frontiers | Research Topics

Artificial intelligence-based medical image automatic diagnosis and prognosis prediction

Edited by
Junchi Yan, Yukun Lai, Yi Xu, Yinqiang Zheng, Zhibin Niu
and Tao Tan

Published in
Frontiers in Physiology
Frontiers in Physics

ISSN 1664-8714
ISBN 978-2-8325-2661-3
DOI 10.3389/978-2-8325-2661-3

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

Artificial intelligence-based medical image automatic diagnosis and prognosis prediction

Topic editors
Junchi Yan — Shanghai Jiao Tong University, China
Yukun Lai — Cardiff University, United Kingdom
Yi Xu — Shanghai Jiao Tong University, China
Yinqiang Zheng — The University of Tokyo, Japan
Zhibin Niu — Tianjin University, China
Tao Tan — Eindhoven University of Technology, Netherlands

Citation
Yan, J., Lai, Y., Xu, Y., Zheng, Y., Niu, Z., Tan, T., eds. (2023). *Artificial intelligence-based medical image automatic diagnosis and prognosis prediction*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2661-3

Table of contents

frontiers | Frontiers in Physics

TYPE Editorial
PUBLISHED 25 May 2023
DOI 10.3389/fphy.2023.1210010

OPEN ACCESS

EDITED AND REVIEWED BY
Ewald Moser,
Medical University of Vienna, Austria

*CORRESPONDENCE
Tao Tan,
✉ taotan@mpu.edu.mo

RECEIVED 21 April 2023
ACCEPTED 10 May 2023
PUBLISHED 25 May 2023

CITATION
Yan J, Lai Y, Xu Y, Zheng Y, Niu Z and Tan T (2023), Editorial: Artificial intelligence-based medical image automatic diagnosis and prognosis prediction. *Front. Phys.* 11:1210010. doi: 10.3389/fphy.2023.1210010

Editorial: Artificial intelligence-based medical image automatic diagnosis and prognosis prediction

Junchi Yan[1], Yukun Lai[2], Yi Xu[3], Yinqiang Zheng[4], Zhibin Niu[5] and Tao Tan[6]*

[1]Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China, [2]School of Computer Science and Informatics, Cardiff University, Cardiff, United Kingdom, [3]Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai, China, [4]Department of Next Generation Artificial Intelligence Research Center for Education and Research in Information Science and Technology, The University of Tokyo, Bunkyo, Japan, [5]College of Intelligence and Computing/School of Media and Communication, Tianjin University, Tianjin, China, [6]Faculty of Applied Sciences, Macao Polytechnic University, Macau, Macao SAR, China

KEYWORDS

artificial intelligence, medical image 3D reconstruction, prognosis prediction, diagnosis, CAD

Editorial on the Research Topic
Artificial intelligence-based medical image automatic diagnosis and prognosis prediction

Computer-assisted diagnosis and prognosis prediction (especially with medical images) consist of a series of long-standing tasks, including classification, regression, segmentation, tracking tasks, etc. Deep learning is one of the most important breakthroughs in the field of artificial intelligence over the last decade. It has achieved great success because of enormously increasing data and computation resources. Not only there has been a constantly growing flow of related research papers, but also substantial progress has been achieved in real-world applications including axillary lymph node (ALN) metastasis status prediction, radiotherapy planning, histological image understanding and retina image recognition.

This Research Topic seeks to present and highlight the latest development on applying advanced deep learning techniques to cover the promising and novel deep learning algorithms in automatic diagnosis and prognosis prediction. It has attracted a fair number of submissions from researchers active in the study of automatic diagnosis and prognosis prediction using medical images or/and other sources of information. After careful peer review, nine manuscripts have finally been selected for publication in this Research Topic, covering the topics including advanced Transfer Learning techniques that transfer knowledge from other tasks or modals, multi-modal Learning techniques that enable multi-modal diagnosis or prognosis prediction, novel multi-task learning framework that enables joint diagnosis and prognosis prediction in a single model and advanced Unsupervised/Semi-Supervised/Weak-Supervised Learning techniques which boost performance with limited annotations.

In the field of cortical surface reconstruction in brain MR, An et al. proposed ResAttn-Recon, a residual self-attention based encoder-decoder framework with skip connections.

They also proposed a truncated and weighted L1 loss function to accelerate network convergence, compared to simply applying the L1 loss function. The average symmetric surface distance (AD) for the inner and outer surfaces is 0.253 ± 0.051 and the average Hausdorff distance (HD) is 0.629 ± 0.186, which is lower than that of DeepCSR, whose absolute distance equals 0.283 ± 0.059 and Hausdorff distance equals 0.746 ± 0.245.

In the area of object detection from Hu et al., it indicated that small object detection is one of the most challenging and important problems especially in medical scenarios. Complementary to the researches with attention to feature extraction and data augmentation of small objects, it proposed a method called pixel level balancing (PLB), which takes into account the number of pixels contained in the detection box as an impact factor to characterize the size of the inspected objects. This factor is then used as a weight in the training loss, so as to improve the accuracy of small object detection. Finally, the PLB operation is applied in the RPN stage of a two-stage detector. The experimental results demonstrated that it can improve the detection effect of small objects and maintain the accuracy of medium and large objects. Overall, the PLB method shows promise in addressing the challenges of small object detection in medical scenarios, particularly in tasks with higher requirements for small objects like blood cell detection.

For vital signs estimation, photoplethysmography (PPG) is a non-invasive method that measures the changes in blood volume in the skin using light. PPG-based devices can estimate blood pressure (BP), heart rate (HR), heart rate variability (HRV), and oxygen saturation (SpO_2) from the PPG signal. However, PPG measurements may be influenced by factors such as subcutaneous fat, skin color, and sex. The paper from Nachman et al. presents a study that compares BP measurements between a PPG-based device and a cuff-based device in different groups of people based on sex, BMI, and skin color. The study found that the PPG-based device had high accuracy and agreement with the cuff-based device across all groups, regardless of their personal characteristics. It also concluded that the PPG-based device can provide valid BP measurements for various populations and enable personalized BP management.

In fundus imaging, a deep learning method from Yao et al. called FunSwin is proposed to analyze diabetic retinopathy grade and macular edema risk based on fundus images. The method uses Swin Transformer, a hierarchical vision transformer, as the main framework, and integrates transfer learning and data augmentation strategies to improve the performance. The method is claimed to outperform other state-of-the-art methods in both binary and multiclass classification tasks on the MESSIDOR dataset, which contains 1,200 fundus images with labels for diabetic retinopathy and age-related macular degeneration.

For histopathological image analysis, Xiao et al. propose a deep learning framework called LAD-GCN for automatic estimation of growth patterns for lung adenocarcinoma diagnosis (LAD). The main idea is to jointly utilize a Graph Convolutional Network (GCN) to extract spatial structure features of cells and a Convolutional Neural Network (CNN) to extract global semantic features from histopathological images. To achieve this, cell nuclei in the images are first segmented using an existing instance segmentation model. By exploiting the complementary information, the proposed method achieves improved performance, as quantitatively validated on a lung adenocarcinoma dataset.

For computed tomography (CT) data analysis, the work by Feng et al. analyzes the influential factors from radiomics signature of pericoronary tissue (PCT) in coronary CT angiography (CCTA) for functional ischemia, measured using CT-derived fractional flow reserve (CCT-FFR). They segmented PCT from CT images and extracted 1,691 radiomic features of each vessel. They then performed feature selection using the Boruta algorithm built on top of random forest classifier to identify most contributive features to functional ischemia. The machine learning derived radiomics signature shows significant association in the study.

In the area of PET imaging, two experimental studies on Flourine-18 fluorodeoxyglucose positron emission tomography/computed tomography (18F-FDG PET/CT) images for diagnosis and monitoring of fatal diseases stand out of the selective reviewing process. The study by Shi et al. aims to investigate the role of radiomics analysis on 18F-FDG PET images for the sake of predicting microvascular invasion (mVI) in hepatocellular carcinoma (HCC), a common liver malignancy that leads to cancer death with a nontrivial probability. It further explores hybrid criteria combining PET/CT and multi-parameter MRI for higher prediction performance. Quantitatively, the 18F-FDG PET image radiomics classifier shows good performance in discriminating HCC with/without mVI, with an AUC of 0.917 (95% CI: 0.824 and 0.970) and 0.771 (95% CI: 0.578 and 0.905), and the hybrid model, which combines radiomics classifier and several key indicators based on contrast-enhanced MRI, yields much improved predictive performance with an AUC of 0.996 (95% CI: 0.939 and 1.000) and 0.953 (95% CI: 0.883 and 1.000). The study by Wang et al. tries to develop and validate 18F-FDG PET/CT image-based radiomics to determine the Ki-67 status of high-grade serous ovarian cancer (HGSOC), a disease with very high risk of recurrence and death. It is found that Radiomics based on Habitat can predict the Ki-67 expression accurately, and the Habitat model can better stratify the prognosis ($p < 0.05$).

In the field of pathology, this study Zhang et al. provides reliable machine learning-based (ML-based) models for predicting the probability of lymph node metastasis (LNM) in kidney cancer patients. The data was extracted from the Surveillance, Epidemiology, and Outcomes (SEER) database from 2010 to 2017, and variables were filtered using the least absolute shrinkage and selection operator (LASSO), univariate and multivariate logistic regression analyses. The independent predictive factors of LNM were identified as pathological grade, liver metastasis, M stage, primary site, T stage, and tumor size. Among six ML algorithms, the XGB model significantly outperformed any other machine learning models with an AUC value of 0.916 in the model validation process, in which M stage, T stage, and pathological grade were the top three important variables. Based on the probability density function (PDF) and clinical utility curve (CUC), the study suggested that 54.6% could be used as a threshold probability for the diagnosis of LNM using the XGB model, which could distinguish about 89% of LNM patients. As a conclusion, the machine learning-based predictive tool can accurately predict the probability of LNM in kidney cancer patients and has promising applications in clinical practice.

All nine papers tackle different but extremely relevant domain issues of artificial intelligence-based medical image automatic diagnosis and prognosis prediction. We believe this Research Topic will raise awareness in the scientific and industry community that a multidisciplinary research path is therefore in need to meet the desire from healthcare providers that are emerging in this field.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

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

Frontiers in Physiology

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fphys.2022.911297

Small Object Detection *via* Pixel Level Balancing With Applications to Blood Cell Detection

Bin Hu[1], **Yang Liu**[2,3]*, **Pengzhi Chu**[1], **Minglei Tong**[4]* and **Qingjie Kong**[5]

[1]*Department of Compute Science and Engineering, School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai, China,* [2]*Department of Dermatology, Shanghai Ninth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China,* [3]*Department of Laser and Aesthetic Medicine, Shanghai Ninth People's Hospital Affiliated to Shanghai Jiao Tong University School of Medicine, Shanghai, China,* [4]*College of Electronics and Information Engineering, Shanghai University of Electric Power, Shanghai, China,* [5]*Riseye Research, Riseye Intelligent Technology (Shanghai) Co., Ltd., Shanghai, China*

Edited by:
Zhibin Niu,
Tianjin University, China

Reviewed by:
Bo Jin,
East China Normal University, China
Xiaobin Zhu,
University of Science and Technology Beijing, China
Min Cao,
Soochow University, China
Changsheng Li,
Beijing Institute of Technology, China

***Correspondence:**
Yang Liu
liuyang870422@126.com
Minglei Tong
tongminglei@shiep.edu.cn

Specialty section:
This article was submitted to Medical Physics and Imaging, a section of the journal Frontiers in Physiology

Received: *02 April 2022*
Accepted: *24 May 2022*
Published: *17 June 2022*

Citation:
Hu B, Liu Y, Chu P, Tong M and Kong Q (2022) Small Object Detection via Pixel Level Balancing With Applications to Blood Cell Detection. Front. Physiol. 13:911297. doi: 10.3389/fphys.2022.911297

Object detection technology has been widely used in medical field, such as detecting the images of blood cell to count the changes and distribution for assisting the diagnosis of diseases. However, detecting small objects is one of the most challenging and important problems especially in medical scenarios. Most of the objects in medical images are very small but influential. Improving the detection performance of small objects is a very meaningful topic for medical detection. Current researches mainly focus on the extraction of small object features and data augmentation for small object samples, all of these researches focus on extracting the feature space of small objects better. However, in the training process of a detection model, objects of different sizes are mixed together, which may interfere with each other and affect the performance of small object detection. In this paper, we propose a method called pixel level balancing (PLB), which takes into account the number of pixels contained in the detection box as an impact factor to characterize the size of the inspected objects, and uses this as an impact factor. The training loss of each object of different size is adjusted by a weight dynamically, so as to improve the accuracy of small object detection. Finally, through experiments, we demonstrate that the size of objects in object detection interfere with each other. So that we can improve the accuracy of small object detection through PLB operation. This method can perform well with blood cell detection in our experiments.

Keywords: medical image detection, object detection, small object, pixel level balance, blood cell detection

1 INTRODUCTION

With the development of artificial intelligence technology, deep learning based on CNN (Convolutional Neural Network) has been widely used in medical image processing field. Using computer aided technology to analyze and process medical images can assist doctors doing qualitative and quantitative analysis of diseases, thereby improving the accuracy and reliability of medical diagnosis greatly. Medical image detection is one of the main tasks in the field of medical image processing. Many medical institutions in the world have rapidly entered this field. Medical image detection has been combined with artificial intelligence technology for a long time. As early as 1993, CNN has been used for lung nodule detection. In 1995, the technology was also applied to detect micro-calcification in mammography.

Medical detection technology has been continuously developed by applying CNN and other deep learning methods (McInerney and Terzopoulos, 1996; Handels et al., 2013; Litjens et al., 2017) to various medical images with different imaging mechanisms. For example, Setio et al. detected lung nodules in 3D chest CT scans and extracted 2D patches in nine different orientations centered on these candidates (Golan et al., 2016). Ross et al. utilized CNN to improve three existing CAD systems for the detection of colonic polyps, sclerosing spinal deformity, and lymphadenopathy in CT imaging (Roth et al., 2016). In recent years, object detection technology has been widely used in pathology (Janowczyk and Madabhushi, 2016), especially in blood cells detection (Yang et al., 2017; Pan et al., 2018; Fujita and Han, 2020). Detecting blood cells can assist diagnosing many kinds of diseases, such as diagnosing breast cancer by detecting mitosis or lymphocytes (Cire ş an et al., 2013; Zhang et al., 2022). Object detection technology is constantly applied into application scenarios of medical image processing, and thus bringing more commercial value.

Object detection is mainly aimed at locating and identifying objects in different positions in the image. In medical image detection, the detection of small objects gets more attentions than the detection of large objects. For example, lesions are identified by detecting whether there are tiny abnormalities in the images, determining whether a patient is likely to develop the disease. In particular, in the detection of blood cells, some types of cells are rarer and smaller than others, but they play an indispensable and key role in the diagnosis of diseases. In such tasks, the large adjacent objects in the image can be ignored, but the detection of small objects is very important. However, object detection for small objects is precisely a more difficult task. Objectively speaking, small objects contain low pheromones and human beings have the disadvantage in recognizing small objects. These reasons bring a higher challenge to the detection task.

Previous research work mainly focused on how to enhance the detection model's extraction of the feature of small objects, and through some methods to solve the unbalance of the samples. In addition, there is also the use of rotation detection with angle factors (Yang et al., 2019a; Yang et al., 2019b; Qian et al., 2019; Yang and Yan, 2020) to better approximate the true position of small objects. However, it is rarely mentioned that under the multi-object detection task, there is a possibility of interference among the multiple objects within one image. In the objective logic of human observation, people can see large objects most intuitively among objects of various sizes, but it is easy to ignore the existence of small objects, let alone the recognition of the small objects. In the detection model, theoretically, it is also necessary to consider the mutual interference between objects of different sizes during the training process, and the contribution of objects with various sizes to the training loss are different. Therefore, it is meaningful to study the interference of large and small objects, and use some means to alleviate such interference and guide the model to the optimization direction of small object detection.

We studied the interference between objects of different sizes in the training process, and designed a weight coefficient called PLB weight to adjust the training effect of the model. The size of the detected object is characterized by the size of the detection rectangle. We use the number of pixels contained in it as the calculation input of the weight coefficient. In the process of model training, this coefficient can be used to dynamically adjust the training loss of each object with various sizes, so that the model can be training to the direction of improving the accuracy of small object detection.

Our innovations and contributions to this work are summarized as follows:

1) Instead of setting fixed empirical values before training, we dynamically set the loss weights for objects of different sizes during the training process of the detection model.
2) With our proposed method PLB, the training trend of two-stage detector can be adjusted and the detection accuracy of small objects can be further improved.
3) Our method can be combined with other methods to improve the detection effect, bringing more potential capacity for some medical applications which need higher detection effect of small objects than bigger ones, such as blood cell detection.

2 RELATED WORK

Current detection models are divided into two categories: two-stage detectors and one-stage detectors. Two-stage detection network is represented by the RCNN series (Girshick et al., 2014; Girshick, 2015; He et al., 2015; Ren et al., 2015; Dai et al., 2016), the second category is represented by the YOLO series (Redmon et al., 2016; Redmon and Farhadi, 2017; Redmon and Farhadi, 2018; Bochkovskiy et al., 2020) and SSD series (Liu et al., 2016; Shrivastava et al., 2016; Fu et al., 2017; Jeong et al., 2017; Li and Zhou, 2017; Shen et al., 2017). Among them, the former adopts the RPN network (Girshick, 2015). When performing localization and recognition tasks, candidate rectangular boxes are proposed in the RPN phase. In the second stage, the candidate proposals are adjusted and the objects in the boxes are identified. One-stage detectors use an end-to-end deep neural network, and the model structure is simpler than two-stage detectors, bring a faster computing speed, so that it is more suitable for some time-sensitive application scenarios. But for the improvement of detection accuracy, one-stage detector usually weaker than two-stage detector.

No matter which detection model is used, the CNN model is used as the feature extractor to obtain the feature space of the train set. With the continuous development of CNN models in recognition tasks, especially the ResNet model (He et al., 2016) and the DenseNet model (Huang et al., 2017), it has been confirmed that the CNN model has a high accuracy and universality for feature extraction in classification tasks. In terms of improving the overall accuracy of the detection model, a deeper CNN can be used as the backbone network to extract the image features (Zhang et al., 2022). Attention mechanism such as SENet (Hu et al., 2018) is used to improve the sensitivity of the model to channel features. The model can be adjusted through EfficientNet (Tan and Le, 2019) by adjusting the depth, width and pixel accuracy in the model to optimize the overall performance, such as EfficientDet (Tan et al., 2020). In addition, other methods such as NMS (Neubeck and Van Gool, 2006) and BN (Li et al., 2019) can be used to optimize the detection model comprehensively.

