Sci.* 13, 12. doi:10.3390/brainsci13010012

Ting, H., and Liu, M. (2023). Multimodal transformer of incomplete mri data for brain tumor segmentation. *IEEE J. Biomed. Health Inf.* 28, 89–99. doi:10.1109/jbhi.2023.3286689

Tong, J., and Wang, C. (2023). A dual tri-path cnn system for brain tumor segmentation. *Biomed. Signal Process. Control* 81, 104411. doi:10.1016/j.bspc.2022.104411

Wang, E., Hu, Y., Yang, X., and Tian, X. (2022a). “Transunet with attention mechanism for brain tumor segmentation on mr images,” in *2022 IEEE international conference on artificial intelligence and computer applications (ICAICA)* (IEEE), 573–577.

Wang, P., Liu, S., and Peng, J. (2022b). “Ast-net: lightweight hybrid transformer for multimodal brain tumor segmentation,” in *2022 26th international conference on pattern recognition (ICPR) (IEEE)*, 4623–4629.

Wang, P., Yang, Q., He, Z., and Yuan, Y. (2023a). Vision transformers in multi-modal brain tumor mri segmentation: a review. *Meta-Radiology* 1, 100004. doi:10.1016/j.metrad.2023.100004

Wang, P., Yang, Q., He, Z., and Yuan, Y. (2023b). Vision transformers in multi-modal brain tumor mri segmentation: a review. *Meta-Radiology* 1, 100004. doi:10.1016/j.metrad.2023.100004

Wang, Y., Chen, J., and Bai, X. (2023c). Gradient-assisted deep model for brain tumor segmentation by multi-modality mri volumes. *Biomed. Signal Process. Control* 85, 105066. doi:10.1016/j.bspc.2023.105066

Wenxuan, W., Chen, C., Meng, D., Hong, Y., Sen, Z., and Jiangyun, L. (2021a). “Transbts: multimodal brain tumor segmentation using transformer,” in *Medical image computing and computer assisted intervention–MICCAI 2021: 24th international conference, strasbourg, France, september 27–october 1, 2021, proceedings, Part I 24* (Springer), 109–119.

Wang, Y., Zhang, Y., Hou, F., Liu, Y., Tian, J., Zhong, C., et al. (2021b). “Modality-pairing learning for brain tumor segmentation,” in *Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries: 6th international workshop, BrainLes 2020, held in conjunction with MICCAI 2020, Lima, Peru* (Springer), 230–240.

Wang, Y., Zhang, Y., Liu, Y., Lin, Z., Tian, J., Zhong, C., et al. (2021c). “Acn: adversarial co-training network for brain tumor segmentation with missing modalities,” in *Medical image computing and computer assisted intervention–MICCAI 2021: 24th international conference, strasbourg, France, september 27–october 1, 2021, proceedings, Part VII 24* (Springer), 410–420.

Wu, X., Yang, X., Li, Z., Liu, L., and Xia, Y. (2024). Multimodal brain tumor image segmentation based on densenet. *Plos one* 19, e0286125. doi:10.1371/journal.pone.0286125

Xing, Z., Yu, L., Wan, L., Han, T., and Zhu, L. (2022). “Nestedformer: nested modality-aware transformer for brain tumor segmentation,” in *International conference on medical image computing and computer-assisted intervention* (Springer), 140–150.

Yang, H., Shen, Z., Li, Z., Liu, J., and Xiao, J. (2021a). “Combining global information with topological prior for brain tumor segmentation,” in *International MICCAI brainlesion workshop* (Springer), 204–215.

Yang, H., Zhou, T., Zhou, Y., Zhang, Y., and Fu, H. (2023). Flexible fusion network for multi-modal brain tumor segmentation. *IEEE J. Biomed. Health Inf.* 27, 3349–3359. doi:10.1109/jbhi.2023.3271808

Yang, Q., and Yuan, Y. (2021b). “Learning dynamic convolutions for multi-modal 3d mri brain tumor segmentation,” in *Brainlesion: glioma, multiple sclerosis, stroke and traumatic brain injuries: 6th international workshop, BrainLes 2020, held in conjunction with MICCAI 2020, Lima, Peru, october 4, 2020* (Springer), 441–451.

Yang, Q., Guo, X., Chen, Z., Woo, P. Y. M., and Yuan, Y. (2022). D 2-net: dual disentanglement network for brain tumor segmentation with missing modalities. *IEEE Trans. Med. Imaging* 41, 2953–2964. doi:10.1109/tmi.2022.3175478

Yeung, M., Sala, E., Schönlieb, C. B., and Rundo, L. (2022). Unified focal loss: generalising dice and cross entropy-based losses to handle class imbalanced medical image segmentation. *Comput. Med. Imaging Graph.* 95, 102026. doi:10.1016/j.compmedimag.2021.102026

Yoo, J., Ahn, N., and Sohn, K. A. (2020). “Rethinking data augmentation for image super-resolution: a comprehensive analysis and a new strategy,” in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 8375–8384.

Zhang, D., Huang, G., Zhang, Q., Han, J., Han, J., and Yu, Y. (2021). Cross-modality deep feature learning for brain tumor segmentation. *Pattern Recognit.* 110, 107562. doi:10.1016/j.patcog.2020.107562

Zhang, J., Qin, Q., Ye, Q., and Ruan, T. (2023). St-unet: Swin transformer boosted u-net with cross-layer feature enhancement for medical image segmentation. *Comput. Biol. Med.* 153, 106516. doi:10.1016/j.compbiomed.2022.106516

Zhang, Y., He, N., Yang, J., Li, Y., Wei, D., Huang, Y., et al. (2022a). “mmformer: multimodal medical transformer for incomplete multimodal learning of brain tumor segmentation,” in *International conference on medical image computing and computer-assisted intervention* (Springer), 107–117.

Zhao, L., Ma, J., Shao, Y., Jia, C., Zhao, J., and Yuan, H. (2022a). Mm-unet: a multimodality brain tumor segmentation network in mri images. *Front. Oncol.* 12, 950706. doi:10.3389/fonc.2022.950706

Zhao, Z., Yang, H., and Sun, J. (2022b). “Modality-adaptive feature interaction for brain tumor segmentation with missing modalities,” in *International conference on medical image computing and computer-assisted intervention* (Springer), 183–192.

Zhou, T., Canu, S., Vera, P., and Ruan, S. (2021). Feature-enhanced generation and multi-modality fusion-based deep neural network for brain tumor segmentation with missing mr modalities. *Neurocomputing* 466, 102–112. doi:10.1016/j.neucom.2021.09.032

Zhou, T., Canu, S., Vera, P., and Ruan, S. (2020). “Brain tumor segmentation with missing modalities via latent multi-source correlation representation,” in *Medical image computing and computer assisted intervention–MICCAI 2020: 23rd international conference, Lima, Peru, october 4–8, 2020, proceedings, Part IV 23* (Springer), 533–541.

Zhou, T. (2023). Modality-level cross-connection and attentional feature fusion-based deep neural network for multi-modal brain tumor segmentation. *Biomed. Signal Process. Control* 81, 104524. doi:10.1016/j.bspc.2022.104524

Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., and Liang, J. (2018). “Unet++: a nested u-net architecture for medical image segmentation,” in *Deep learning in medical image analysis and multimodal learning for clinical decision support: 4th international workshop, DLMIA 2018, and 8th international workshop, ML-CDS 2018, held in conjunction with MICCAI 2018, granada, Spain, september 20, 2018, proceedings 4*. Springer, 3–11.

Zhu, Z., He, X., Qi, G., Li, Y., Cong, B., and Liu, Y. (2023a). Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal mri. *Inf. Fusion* 91, 376–387. doi:10.1016/j.inffus.2022.10.022

Zhu, Z., He, X., Qi, G., Li, Y., Cong, B., and Liu, Y. (2023b). Brain tumor segmentation based on the fusion of deep semantics and edge information in multimodal mri. *Inf. Fusion* 91, 376–387. doi:10.1016/j.inffus.2022.10.022

Žlahtič, B., Završnik, J., Blažun Vošner, H., and Kokol, P. (2024). Transferring black-box decision making to a white-box model. *Electronics* 13, 1895. doi:10.3390/electronics13101895

frontiers | Frontiers in Bioengineering and Biotechnology

TYPE Brief Research Report
PUBLISHED 02 August 2024
DOI 10.3389/fbioe.2024.1369514

Check for updates

OPEN ACCESS

EDITED BY
Laura Cercenelli,
University of Bologna, Italy

REVIEWED BY
Martin L. Tomov,
Emory University, United States
Michaela Servi,
University of Florence, Italy
Camilla Gironi,
University of Bologna, Italy

*CORRESPONDENCE
Simona Celi,
s.celi@ftgm.it

RECEIVED 12 January 2024
ACCEPTED 15 July 2024
PUBLISHED 02 August 2024

CITATION
Capellini K, Ait-Ali L, Pak V, Cantinotti M, Murzi M, Vignali E, Fanni BM, Clemente A, Celi S and Gasparotti E (2024), Three-dimensional printed models as an effective tool for the management of complex congenital heart disease.
Front. Bioeng. Biotechnol. 12:1369514.
doi: 10.3389/fbioe.2024.1369514

COPYRIGHT
© 2024 Capellini, Ait-Ali, Pak, Cantinotti, Murzi, Vignali, Fanni, Clemente, Celi and Gasparotti. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Three-dimensional printed models as an effective tool for the management of complex congenital heart disease

Katia Capellini[1], Lamia Ait-Ali[2], Vitali Pak[3], Massimiliano Cantinotti[4], Michele Murzi[5], Emanuele Vignali[1], Benigno Marco Fanni[1], Alberto Clemente[6], Simona Celi[1]* and Emanuele Gasparotti[1]

[1]BioCardioLab, Bioengineering Unit, Fondazione Toscana Gabriele Monasterio, Massa, Italy, [2]Institute of Clinical Physiology, CNR, Massa, Italy, [3]Department of Pediatric Cardiac Surgery, Fondazione Toscana Gabriele Monasterio, Massa, Italy, [4]Department of Pediatric Cardiology, Fondazione Toscana Gabriele Monasterio, Massa, Italy, [5]Department of Adult Cardiac Surgery, Fondazione Toscana Gabriele Monasterio, Massa, Italy, [6]Department of Clinical Imaging, Fondazione Toscana Gabriele Monasterio, Pisa, Italy

Introduction: Three-dimensional printed models are widely used in the medical field for surgical and interventional planning. In the context of complex cardiovascular defects such as pediatric congenital heart diseases (CHDs), the adoption of 3D printed models could be an effective tool to improve decision-making. In this paper, an investigation was conducted into the characteristics of 3D printed models and their added value in understanding and managing complex pediatric congenital heart disease, also considering the associated cost.

Methods: Volumetric MRI and CT images of subjects with complex CHDs were retrospectively segmented, and the associated 3D models were reconstructed. Different 3D printing technologies and materials were evaluated to obtain the 3D printed models of cardiac structures. An evaluation of time and costs associated with the 3D printing procedure was also provided. A two-level 3D printed model assessment was carried out to investigate the most suitable 3D printing technology for the management of complex CHDs and the effectiveness of 3D printed models in the pre-surgical planning and surgical strategies' simulations.

Results: Among the different techniques, selective laser sintering resulted to be the most suitable due to its reduced time and cost and for the positive clinical feedback (procedure simulation, surface finish, and reproduction of details).

Conclusion: The adoption of 3D printed models contributes as an effective tool in the management of complex CHDs, enabling planning and simulations of surgical procedures in a safer way.

KEYWORDS

3D printing, 3D models, congenital heart disease, pediatric surgery, pre-planning

1 Introduction

Three-dimensional printing technology in medicine has developed very rapidly in recent years due to the several approaches where it can be applied and the provided advantages (Bozkurt and Karayel, 2021; Javaid et al., 2022). In the cardiovascular field, 3D printed models are used for several purposes such as medical students' education and training, patients' communications, planning of surgical and percutaneous interventions, and applications in mock circulatory systems to investigate fluid dynamics *in vitro* (Medero et al., 2017; Vukicevic et al., 2017; Buonamici et al., 2022; Santoro et al., 2022; Vignali et al., 2022; Fanni et al., 2023; Masoumkhani et al., 2023; Stomaci et al., 2023).

This great development is strictly correlated with the advance in technology of volumetric medical imaging, and the resolution and signal-to-noise ratio are indeed crucial aspects to obtain an accurate 3D printed model. A certain demarcation between the structures of interest and the other image regions is one of the main starting points in the process of 3D model reconstruction (Fan et al., 2019). Although 3D echocardiography concerns mostly 3D models of cardiac valves and small defects due to a limited field of view, computed tomography (CT) and magnetic resonance imaging (MRI) are the most diffused acquisition techniques suitable for 3D model generation of several cardiovascular pathological anatomies. CT images are characterized by high spatial resolution and contrast; however, their adoption is not recommended for pediatric subjects, given the necessity to inject an iodinated agent to enhance contrast and the ionizing nature of the radiations involved. On the other hand, MRI does not involve ionizing radiation and presents a good tissue to blood contrast but is affected by a poorer spatial resolution and more artifacts compared to CT (Celi et al., 2021).

Starting from the elaboration of these aforementioned medical images, 3D printed models can be realized with several kinds of 3D printing technologies and materials, resulting in different characteristics in terms of quality, color, opacity, deformability, time, and costs (Gharleghi et al., 2021).

The most commonly used 3D printing technologies include fused deposition modeling (FDM), stereolithography (SLA), and selective laser sintering (SLS).

Regarding FDM technology, the material is deposited in layers to form a 3D object by adopting a thermoplastic filament melted from a heated nozzle (Awasthi and Banerjee, 2021). FDM is adopted in several works evaluating the inclusion of 3D printed models in the surgical planning of cardiac defects (Valverde et al., 2017; Capellini et al., 2020).

The SLA technique consists in a UV laser light that induces a polymerization process with a tank filled with photo-polymeric resin. SLA technology allows for only one material printing at a time, and external and internal supports are required to prevent model collapse. Rigid or soft resin materials can be used, allowing the adoption of SLA in cases of training (Mafeld et al., 2017) and pre-planning for the percutaneous procedure (Capellini et al., 2020; Kaufmann et al., 2023).