The cost of manual labeling of medical datasets is more expensive, and the acquisition and labeling of datasets is more difficult than other scenarios. Therefore, for some incompletely labeled datasets, some methods are also needed to improve the accuracy of object detection. Unsupervised active learning methods can be applied into this task to improve detection performance (Changsheng et al., 2019). Such as Active Learning Matrix Sketching (ALMS) (Li et al., 2021) which is used to do simultaneous sample and feature selection in an unsupervised setting. These methods aim to improve the effectiveness of the latent feature space (Li et al., 2022), so that the detection model can achieve more stable and good performance.

Improving the detection accuracy of small objects is a more difficult challenge. For the detection improvement of small target objects, rotation detection is also an effective method. Traditional detection models generally use horizontal rectangular boxes as labels for localization tasks. However, for small objects, the rotating detection boxes with an angle can more closely approximate the real position (Yang et al., 2021a; Yang et al., 2021b; Yang and Yan, 2022). Small objects have higher sensitivity requirements to position, and rotation detection can bring better training effects for the detection of small objects (Yang et al., 2020; Yang et al., 2021c; Yang et al., 2022).

In the detection model based on deep neural network, with the deepening of the network, the image features can be extracted better to fit our detection task. But in the feature space at the end of the model, the represented receptive field is getting bigger and bigger, while the features corresponding to the small objects may disappear. FPN network (Lin et al., 2017a; Pan et al., 2018) is a good solution to this problem. In this network structure, middle layers in the feature extraction process are reserved and combined with the upper and lower layers, so that the feature of small objects will not disappear with the deepening of the network. Finally multi-layer feature vectors are obtained by FPN. Among them, the low-dimensional feature has a smaller receptive field for small objects which is biased towards the shape features of the object, while the high-dimensional feature has a larger receptive field which is biased towards the semantic features. The FPN network improves the feature extraction effect especially for small objects as it can retain more features. As a result, the FPN network can effectively improve the detection accuracy of small objects.

Different kinds of imbalances within the training samples are also the reasons for the difficulty of detecting small objects. These imbalances mainly include the imbalance of the object categories, and the imbalance of the proportion of small objects and large objects in the samples. In addition, the imbalance between the foreground and background is also an important factor that disturbs the training effect. In the detection model, the corresponding weights can be set for each category in the data set through a prior knowledge, and the loss in the training process is weighted to adjust for the category imbalance of samples. For the spatial imbalance of the detection task, some data augmentation techniques (Pan et al., 2018; Kisantal et al., 2019) can be used to deal with this problem. For example, copy the small objects and paste them at different positions in the image multiple times to increase the proportion of small objects. Besides, data augmentation of training samples can also be performed through image fusion (Li et al., 2013; Xu et al., 2013) and image adversarial generation (Fang et al., 2020) techniques. This expansion method can alleviate the imbalance of samples in space. For the imbalance between foreground and background, the weight of difficult and easy samples can be adjusted reference to the theoretical method Focal Loss (Lin et al., 2017b). The main idea of this method is to use an appropriate function to measure the contribution of hard-to-classify and easy-to-classify samples to the total loss for a better training effect.

3 PROPOSED METHOD

3.1 Overall Structure of Our Model

Figure 1 shows the structure of our detection model. In our new model, the network structure of the model is mostly like a general faster-rcnn structure. Our design is that the "RPN Header" module and the "ROI Header" module in the detection will output the coordinate of the detect box which will be put in the calculation process of the loss function. We modified the calculation of the loss function in a general two-stage detection model, using the size of the "Detect box" as a weight factor for the training loss contributed by each detect box. Through the computational design of the weight factor, we can appropriately adjust the loss contribution of each object with different sizes, and then improve the detection effect of small objects during the training process.

In the structure of our detection model, there are four loss functions that can be adjusted in this way, which correspond to the training effect of all classification task and localization task in the two-stage detection model. As the four new loss function shown in the right of **Figure 1**, these components use the weight factors to adjust the contribution of each detect box to the corresponding original loss.

The total loss of the detection is calculated by weighting the above four components and setting a certain weight coefficient for the original Smooth L1 Loss (Ren et al., 2015) and Cross Entropy Loss (Ren et al., 2015). For different application scenarios, the PLB (Pixel Level Balancing) operation can selectively adopt a combination strategy of these four new loss functions.

In our detection model, the training loss for each image is defined as:

$$L(p_i, t_i) = \frac{1}{N_{cls\_rpn}}\sum_i^n Plb_i L_{cls\_rpn}(p_i, p_i^*) + \lambda \frac{1}{N_{reg\_rpn}}\sum_i^n p_i^* Plb_i L_{reg\_rpn}(t_i, t_i^*) + \frac{1}{N_{cls\_roi}}\sum_i^n Plb_i L_{cls\_roi}(p_i, p_i^*) + \lambda \frac{1}{N_{reg\_roi}}\sum_i^n Plb_i L_{reg\_roi}(t_i, t_i^*) \tag{1}$$

where p_i and t_i are the predicted category and position results, λ is a parameter to weight the classification and the localization task.

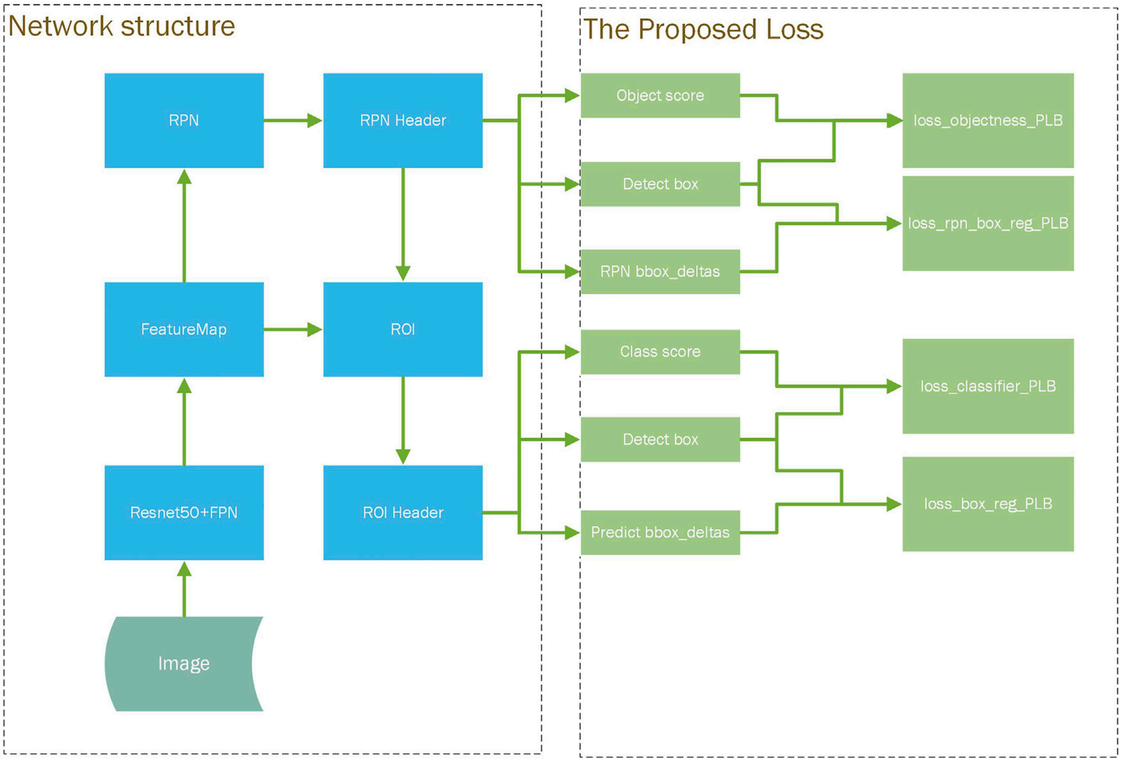


FIGURE 1 | Overview of our approach design.

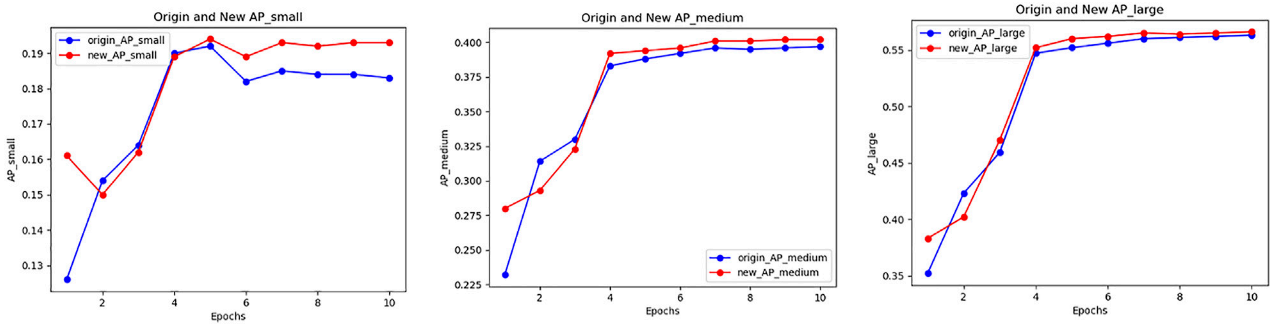


FIGURE 2 | Accuracy variation of small, medium and large object in PLB1C.

In the RPN stage, only the loss of the foreground object is calculated. The total loss function can optionally use the new loss function weighted by PLB weight factors to replace the original loss function components. If using the original loss function, just set the PLB weights to 1.

3.2 Design of Pixel Level Balance Factor

In the training process of the detection model, due to the different sizes of the objects, the sensitivity of the training to the size of the objects is different, and there exists a potential mutual interference. Pixel level balance refers to adjust the weight coefficients for the training loss caused by each object under inspection when multiple objects appear in the same image, and considering their different sizes as a factor to change their mutual interference. In particular, it can be assumed that large objects will adversely affect the detection of small objects, so that in the model, the detection accuracy of small objects is further reduced. On the contrary, we can actively guide the model to change towards the optimization of small object detection by adjusting the weight coefficient of each inspected object.

The number of pixels of the inspected object can be used to measure the sensitivity of its size to detection. During the training process, the size of each object is measured by the specific rectangular box area. For the selection of the rectangular box, the predicted box in the model training process can be used, or the ground-truth labeled box that best matches the candidate box can be used, they can be used as the representation of the object size.

The pixel level balance factor is defined as follows:

$$\text{area\_mean} = \frac{\sum_{\text{i}}^{\text{n}} \text{box\_area}}{\text{n}} \tag{2}$$

$$\text{PLB\_weight} = \frac{\text{area\_mean}*2}{\text{area\_predict} + \text{area\_mean}} \tag{3}$$

where n is the number of detection boxes after filtered for the loss in this training, "area_predict": the area of the predicted box or labeled box.

If the contribution of the loss is determined according to the number of pixels, it can be considered that an object of average size has a balance factor of 1. Taking the detection accuracy of small objects as the goal, in the above formula, when "area_predict" approaches to 0, PLB value is equal to 2, which increases the weight coefficient of small objects; when "area_predict" approaches to the largest object among the n inspected objects, assumed that sizes of the rest objects is close to 0, then

$$\text{PLB\_weight} = \frac{2*\frac{\text{area\_max}}{\text{n}}}{\frac{\text{area\_max}}{\text{n}} + \text{area\_max}} = \frac{2}{\text{n}+1} \tag{4}$$

Obviously, when there is only one object under detection, that means n is 1, then PLB weight is 1. Larger the object is, the PLB weight of the object is getting smaller.

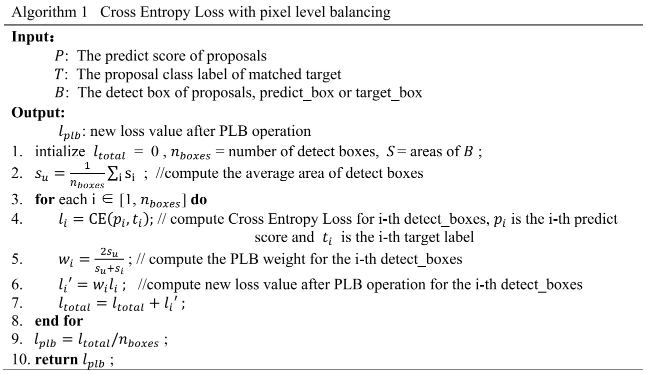
The value scope of the PLB factor is in (2/n+1, 2), and the object with the average size has a corresponding weight of 1. It can be considered that after adding the pixel balance factor as a weight, each inspected object, regardless of its size, will contribute equally to the loss function. Using such a design, the training of the new model has a better effect compared to the original method on optimizing the detection accuracy of small objects.

3.3 Loss Function Combined With Pixel Level Balance

Taking "faster_resnet50_fpn" as a basic model, for its classification loss and border regression loss, PLB operations can be integrated in the four loss functions.

For the loss of classification, we still use the Cross Entropy Loss function as the loss standard of the model. But we need to calculate the pixel level balancing factor according to the size of the object corresponding to each detected box. Then we use it as the weight for multiple classification loss in a batch of images. The pixel level balancing factor can be calculated by the predicted box or its corresponding labeled box. The implementation logic of the function is shown in Algorithm 1.

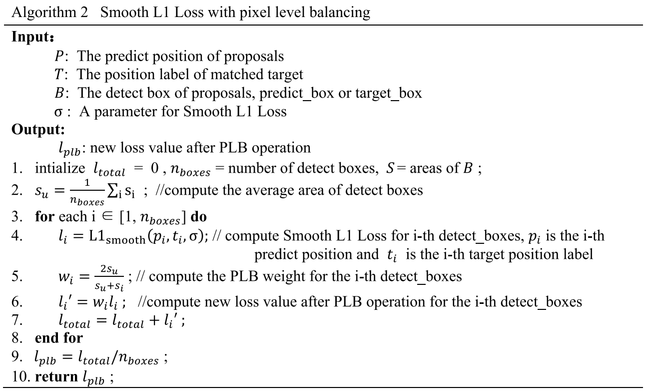
Algorithm 1. Cross Entropy Loss With Pixel Level Balancing



```
Algorithm 1   Cross Entropy Loss with pixel level balancing
Input:
    $P$: The predict score of proposals
    $T$: The proposal class label of matched target
    $B$: The detect box of proposals, predict_box or target_box
Output:
    $l_{plb}$: new loss value after PLB operation
intialize $l_{total} = 0$, $n_{boxes}$ = number of detect boxes, $S$ = areas of $B$ ;
$s_u = \frac{1}{n_{boxes}} \sum_{\text{i}} \text{s}_{\text{i}}$ ; //compute the average area of detect boxes
for each i $\in [1, n_{boxes}]$ do
    $l_i = \text{CE}(p_i, t_i)$; // compute Cross Entropy Loss for i-th detect_boxes, $p_i$ is the i-th predict score and $t_i$ is the i-th target label
    $w_i = \frac{2s_u}{s_u+s_i}$ ; // compute the PLB weight for the i-th detect_boxes
    $l_i' = w_i l_i$ ; //compute new loss value after PLB operation for the i-th detect_boxes
    $l_{total} = l_{total} + l_i'$ ;
end for
$l_{plb} = l_{total}/n_{boxes}$ ;
return $l_{plb}$ ;
```

For the border regression loss function, Smooth L1 Loss is also used as the loss standard of the model, but it is necessary to calculate the pixel level balance factor according to the size of objects corresponding to each predicted box, and then calculate the pixel level balance factor of multiple objects of different sizes within a batch of images. Then we use it as the weight for localization loss for each object in our new function. The specific implementation process is shown in Algorithm 2.

Algorithm 2. Smooth L1 Loss With Pixel Level Balancing



```
Algorithm 2   Smooth L1 Loss with pixel level balancing
Input:
    $P$: The predict position of proposals
    $T$: The position label of matched target
    $B$: The detect box of proposals, predict_box or target_box
    $\sigma$ : A parameter for Smooth L1 Loss
Output:
    $l_{plb}$: new loss value after PLB operation
intialize $l_{total} = 0$, $n_{boxes}$ = number of detect boxes, $S$ = areas of $B$ ;
$s_u = \frac{1}{n_{boxes}} \sum_{\text{i}} \text{s}_{\text{i}}$ ; //compute the average area of detect boxes
for each i $\in [1, n_{boxes}]$ do
    $l_i = \text{L1}_{\text{smooth}}(p_i, t_i, \sigma)$; // compute Smooth L1 Loss for i-th detect_boxes, $p_i$ is the i-th predict position and $t_i$ is the i-th target position label
    $w_i = \frac{2s_u}{s_u+s_i}$ ; // compute the PLB weight for the i-th detect_boxes
    $l_i' = w_i l_i$ ; //compute new loss value after PLB operation for the i-th detect_boxes
    $l_{total} = l_{total} + l_i'$ ;
end for
$l_{plb} = l_{total}/n_{boxes}$ ;
return $l_{plb}$ ;
```

It can be seen that in the new loss function, the prediction information of the box is also added to the calculation of the classification loss, so that the classification loss and the width and height of the detected box have a certain correlation. The border regression loss will focus on the coordinate position of the border and its width and height at the same time.

3.4 Proposed Framework MindSpore

We implement our PLB method in PyTorch for research and exploration. At the same time, we recommend using an implementation version under the MindSpore framework as the final application. MindSpore is an enterprise-level application framework based on Huawei's AI ecosystem. It has been used by Huawei in the medical field, and has open sourced the code of many detection models. This framework is an open-sourced product in the AI field that Huawei has been promoting in the past two years. Based on the hardware environment of Huawei's Ascend series chips, it optimizes a large number of

TABLE 1 | Training accuracy effect of PLB2C (IOU = 0.50:0.95) on PyTorch.

	MAP	AP_small	AP_ Medium	AP_large
Original faster_rcnn	48.9	18.3	39.7	56.3
PLB2C (default predict box)	46.6 (−4.70%)	19.2 (+4.92%)	38.8 (−2.27%)	53.9 (−4.26%)
PLB2C (matched labeled box)	47.7 (−2.45%)	16.8 (−8.20%)	38.3 (−3.53%)	55.1 (−2.13%)

calculations in the model, speeds up the training and inference of the model. Due to the framework is easy to develop, efficient for execution and its full scene coverage, it can bring higher application value to our PLB method.

The implementation of our method in the MindSpore framework is basically the same as that under PyTorch, but we need to do some extra processing on the PLB weight computation. In order to prevent that all predict box areas may be zero during the training process, the area value of the predict box need to add by 1 to avoid division by zero exceptions.

4 EXPERIMENT RESULTS ANALYSIS

We use the dataset Pascal VOC2007 to explore the right way to apply our PLB methods, and use the BCCD blood cell detection dataset (Banik et al., 2020) to verify the effectiveness of PLB for medical image detection. Through exploration the effectiveness of our methods on Pascal VOC 2007, we verify it on BCCD datasets with our implementations both on PyTorch and MindSpore (https://www.MindSpore.cn/en). In our comparison experiment group, we use faster-rcnn model as the reference, chose resnet50 and FPN as the backbone network, and basically set the default values in the PyTorch library for its hyper-parameters, the SGD optimizer with momentum = 0.9, and the initial learning rate lr = 0.005, the adjustment step size of the learning rate step_size = 3.

The source code was released at: https://gitee.com/hubindijun/faster-rcnn-plb-MindSpore.git (MindSpore version); https://github.com/hubindijun/faster-rcnn-plb-PyTorch.git (PyTorch version).

4.1 Exploration of PLB in Natural Image Detection

The experiment uses the Pascal VOC2007 dataset (5011 images for training and 4,952 images for validation, 20 different categories). Then we evaluate our method with the coco evaluation standard. Finally, we mainly focus on the MAP and the detection accuracy of objects with different sizes to analyze the effect of PLB. The area range of small objects is (0,32*32), the area range of medium-sized objects is [32*32,96*96] and the area range of large objects is greater than 96*96, using pixel point number as the unit of object size. In the dataset, the ratio of small, medium and large objects is 845:2,698:4,301 in the training dataset, while the ratio in the testing dataset is 909:2,706:4,203.

We selectively perform PLB operations on different parts of the loss function. When the training epoch is 10, both the original model and new one reach a status of convergence.

PLB method in the four different loss components are named as follows, all of the four PLB operations use predict box as the default standard for size representation.

PLB1C: PLB in the first RPN stage of the detection model within coarse-grained classification loss;

PLB1B: PLB in the first RPN stage of the detection model within bounding box regression loss;

PLB2C: PLB in the second stage of the detection model within fine-grained classification loss;

PLB2B: PLB in the second stage of the detection model within further bounding box regression loss.

4.1.1 The Selection of Predict Box or Matched Labeled Box for Size Representation

Firstly, we conduct the two experiments about PLB2C with default predict box and matched labeled box as the standard for size representation. The training accuracy effect are showed in **Table 1**. The results of the PLB2C shows that only use PLB operation in the fine-grained classification loss can significantly improve the detection accuracy of small objects, but the overall accuracy of the model is reduced. PLB2C means higher requirements for small objects and reduces the expectation of the detection effect of medium and large objects. Although the detection accuracy of small objects gets improved, the detection accuracy of medium and large ones will decrease. Finally, due to the proportion of small objects is relatively small in the dataset, the overall detection accuracy will also decrease in the training.

However, after replacing predict box with matched labeled box as the representation of the object size, the detection effect is reduced, even the detection effect for small objects is reduced by 8.2% as shows in **Table 1**. We can draw a conclusion that compare to the predict box, the matched labeled box is not suitable for the representation of the object size in the model training process.

4.1.2 Ablation Experiments Analysis

Through our design and experiments, we summarize the detection effects of each scheme on the accuracy of small objects, as shown in **Table 2**. Comparing the results of each scheme, all of the PLB methods can obviously improve the detection accuracy of small objects. However, the detect effect of the PLB methods is different for medium and large objects.

With method PLB1C or PLB1B, the overall effect of the model keeps well, especially the detection accuracy of small objects has a significant improvement. Meanwhile, the methods have little impact on medium and large objects. The training accuracy tendency of PLB1C is shown in **Figure 2.**

PLB operations in the second stage also improve the detection accuracy of small objects, as the results of PLB2C and PLB2B.

TABLE 2 | Comparison of the PLB methods (IOU = 0.50:0.95) on PyTorch.

PLB1C	PLB1B	PLB2C	PLB2B	MAP	AP_small	AP_medium	AP_large
Not used	-	-	-	48.9	18.3	39.7	56.3
✓	-	-	-	49.3	19.3 (+5.46%)	40.2	56.5
-	✓	-	-	49.2	19.1 (+4.37%)	39.9	56.8
-	-	✓	-	46.6	19.2 (+4.92%)	38.8	53.9
-	-	-	✓	48.7	19.0 (+3.83%)	39.2	56.2
✓	-	✓		49.3	18.6 (+1.64%)	39.9	56.7
✓	✓	-	-	49.1	19.3 (+5.46%)	39.9	56.5

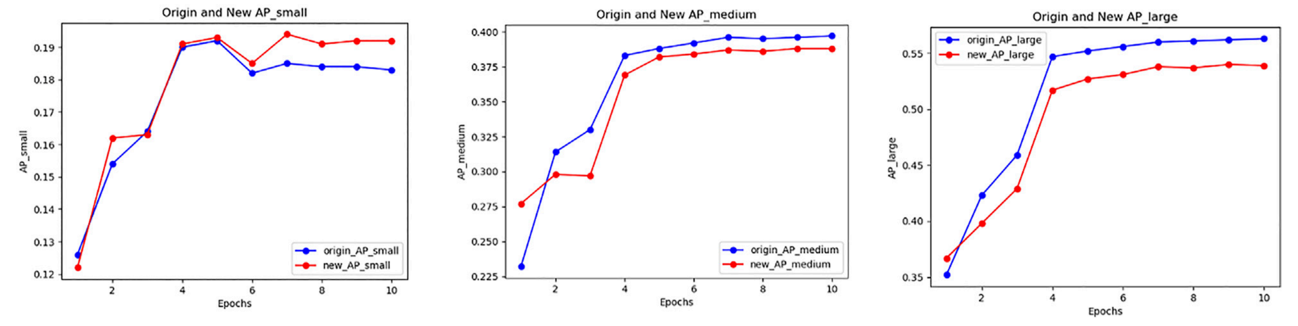


FIGURE 3 | Accuracy variation of small, medium and large object in PLB2C.

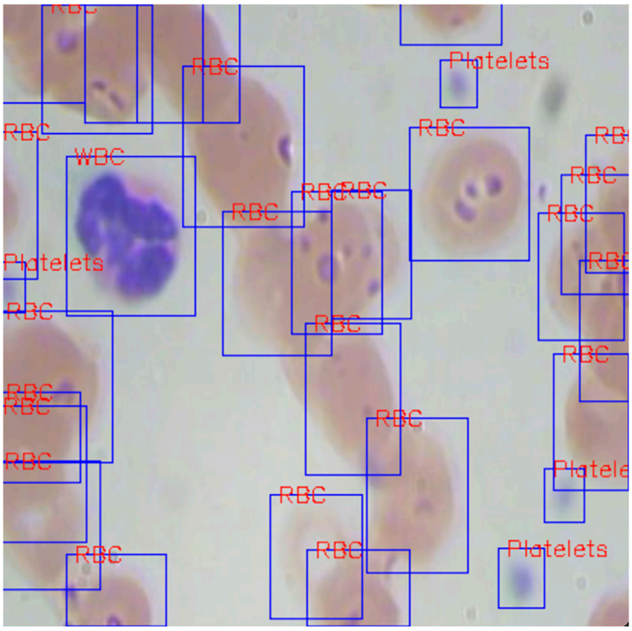


FIGURE 4 | Blood cell detection example from the BCCD data set. The red tags denote different classifications of the detected objects.

However, that methods have negative impact on detection of medium and large objects. Due to there are more samples of medium and large objects than small ones, the overall detection accuracy is not well. As the accuracy tendency of PLB2C shows us in **Figure 3**.

TABLE 3 | Accuracy effect with PLB in BBCD Dataset (IOU = 0.50:0.95) on PyTorch.

	MAP	AP_small	AP_ Medium	AP_large
Origin model	63.1	47.8	64.9	50.2
PLB1C	63.6	49.7 (+3.97%)	65.2 (+0.46%)	49.6 (–1.20%)
PLB1B	63.1	48.6 (+1.67%)	64.0 (–1.40%)	50.6 (+0.80%)
PLB1C + PLB1B	63.7	49.0 (+2.51%)	66.6 (+2.62%)	50.2 (+0%)

TABLE 4 | Accuracy effect with PLB in BBCD Dataset (IOU = 0.50:0.95) on MindSpore.

	MAP	AP_small	AP_ Medium	AP_large
Origin model	62.9	48.7	65.4	48.0
PLB1C	63.3	51.4 (+5.54%)	65.6 (+0.31%)	48.3 (+0.63%)
PLB1B	63.1	49.6 (+1.85%)	64.2 (–1.83%)	50.3 (+4.79%)
PLB1C + PLB1B	63.6	50.4 (+3.49%)	66.6 (+1.83%)	49.2 (+2.50%)

Through the different results of the PLB methods, we can draw a conclusion that in the training process of the two-stage detector, the PLB operation utilized in the RPN stage can improve the detection effect of small objects and maintain the accuracy of medium and large objects. The purpose of PLB operation is to balance the contribution to the model loss of objects which have different sizes during the training process. In particular, we can adjust the design of the PLB factor so that the training of the model is transformed towards the detection accuracy improvement of small objects.

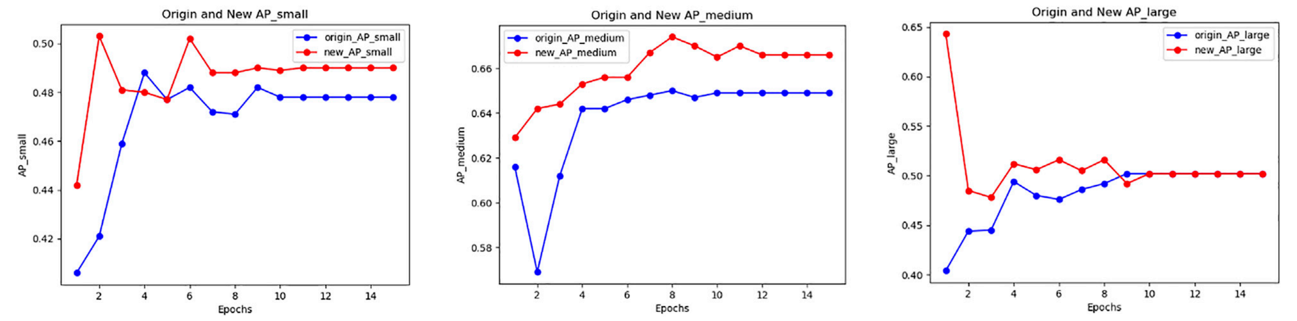


FIGURE 5 | Accuracy variation of PLB1C combined with PLB1B within BCCD dataset.

So how does PLB methods be more effective in the first stage? We guess that in the first stage of the training process, the coarse-grained classification task is mainly to classify the inspected objects as foreground or background, among them the background will not be included in the subsequent loss calculation. The smaller the object is, the easier it is to be misclassified as a background. So that in the second stage of detection, it is no longer involved in training process. Therefore, PLB method has a relatively obvious effect in the RPN stage of the two-stage detection model. Moreover, compared with the transformation of border regression loss with PLB, the loss transformation effect for the classification with PLB is more effective.

4.2 Practice of PLB in Medical Image Detection

We have verified that PLB has a certain adjustment effect for object detection, and using PLB in the RPN stage is more effective. Our design can also be used in specific medical application scenarios, such as routine blood testing and breast cancer diagnosis through lymphocyte detection. All of these application scenarios are based on detecting and measuring various types of blood cells to assist disease diagnosis. We use the BCCD data set (765 pictures for training, and 73 pictures for evaluating) to check the effect of PLB in cell detection. **Figure 4** shows the blood cell detection in general, where there are three different cell types, in which the platelet size is relatively small and hard to detect.

We carried out three sets of experiments with this dataset. Using PLB with the model only in the coarse-grained classification loss, or only in the RPN border regression loss, or both of them at the same time, to demonstrate the effectiveness of it in medical application scenarios. When the training epochs is 15, both the old and new models reach to convergence. We conduct the experiments with our implementation both on PyTorch version and MindSpore version. **Tables 3** and **Table 4** shows the final results of pixel level balance respectively to these three experiments.

Experimental results show that the PLB methods can effectively improve the detection effect of small objects in the process of medical image detection tasks. When using PLB in the two loss functions in the RPN stage at the same time, the overall detection effect is improved, especially for the detection accuracy of small objects. **Figure 5** shows the accuracy variation of small, medium and large object when using PLB in the two loss components in the RPN stage.

5 CONCLUSIONS AND OUTLOOK

In this paper, we have proposed pixel level balance different from previous research, which focuses in the correlation of large and small objects in the training process. This method can be combined with other effective methods to improve small object detection, such as FPN network to improve the feature extraction, or data augmentation on the input dataset samples, etc. In some specific application scenarios, pixel level balance can provide more special effects. Obviously, in a train model, we can improve the detection accuracy of small objects while ignoring the large one by modifying the design of pixel level balance factors.

Pixel level balance can perform well in the problem of higher requirements for small objects in medical image detection. In our experiments, the effectiveness of this method for blood cell detection tasks has been demonstrated. It can be used in more other medical detection tasks in the future and achieve more development space or commercial value to medical image detection technology.

For future work, to alleviate the strong label requirement for deep learning-based detection, we would like to explore the possible way of applying visual matching-based approaches (Jiang et al., 2021a) for object detection and recognition. One promising technique is adopting graph matching with (higher-order) structure information (Yan et al., 2018) which can be more generalizable to new objects, and the detection may be performed in a joint matching fashion with multiple candidate objects with different techniques from heuristic optimization (Yan et al., 2015; Yan et al., 2016a) to dynamic programming based one (Jiang et al., 2021b). Moreover, the recently developed deep learning-based graph matching models (Wang et al., 2020; Wang et al., 2021) can also be explored which can better model the visual features for matching and object recognition. Readers are referred to the survey papers for more comprehensive study of these areas, in terms of both traditional learning-free methods (Yan et al., 2016b) as well as deep learning models (Yan et al., 2020). The hope is that a more structure information can be effectively used for object detection, against outliers, deformation, occlusion and other noise.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: Pascal VOC2007 http://host.robots.ox.ac.uk/pascal/VOC/voc2007/VOCtrainval_06-Nov-2007. tar BCCD https://public.roboflow.com/object-detection/bccd/4/download.

AUTHOR CONTRIBUTIONS

BH and YL contributed to the conception and design of the study. BH wrote the first draft of the manuscript. PC, MT and QK reviewed and revised the manuscript. All authors contributed to manuscript revision and approved the submitted version.

FUNDING

This research was partly funded by the Interdisciplinary Program of Shanghai Jiao Tong University (YG2021QN67), CAAI-Huawei MindSpore Open Fund (CAAIXSJLJJ-2020-022A).

ACKNOWLEDGMENTS

The authors would like to appreciate the Student Innovation Center of SJTU for providing GPUs.

REFERENCES

Banik, P. P., Saha, R., and Kim, K.-D. (2020). An Automatic Nucleus Segmentation and CNN Model Based Classification Method of White Blood Cell. *Expert Syst. Appl.* 149, 113211. doi:10.1016/j.eswa.2020.113211

Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M. (2020). *Yolov4: Optimal Speed and Accuracy of Object Detection*, Ithaca: Cornell University, http://arXiv.org/abs/2004.10934.

Changsheng, C., Xiangfeng Wang, X., Weishan Dong, W., Junchi Yan, J., Qingshan Liu, Q., and Hongyuan Zha, H. (2019). Joint Active Learning with Feature Selection via Cur Matrix Decomposition. *IEEE Trans. Pattern Anal. Mach. Intell.* 41 (6), 1382–1396. doi:10.1109/TPAMI.2018.2840980

Cireşan, D. C., Giusti, A., Gambardella, L. M., and Schmidhuber, J. (2013). Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks." in International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Berlin, Heidelberg, September, 2013 Editors H. Ishikawa, C. Lin Liu, T. Pajdla, and S. Jianbo 411 Berlin Heidelberg: Springer.

Dai, J., Li, Y., He, K., and Sun, J. (2016). "R-fcn: Object Detection via Region-Based Fully Convolutional Networks," in Advances in Neural Information Processing Systems, Barcelona, Spain, December, 2021 29, 379–387.

Fang, Y., Deng, W., Du, J., and Hu, J. (2020). Identity-aware CycleGAN for Face Photo-Sketch Synthesis and Recognition. *Pattern Recognit.* 102, 107249. doi:10.1016/j.patcog.2020.107249

Fu, C. Y., Liu, W., Ranga, A., Tyagi, A., and Berg, A. C. (2017). *Dssd: Deconvolutional Single Shot Detector*. Ithaca: Cornell University. http://arXiv.org/abs/arXiv:1701.06659.

Fujita, S., and Han, X. H. (2020). "Cell Detection and Segmentation in Microscopy Images with Improved Mask R-CNN," in Proceedings of the Asian Conference on Computer Vision, December, 2020. Online Meeting.

Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2014). "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, June, 2014 (New Jersey: IEEE), 580–587. doi:10.1109/cvpr.2014.81

Girshick, R. (2015). "Fast R-Cnn," in Proceedings of the IEEE International Conference on Computer Vision, Santiago, December, 2015 (New Jersey: IEEE), 1440–1448. doi:10.1109/iccv.2015.169

Golan, R., Jacob, C., and Denzinger, J. (2016). "Lung Nodule Detection in CT Images Using Deep Convolutional Neural Networks," in 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, July, 2016 (New Jersey: IEEE), 243–250. doi:10.1109/ijcnn.2016.7727205

Handels, H., Mersmann, S., Palm, C., Tolxdorff, T., Wagenknecht, G., and Wittenberg, T. (2013). Viewpoints on Medical Image Processing: from Science to Application. *Curr. Med. Imaging* 9 (2), 79. doi:10.2174/1573405611309020002

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 (USA: IEEE), 770–778. doi:10.1109/cvpr.2016.90

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 37 (9), 1904–1916. doi:10.1109/tpami.2015.2389824

Hu, J., Shen, L., and Sun, G. (2018). "Squeeze-and-excitation Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake, June, 2018 (USA: IEEE), 7132–7141. doi:10.1109/cvpr.2018.00745

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, Honolulu, HI, July, 2017 (USA: IEEE), 4700–4708. doi:10.1109/cvpr.2017.243

Janowczyk, A., and Madabhushi, A. (2016). Deep Learning for Digital Pathology Image Analysis: A Comprehensive Tutorial with Selected Use Cases. *J. Pathol. Inf.* 7, 29. doi:10.4103/2153-3539.186902

Jeong, J., Park, H., and Kwak, N. (2017). *Enhancement of SSD by Concatenating Feature Maps for Object Detection*. Ithaca: Cornell University. http://arXiv.org/abs/arXiv:1705.09587.

Jiang, X., Fan, A., Jiang, J., and Yan, J. (2021). Image Matching from Handcrafted to Deep Features: A Survey. *Int. J. Comput. Vis.* 129, 23–79. doi:10.1007/s11263-020-01359-2

Jiang, Z., Wang, T., and Yan, J. (2021). Unifying Offline and Online Multi-Graph Matching via Finding Shortest Paths on Supergraph. *IEEE Trans. Pattern Anal. Mach. Intell.* 43 (10), 3648–3663. doi:10.1109/tpami.2020.2989928

Kisantal, M., Wojna, Z., Murawski, J., Naruniec, J., and Cho, K. (2019). *Augmentation for Small Object Detection*, Ithaca: Cornell University, http://arXiv.org/abs/arXiv:1902.07296.