SLS involves the powder bed fusion technology, which utilizes the laser energy to heat and melt powder particles (Lekurwale et al., 2022). SLS technology allows for only one material printing at a time, and it does not require the adoption of external and internal supports during the fabrication process (Kappanayil et al., 2017). It is possible to adopt opaque rigid or soft materials.

Congenital heart disease (CHD) is one of the most widespread congenital pathology in infants (Wu et al., 2020). In particular, complex CHDs consist in the concomitant presence of more morphological defects (Festa et al., 2023), and they are characterized by uncommon anatomic relations; often, they require complex surgical approaches. The role of 3D models in the CHD understanding and surgery is well-established in the literature, together with their utility in teaching, training, and communication (Capellini et al., 2020; Awori et al., 2021; Karsenty et al., 2021). Although there are some studies in the literature that highlight the benefits of 3D virtual models in the field of CHDs (Ong et al., 2018; Awori et al., 2023), the adoption of 3D printed models is also able to provide a tactile response and more realistic understanding of depth and anatomic relationships with surgeons (Costello et al., 2015; Celi et al., 2021; Illi et al., 2022).

However, clinicians' feedback arising from both the effectiveness in the decision-making procedure for pediatric complex CHDs and the different kinds of 3D printing technologies and materials needs further investigation. This study aims to assess the additional value of 3D printed models in the pre-planning of complex CHDs as well, with respect to different 3D printing techniques and materials.

2 Materials and methods

2.1 Image data

Ten pre-surgical image datasets of patients (five male subjects and five female subjects with an average age of 4 years) presenting complex CHDs scheduled for the surgical procedure were retrospectively analyzed in this study. In particular, two volumetric CT and eight MRI datasets acquired with 320-detector scanner (Toshiba Aquilion One, Toshiba, Japan) and 3 Tesla scanners (Ingenia, Philips Medical Systems, Netherlands), respectively, were considered. Examples of CT and MR datasets with the associated volume rendering representations are shown in Figures 1A, B, respectively.

2.2 Image processing

After an operation of image cropping to reduce the dataset volume to the region of interest, a segmentation process is required to obtain the 3D model of the anatomical structures. All the segmentation procedures described below were implemented by using 3D Slicer, a free and open-source software application (www.slicer.org) for the analysis and segmentation of biomedical images (Kikinis et al., 2013). A threshold algorithm was initially applied for the segmentation of both CT and MRI datasets, given the presence of contrast between blood in the vessels and heart chambers and non-vascular tissues. Secondary, a semi-automatic technique based on the region growing algorithm was applied together with a final phase of manual editing, often consisting in a slice by slice segmentation, due to the difficulty for automatic methods to accurately identify regions with complex defects or artifacts. This procedure was carried out by highly specialized biomedical engineers in the cardiovascular

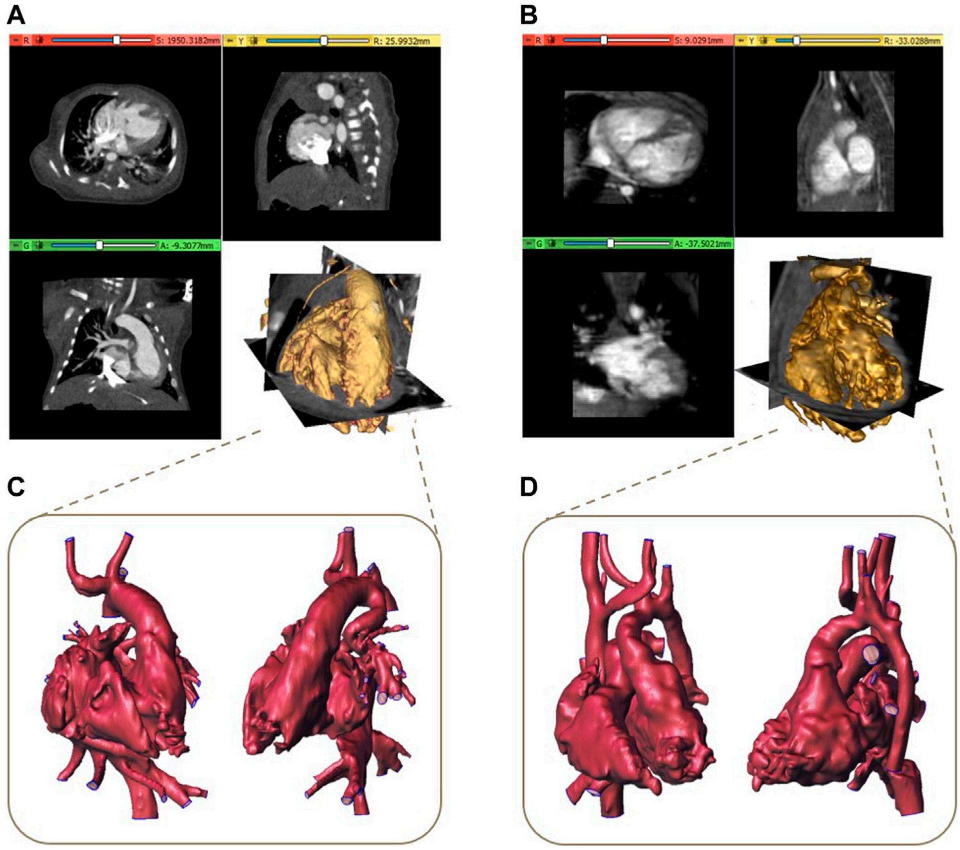


FIGURE 1
Examples of CT **(A)** and MR dataset **(B)** with the associated volume rendering visualization and the corresponding reconstructed 3D models **(C, D)**.

field with specific attention to cardiac structures with more than 5 years of experience in complex CHD segmentation with the support of clinicians with expertise in cardiac imaging. Starting from the obtained segmentation binary mask, a preliminary 3D model was reconstructed and then refined by performing detection and correction of any mesh holes, a removal of non-manifold edges and islands. Examples of 3D final models are reported in Figures 1C, D. To obtain the final 3D model suitable for printing, for all the models, the surfaces were thickened of 0.8 mm outward from the mesh to maintain the actual dimensions of the reconstructed blood cavities.

2.3 Three-dimensional printing technologies

Three different types of 3D printing technology were tested in this study: FDM, SLA, and SLS.

First, a subset of three 3D models was realized with all aforementioned 3D printing technologies to perform an accurate comparison and evaluation of the most suitable technique for the complex CHDs. Finally, the remaining 3D models were realized with the chosen technique. Flexible materials with a comparable shore, approximately 80 A, were chosen, given its availability for each considered 3D printing technology and its applications for the realization of anatomical models (Gómez-Ciriza et al., 2021; Lau et al., 2021; Lau and Sun, 2022; Kaufmann et al., 2023). Regarding the FDM technology, a G-code file was generated by using slicing SSI software (3ntr, Oleggio, Italy), allowing the definition of the model position and orientation on the printing plate and the selection of supports necessary for the printing. An A4v4 printer (3ntr, Oleggio, Italy) was adopted (printing volume: $30.0 \times 17.1 \times 20.0$ cm^3). The model was printed in a thermoplastic polyurethane printing filament (TPU Elasto85) with shore 85 A. With regards to the printing parameters, the thickness of the layer was chosen according to the selected material. An adaptable layer ranging from 0.15 mm to 0.25 mm was adopted together with an infill of 100%, given the thickness of the model. A polyvinyl alcohol (PVA) material (SSU04) was used as support, given its water-soluble properties. The post-processing steps consisted in the manual removal of the external support and then by using a tank filled of warm water with ultrasound enabled to dissolve the internal support. The FDM technology needs approximately 310×90 cm to house the printer, the ultrasound washer, and a workbench for the manual operations. The expanses of this technique include the equipment cost (evaluated equal 0.8 €/h), the material cost (Elasto85 = 104 €/kg, SSU04 = 121 €/kg), and the post-processing equipment cost (estimated 0.1 €/h).