Li, C., Li, R., Yuan, Y., Wang, G., and Xu, D. (2021). Deep Unsupervised Active Learning via Matrix Sketching. *IEEE Trans. Image Process.* 30, 9280–9293. doi:10.1109/tip.2021.3124317

Li, C., Ma, H., Yuan, Y., Wang, G., and Xu, D. (2022). Structure Guided Deep Neural Network for Unsupervised Active Learning. *IEEE Trans. Image Process.*, 2767–2781. doi:10.1109/tip.2022.3161076

Li, Q., Du, J., Song, F., Wang, C., Liu, H., and Lu, C. (2013). "Region-Based Multi-Focus Image Fusion Using the Local Spatial Frequency," in 2013 25th Chinese Control and Decision Conference (CCDC), Guiyang, GuiZhou, May, 2013 (USA: IEEE), 3792.

Li, X., Chen, S., Hu, X., and Yang, J. (2019). "Understanding the Disharmony between Dropout and Batch Normalization by Variance Shift," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (USA: IEEE), 2682–2690. doi:10.1109/cvpr.2019.00279

Li, Z., and Zhou, F. (2017). "*FSSD: Feature Fusion Single Shot Multibox Detector*," in (Ithaca: Cornell University).hppt://arXiv.org/abs/arXiv:1712.00960.

Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., and Belongie, S. (2017a). "Feature Pyramid Networks for Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, July, 2017 (USA: IEEE), 2117–2125. doi:10.1109/cvpr.2017.106

Lin, T. Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017b). "Focal Loss for Dense Object Detection," in Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, October, 2017 (USA: IEEE), 2980–2988. doi:10.1109/iccv.2017.324

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., and Ghafoorian, M. (2017). A Survey on Deep Learning in Medical Image Analysis. *Med. image Anal.* 42, 60–88. doi:10.1016/j.media.2017.07.005

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). "SSD: Single Shot MultiBox Detector," in European Conference on Computer Vision, Springer, Cham, October, 2016 (Cham: Springer), 21–37. doi:10.1007/978-3-319-46448-0_2

McInerney, T., and Terzopoulos, D. (1996). Deformable Models in Medical Image Analysis: a Survey. *Med. image Anal.* 1 (2), 91–108. doi:10.1016/s1361-8415(96)80007-7

Neubeck, A., and Van Gool, L. (2006). Efficient Non-maximum Suppression." in 18th 457 International Conference on Pattern Recognition (ICPR'06), Hong Kong, China, August, 2006. *Sarcoidosis* 3, 101.

Pan, X., Yang, D., Li, L., Liu, Z., Yang, H., and Cao, Z. (2018). Cell Detection in Pathology and Microscopy Images with Multi-Scale Fully Convolutional Neural Networks. *World Wide Web* 21 (6), 1721–1743. doi:10.1007/s11280-017-0520-7

Qian, W., Yang, X., Peng, S., Guo, Y., and Yan, J. (2019). *Learning Modulated Loss for Rotated Object Detection*, http://arXivarXiv:1911.08299. United States: arxiv.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You Only Look once: Unified, Real-Time Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, June, 2016 (USA: IEEE), 779–788. doi:10.1109/cvpr.2016.91

Redmon, J., and Farhadi, A. (2017). "YOLO9000: Better, Faster, Stronger, " in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, July, 2017 (USA: IEEE), 7263–7271. doi:10.1109/cvpr.2017.690

Redmon, J., and Farhadi, A. (2018) *Yolov3: An Incremental Improvement*. Ithaca: Cornell University, hppt://arXiv arXiv:1804.02767. .

Ren, S., He, K., Girshick, R., and Sun, J. (2015). Faster R-Cnn: Towards Real-Time Object Detection with Region Proposal Networks. *Adv. neural Inf. Process. Syst.* 28, 91–99. doi:10.5555/2969239.2969250

Roth, H. R., Lu, L., Liu, J., Yao, J., Seff, A., Cherry, K., et al. (2016). Improving Computer-Aided Detection UsingConvolutional Neural Networks and Random View Aggregation. *IEEE Trans. Med. Imaging* 35 (5), 1170–1181. doi:10.1109/tmi.2015.2482920

Shen, Z., Liu, Z., Li, J., Jiang, Y. G., Chen, Y., and Xue, X. (2017). "Dsod: Learning Deeply Supervised Object Detectors from Scratch," in Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, October, 2017 (USA: IEEE), 1919–1927. doi:10.1109/iccv.2017.212

Shrivastava, A., Sukthankar, R., Malik, J., and Gupta, A. (2016). *Beyond Skip Connections: Top-Down Modulation for Object Detection*. Ithaca: Cornell University.http://arXiv.org/abs/arXiv:1612.06851

Tan, M., and Le, Q. (2019). ". Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks," in International Conference on Machine Learning, California, United States, June, 2019 (USA: IEEE), 6105.

Tan, M., Pang, R., and Le, Q. V. (2020). "Efficientdet: Scalable and Efficient Object Detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, June, 2020 (USA: IEEE), 10781–10790. doi:10.1109/cvpr42600.2020.01079

Wang, R., Yan, J., and Yang, X. (2020). Combinatorial Learning of Robust Deep Graph Matching: an Embedding Based Approach. *IEEE Trans. Pattern Anal. Mach. Intell.* 42 (6), 1–16. doi:10.1109/TPAMI.2020.3005590

Wang, R., Yan, J., and Yang, X. (2021). Neural Graph Matching Network: Learning Lawler's Quadratic Assignment Problem with Extension to Hypergraph and Multiple-Graph Matching." in *IEEE Transactions on Pattern Analysis and Machine Intelligence*. USA: IEEE.

Xu, L., Du, J., and Li, Q. (2013). Image Fusion Based on Nonsubsampled Contourlet Transform and Saliency-Motivated Pulse Coupled Neural Networks. *Math. Problems Eng.* 2013 (3), 1–10. doi:10.1155/2013/135182

Yan, J., Cho, M., Zha, H., Yang, X., and Chu, S. M. (2016a). Multi-Graph Matching via Affinity Optimization with Graduated Consistency Regularization. *IEEE Trans. Pattern Anal. Mach. Intell.* 38 (6), 1228–1242. doi:10.1109/tpami.2015.2477832

Yan, J., Li, C., Li, Y., and Cao, G. (2018). Adaptive Discrete Hypergraph Matching. *IEEE Trans. Cybern.* 48 (2), 765–779. doi:10.1109/tcyb.2017.2655538

Yan, J., Wang, J., Zha, H., Yang, X., and Chu, S. (2015). Consistency-Driven Alternating Optimization for Multigraph Matching: A Unified Approach. *IEEE Trans. Image Process.* 24 (3), 994–1009. doi:10.1109/tip.2014.2387386

Yan, J., Yang, S., and Hancock, E. (2020). *Learning Graph Matching and Related Combinatorial Optimization Problems, International Joint Conferences on Artificial Intelligence*. USA: ACM, 4988

Yan, J., Yin, X., Lin, W., Deng, C., Zha, H., and Yang, X. (2016b). A Short Survey of Recent Advances in Graph Matching. ACM International Conference on Multimedia Retrieval, New York, NY, June, 2016, 167–174. doi:10.1145/2911996.2912035

Yang, S., Fang, B., Tang, W., Wu, X., Qian, J., and Yang, W. (2017). "December) Faster R-CNN Based Microscopic Cell Detection," in *2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)* (New Jersey: IEEE), 345.

Yang, X., Hou, L., Zhou, Y., Wang, W., and Yan, J. (2021c). "Dense Label Encoding for Boundary Discontinuity Free Rotation Detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June, 2021 (USA: IEEE), 15819–15829. Online Meeting. doi:10.1109/cvpr46437.2021.01556

Yang, X., Liu, Q., Yan, J., Li, A., Zhang, Z., and Yu, G. (2019b). "R3det: Refined Single-Stage Detector with Feature Refinement for Rotating Object," in Advances in Neural Information Processing Systems, December, 2021. Online Meeting 2 (4), 18381–18394. doi:10.48550/ARXIV.1911.08299

Yang, X., and Yan, J. (2020). "Arbitrary-oriented Object Detection with Circular Smooth Label," in European Conference on Computer Vision (Cham: Springer), 677–694. doi:10.1007/978-3-030-58598-3_40

Yang, X., Yan, J., Ming, Q., Wang, W., Zhang, X., and Tian, Q. (2021a). "Rethinking Rotated Object Detection with Gaussian Wasserstein Distance Loss," in International Conference on Machine Learning, July, 2021 (New York, NY: PMLR), 11830. Online Meeting.

Yang, X., and Yan, J. (2022). On the Arbitrary-Oriented Object Detection: Classification Based Approaches Revisited. *Int. J. Comput. Vis.* 130 (5), 1340–1365. doi:10.1007/s11263-022-01593-w

Yang, X., Yan, J., Yang, X., Tang, J., Liao, W., and He, T. (2020). *Scrdet++: Detecting Small, Cluttered and Rotated Objects via Instance-Level Feature Denoising and Rotation Loss Smoothing*. Ithaca: Cornell University.http://arXiv.org/abs/arXiv:2004.13316.

Yang, X., Yang, J., Yan, J., Zhang, Y., Zhang, T., Guo, Z., et al. (2019a). "Scrdet: Towards More Robust Detection for Small, Cluttered and Rotated Objects," in Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, South Korea, October, 2019 (New Jersey: IEEE), 8232–8241. doi:10.1109/iccv.2019.00832

Yang, X., Yang, X., Yang, J., Ming, Q., Wang, W., Tian, Q., et al. (2021b). Learning High-Precision Bounding Box for Rotated Object Detection via Kullback-Leibler Divergence. *Adv. Neural Inf. Process. Syst.* 34. doi:10.48550/ARXIV.2106.01883

Yang, X., Zhou, Y., Zhang, G., Yang, J., Wang, W., Yan, J., et al. (2022). *The KFIoU Loss for Rotated Object Detection*. Ithaca: Cornell University, http://arXiv.org/abs/arXiv:2201.12558.

Zhang, X., Zhu, X., Tang, K., Zhao, Y., Lu, Z., and Feng, Q. (2022). DDTNet: A Dense Dual-Task Network for Tumor-Infiltrating Lymphocyte Detection and Segmentation in Histopathological Images of Breast Cancer. *Med. Image Anal.* 78, 102415. doi:10.1016/j.media.2022.102415

Conflict of Interest: Author QK was employed by Riseye Intelligent Technology (Shanghai) Co., Ltd.

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.

frontiers | Frontiers in Physiology

ORIGINAL RESEARCH
published: 29 June 2022
doi: 10.3389/fphys.2022.911544

Influence of Sex, BMI, and Skin Color on the Accuracy of Non-Invasive Cuffless Photoplethysmography-Based Blood Pressure Measurements

Dean Nachman[1,2†], *Arik Eisenkraft*[2,3†], *Nir Goldstein*[3], *Arik Ben-Ishay*[3], *Meir Fons*[3], *Roei Merin*[3] and *Yftach Gepner*[4*]

[1]Heart Institute, Hadassah Ein Kerem Medical Center, Jerusalem, Israel, [2]Institute for Research in Military Medicine, Faculty of Medicine, The Hebrew University of Jerusalem and the Israel Defense Force Medical Corps, Jerusalem, Israel, [3]Biobeat Technologies LTD., Petach Tikva, Israel, [4]Department of Epidemiology and Preventive Medicine, School of Public Health, Sackler Faculty of Medicine and Sylvan Adams Sports Institute, Tel Aviv University, Tel-Aviv, Israel

OPEN ACCESS

Edited by:
Yi Xu,
Shanghai Jiao Tong University, China

Reviewed by:
Liqun Zhao,
Shanghai First People's Hospital, China
Wen-Yi Yang,
Shanghai Jiao Tong University, China

***Correspondence:**
Yftach Gepner
gepner@tauex.tau.ac.il

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Medical Physics and Imaging, a section of the journal Frontiers in Physiology

Received: *02 April 2022*
Accepted: *23 May 2022*
Published: *29 June 2022*

Citation:
Nachman D, Eisenkraft A, Goldstein N, Ben-Ishay A, Fons M, Merin R and Gepner Y (2022) Influence of Sex, BMI, and Skin Color on the Accuracy of Non-Invasive Cuffless Photoplethysmography-Based Blood Pressure Measurements. Front. Physiol. 13:911544. doi: 10.3389/fphys.2022.911544

Vital signs obtained by photoplethysmography-based devices might be influenced by subcutaneous fat and skin color. This observational comparison study aimed to test the accuracy of blood pressure (BP) measurements between a photoplethysmography-based device and cuff-based BP device in ambulatory individuals, coming for a routine BP checkup. Systolic BP (SBP) and diastolic BP (DBP) measurements were stratified based on sex, BMI (<25; $25 \leq \text{BMI}<30$; $30 \leq \text{kg/m}^2$), and skin color (types 1–3 and 4–6 by the Fitzpatrick scale). A total of 1548 measurements were analyzed. Correlations of SBP and DBP between the devices among males/females were between 0.914–0.987 (p < 0.001), and Bland-Altman analysis showed a bias of less than 0.5 mmHg for both sexes. Correlations of SBP and DBP between the devices among BMI groups were between 0.931–0.991 (p < 0.001), and Bland-Altman analysis showed a bias of less than 1 mmHg for all. Correlations of SBP and DBP between the devices among the skin color groups were between 0.936–0.983 (p < 0.001), and Bland-Altman analysis showed a bias of less than 1 mmHg for all. This study shows similar and high agreements between BP measurements obtained using a PPG-based non-invasive cuffless BP device and a cuff-based BP device across sex, BMI, and skin color groups.

Keywords: blood pressure, photoplethysmography, age, BMI, sex

INTRODUCTION

Collecting physiological vital signs, such as blood pressure and blood oxygen saturation, is regarded as a basic component in the clinical assessment of individuals. One of the challenges in collecting these vital signs is the accuracy of the measurements in individuals that are overweight or obese and individuals with dark skin color, as measurements seem to be less reliable (Ries et al., 1985; Ries et al., 1989; Jubran and Tobin, 1990; Zeballos and Weisman, 1991; Young, 1995; O'Brien, 1996; Palatini and Parati, 2011; Palatini et al., 2019). This is still an ongoing debate with conflicting evidence and is especially relevant in devices employing photoplethysmography (PPG) technology. As the US population is comprised of about 13.4% African-Americans (see https://www.census.gov/), and 42.4% of the US population is obese (see https://www.cdc.gov/obesity/data/adult.html), verifying the accuracy of such devices in these sub-populations is

TABLE 1 | Demographic data of the participants.

Characteristic	**Mean ± SD**		
Age (years)	35.1 ± 23.8		
Sex (M/F)	592/467		
BMI (kg/m^2)	24.1 ± 4.7		
Fitzpatrick	3.3 ± 1.5		
Total samples			
	n (%)	Males (%)	Age (years)
	1548 (100%)	664 (42.9%)	35.1 ± 23.8
BMI			
BMI<25	1057 (68.3%)	417 (39.4%)	29.1 ± 21.4
25 ≤ BMI<30	346 (22.4%)	168 (48.4%)	43.8 ± 24.3
30 ≤ BMI	145 (9.4%)	79 (54.5%)	48.6 ± 23.5
Skin color type			
Fitzpatrick 1–3	936 (60.5%)	420 (44.9%)	35.2 ± 23.8
Fitzpatrick 4–6	612 (39.5%)	244 (39.9%)	34.9 ± 23.8

Body mass index (BMI): overweight defined as 25 ≤BMI <30, and obese defined as 30 ≤BMI. The Fitzpatrick color scale: Type 1—always burns, never tans, palest, can have freckles; Type 2—usually burns, tans minimally, light-colored but darker than fair; Type 3—sometimes mild burn, tans uniformly, golden honey or olive; Type 4—burns minimally, always tans well, moderate brown; Type 5—very rarely burns, tans very easily, dark brown; and Type 6—never burns, deeply pigmented dark brown to darkest brown.

highly important to allow their general use. Moreover, though sex is regarded as a fundamental aspect of human physiology, it is usually not considered in the design of studies, or in developing personalized medical strategies (Miller, 2014).

In the medical literature, several research groups have found that skin pigmentation does not affect the bias or precision of pulse oximetry, and has no clinically significant effect on pulse oximetry signal quality (Bothma et al., 1996; Adler et al., 1998; Pipek et al., 2021). In a recent study in patients with COVID-19 pneumonitis, measurement of SpO_2 was not affected by patient ethnicity to a clinically significant degree (Wiles et al., 2021). Other research groups have found that several pulse oximeters overestimated arterial oxygen saturation during hypoxia in dark-skinned individuals (Bickler et al., 2005; Feiner et al., 2007); in a study of two large cohorts, dark-skinned patients had nearly three times the frequency of occult hypoxemia that was not detected by pulse oximetry as White patients (Sjoding et al., 2020). According to The US Food and Drug Administration's standards for studies undertaken to test the accuracy of pulse oximeters, they should include among other components a range of skin pigmentations, including at least two darkly pigmented subjects or 15% of the study population, whichever is larger (US Food and Drug Administration, 2013).

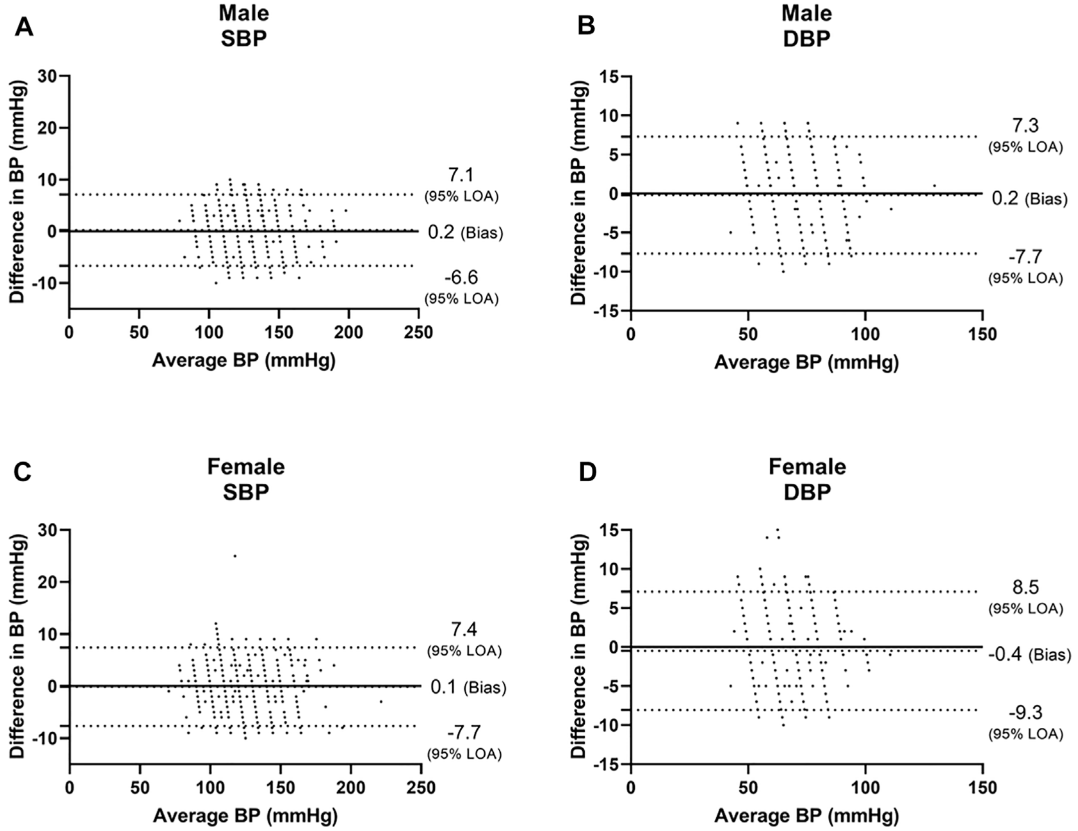


FIGURE 1 | Bland-Altman analysis of males and females. SBP systolic blood pressure, DBP diastolic blood pressure.