Three-dimensional printing with SLA technology was allowed by using Preform software (Formlabs, Somerville, United States). The position and the orientation of the model were defined to allow the easier removal of supports and guarantee the highest quality of the printed details. The SLA models were fabricated through Form

3BL printer (Formlabs, Somerville, United States), characterized by a maximum printing volume of $33.5 \times 20.0 \times 30.0\ cm^3$. The model was printed in a deformable resin (Flexible 80 A) with a shore value of 80 A. Each model was realized with a layer thickness of 0.1 mm, compatible with the selected material. Post-processing of SLA consisted of cleaning the model from uncured resin by a washing procedure using isopropyl alcohol, then curing in an ultraviolet and warm environment, and finally the removal of the outer and inner supports by using proper scissors.

The SLA requires approximately 450×90 cm of space to accommodate the printer machine, the washer, the curer, and a workbench for the post-processing procedures. The equipment cost of SLA was evaluated to be equal to 1.25 €/h for the printing process and the material cost of Flexible 80 A was 242.8 €/kg, whereas the equipment cost for the post-processing was estimated to be equal to 0.42 €/h.

The SLS technology was defined by using Sinterit Studio Advanced (Sinterit, Krakow, Poland), where the position and the orientation of each model were set to guarantee the highest quality of the model and reduce the printing time. The SLS CHD models were realized through a Lisa PRO printer (Sinterit, Krakow, Poland) with a printing volume of $11 \times 15 \times 25\ cm^3$. The material used was Flexa Bright, with 79 A of shore. Each model was realized with a layer thickness of 0.125 mm. SLS post-processing consisted of cleaning the machine using the unsintered powder, removing the powder deposited on model surfaces during printing by vacuum suction, air sanding, and finally cleaning with soap and water. The SLS needs approximately 450×90 cm to host the printer, the powder sieve, the sandblaster, the vacuum cleaner, and a workbench for the manual operations. The equipment cost of SLS was evaluated to be equal to 1.7 €/h, whereas the material cost of Flexa Bright was 250 €/kg and the post-processing equipment cost was estimated to be 0.36 €/h.

Supplementary Table S1 summarizes the main features of 3D printing technologies adopted in this work. Regarding the 3D printing equipment, the hourly costs were calculated considering a lifespan of 5 years (15,000 h). In addition to the abovementioned costs, the cost of the printer operator associated with manual activities performed before and after the printing process was also considered. The operator involved in this study had more than 1 year of experience in the 3D printing technologies applied for cardiovascular anatomies, and the associated hourly labor cost was 20 €/h (it is important to report that this cost depends on each country and legislation). The total print time, encompassing setup time (for G-code generation and printer preparation), printing time, and post-processing time were assessed.

2.4 Evaluation of 3D printed models' effectiveness

A two-level 3D printing effectiveness evaluation was carried out. The effectiveness was evaluated by a clinical team that did not take part in the pre-planning and surgical phases of the selected cases. The team was composed by a total of six physicians, of which three were cardiologists and three were pediatric surgeons. The first level of evaluation assessed which adopted 3D printing technologies and materials were the most suitable for the management of complex CHDs. A subset of 3D models, including the most representative cardiac defects, such as transposition of the great arteries, crisscross heart, and ventricular and atrial septal defects, was chosen for the evaluation and printed with the considered technologies. In particular, the clinical team assessed three different parameters:

- Surface—surface finishing.
- Details—the capability to replicate anatomical details with high accuracy.
- Behavior—the material feedback for the simulation of the surgical procedure, considering the level of deformability and the behavior of the model to be cut with the scalpel.

Each member of the clinical team was asked to assign a score from 1 to 5 for these parameters for each printed model. In this way, each technology was evaluated on the basis of 54 scores (three parameters × three models × six clinicians), with 18 scores for each parameter, arising from the clinicians rate to each printed model.

The 3D printing technology characterized by the best scores in the first assessment was selected to realize all 10 cases included in the second level of evaluation. This approach allowed cost and time savings by avoiding the adoption of 3D printing technologies that received lower ratings from clinicians. For each model, the pre-planning of the surgical procedure was carried out by the clinical team. In particular, in this phase, the surgical strategy was defined considering for each case, at first, only the available clinical images (a); then, a strategy reassessment was performed including the 3D printed model as well (b). In this way, the impact of 3D printing on the management of complex CHDs was estimated by investigating whether and how the additional information provided by the 3D printed model modify the course of surgical planning and eventually the surgical decision. In particular, the clinical team evaluated the possibility to better understand the pathology as well by exploring internal heart chambers from different views, the time of pre-planning, and the potential improvement in the communication between clinicians. The evaluation was performed on all CHD models. The clinical team was required to assign a rating ("don't know," "worsen," "irrelevant," "quite relevant," "relevant," and "very relevant") in a specific survey.

3 Results

3.1 Image processing

Regarding the segmentation process, manual editing was necessary in most of the MRI datasets, especially in the cases of newborns due to the complexity of structures and presence of motion artifacts. However, for the CT datasets, which required further segmentation phase after the threshold algorithm application, semi-automatic methods limited to restricted volume regions were preferred due to the higher spatial resolution and size of datasets. The 3D models were successfully reconstructed for all selected image datasets, and the segmentation times requested for the generation of 3D models are reported in Table 1. The time values varied depending on the complexity of CHDs, the image contrast

TABLE 1 Image processing and 3D printing characteristics. "–" means no printed model for the technique.

Case	Image modality	Segmentation time (h)	Printing volume (cm^3)	Printing time (h)		
				FDM	SLA	SLS
1	MR	7.3	8.6 × 6.2 × 11.2	19.3	19.5	16.0
2	CT	9.2	6.5 × 5.8 × 9.7	10.7	15.0	10.3
3	MR	10.4	8.6 × 6.9 × 10.2	17.0	17.1	14.0
4	MR	4.5	9.8 × 6.7 × 12.5	–	–	19.1
5	MR	5.3	10.5 × 9.3 × 14.3	–	–	23.5
6	MR	8.1	10.6 × 7.0 × 13.0	–	–	21.2
7	MR	10.4	8.9 × 6.7 × 7.6	–	–	13.0
8	CT	8.5	8.4 × 6.1 × 8.5	–	–	14.5
9	MR	7.5	10.6 × 9.3 × 13.2	–	–	25.0
10	MR	7.3	14.4 × 10.5 × 9.6	–	–	20.0