TABLE 2 | Level of agreement between the PPG-based wearable measurements across commonly accepted standard.

	<5 mmHg		5–7 mmHg		7–10 mmHg		>10 mmHg	
	SBP	DBP	SBP	DBP	SBP	DBP	SBP	DBP
Sex								
Male	79.2%	80.1%	15.9%	13.7%	4.9%	6.2%	0.0%	0.0%
Female	76.9%	80.0%	17.1%	14.7%	5.6%	4.9%	0.3%	0.3%
Weight category								
Normal weight	77.0%	79.7%	17.7%	14.9%	5.0%	5.3%	0.3%	0.2%
Overweight	80.7%	79.6%	14.2%	13.9%	5.1%	6.2%	0.0%	0.4%
Obese	76.3%	83.1%	16.1%	11.9%	7.6%	5.1%	0.0%	0.0%
Skin color								
Fitzpatrick 1–3	77.3%	80.8%	17.4%	13.9%	5.3%	5.1%	0.0%	0.2%
Fitzpatrick 1–3	78.8%	78.8%	15.4%	14.9%	5.3%	6.0%	0.5%	0.2%

Number in the table are percentage of measurements with a delta of less than 5, 5–7, 7–10 or above 10 mmHg between the reference device and the PPG-based wearable measurements. Body mass index (BMI): overweight defined as 25 ≤BMI <30, and obese defined as 30 ≤BMI. The Fitzpatrick color scale: Type 1—always burns, never tans, palest, can have freckles; Type 2—usually burns, tans minimally, light-colored but darker than fair; Type 3—sometimes mild burn, tans uniformly, golden honey or olive; Type 4—burns minimally, always tans well, moderate brown; Type 5—very rarely burns, tans very easily, dark brown; and Type 6—never burns, deeply pigmented dark brown to darkest brown.

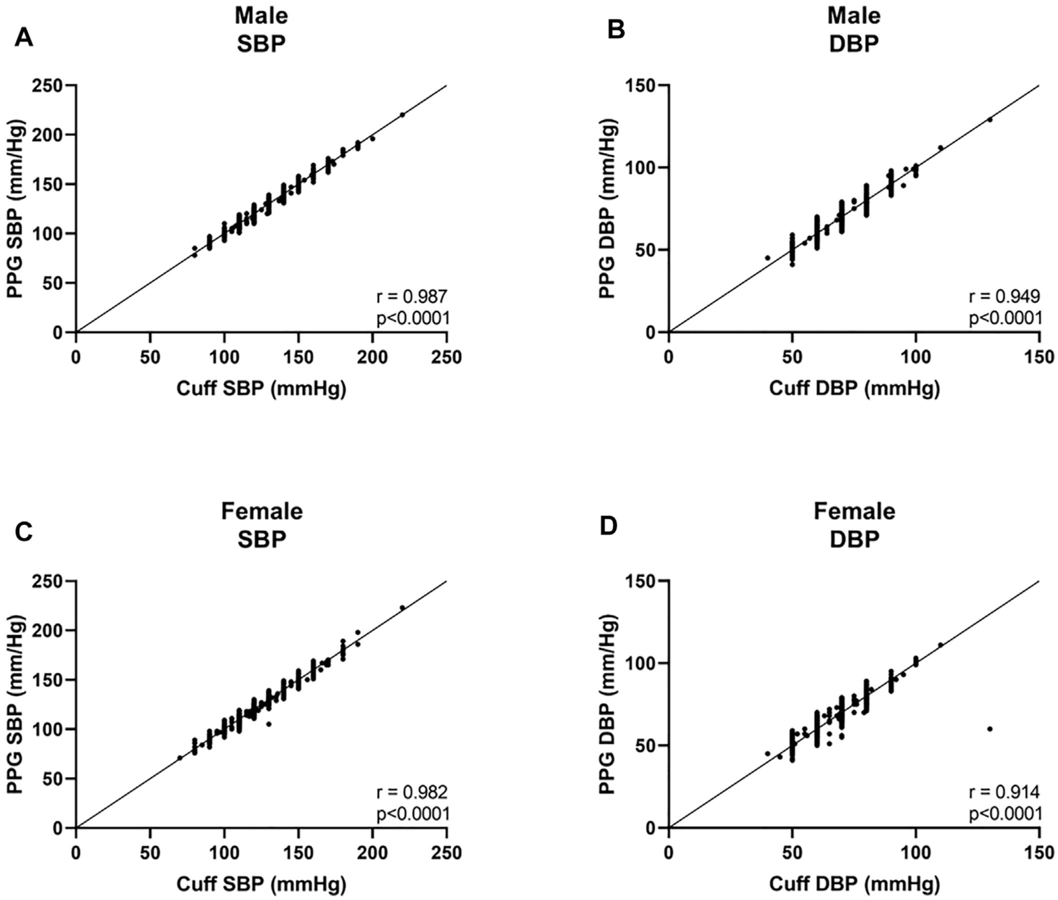


FIGURE 2 | Correlation curves of males and females. SBP systolic blood pressure, DBP, diastolic blood pressure.

In recent years, exciting algorithm and hardware improvements in PPG technology allowed the development and validation of non-invasive monitoring methods of blood pressure and other advanced hemodynamic parameters (Nachman et al., 2020a; Nachman et al., 2020b; Nachman et al., 2021). Whether these devices provide accurate measurements of such parameters in both sexes, in various skin tones, and body girth was not fully studied yet. In the current study, we aimed to test the accuracy of blood pressure

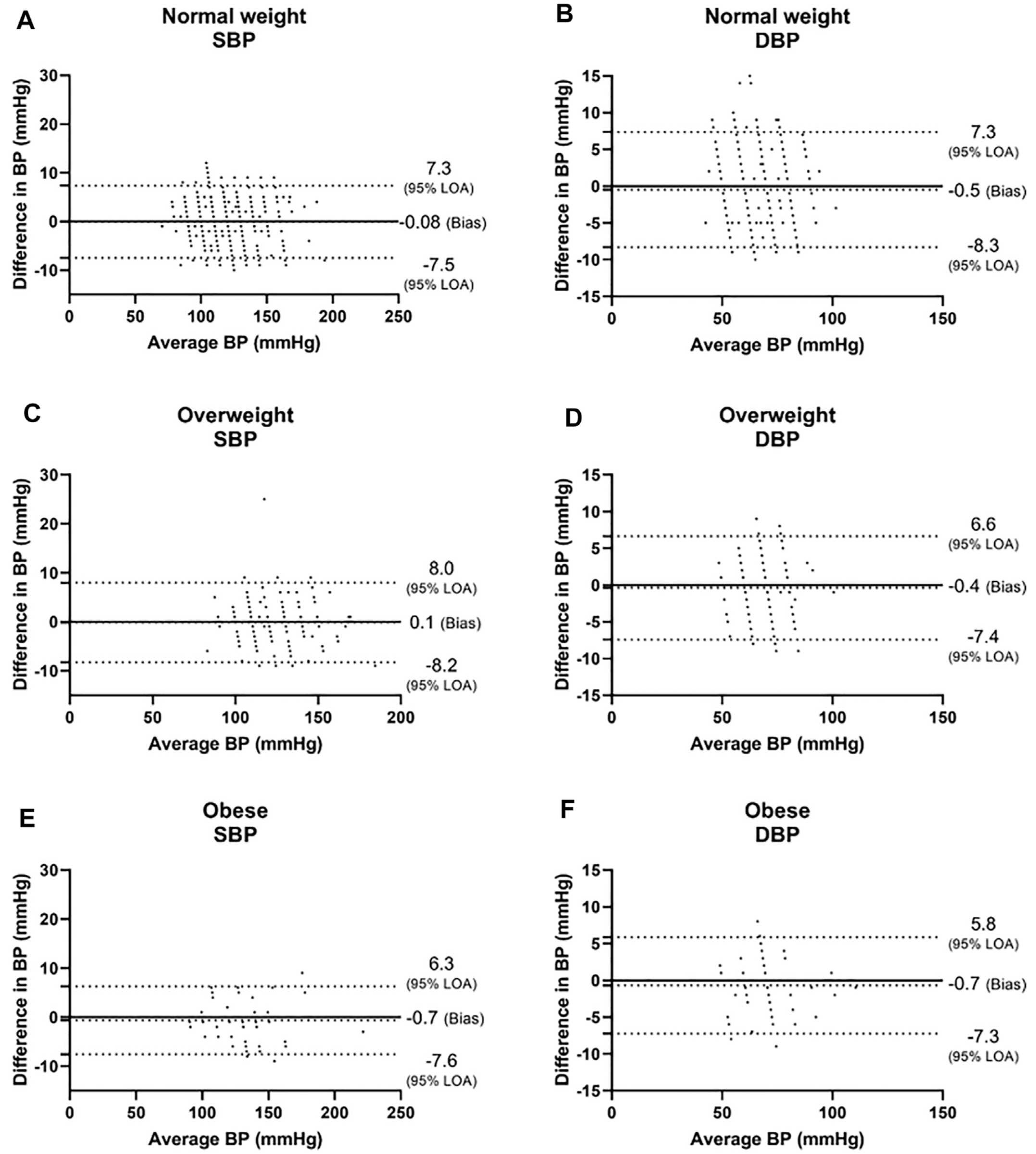


FIGURE 3 | Bland-Altman analysis of the body mass index (BMI) groups. Normal weight was defined as BMI<25, overweight was defined as 25 ≤BMI <30, and obese was defined as 30 ≤BMI. LOA level of agreement, SBP systolic blood pressure, DBP diastolic blood pressure.

(BP) measurements using a novel PPG-based device stratifying the data based on sex, BMI, and skin color, in a field setting.

METHODS

Study Population

This was an observational comparative field study. Male and female volunteers of all ages were recruited as they arrived for a BP checkup provided by skilled evaluators from the Israeli National Emergency Medical Services (EMS) in designated BP measurement public screening stations deployed across the country. Each participant has signed an informed consent form as defined by the Institutional Review Board of the Tel-Aviv Medical Center, Tel-Aviv, Israel (0032-15-TLV).

Study Protocol

Reference calibration BP measurements using a cuff-based device (Welch Allyn DuraShock DS65 hand sphygmomanometer, Skaneateles Falls, NY 13153, United States) were taken from each participant and used as a baseline calibration value for the PPG-based wearable device. Next, the wearable device was

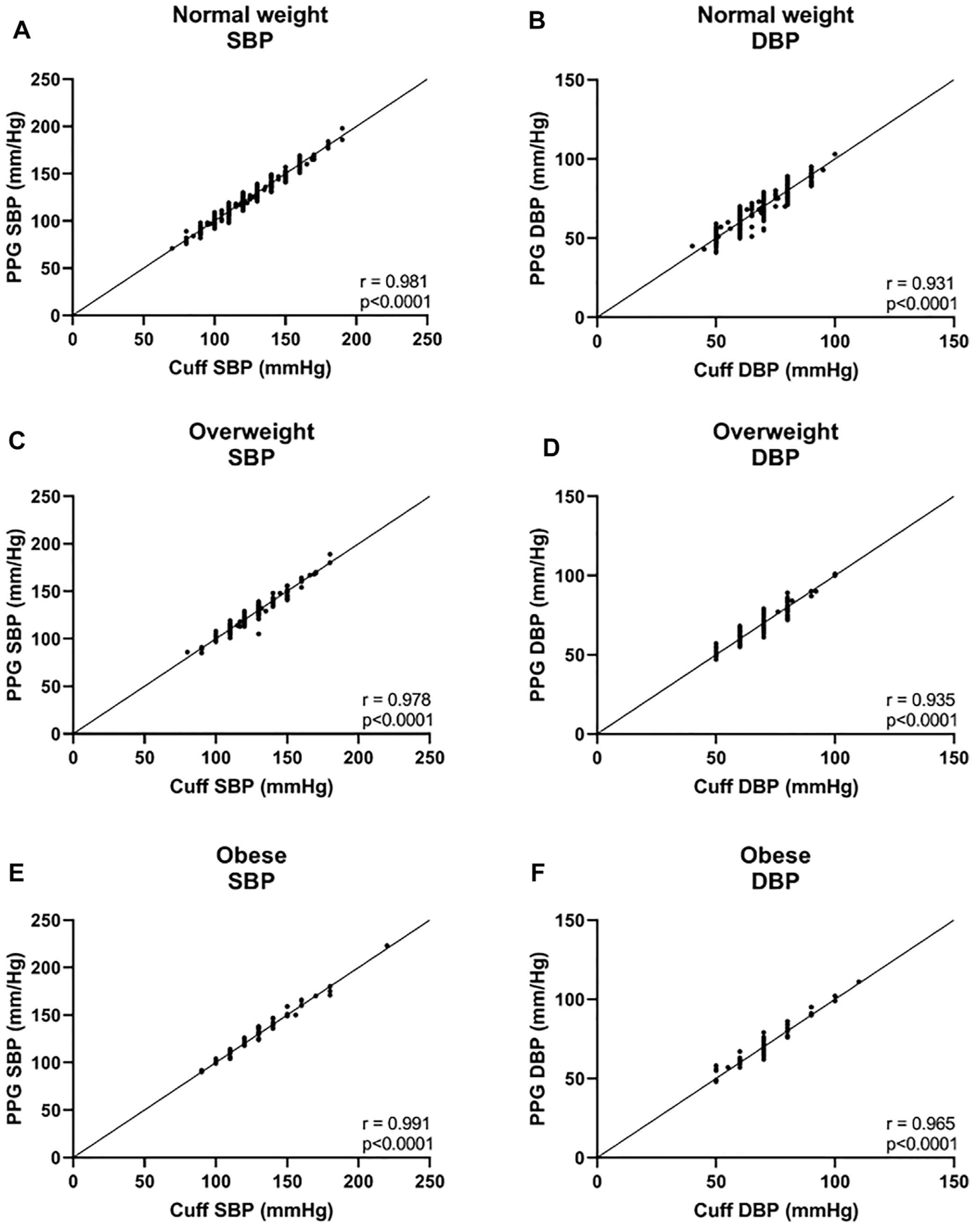


FIGURE 4 | Correlation curves of the body mass index (BMI) groups. Normal weight was defined as BMI <25, overweight was defined as 25 ≤BMI <30, and obese was defined as 30 ≤BMI. SBP systolic blood pressure, DBP diastolic blood pressure.

attached to each of the subjects by the EMS personnel and left for 30 s to acquire a PPG signal reading while the volunteers were at rest. This was repeated on the other hand to show there are no differences. After obtaining a signal, BP was measured concomitantly in all subjects using both devices - the standard cuff-based device previously mentioned on one hand and the PPG-based device on the opposite arm. Since the cuff-based device occludes blood flow to the arm, PPG measurements cannot be obtained if the device is located on that same arm. When comparing the PPG-based device to the cuff-based BP device, and since it takes time between the systolic and the diastolic measurements when using the cuff-based device, the systolic value of the PPG-based device was recorded once the systolic value was reached in the cuff-based device, and once the diastolic value was reached in the cuff-based device, the diastolic measurement in the PPG-based device was recorded. Adverse

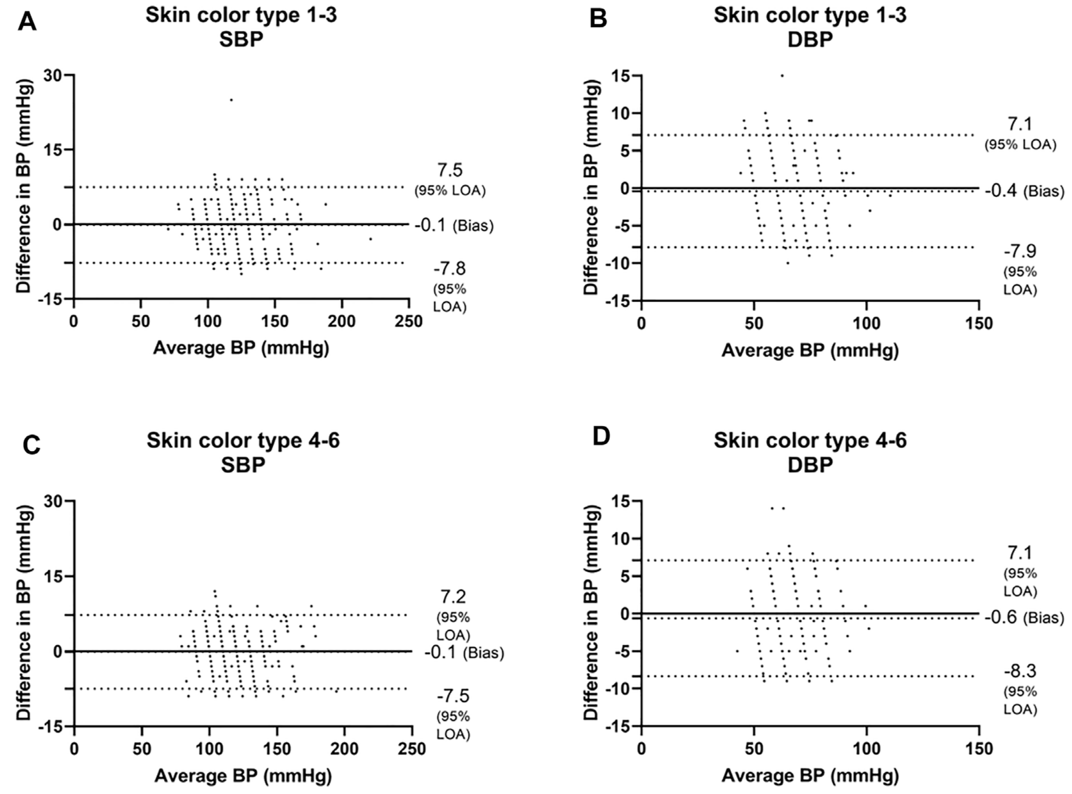


FIGURE 5 | Bland-Altman analysis of the body's skin color is based on the Fitzpatrick color scale (Fitzpatrick, 1988). Type 1—always burns, never tans, palest, can have freckles; Type 2—usually burns, tans minimally, light-colored but darker than fair; Type 3—sometimes mild burn, tans uniformly, golden honey or olive; Type 4—burns minimally, always tans well, moderate brown; Type 5—very rarely burns, tans very easily, dark brown; and Type 6—never burns, deeply pigmented dark brown to darkest brown. SBP systolic blood pressure, DBP diastolic blood pressure.

events were recorded by the research team. The evaluators were also recording basic demographic details of each participant, including height, weight, and based on the Fitzpatrick color scale provided to each, also determined the skin color type of each participant. There was one evaluator per patient, as the aim was to sustain the local routine blood pressure check-up protocol.

The PPG-Based Device

The skin-attached device (BB-613WP, Biobeat Technologies Ltd., Petah Tikva, Israel) is indicated for use in measuring and displaying pulse oxymetry and blood pressure values, as well as other advanced cardio-pulmonary parameters, using reflective PPG. Following transmission of several specific wavelengths of light, a detector measures the changing absorbance resulting from the pulsating arterial blood at each of the wavelengths. When using the device, a preliminary calibration step is performed using an FDA-cleared cuff-based BP device. The calibration value is entered into a user's application, valid for up to 3 months (relevant supportive clinical data has been presented by the company and accepted by the FDA), after which a new calibration measurement should be taken and introduced into the application using the same method. From that moment on, the PPG measurements are analyzed using pulse wave transit time (PWTT).

Statistical Analysis

Participants were stratified by sex, by BMI (BMI<25, defined as normal weight; 25 ≤BMI<30, defined as overweight; and 30 ≤BMI, defined as obese; based on the World Health Organization (WHO) definitions, https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight), and skin color based on the Fitzpatrick scale, divided between type 1–3 and type 4–6 (Fitzpatrick, 1988) (stratifications based on USFDA, 2013). Correlation analysis was performed using Pearson's correlation, and agreement was evaluated based on the Bland-Altman method using 95% limits of agreement (LOA). *p*-values were set at 0.05. Data analysis was performed using GraphPad Prism 8 and presented as mean ± standard deviation.

RESULTS

Table 1 includes demographic data and characteristics of the participants, stratified by BMI and skin color. No adverse events were recorded during the study. A total of 1548 samples recorded from 1057 participants were included in the analysis. For SBP, the Bland-Altman analysis of the whole group showed that the PPG had a bias of -0.02 ± 3.7 mmHg with 7.3 and −7.2 mmHg 95%

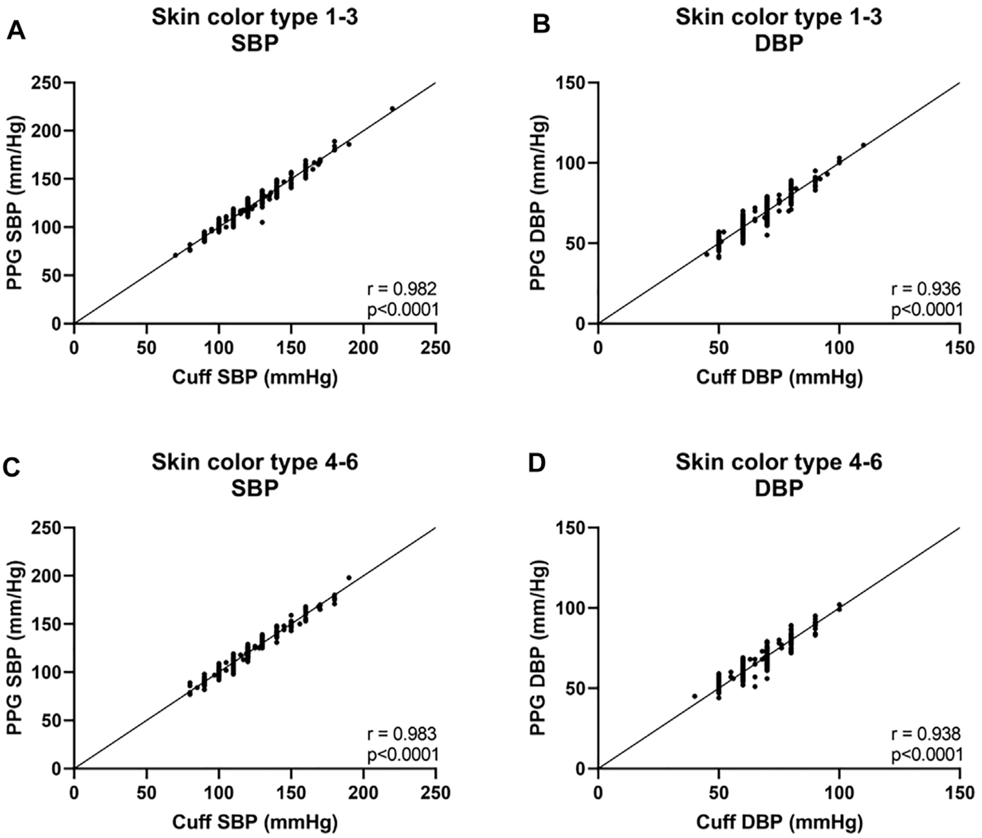


FIGURE 6 | Correlation curves of the body's skin color are based on the Fitzpatrick color scale (Fitzpatrick, 1988). Type 1—always burns, never tans, palest, can have freckles; Type 2—usually burns, tans minimally, light-colored but darker than fair; Type 3—sometimes mild burn, tans uniformly, golden honey or olive; Type 4—burns minimally, always tans well, moderate brown; Type 5—very rarely burns, tans very easily, dark brown; and Type 6—never burns, deeply pigmented dark brown to darkest brown. SBP systolic blood pressure, DBP diastolic blood pressure.

LOA, and Pearson's correlation (r) of 0.985. For DBP, the Bland-Altman analysis of the whole group showed that the PPG had a bias of −0.3 ± 4.2 mmHg with 8.0 and −8.6 mmHg 95% LOA, r = 0.931.

Among males, we found a bias of 0.2 ± 3.5 mmHg with 7.1 and −6.6 mmHg 95% LOA for SBP, with r = 0.987 (**Figures 1A**, **2A**), and a bias of 0.2 ± 3.8 mmHg with 7.3 and −7.7 mmHg 95% LOA for DBP, with r = 0.949 (**Figures 1B**, **2B**). Among females, we found a bias of 0.1 ± 3.8 mmHg with 7.4 and −7.7 mmHg 95% LOA for SBP, with r = 0.982 (**Figures 1C**, **2C**), and a bias of −0.4 ± 3.9 mmHg with 8.5 and −9.3 mmHg 95% LOA for DBP, with r = 0.914 (**Figures 1D**, **2D**).

In the three BMI groups, for SBP, a bias of −0.08 ± 3.8 mmHg with 7.3 and −7.5 mmHg 95% LOA was found in the normal weight group, r = 0.981; 0.1 ± 4.1 mmHg with 8.0 and −8.2 mmHg 95% LOA in the overweight group, r = 0.978; and −0.7 ± 3.5 mmHg with 6.3 and −7.6 mmHg 95% LOA in the obese group, r = 0.991 (**Figures 3A,C,E**, and **Figures 4A,C,E**). For DBP, a bias of −0.5 ± 4.0 mmHg with 7.3 and −8.3 mmHg 95% LOA in the normal weight group, r = 0.931; −0.4 ± 3.6 mmHg with 6.6 and −7.4 mmHg 95% LOA in the overweight group, r = 0.935; and −0.7 ± 3.3 mmHg with 5.8 and −7.3 mmHg 95% LOA in the obese group, r = 0.965 (**Figures 3B,D,F**, and **Figures 4B,D,F**).

Next, we analyzed the data of the two groups defined by the Fitzpatrick scale. For SBP, the Bland-Altman analysis showed that the PPG has a bias of −0.1 ± 3.9 mmHg with 7.4 and −7.4 mmHg 95% LOA in type 1–3 group, r = 0.982; and −0.1 ± 3.8 mmHg with 7.2 and −7.5 mmHg 95% LOA in type 4–6 group, r = 0.983 (**Figures 5A,C** and **Figures 6A,C**). For DBP, the Bland-Altman analysis showed that the PPG has a bias of −0.4 ± 3.8 mmHg with 7.1 and −7.9 mmHg 95% LOA in type 1–3 group, r = 0.936; and −0.6 ± 3.9 mmHg with 7.1 and −8.3 mmHg 95% LOA in type 4–6 group, r = 0.938 (**Figures 5B,D** and **Figures 6B,D**).

We also tested the level of agreement between the PPG-based wearable measurements across commonly accepted standard (Table 2). Overall, we found that the level of agreement between divides for all sub-population was ~80% within 5 mmHg, 94% within 7 mmHg, and 99% in the range of 10 mmHg.

DISCUSSION

In this study, we found a high level of accuracy of BP obtained by cuffless PPG-based BP device as compared to cuff-based BP measurements, an agreement that was not influenced by sex, BMI, or skin color.

Hypertension is a leading risk factor for cardiovascular morbidity and mortality in industrialized and developing countries. Despite numerous developments in treatment, one-third of patients are considered to suffer from uncontrolled BP. The diagnosis and continuous management of BP rely on accurate BP measurements. Although BP measurements using sphygmomanometer are common and acceptable practice they suffer from detrimental limitations, including the provision of only point measurements, they are cumbersome, uncomfortable, and in many patients, the measurements are unreliable due to technical issues (Kronish et al., 2017; Ruzicka and Hiremath, 2017).

According to the WHO, in 2016, 39% of adults aged 18 years and over (39% of men and 40% of women) were overweight, and about 13% of the world's adult population (11% of men and 15% of women) were obese (see URL in the Statistical analysis section). BMI and skin color are regarded as two important personal characteristics when looking at BP measurements, particularly as African Americans and obese are at higher risk of hypertension (Muntner et al., 2018; Chrysant, 2019; Ferdinand et al., 2020).

Novel PPG-based devices may provide continuous BP measurements with ease, making long-term BP monitoring and tailored therapy adjustments a reality (Nachman et al., 2021). Some PPG-based oximeters demonstrated unreliable measurements on people with darker skin tone (Ries et al., 1985; Ries et al., 1989; Jubran and Tobin, 1990; Zeballos and Weisman, 1991; Young, 1995). It is also accepted that cuff-based BP measurements are influenced by an individual's BMI. Thus, these two personal characteristics could lead to biased and less accurate measurements and have an impact on the quality of care, and it is of utmost importance to assure accurate measurement when using a PPG-based device.

In the current study, we compared BP measurements using a novel PPG-based non-invasive cuffless BP device with a gold standard non-invasive cuff-based BP device, and when stratifying the participants based on sex, BMI, and skin color, we found a high correlation and accordance between BP measurements obtained using the two devices in all sub-groups. We have previously shown that the device stands for the ANSI/AAMI/ISO 81060-2 requirements, for intermittent BP measurement, as defined in Stergiou et al., 2018 (Nachman et al., 2020a).

The PPG-based device may allow transformative monitoring options with substantial clinical impacts such as improving BP control and diagnosis. Our group demonstrated the accuracy of the PPG-based device for continuous 24-h BP measurement in a small pilot study and validation studies are now undertaken by our group to show the accordance with other physiological parameters recorded using the same PPG-based devices, including more advanced hemodynamic parameters such as stroke volume, cardiac output, and systemic vascular resistance.

Future studies in a large population in which several BP measurements will be taken for each participant along several time points, rather than single time points, could further strengthen the observations of this study.

A limitation of this study is that we did not follow an accepted standard such as the ANSI/AAMI/ISO 81060-2 (Stergiou et al., 2018), however the standard deals with cuff-based devices, and there are currently no accepted standards beyond unofficial recommendations regarding validation of cuff-less devices. Moreover, this was a field study in which we did not want to interfere too much with the routine. For this reason, we had only one evaluator per patient measuring the BP and assessing the skin type color, and not two. This also allowed us to collect data from more people in a relatively short period and a real-world setting.

CONCLUSION

We have found that the PPG-based BP device provides valid measurements regardless of sex, skin tone, or BMI. These findings provide confidence in the generalizability of such technology, helping in paving the way to a future of seamless, personalized BP management.

REPORTING SUMMARY

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

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 Institutional Review Board of the Tel-Aviv Medical Center, Tel-Aviv, Israel (0032-15-TLV). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

Conception of the work: AE, NG, and YG. Design of the work: AE, NG. Acquisition of the data: DN, AE, NG, AB-I, MF, RM. Analysis of the data: AE, NG, and YG. Interpretation of the data: DN, AE, NG, YG. Drafted or substantially revised the paper: DN, AE, NG, YG. All authors have read and agreed to the published version of the manuscript.

REFERENCES

Adler, J. N., Hughes, L. A., Vtvilecchia, R., and Camargo, C. A., Jr. (1998). Effect of Skin Pigmentation on Pulse Oximetry Accuracy in the Emergency Department. *Acad. Emerg. Med.* 5, 965–970. doi:10.1111/j.1553-2712.1998.tb02772.x

Bickler, P. E., Feiner, J. R., and Severinghaus, J. W. (2005). Effects of Skin Pigmentation on Pulse Oximeter Accuracy at Low Saturation. *Anesthesiology* 102, 715–719. doi:10.1097/00000542-200504000-00004

Bothma, P. A., Joynt, G. M., Lipman, J., Hon, H., Mathala, B., Scribante, J., et al. (1996). Accuracy of Pulse Oximetry in Pigmented Patients. *S. Afr. Med. J.* 86, 594–596.

Chrysant, S. G. (2019). Pathophysiology and Treatment of Obesity-Related Hypertension. *J. Clin. Hypertens.* 21, 555–559. doi:10.1111/jch.13518

Feiner, J. R., Severinghaus, J. W., and Bickler, P. E. (2007). Dark Skin Decreases the Accuracy of Pulse Oximeters at Low Oxygen Saturation: The Effects of Oximeter Probe Type and Gender. *Anesth. Analgesia* 105, S18–S23. doi:10.1213/01.ane.0000285988.35174.d9

Ferdinand, D. P., Nedunchezhian, S., and Ferdinand, K. C. (2020). Hypertension in African Americans: Advances in Community Outreach and Public Health Approaches. *Prog. Cardiovasc. Dis.* 63, 40–45. doi:10.1016/j.pcad.2019.12.005

Fitzpatrick, T. B. (1988). The Validity and Practicality of Sun-Reactive Skin Types I Through VI. *Arch. Dermatol.* 124, 869–871. doi:10.1001/archderm.1988.01670060015008

Jubran, A., and Tobin, M. J. (1990). Reliability of Pulse Oximetry in Titrating Supplemental Oxygen Therapy in Ventilator-Dependent Patients. *Chest* 97, 1420–1425. doi:10.1378/chest.97.6.1420

Kronish, I. M., Kent, S., Moise, N., Shimbo, D., Safford, M. M., Kynerd, R. E., et al. (2017). Barriers to Conducting Ambulatory and Home Blood Pressure Monitoring During Hypertension Screening in the United States. *J. Am. Soc. Hypertens.* 11, 573–580. doi:10.1016/j.jash.2017.06.012

Miller, V. M. (2014). Why Are Sex and Gender Important to Basic Physiology and Translational and Individualized Medicine? *Am. J. Physiol. Heart. Circ. Physiol.* 306, H781–H788. doi:10.1152/ajpheart.00994.2013

Muntner, P., Carey, R. M., Gidding, S., Jones, D. W., Taler, S. J., Wright, J. T., et al. (2018). Potential US Population Impact of the 2017 ACC/AHA High Blood Pressure Guideline. *J. Am. Coll. Cardiol.* 71, 109–118. doi:10.1016/j.jacc.2017.10.073

Nachman, D., Constantini, K., Poris, G., Wagnert-Avraham, L., Gertz, S. D., Littman, R., et al. (2020b). Wireless, Non-Invasive, Wearable Device for Continuous Remote Monitoring of Hemodynamic Parameters in a Swine Model of Controlled Hemorrhagic Shock. *Sci. Rep.* 10, 17684. doi:10.1038/s41598-020-74686-6

Nachman, D., Gepner, Y., Goldstein, N., Kabakov, E., Ishay, A. B., Littman, R., et al. (2020a). Comparing Blood Pressure Measurements Between a Photoplethysmography-Based and a Standard Cuff-Based Manometry Device. *Sci. Rep.* 10, 16116. doi:10.1038/s41598-020-73172-3

Nachman, D., Gilan, A., Goldstein, N., Constantini, K., Littman, R., Eisenkraft, A., et al. (2021). Twenty-Four-Hour Ambulatory Blood Pressure Measurement Using a Novel Non-Invasive, Cuff-Less, Wireless Device. *Am. J. Hypertens.* 34, 1171–1180. doi:10.1093/ajh/hpab095

O'Brien, E. (1996). Review: A Century of Confusion: Which Bladder for Accurate Blood Pressure Measurement? *J. Hum. Hypertens.* 10, 565–572.

Palatini, P., Benetti, E., Fania, C., and Saladini, F. (2019). Only Troncoconical Cuffs Can Provide Accurate Blood Pressure Measurements in People with Severe Obesity. *J. Hypertens.* 37, 37–41. doi:10.1097/hjh.0000000000001823

Palatini, P., and Parati, G. (2011). Blood Pressure Measurement in Very Obese Patients: A Challenging Problem. *J. Hypertens.* 29, 425–429. doi:10.1097/hjh.0b013e3283435b65

Pipek, L. Z., Nascimento, R. F. V., Acencio, M. M. P., and Teixeira, L. R. (2021). Comparison of SpO2 and Heart Rate Values on Apple Watch and Conventional Commercial Oximeters Devices in Patients with Lung Disease. *Sci. Rep.* 11, 18901. doi:10.1038/s41598-021-98453-3

Ries, A. L., Farrow, J. T., and Clausen, J. L. (1985). Accuracy of Two Ear Oximeters at Rest and During Exercise in Pulmonary Patients. *Am. Rev. Respir. Dis.* 132, 685–689. doi:10.1164/arrd.1985.132.3.685

Ries, A. L., Prewitt, L. M., and Johnson, J. J. (1989). Skin Color and Ear Oximetry. *Chest* 96, 287–290. doi:10.1378/chest.96.2.287

Ruzicka, M., and Hiremath, S. (2017). Accuracy-Limiting Factor of Home Blood Pressure Monitors? *Am. J. Hypertens.* 30, 661–664. doi:10.1093/ajh/hpx056

Sjoding, M. W., Dickson, R. P., Iwashyna, T. J., Gay, S. E., and Valley, T. S. (2020). Racial Bias in Pulse Oximetry Measurement. *N. Engl. J. Med.* 383, 2477–2478. doi:10.1056/nejmc2029240

Stergiou, G. S., Alpert, B., Mieke, S., Asmar, R., Atkins, N., Eckert, S., et al. (2018). A Universal Standard for the Validation of Blood Pressure Measuring Devices: Association for the Advancement of Medical Instrumentation/European Society of Hypertension/International Organization for Standardization (AAMI/ESH/ISO) Collaboration Statement. *J. Hypertens.* 36, 472–478. doi:10.1097/hjh.0000000000001634

US Food and Drug Administration (2013). Pulse Oximeters - Premarket Notification Submissions [510(k)s]: Guidance for Industry and Food and Drug Administration Staff. Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pulse-oximeters-premarket-notification-submissions-510ksguidance-industry-and-food-and-drug (Last Accessed on 11/25/2021).