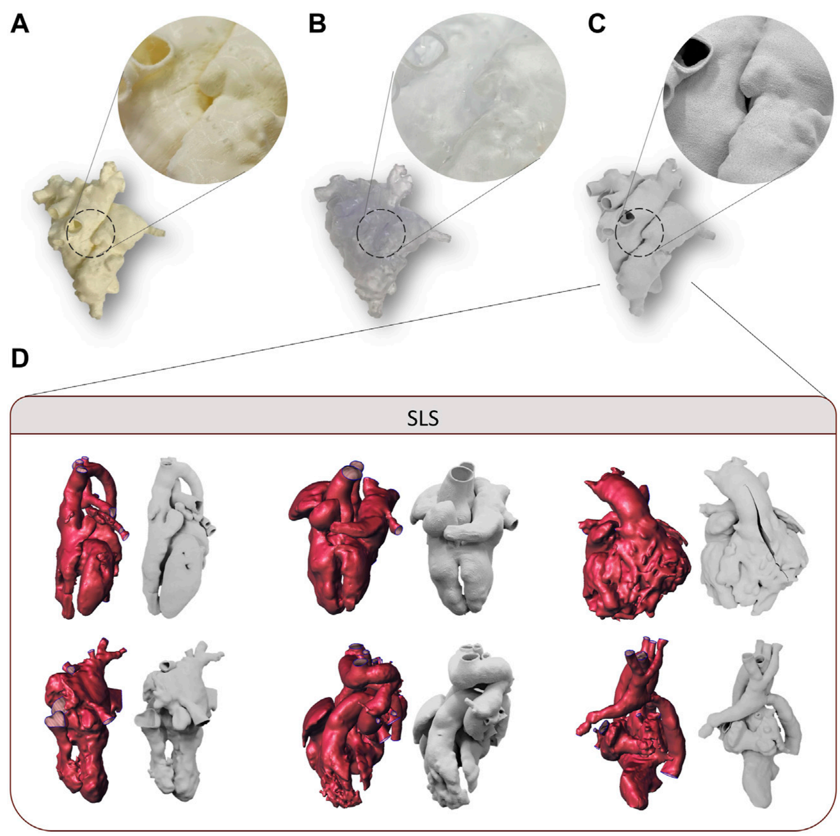


FIGURE 2
Examples of the 3D printed model of case 3 by using FDM **(A)**, SLA, **(B)** and SLS **(C)** for first level of evaluation, with detail magnification. Examples of 3D models and the corresponding 3D printed models of cases manufactured with SLS and used for second level of evaluation **(D)**.

and resolution, and the type of adopted segmentation algorithm, ranging between 4.5 and 10.4 h. Table 1 also reports the printing volume associated with the segmented 3D models, and it is possible to observe that the dimensions of the considered models allow the adoption of the considered printing technologies.

3.2 Three-dimensional printing

FDM technology allowed the printing of all selected models without complications related to anatomical complexity, reproducing both the external and the internal structures without

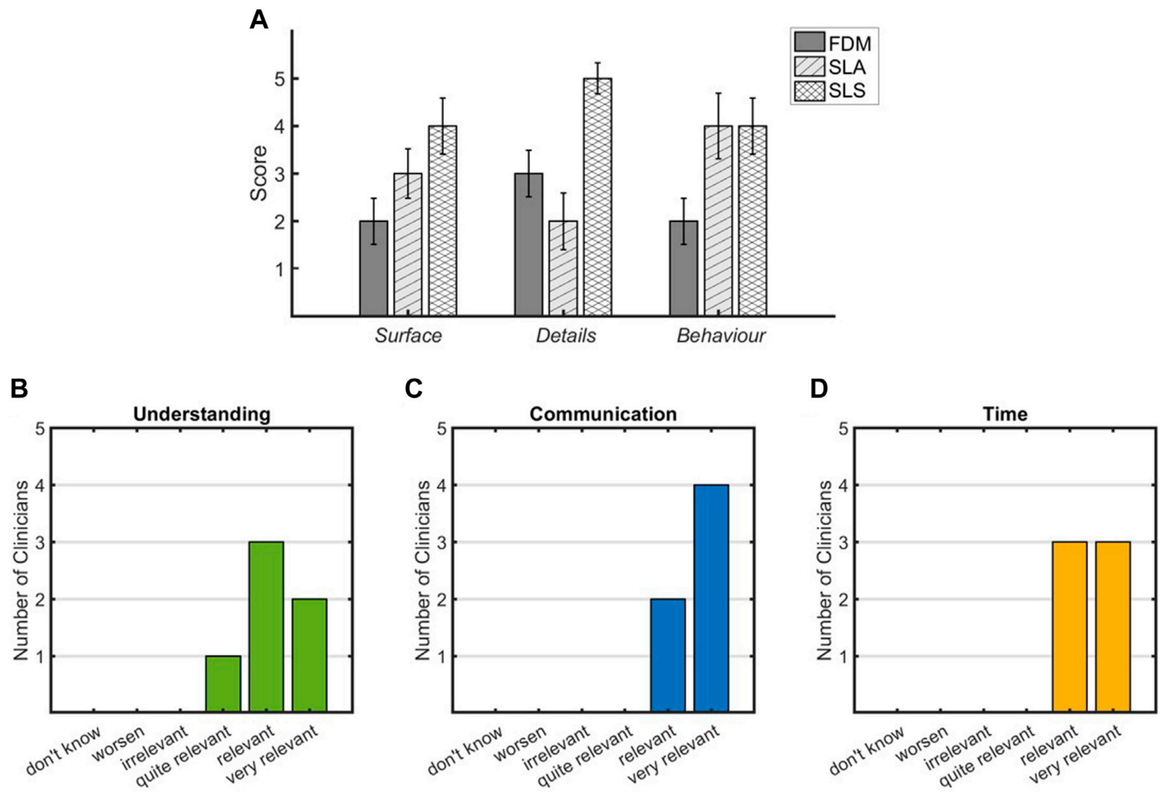


FIGURE 3
Reports of the clinical team evaluation for 3D printing technologies (first level of evaluation) **(A)** and for the impact of 3D printed models on pre-planning (second level of evaluation) **(B–D)**.

macroscopic defects (Figure 2A). The printing time to realize the models with FDM technology increased with the increasing printing volume of the model, showing a mean value of 15.7 h (Table 1). The post-processing time was approximately 6 h, considering a washing time of 5 h, and it is independent from the printing volume in the range of complex CHDs considered. The cost related to the FDM process is proportional to the printing volume with a mean value of 77 €.

SLA technology allowed the printing of all selected models, although some critical issues arose in the case of small complex structures characterizing the CHDs. In fact, for all the models analyzed, the main issues were found in the post-processing phase. In particular, the extraction of supports without damaging the model was made difficult by the need to use high-density supports for internal structures, the small size of the models, and the lack of holes of access usable for the scissors. In addition, given the presence of high-density support, support imprints on the surface of the models were observed (Figure 2B). The printing time of SLA technology increased with the increasing printing volume of the model, showing a mean value of 17 h (Table 1). The post-processing time was approximately 2 h, considering 30 min for model washing and curing, and the support manual remotion is dependent on the printing volume and the degree of complexity of the CHDs. The cost related to the SLA process is proportional to the printing volume, with a mean value of 81€.

The realization of all 3D models was feasible through the SLS printer. The relatively small dimensions of the printing plate, in fact, did not represent a limitation for the printing because of the small size of pediatric heart models. SLS was able to reproduce both the external and the internal structures of the 3D model without macroscopic defects, with accurate adhesion of the printer layers (Figure 2C). As observed for the other two technologies, SLS exhibited printing time proportional to the model volume, with a mean value of 13.4 h (Table 1). The post-processing time was independent from the model weight, and it was equal to 1 h. The mean cost related to SLS was equal to 56€. Some examples of 3D printed models by SLS are reported in Figure 2D.