Wiles, M. D., El-Nayal, A., Elton, G., Malaj, M., Winterbottom, J., Gillies, C., et al. (2021). The Effect of Patient Ethnicity on the Accuracy of Peripheral Pulse Oximetry in Patients with COVID-19 Pneumonitis: A Single-Centre, Retrospective Analysis. *Anaesthesia* 77, 143–152. doi:10.1111/anae.15581

Young, K. (1995). A Study of the Accuracy of Pulse Oximetry Among Varying Skin Pigmentation Populations. *Am. J. Crit. Care.* 4, 256.

Zeballos, R. J., and Weisman, I. M. (1991). Reliability of Noninvasive Oximetry in Black Subjects During Exercise and Hypoxia. *Am. Rev. Respir. Dis.* 144, 1240–1244. doi:10.1164/ajrccm/144.6.1240

Conflict of Interest: Authors AE, NG, AB-I, MF, and RM are employed by Biobeat Technologies Ltd.

The remaining 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.

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 25 July 2022
DOI 10.3389/fphys.2022.961386

EDITED BY
Tao Tan,
Eindhoven University of Technology, Netherlands

REVIEWED BY
Siyi Xun,
Macao Polytechnic University, Macao SAR, China
Qingchuan Li,
National University of Defense Technology, China
Qinhang Xu,
National University of Defense Technology, China

*CORRESPONDENCE
Guoxu Zhang,
zhangguoxu_502@163.com
Zhiguo Wang,
wangzhiguo5778@163.com

SPECIALTY SECTION
This article was submitted to Medical Physics and Imaging, a section of the journal Frontiers in Physiology

RECEIVED 04 June 2022
ACCEPTED 28 June 2022
PUBLISHED 25 July 2022

CITATION
Yao Z, Yuan Y, Shi Z, Mao W, Zhu G, Zhang G and Wang Z (2022), FunSwin: A deep learning method to analysis diabetic retinopathy grade and macular edema risk based on fundus images. *Front. Physiol.* 13:961386.
doi: 10.3389/fphys.2022.961386

FunSwin: A deep learning method to analysis diabetic retinopathy grade and macular edema risk based on fundus images

Zhaomin Yao[1,2], Yizhe Yuan[3], Zhenning Shi[3], Wenxin Mao[3], Gancheng Zhu[3], Guoxu Zhang[1,2]* and Zhiguo Wang[1,2]*

[1]College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, Liaoning, China, [2]Department of Nuclear Medicine, General Hospital of Northern Theater Command, Shenyang, Liaoning, China, [3]College of Computer Science and Technology, Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun, Jilin, China

Diabetic retinopathy (DR) and age-related macular degeneration (AMD) are forms of degenerative retinal disorders that may result in vision impairment or even permanent blindness. Early detection of these conditions is essential to maintaining a patient's quality of life. The fundus photography technique is non-invasive, safe, and rapid way of assessing the function of the retina. It is widely used as a diagnostic tool for patients who suffer from fundus-related diseases. Using fundus images to analyze these two diseases is a challenging exercise, since there are rarely obvious features in the images during the incipient stages of the disease. In order to deal with these issues, we have proposed a deep learning method called FunSwin. The Swin Transformer constitutes the main framework for this method. Additionally, due to the characteristics of medical images, such as their small number and relatively fixed structure, transfer learning strategy that are able to increase the low-level characteristics of the model as well as data enhancement strategy to balance the data are integrated. Experiments have demonstrated that the proposed method outperforms other state-of-the-art approaches in both binary and multiclass classification tasks on the benchmark dataset.

KEYWORDS

fundus image, diabetic retinopathy, macular edema, disease stage prediction, swin transformer

1 Introduction

An eyeball is an impressively ingenious structure, with an optical system that mimics a traditional camera, and the fundus which functions as the photographic plate of the camera, allows one to see the dynamic of the blood circulation and the health status of the human body (Musadiq et al., 2003; Lee and Szema, 2005). For example, various characteristics of certain complications of diabetes, hypertension, coronary heart disease, and kidney disease can be identified in the fundus (Nathan, 1993; Wong et al., 2001; Grunwald et al., 2010; Xu et al., 2021). Presently, fundus photography is a commonly used method for screening the fundus. This technique enables visual perception of structure, which allows us to determine if there is any abnormality in the fundus (Yannuzzi et al., 2004).

Diabetic retinopathy (DR) and age-related macular degeneration (AMD) are two ophthalmic diseases that can be diagnosed through fundus photographs. Basic clinical manifestations of DR will appear on fundus images as neovascularization, capillary hemangiomas, vasodilation, hemorrhage, and occlusion of capillaries and arterioles (Stitt et al., 2016), whereas the basic manifestations of AMD will appear on fundus images as mainly the alteration of fundus macula (Zarbin et al., 2014). Unfortunately, in the early stages of the disease there may not be obvious clinical symptoms evident in the fundus image, making diagnosis challenging (Agurto et al., 2011).

Deep learning has made great strides in medical image diagnosis over the last decade. In particular, a number of deep neural networks have been modified and applied for detecting diseases related with fundus images in recent years. For example, several structural features of biological damage, such as blood vessels, fundus hemorrhage, and exudate, are added to advanced neural networks to train classification models based on artificially designed features (Alqudah et al., 2018; Ghani et al., 2019; Dong et al., 2021). These neural frameworks can also be trained using simple image characteristics such as pixel intensities (Das et al., 2021; Kanimozhi et al., 2021; Selçuk et al., 2022). Besides focusing on feature innovations, scientists will also focus on methodological innovations, such as developing a high performance deep neural network and integrating different machine learning algorithms with ensemble models (Sikder et al., 2021; Du et al., 2022).

It is true that these state-of-the-art methods have provided good results, however, many of them do not offer a diagnosis of disease staging, at the same time, for the detection of the above two diseases, they must be enhanced. Furthermore, the diagnostic performance of the models needs to be enhanced for the above two diseases. To address the above two pivotal questions, we propose a deep learning method based on Swin transformer. The main contributions of this work are summarized as follows:

For one thing, the appropriate benchmark model and other modules are selected and optimized for integrating a suitable deep learning framework for analyzing fundus images in accordance with the specific research objectives. For another, a series of highly reliable preprocessing operations are implemented based on the properties of the fundus images, while ensuring the integrity of the distribution of the data, thereby enhancing the accuracy of the resulting prediction. Finally, the ImageNet-based transfer learning mode is used as the basis training model in order to obtain sufficient low-level features for the learning. Consequently, when the model is fused with high-dimensional features, the model can be perceived more clearly (particularly some potential disease classification bases) and classification accuracy can be improved.

The paper is organized as follows: In Section 2, we provide an overview of the datasets and a brief description of the methods. In Section 3, we present experimental results and conclusions based on these results. Finally, in Section 4, we conclude the paper with a brief summary.

2 Materials and methods

As shown in Figure 1, This study involved four major stages: Dataset curation, data preprocessing, model training and prediction. The collected fundus images are first cut into squares and normalized to the same size, then the samples are balanced based on the number of each class. In addition, mixup and cutmix are used to further process the data. Since the Swin Transformer demonstrates excellent performance on other medical image classification problems, this framework is used in this study, and its parameters are adjusted based on the model's performance. Finally, the binanry and mutil classification performance of the optimized will be evaluated by the evaluation metics. As a reminder, "Binary Classification" refers to the classification of health and disease, whereas "Multiclassification" refers to the classification of health and disease at different levels. The details of each process are described in the following sections.

2.1 Benchmark dataset

In this article, we used data from the MESSIDOR dataset (Decencière et al., 2014) that is available online for public use. There are 1,200 color numerical images of the posterior pole of the eye in this dataset, which correspond to two diseases: Diabetic retinopathy (4 levels, 0–3) and macular edema (3 levels, 0–2), where level 0 corresponds to a healthy subject. Medical experts have provided diagnoses for each image. Full details are available in Supplementary Tables S1, S2.

2.2 Data preprocessing

Firstly, only the middle part of the image is intercepted, which contains all pixels within the field of view, in order to

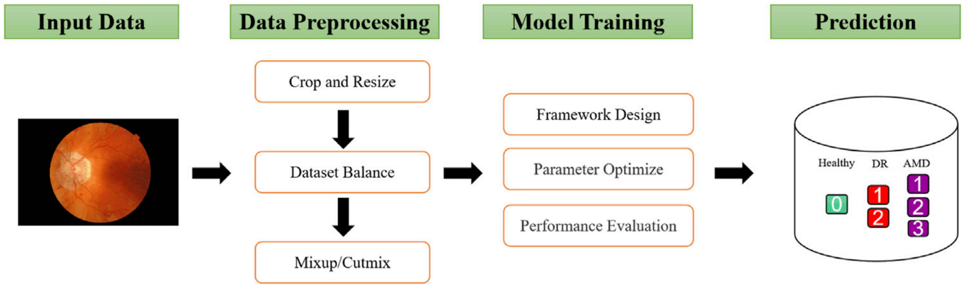


FIGURE 1
Overview of the proposed methodology.

lessen the interference caused by large areas of black background. And the size of these small crops is set to 960 × 960. As a second step, the simplest data enhancement methods (rotation and mirroring) are used to increase the number of minority samples, eliminating the effects of categories imbalance and maintaining an equal distribution of the dataset. For example, As regards diabetic retinopathy, the sample sizes for 0–3 levels are 546, 153, 247, and 254 respectively. On level 1 samples, we use 180-degree rotation, mirroring, mirroring+180-degree rotation to expand them to three times their original size. While level 2 and 3 samples are expanded by 180-degree rotation and mirroring. This procedure allows each subclass to attain an approximate balance in the number of samples and eliminates the impact of imbalanced categories. Lastly, the internal mix-up and cut-mix methods of the network (Liu et al., 2021) will also be used to optimize the performance of the Swin Transformer. CutMix and MixUp enable us to create inter-class examples. CutMix randomly interpolates the pixel values between two images and places fragments of one image over another, while MixUp randomly interpolates the pixels between two images. The two processes prevent the model from being overfitted to the training distribution and improve its likelihood of being able to generalize to examples outside of the distribution. A further benefit of CutMix is that it prevents a model from over relying on any particular feature when it is performing its classifications.

2.3 Model training

2.3.1 The pipeline of the framework

The deep learning framework implemented in this study consists of three components, which are the backbone, the neck, and the head.

As illustrated in Figure 2, the Swin Transformer is served as the framework's backbone. Its structure is reminiscent of a convolutional hierarchy and the resolution is reduced by half, while the number of channels is doubled. The first patch partition divides the image into a series of blocks, followed by four stages, each of which contains two parts: patch merging (the first block is linear) and Swin Transformer block. Patch merging is similar to pooling; however, it does not lose information in the process.

As depicted in Figure 3, Swin Transformer Block is basically similar to a common transformer block except that it uses window multi-head self-attention (W-MSA) and mobile window multi-head self-attention (SW-MSA) to replace multi-head self-attention (MSA) module. With this moving-window method, self-focused computations are limited to a non-overlapping local window, allowing for inter-window connectivity. Moreover, this hierarchical converter is capable of modeling images of various sizes and has linear computation complexity. As a result of these features, Swin converter is highly competitive in handling a wide variety of visual tasks (Zhang et al., 2021; Jiang et al., 2022).

Global Average Pooling (GAP) composes the neck of the framework. There are several advantages of GAP over traditional fully ensemble layers. One is that it is more suitable for convolutional structures by improving the compatibility of function maps and categories, another is that there are no parameters to adjust in the global media collection, meaning that overestimation at the global level can be reduced. It contributes to the achievement of good results in many network structures for medical data (Bien et al., 2018; Valan et al., 2019; Shahhosseini et al., 2021).

Linear CLS is the head of the framework. Using this module makes the model relatively simple and easier to train since the mapping between features and categories is clearly visible. Its loss function is described as follows:

$$Loss = -\sum_{i=1}^{n} y(i)\,log\left(p(x_i)\right) \tag{1}$$

Where $p(x_i)$ is the result of the model output computed by softmax and $y(i)$ is calculated as follows:

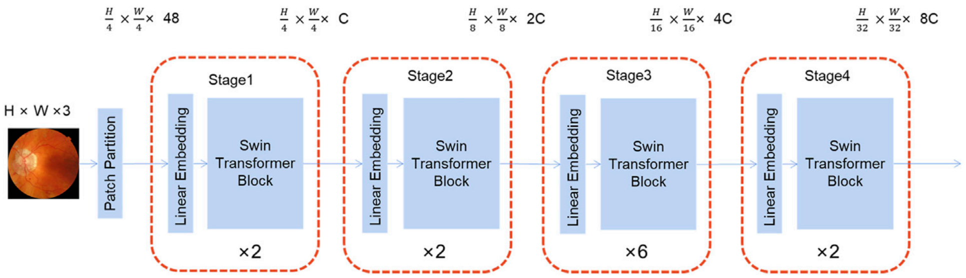


FIGURE 2
The framework of the Swin transformer.

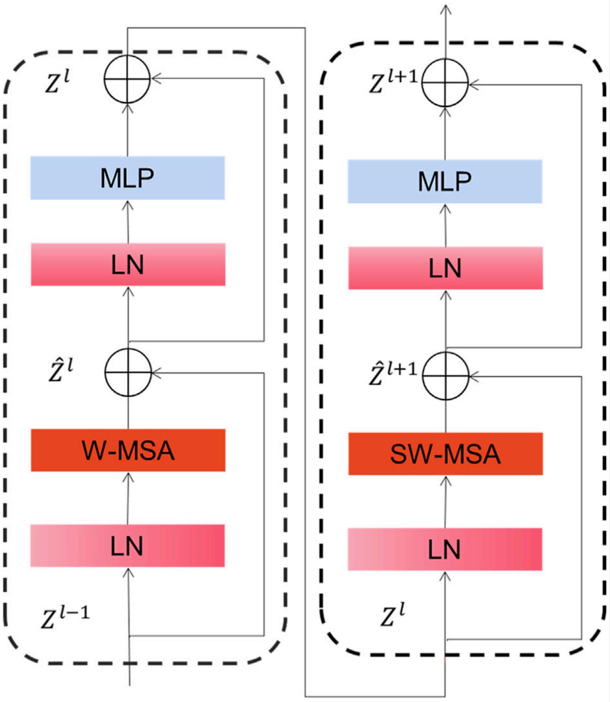


FIGURE 3
The details of the Swin Transformer Block.

$$y(i) = \begin{cases} \frac{\alpha}{n}, & i \neq class \\ 1 - \alpha + \frac{\alpha}{n}, & i = class \end{cases} \tag{2}$$

Where n is the number of categories, i is the predicted label, and class is the current real category. Where α is smoothing coefficient, which we set to 0.1, as in (Liu et al., 2021).

2.3.2 Parameter setting

This framework has the following parameters: the batch size is set to 32, the epoch is set to 600, and the initialization policy of the CLS head is TruncNormal, which has a standard deviation of 0.02. Additionally, AdamW is used as the optimizer with a learning rate of 0.0001 and decay rate of 0.05. All other parameters are set to the default values.

2.3.3 Performance metrics

Performance metrics have been employed to assess the predictive performance of our models, including sensitivity, specificity, accuracy, and F1-score. In medical image analysis, these evaluation measurements are well established and they have been used in the benchmark studies on the diagnosis of fundus-related diseases as well. The metrics are calculated as follows.

$$\begin{cases} Acc = \dfrac{TP + TN}{TP + FP + TN + FN} \\ SN = RE = \dfrac{TP}{TP + FN} \\ SP = \dfrac{TN}{TN + FP} \\ PR = \dfrac{TP}{TP + FP} \\ F1 - Score = \dfrac{2 \times PR \times RE}{PR + RE} \end{cases} \tag{3}$$

The variables TP, FP, TN and FN represent the true positive, false positive, true negative, and false negative values, respectively. RE and PR represent recall and precision.

3 Results and discussions

3.1 The result of binary classification on diabetic retinopathy and macular edema

This paper compares nine state-of-the-art methods that have been widely used in medical imaging in recent years, which are Conformer (Peng et al., 2021), ConvNeXt (Liu et al., 2022), HRnet (Wang et al., 2020), Vgg 11 (Simonyan

TABLE 1 Comparing the existing methods of binary classification for diabetic retinopathy.

Methods	Accuracy	Sensitivity	Specificity	F1-score
Conformer	0.8855	0.8817	0.8963	0.9193
Convnext	0.8267	**0.9613**	0.4451	0.8913
HRnet	0.8378	0.8796	0.7195	0.8891
Vgg11	0.8156	0.8538	0.7073	0.8725
Mlp Mixer	0.8585	0.9226	0.6768	0.9060
Res2net50	0.8267	0.8452	0.7744	0.8782
Shufflenet_v1	0.8045	0.8796	0.5915	0.8693
T2T Vit	0.8553	0.8323	**0.9207**	0.8948
Vit Transformer	0.8076	0.9419	0.4268	0.8786
Our Method	**0.9062**	0.9376	0.8171	**0.9366**

TABLE 2 Comparing the existing methods of binary classification for macular edema.

Methods	Accuracy	Sensitivity	Specificity	F1-score
Conformer	0.9745	0.9755	0.9727	0.9801
Convnext	0.9320	0.8962	**0.9966**	0.9443
HRnet	0.9453	0.9283	0.9761	0.9563
Vgg11	0.7947	0.7283	0.9147	0.8204
Mlp Mixer	0.9441	0.9453	0.9420	0.9561
Res2net50	0.9648	0.9566	0.9795	0.9722
Shufflenet_v1	0.9648	0.9698	0.9556	0.9726
T2T Vit	0.9587	0.9377	**0.9966**	0.9669
Vit Transformer	0.9611	0.9604	0.9625	0.9695
Our Method	**0.9866**	**0.9868**	0.9863	**0.9896**

TABLE 3 Comparing the existing methods of multi-classification for diabetic retinopathy.

Methods	Accuracy	Sensitivity	Specificity	F1-score
Conformer	0.7704	0.7524	0.9122	0.7691
Convnext	0.7123	0.6754	0.8770	0.6999
HRnet	0.7374	0.7031	0.8954	0.7311
Vgg11	0.6211	0.5630	0.8547	0.6013
Mlp Mixer	0.7248	0.7121	0.8942	0.7246
Res2net50	0.6352	0.5568	0.8665	0.6054
Shufflenet_v1	0.6101	0.6500	0.8383	0.6026
T2T Vit	0.6950	0.7023	0.8679	0.6845
Vit Transformer	0.7563	0.7268	0.9150	0.7490
Our Method	**0.8412**	**0.8154**	**0.9413**	**0.8400**

TABLE 4 Comparing the existing methods of multi-classification for macular edema.