3.3 Evaluation of 3D printed models' effectiveness

Regarding the comparison of the tested 3D printing technologies, the mean value of the scores assigned by the clinical team is shown in Figure 3A. The team assigned the highest value of *Surface* score to the SLS technique as it presented satisfactory layer adhesion and good surface finish. SLA achieved a *Surface* rating of 3 as the imprint of the removed supports was present on some regions of the external surface. The FDM reached the lower value as it presents a coarse surface with visible layers. In terms of the *Details* parameter, the SLS achieved the highest value as the anatomical details were reproduced with high quality on both the external and inner surfaces. The FDM model showed a lower quality of details on the internal surfaces, while the SLA model reached the lowest value of 2 as the inspection of the internal details was limited by the presence of the support and the transparency of the model. The last parameter was *Behavior*; the SLS and SLA scored the highest because the elasticity offered allowed surgical incision and inspection without damaging the model. In addition, the material compliance was more similar to cardiac tissue than the FDM material when handled by the surgeon. The lower score for FDM (equal to 2) was due to the stiffer behavior of the material, which made the cutting phase more difficult.

As concerns for the second-level assessment, the survey was completed by the clinicians for all cases. Figures 3B–D report the results in terms of rating assigned to each parameter. Regarding the understanding of individual complex CHDs, all experts agreed that SLS-printed models provide an increase in the comprehension of anatomical details and relations that characterize the pathology, as reported in Figure 3B, where five out of six clinicians assigned a rate of "relevant" and "very relevant." All clinicians were agreed on the relevance of the SLS-printed model in the communication among physicians during the decision-making procedure (Figure 3C). The inclusion of the 3D printed model in the clinical discussion showed a relevant impact on the time needed for the decision-making procedure that resulted to be reduced (Figure 3D). The importance of performing strategies' simulations on the 3D printed model was highlighted in eight cases, and for 4 out of 10 cases, the surgical strategy resulted different, compared to that performed based only on clinical images.

4 Discussion

In this work, an investigation of the main diffused 3D printing technologies and added value of 3D printed models in the management of complex CHDs was carried out. In the last few years, the adoption of 3D virtual and printed models in the cardiac field has significantly increased (Gasparotti et al., 2019; Gardin et al., 2020; Ma et al., 2021). In particular, their usage can be crucial in the understanding, pre-planning, and simulation of surgical procedures of pediatric complex CHDs.

In our study, 10 complex CHD anatomies were analyzed starting from the segmentation of medical images, up to 3D printing realization and clinical evaluation to assess the applicability of the 3D printed models in pre-planning. The clinical images play a central role in the process of 3D model realization, in particular MR and CT acquisitions are the most used in the cases of complex CHDs. In addition to the advantages and drawbacks associated with the above cited imaging modalities and already described in Section 1, the proper balance between spatial resolution and image contrast is a crucial aspect. A good resolution in terms of small voxel size allows the visualization and reconstruction of small structures details but involves a decrease in the signal-to-noise ratio with consequent difficulties associated with segmentation algorithm application (Otton et al., 2017).

In the context of a clinical application, the total print time is a parameter that characterizes the availability of the printed model, following a clinical request. All considered techniques result in compliance with the clinical practice, allowing the model to be available to the clinician within 24 h on average after the segmentation procedure (7.8 h). Among the techniques, the SLS resulted to be the fastest method, requiring 34% and 28% less time with respect to FDM and SLA. This is related to the smaller amount of material used for printing due to the absence of supports.

Moreover, the spaces required by the 3D printing techniques inside a clinical environment have to be taken into account. The considered 3D printing technologies require a comparable amount of space, and a dedicated room should be needed for any of them due to the characteristics of materials involved in setup, printing, and post-processing phases. In addition, the cost associated with 3D printing is an important aspect discussed in the literature (Lau et al., 2019; Yoo et al., 2021) since it represents one of the discriminating factors for the inclusion of this procedure in the clinical routine management of CHDs. Other technologies are widespread in the field of 3D printing such as PolyJet technology by Stratasys that permits the realization of 3D models varying the shore values of the printed structures. In particular, the Digital Anatomy printer is used to create cardiac structures with different stiffness. Nevertheless, these other technologies have an equipment acquisition cost of over 70 k€ (Chae et al., 2015; Chen et al., 2021), while in this work, we focused on 3D printer machines available in our center, characterized by a cost lower than 30 k€. The tested 3D printing technologies were selected, taking into account both the technical characteristics associated with the printing models and the associated cost-effectiveness that is fundamental in a clinical scenario. From the evaluation of the costs of the 3D printing technologies tested in our work, it appears that, even if the cost of producing a model is similar for the analyzed technologies and therefore compatible with clinical application, the SLS process is the cost-effective method with a cost-benefit ratio 28% and 31% lower than FDM and SLA, respectively. Although SLS has the highest cost of material per kilogram, this technology minimizes material waste, and the costs for post-processing are the lowest due to the absence of supports.

The accuracy in the delineating anatomical details is an explored factor for the 3D printing in CHDs. The clinical team attested the low resolution in the reproduction of complex and small anatomical details of FDM printed models, together with the worst tactile feedback due to both the surface roughness and the response to the surgical incision. Although a good resolution of 3D printed models was obtained by SLA technology, the impossibility of a perfect removal of the imprints of external supports negatively affected the surface finish (Figure 3A). Moreover, the simulation of the surgical procedure and the exploration of internal cavities was not always feasible or positively evaluated by the surgical team due to the presence of internal residual supports. On the other hand, the tactile feedback and the response to surgical incision were adequate for all clinicians. Similar behavior was recorded for SLS models, which provided, in addition, the best model quality both in terms of surface finish and accuracy in anatomical detail reproduction. Moreover, the low value of the resulting standard deviation (Figure 3A) denotes an agreement within the clinicians in the evaluation of 3D printing techniques. On the basis of the above reported results, the 3D printed models obtained with SLS technology turned out to be the most suitable for the surgical planning of complex CHDs.

The surgical pre-planning phase of complex CHDs is challenging and time-consuming due to anatomies characterized by uncommon anatomic relations, morphological abnormalities, and often very small structures that can be very dissimilar among the patients (Festa et al., 2023). The inclusion of SLS-based 3D printed models in the pre-planning allows the 3D visualization of the anatomical structures from different perspectives, the manipulation of 1:1 scaled geometries, the exploration of the heart chambers from a surgical point of view, the visualization of the vessels, and chamber spatial relationships. All these aspects have a relevant impact on the decision-making procedure and surgical strategy, bringing an added value in terms of understanding, communication, and time-saving as arisen from clinicians' response (Figures 3B–D). Regarding the time for the 3D printed model realization, the duration of the segmentation phase is a crucial step that is significantly dependent on the quality of image dataset, the complexity of investigated scenario, and the competences and experience of the multidisciplinary team. The resulting segmentation times in this study were obtained by highly specialized biomedical engineers in the cardiovascular field focused on cardiac structures with more than 5 years of experience. Even if the artificial intelligence applied to image processing is an emerging instrument to obtain an automatic and faster segmentation mask (Garzia et al., 2023), its adoption in the complex CHDs is challenging (Karimi-Bidhendi et al., 2020; Yao et al., 2023) due to the structure relations and abnormalities that can be different for each case. Although the 3D printed models turn out to be a powerful tool to improve the pre-planning procedure for the complex CHDs, their inclusion in the clinical setting is still limited due to the associated costs and time. These aspects could be significantly improved if the facilities required by the 3D printing procedure are present inside the clinical structures without using external service.

5 Conclusion

Three-dimensional printing represents an effective instrument for the management of complex CHDs. Among the main diffused technologies, the SLS resulted to be the most suitable for the understanding and planning of surgical procedures, especially for the accuracy of detail reproduction and for the best tactile feedback. The effectiveness of 3D printed models obtained with SLS technology arose from clinicians' response to a specific satisfaction survey. The presented work demonstrated the improvement provided by 3D printed models in the process of decision-making in the field of complex CHDs.

Data availability statement

The datasets presented in this article are not readily available because the raw data contains sensitive data. Requests to access the datasets should be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by Comitato Etico della Regione Toscana–Pediatrico/PO1007. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin.

Author contributions

KC: Conceptualization, Methodology, Visualization, Writing–original draft, Writing–review and editing. LA-A: Validation, Writing–review and editing. VP: Validation, Writing–review and editing. MC: Conceptualization, Validation, Writing–review and editing, Funding acquisition. MM: Validation, Writing–review and editing. EV: Writing–review and editing. BF: Writing–review and editing. AC: Writing–review and editing. SC: Conceptualization, Methodology, Visualization, Writing–original draft, Writing–review and editing, Funding acquisition. EG: Conceptualization, Methodology, Visualization, Writing–original draft, Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the 3D VIRTUAL BABY HEART project (2018–2020) funded by the Italian Ministry of Health (grant number: GR-2016-02365072), Italy.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2024.1369514/full#supplementary-material

References

Awasthi, P., and Banerjee, S. S. (2021). Fused deposition modeling of thermoplastic elastomeric materials: challenges and opportunities. *Addit. Manuf.* 46, 102177. doi:10.1016/j.addma.2021.102177

Awori, J., Friedman, S. D., Chan, T., Howard, C., Seslar, S., Soriano, B. D., et al. (2021). 3d models improve understanding of congenital heart disease. *3D Print. Med.* 7, 26–29. doi:10.1186/s41205-021-00115-7

Awori, J., Friedman, S. D., Howard, C., Kronmal, R., and Buddhe, S. (2023). Comparative effectiveness of virtual reality (vr) vs 3d printed models of congenital heart disease in resident and nurse practitioner educational experience. *3D Print. Med.* 9, 2–8. doi:10.1186/s41205-022-00164-6

Bozkurt, Y., and Karayel, E. (2021). 3d printing technology; methods, biomedical applications, future opportunities and trends. *J. Mater. Res. Technol.* 14, 1430–1450. doi:10.1016/j.jmrt.2021.07.050

Buonamici, F., Mussi, E., Santarelli, C., Carrabba, N., Stefano, P., Marchionni, N., et al. (2022). Modelling and fabrication procedure for a 3d printed cardiac model-surgical planning of left ventricular aneurysm. *MethodsX* 9, 101822. doi:10.1016/j.mex.2022.101822

Capellini, K., Tripicchio, P., Vignali, E., Gasparotti, E., Ali, L. A., Cantinotti, M., et al. (2020). 3d printing and 3d virtual models for surgical and percutaneous planning of congenital heart diseases. *VISIGRAPP* 3 (IVAPP), 281–287. doi:10.5220/0009160702810287

Celi, S., Gasparotti, E., Capellini, K., Vignali, E., Fanni, B. M., Ali, L. A., et al. (2021). 3d printing in modern cardiology. *Curr. Pharm. Des.* 27, 1918–1930. doi:10.2174/18734286mta3fntysx

Chae, M. P., Rozen, W. M., McMenamin, P. G., Findlay, M. W., Spychal, R. T., and Hunter-Smith, D. J. (2015). Emerging applications of bedside 3d printing in plastic surgery. *Front. Surg.* 2, 25. doi:10.3389/fsurg.2015.00025

Chen, J. V., Dang, A. B., and Dang, A. (2021). Comparing cost and print time estimates for six commercially-available 3d printers obtained through slicing software for clinically relevant anatomical models. *3D Print. Med.* 7, 1–14. doi:10.1186/s41205-020-00091-4

Costello, J. P., Olivieri, L. J., Su, L., Krieger, A., Alfares, F., Thabit, O., et al. (2015). Incorporating three-dimensional printing into a simulation-based congenital heart disease and critical care training curriculum for resident physicians. *Congenit. heart Dis.* 10, 185–190. doi:10.1111/chd.12238

Fan, Y., Wong, R. H., and Lee, A. P.-W. (2019). Three-dimensional printing in structural heart disease and intervention. *Ann. Transl. Med.* 7, 579. doi:10.21037/atm.2019.09.73

Fanni, B. M., Gasparotti, E., Vignali, E., Capelli, C., Positano, V., and Celi, S. (2023). An integrated *in-vitro* and in-silico workflow to study the pulmonary bifurcation hemodynamics. *Comput. Fluids* 260, 105912. doi:10.1016/j.compfluid.2023.105912

Festa, P., Ciancarella, P., Ait Ali, L., and Secinaro, A. (2023). "Complex congenital heart diseases," in *Case-based atlas of cardiovascular magnetic resonance* (Springer), 317–338.

Gardin, C., Ferroni, L., Latremouille, C., Chachques, J. C., Mitrečić, D., and Zavan, B. (2020). Recent applications of three dimensional printing in cardiovascular medicine. *Cells* 9, 742. doi:10.3390/cells9030742

Garzia, S., Scarpolini, M. A., Mazzoli, M., Capellini, K., Monteleone, A., Cademartiri, F., et al. (2023). Coupling synthetic and real-world data for a deep learning-based segmentation process of 4d flow mri. *Comput. Methods Programs Biomed.* 242, 107790. doi:10.1016/j.cmpb.2023.107790

Gasparotti, E., Vignali, E., Losi, P., Scatto, M., Fanni, B. M., Soldani, G., et al. (2019). A 3d printed melt-compounded antibiotic loaded thermoplastic polyurethane heart valve ring design: an integrated framework of experimental material tests and numerical simulations. *Int. J. Polym. Mater. Polym. Biomaterials* 68, 1–10. doi:10.1080/00914037.2018.1525717

Gharleghi, R., Dessalles, C. A., Lal, R., McCraith, S., Sarathy, K., Jepson, N., et al. (2021). 3d printing for cardiovascular applications: from end-to-end processes to emerging developments. *Ann. Biomed. Eng.* 49, 1598–1618. doi:10.1007/s10439-021-02784-1

Gómez-Ciriza, G., Gómez-Cía, T., Rivas-González, J. A., Velasco Forte, M. N., and Valverde, I. (2021). Affordable three-dimensional printed heart models. *Front. Cardiovasc. Med.* 8, 642011. doi:10.3389/fcvm.2021.642011

Illi, J., Bernhard, B., Nguyen, C., Pilgrim, T., Praz, F., Gloeckler, M., et al. (2022). Translating imaging into 3d printed cardiovascular phantoms: a systematic review of applications, technologies, and validation. *Basic Transl. Sci.* 7, 1050–1062. doi:10.1016/j.jacbts.2022.01.002

Javaid, M., Haleem, A., Singh, R. P., and Suman, R. (2022). 3d printing applications for healthcare research and development. *Glob. Health J.* 6, 217–226. doi:10.1016/j.glohj.2022.11.001

Kappanayil, M., Koneti, N. R., Kannan, R. R., Kottayil, B. P., and Kumar, K. (2017). Three-dimensional-printed cardiac prototypes aid surgical decision-making and preoperative planning in selected cases of complex congenital heart diseases: early experience and proof of concept in a resource-limited environment. *Ann. Pediatr. Cardiol.* 10, 117. doi:10.4103/apc.apc_149_16

Karimi-Bidhendi, S., Arafati, A., Cheng, A. L., Wu, Y., Kheradvar, A., and Jafarkhani, H. (2020). Fully-automated deep-learning segmentation of pediatric cardiovascular magnetic resonance of patients with complex congenital heart diseases. *J. Cardiovasc. magnetic Reson.* 22, 80. doi:10.1186/s12968-020-00678-0

Karsenty, C., Guitarte, A., Dulac, Y., Briot, J., Hascoet, S., Vincent, R., et al. (2021). The usefulness of 3d printed heart models for medical student education in congenital heart disease. *BMC Med. Educ.* 21, 480–488. doi:10.1186/s12909-021-02917-z

Kaufmann, R., Deutschmann, M., Meissnitzer, M., Scharinger, B., Hergan, K., Vötsch, A., et al. (2023). Manufacturing flexible vascular models for cardiovascular surgery planning and endovascular procedure simulations: an approach to segmentation and post-processing with open-source software and end-user 3d printers. *Int. J. Bioprinting* 9, 669. doi:10.18063/ijb.v9i2.669

Lau, I., Gupta, A., and Sun, Z. (2021). Clinical value of virtual reality versus 3d printing in congenital heart disease. *Biomolecules* 11, 884. doi:10.3390/biom11060884

Lau, I., and Sun, Z. (2022). The role of 3d printed heart models in immediate and long-term knowledge acquisition in medical education. *Rev. Cardiovasc. Med.* 23, 1–9. doi:10.31083/j.rcm2301022

Lau, I., Wong, Y. H., Yeong, C. H., Aziz, Y. F. A., Sari, N. A. M., Hashim, S. A., et al. (2019). Quantitative and qualitative comparison of low-and high-cost 3d-printed heart models. *Quantitative imaging Med. Surg.* 9, 107–114. doi:10.21037/qims.2019.01.02

Lekurwale, S., Karanwad, T., and Banerjee, S. (2022). Selective laser sintering (sls) of 3d printlets using a 3d printer comprised of ir/red-diode laser. *Ann. 3D Print. Med.* 6, 100054. doi:10.1016/j.stlm.2022.100054

Ma, Y., Ding, P., Li, L., Liu, Y., Jin, P., Tang, J., et al. (2021). Three-dimensional printing for heart diseases: clinical application review. *Bio-design Manuf.* 4, 675–687. doi:10.1007/s42242-021-00125-8

Mafeld, S., Nesbitt, C., McCaslin, J., Bagnall, A., Davey, P., Bose, P., et al. (2017). Three-dimensional (3d) printed endovascular simulation models: a feasibility study. *Ann. Transl. Med.* 5, 42. doi:10.21037/atm.2017.01.16

Masoumkhani, F., Fallah, A., Amani-Beni, R., Mohammadpour, H., Shahbazi, T., Bakhshi, A., et al. (2023). 3d printing for cardiovascular surgery and intervention: a review article. *Curr. Problems Cardiol.*, 102086. doi:10.1016/j.cpcardiol.2023.102086

Medero, R., García-Rodríguez, S., François, C. J., and Roldán-Alzate, A. (2017). Patient-specific *in vitro* models for hemodynamic analysis of congenital heart disease–additive manufacturing approach. *J. Biomechanics* 54, 111–116. doi:10.1016/j.jbiomech.2017.01.048

Ong, C. S., Krishnan, A., Huang, C. Y., Spevak, P., Vricella, L., Hibino, N., et al. (2018). Role of virtual reality in congenital heart disease. *Congenit. heart Dis.* 13, 357–361. doi:10.1111/chd.12587

Otton, J. M., Birbara, N. S., Hussain, T., Greil, G., Foley, T. A., and Pather, N. (2017). 3d printing from cardiovascular ct: a practical guide and review. *Cardiovasc. Diagnosis Ther.* 7, 507–526. doi:10.21037/cdt.2017.01.12

Santoro, G., Rizza, A., Pizzuto, A., Berti, S., Cuman, M., Gasparotti, E., et al. (2022). Transcatheter treatment of ascending aorta pseudoaneurysm guided by 3d-model technology. *Case Rep.* 4, 343–347. doi:10.1016/j.jaccas.2022.01.005

Stomaci, T., Buonamici, F., Gelati, G., Meucci, F., and Carfagni, M. (2023). 3d-printed models for left atrial appendage occlusion planning: a detailed workflow. *Rapid Prototyp. J.* 29, 74–81. doi:10.1108/rpj-10-2022-0351

Valverde, I., Gomez-Ciriza, G., Hussain, T., Suarez-Mejias, C., Velasco-Forte, M. N., Byrne, N., et al. (2017). Three-dimensional printed models for surgical planning of complex congenital heart defects: an international multicentre study. *Eur. J. Cardio-Thoracic Surg.* 52, 1139–1148. doi:10.1093/ejcts/ezx208

Vignali, E., Gasparotti, E., Mariotti, A., Haxhiademi, D., Ait-Ali, L., and Celi, S. (2022). High-versatility left ventricle pump and aortic mock circulatory loop development for patient-specific hemodynamic *in vitro* analysis. *ASAIO J.* 68, 1272–1281. doi:10.1097/mat.0000000000001651

Vukicevic, M., Mosadegh, B., Min, J. K., and Little, S. H. (2017). Cardiac 3d printing and its future directions. *JACC Cardiovasc. Imaging* 10, 171–184. doi:10.1016/j.jcmg.2016.12.001

Wu, W., He, J., and Shao, X. (2020). Incidence and mortality trend of congenital heart disease at the global, regional, and national level. *Medicine* 99, e20593–e22017. doi:10.1097/md.0000000000020593

Yao, Z., Xie, W., Zhang, J., Yuan, H., Huang, M., Shi, Y., et al. (2023). Graph matching and deep neural networks based whole heart and great vessel segmentation in congenital heart disease. *Sci. Rep.* 13, 7558–7611. doi:10.1038/s41598-023-34013-1

Yoo, S. J., Hussein, N., Peel, B., Coles, J., Arsdell, G. S. v., Honjo, O., et al. (2021). 3d modeling and printing in congenital heart surgery: entering the stage of maturation. *Front. Pediatr.* 9, 621672. doi:10.3389/fped.2021.621672

Frontiers in
Bioengineering and Biotechnology

Accelerates the development of therapies, devices, and technologies to improve our lives

A multidisciplinary journal that accelerates the development of biological therapies, devices, processes and technologies to improve our lives by bridging the gap between discoveries and their application.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

frontiers | Research Topics