Methods	Accuracy	Sensitivity	Specificity	F1-score
Conformer	0.9733	0.9733	0.9865	0.9733
Convnext	0.8214	0.8147	0.9104	0.8091
HRnet	0.8761	0.8725	0.9379	0.8745
Vgg11	0.6136	0.6035	0.8056	0.5753
Mlp Mixer	0.9210	0.9202	0.9602	0.9209
Res2net50	0.8882	0.8849	0.9443	0.8872
Shufflenet_v1	0.9635	0.9638	0.9816	0.9636
T2T Vit	0.9174	0.9144	0.9583	0.9160
Vit Transformer	0.9514	0.9510	0.9754	0.9514
Our Method	**0.9866**	**0.9866**	**0.9932**	**0.9866**

and Zisserman, 2014), Mlp-Mixer (Tolstikhin et al., 2021), Res2Net 50 (Gao et al., 2019), ShuffleNet V1 (Zhang et al., 2018), T2T Vit (Yuan et al., 2021) and Vit Transformer (Dosovitskiy et al., 2020).

As depicted in Table 1, the proposed method yielded the best results for diabetic retinopathy on both the accuracy and F1-score, and for sensitivity, our effect is second only to that of VIT Transformer, which is only 0.43% worse. With regard to specificity, our effect ranks third, which is 10.36% less than that of the highest-performing model. Similarly, for macular edema, as illustrated in Table 2, accuracy, F1-score, and sensitivity of the proposed model have reached the highest values, 98.66, 98.96, and 98.68%, respectively, while their specificity is 1% lower than the best model.

The relationship between specificity and sensitivity is often asymmetric, so it is very challenging to make sure both will produce positive results. We have maintained that the objective of this project is to better screen out patients with diseases, therefore in terms of method design and model training, we have sought a higher degree of sensitivity. Perhaps this explains why our model is less specific than other models in binary diagnostics of these disorders.

3.2 The results of multi-classification on diabetic retinopathy and macular edema

Having to deal with the problem of multi-classification, macro-average is used to calculate these indicators. According to this principle, increasing the proportion of each category of images will increase the weight of that category. The final result of the indicator is the sum of the results obtained from multiplying the corresponding indicator results of each subcategory by their respective weights.

Table 3 and Table 4 illustrate that the proposed method is superior to other models in various indicators with regard to the

TABLE 5 Performance of binary classification before and after data augmentation.

	Methods	Accuracy	Sensitivity	Specificity	F1-score
Diabetic Retinopathy	No Augmentation	0.5444	1.0	0	0.7050
	Augmentation	0.9062	0.9376	0.8171	0.9366
Macular Edema	No Augmentation	0.9389	0.7941	0.9726	0.8308
	Augmentation	0.9866	0.9868	0.9863	0.9896

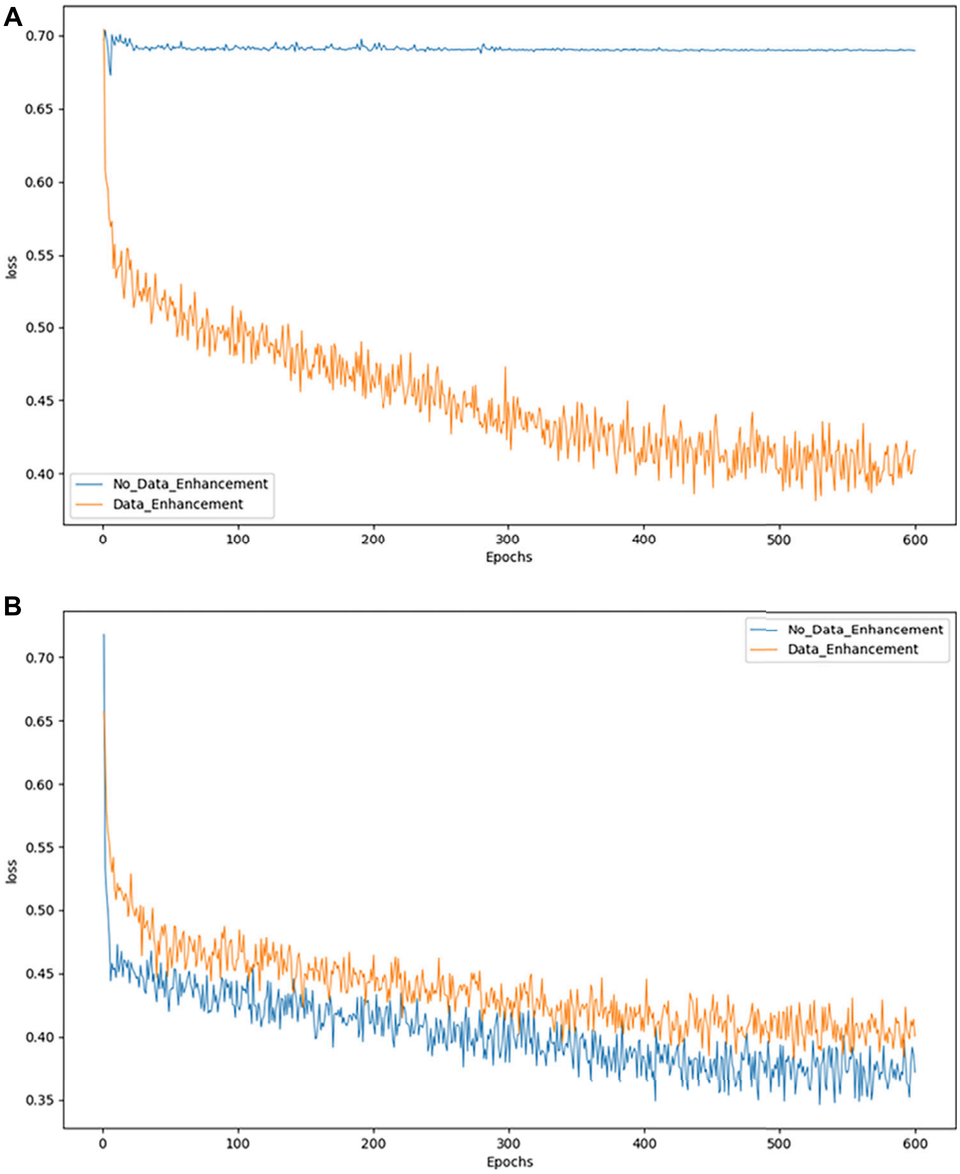


FIGURE 4
Model convergence performance of binary classification.

multiclassification problem of diabetic retinopathy and macular edema. In comparison with the binary classification problem, our model does not demonstrate any reduction in the multi-classification problem. It may be that there are only three subcategories to classify, or that there are distinct features which separate subcategories.

3.3 Performance of data enhancement on the model

The data augmentation test was conducted on the binary classification. Table 5 summarizes the changes to evaluation indicators of the proposed method before and after data enhancement. Following data enhancement, almost all indicators were improved. There are two main factors contributing to this: The first is the expansion of minorities by rotating and flipping against class imbalances; this allows the data to be more balanced, reduces the impact of unbalanced data on the model, and enhances its performance. Additionally, fundus images the physiological structure reflected by fundus images is relatively fixed, that is, the distribution of segmentation targets in fundus images is essentially regular, and the semantic understanding of these targets is rather straightforward. So, low-resolution information provides specific features that are necessary for target object recognition. Although the model has gained sufficient low-level features from migration learning, there are still only a limited number of original images available for input, which means that the enhanced images compensate for the lack of original data.

3.4 Performance of transfer learning on the model

An assessment of transfer learning was conducted on the binary classification. Accordingly, the accuracy, sensitivity, specificity, and F1-score values of diabetic retinopathy are respectively 0.7393, 1, 0 and 0.8501. Similarly, these indicators of macular edema are 0.6457, 1, 0, and 0.7847. ACC and F1 of both diseases increased significantly following the addition of transfer learning.

3.5 Convergence of the model

As an example, we exploit the binary classification problem to demonstrate the effects of model convergence. Figures 4A,B show the convergence of the model in DR and MD respectively. The abscissa represents the number of epochs used, while the ordinate represents the value of the loss suffered by the epoch. As shown in Figure 4A, the model converges more effectively only after data augmentation. While in 4B the model without data augmentation achieves better convergence, it is not evident from the actual test results.

Loss declines may not be apparent when the model cannot solve the problem of category imbalance. Accordingly, we can consider balancing the data set to resolve this issue. However, when the number of pictures increases because of a data balance, the overall loss may increase slightly, although this may be due to the increase in images, which does not adversely affect our classification accuracy.

3.6 Running environment and time cost

The experiments were conducted using a computing server with an Intel i9-11900K CPU, an NVIDIA RTX-3090 GPU, and Kingston 32 GB memory. Ubuntu 18.04 is the operating system of the server. In general, the training time for a picture is 0.02 s and the testing time is 0.007 s.

4 Conclusion

In this project, a method referred to as FunSwin is proposed as a means to solve the problem of grading diabetic retinopathy and estimating macular edema risk using fundus images. The basic framework for the method is Swin Transformer, with some modules based on some features of medical data to improve performance. In comparison to the existing studies on this benchmark dataset, FunSwin was found to outperform the existing studies in binary classifications and multi-classifications of these two diseases. Furthermore, as regards binary classification, when each subcategory of disease is given the same amount of training data, i.e., assuming that all data for each subcategory is balanced, then the binary classification effect of the model will still be improved. The study however, may need further evaluations in the clinical practices. There have been very few clinical studies on AI-based retinal diseases due to a variety of challenges, such as regulatory requirements and the annotations of experienced clinicians. Additionally, there is no specific evidence that these fundus-related symptoms are directly connected to particular diseases. After receiving ethical approval and the accumulating of a large, well-annotated dataset, this limitation of the proposed method will be resolved in a future study. In future studies, we also intend to use FunSwin in treating other retinal disorders, such as stroke, heart disease, etc.

Data availability statement

Publicly available datasets were analyzed in this study. The data can be found at: https://www.adcis.net/en/third-party/messidor/[MESSIDOR DATASET].

Author contributions

ZW and GuZ designed the project, ZS, WM, and GaZ collected and preprocessed the datasets, YY and ZY carried out the coding of the computational analysis, ZY and YY drafted the manuscript, ZW and GuZ revised and polished the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Key Research and Development Program of Liaoning Provence (2019JH2/10300010).

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphys.2022.961386/full#supplementary-material

References

Agurto, C., Barriga, E. S., Murray, V., Nemeth, S., Crammer, R., Bauman, W., et al. (2011). Automatic detection of diabetic retinopathy and age-related macular degeneration in digital fundus images. *Invest. Ophthalmol. Vis. Sci.* 52 (8), 5862–5871. doi:10.1167/iovs.10-7075

Alqudah, A. M., Hiam, A., Isam, A., and Alaa, B. (2018). Employing image processing techniques and artificial intelligence for automated eye diagnosis using digital eye fundus images. *J. Biomimetics, Biomaterials Biomed. Eng.* 39, 40–56. doi:10.4028/www.scientific.net/jbbbe.39.40

Bien, N., Rajpurkar, P., Ball, R. L., Irvin, J., Park, A., Jones, E., et al. (2018). Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. *PLoS Med.* 15 (11), e1002699. doi:10.1371/journal.pmed.1002699

Das, S., Kharbanda, K., M, S., Raman, R., and D, E. D. (2021). Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy. *Biomed. Signal Process. Control* 68, 102600. doi:10.1016/j.bspc.2021.102600

Decencière, E., Xiwei, Z., Guy, C., Bruno, L., Béatrice, C., Caroline, T., et al. (2014). Feedback on a publicly distributed image database: the messidor database. *Image Analysis Stereol.* 33 (3), 231–234. doi:10.5566/ias.1155

Dong, L., Yang, Q., Zhang, R. H., and Wei, W. B. (2021). Artificial intelligence for the detection of age-related macular degeneration in color fundus photographs: a systematic review and meta-analysis. *EclinicalMedicine* 35, 100875. doi:10.1016/j.eclinm.2021.100875

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.

Du, J., Zou, B., Ouyang, P., and Zhao, R. (2022). Retinal microaneurysm detection based on transformation splicing and multi-context ensemble learning. *Biomed. Signal Process. Control* 74, 103536. doi:10.1016/j.bspc.2022.103536

Gao, S.-H., Cheng, M. M., Zhao, K., Zhang, X. Y., Yang, M. H., and Torr, P. H. S. (2019). Res2net: A new multi-scale backbone architecture. *IEEE Trans. Pattern Anal. Mach. Intell.* 43 (2), 652–662. doi:10.1109/TPAMI.2019.2938758

Ghani, A., See, C. H., Sudhakaran, V., Ahmad, J., and Abd-Alhameed, R. (2019). Accelerating retinal fundus image classification using artificial neural networks (ANNs) and reconfigurable hardware (FPGA). *Electronics* 8 (12), 1522. doi:10.3390/electronics8121522

Grunwald, J. E., Alexander, J., Maguire, M., Whittock, R., Parker, C., McWilliams, K., et al. (2010). Prevalence of ocular fundus pathology in patients with chronic kidney disease. *Clin. J. Am. Soc. Nephrol.* 5 (5), 867–873. doi:10.2215/cjn.08271109

Jiang, Y., Zhang, Y., Lin, X., Dong, J., Cheng, T., and Liang, J. (2022). SwinBTS: a method for 3D multimodal brain tumor segmentation using swin transformer. *Brain Sci.* 12 (6), 797. doi:10.3390/brainsci12060797

Kanimozhi, J., Vasuki, P., and Roomi, S. M. (2021). Fundus image lesion detection algorithm for diabetic retinopathy screening. *J. Ambient. Intell. Hum. Comput.* 12 (7), 7407–7416. doi:10.1007/s12652-020-02417-w

Lee, L. P., and Szema, R. (2005). Inspirations from biological optics for advanced photonic systems. *Science* 310 (5751), 1148–1150. doi:10.1126/science.1115248

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. doi:10.1109/iccv48922.2021.00986

Liu, Z., Mao, H., Wu, C-Y., Feichtenhofer, C., Darrell, T., and Xie, S. (2022). A ConvNet for the 2020s. arXiv preprint arXiv:2201.03545.

Musadiq, M., Patsoura, E., Hughes, S., and Yang, Y. C. (2003). Measurements of linear dimensions on fundus photographs: comparison between photographic film and digital systems. *Eye* 17 (5), 619–622. doi:10.1038/sj.eye.6700418

Nathan, D. M. (1993). Long-term complications of diabetes mellitus. *N. Engl. J. Med.* 328 (23), 1676–1685. doi:10.1056/nejm199306103282306

Peng, Z., Huang, W., Gu, S., Xie, L., Wang, Y., Jiao, J., et al. (2021). "Conformer: Local features coupling global representations for visual recognition," in Proceedings of the IEEE/CVF International Conference on Computer Vision, 367–376. doi:10.1109/iccv48922.2021.00042

Selçuk, T., Beyoğlu, A., and Alkan, A. (2022). Automatic detection of exudates and hemorrhages in low-contrast color fundus images using multi semantic convolutional neural network. *Concurrency Comput. Pract. Exp.* 34 (6), e6768. doi:10.1002/cpe.6768

Shahhosseini, M., Hu, G., Khaki, S., and Archontoulis, S. V. (2021). Corn yield prediction with ensemble CNN-DNN. *Front. Plant Sci.* 12, 709008. doi:10.3389/fpls.2021.709008

Sikder, N., Masud, M., Bairagi, A. K., Arif, A. S. M., Nahid, A.-A., and Alhumyani, H. A. (2021). Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images. *Symmetry* 13 (4), 670. doi:10.3390/sym13040670

Simonyan, K., and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Stitt, A. W., Curtis, T. M., Chen, M., Medina, R. J., McKay, G. J., Jenkins, A., et al. (2016). The progress in understanding and treatment of diabetic retinopathy. *Prog. Retin. Eye Res.* 51, 156–186. doi:10.1016/j.preteyeres.2015.08.001

Tolstikhin, I. O., Houlsby, N., Kolesnikov, A., Beyer, L., Zhai, X., Unterthiner, T., et al. (2021). Mlp-mixer: An all-mlp architecture for vision. *Adv. Neural Inf. Process. Syst.* 34, 24261–24272.

Valan, M., Makonyi, K., Maki, A., Vondráček, D., and Ronquist, F. (2019). Automated taxonomic identification of insects with expert-level accuracy using effective feature transfer from convolutional networks. *Syst. Biol.* 68 (6), 876–895. doi:10.1093/sysbio/syz014

Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., et al. (2020). Deep high-resolution representation learning for visual recognition. *IEEE Trans. Pattern Analysis Mach. Intell.* 43 (10), 3349–3364.

Wong, T. Y., Klein, R., Klein, B. E. K., Tielsch, J. M., Hubbard, L., and Nieto, F. J. (2001). Retinal microvascular abnormalities and their relationship with hypertension, cardiovascular disease, and mortality. *Surv. Ophthalmol.* 46 (1), 59–80. doi:10.1016/s0039-6257(01)00234-x

Xu, F., Wan, C., Zhao, L., You, Q., Xiang, Y., Zhou, L., et al. (2021). Predicting central serous chorioretinopathy recurrence using machine learning. *Front. Physiol.* 12, 649316. doi:10.3389/fphys.2021.649316

Yannuzzi, L. A., Ober, M. D., Slakter, J. S., Spaide, R. F., Fisher, Y. L., Flower, R. W., et al. (2004). Ophthalmic fundus imaging: today and beyond. *Am. J. Ophthalmol.* 137 (3), 511–524. doi:10.1016/j.ajo.2003.12.035

Yuan, L., Chen, Y., Wang, T., Yu, W., Shi, Y., Jiang, Z., et al. (2021). "Tokens-to-token vit: Training vision transformers from scratch on imagenet," in Proceedings of the IEEE/CVF International Conference on Computer Vision. doi:10.1109/iccv48922.2021.00060

Zarbin, M. A., Casaroli-Marano, R. P., and Rosenfeld, P. J. (2014). Age-related macular degeneration: clinical findings, histopathology and imaging techniques. *Cell-Based Ther. Retin. Degener. Dis.* 53, 1–32. doi:10.1159/000358536

Zhang, X., Zhou, X., Lin, M., and Sun, J. (2018). "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in Proceedings of the IEEE conference on computer vision and pattern recognition, 6848–6856. doi:10.1109/cvpr.2018.00716

Zhang, Z., Gong, Z., Hong, Q., and Jiang, L. (2021). "Swin-transformer based classification for rice diseases recognition," in 2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI), 153–156 (IEEE). doi:10.1109/cisai54367.2021.00036

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 10 August 2022
DOI 10.3389/fphys.2022.946099

LAD-GCN: Automatic diagnostic framework for quantitative estimation of growth patterns during clinical evaluation of lung adenocarcinoma

Wei Xiao[1]*, Yanyun Jiang[2], Zhigang Yao[3], Xiaoming Zhou[1], Xiaodan Sui[2] and Yuanjie Zheng[2]

[1]Shandong Provincial Hospital, Shandong University, Jinan, China, [2]School of Information Science and Engineering, Shandong Normal University, Jinan, China, [3]Department of Pathology, Shandong Provincial Hospital Affiliated to Shandong First Medical University Department of Pathology, Shandong Provincial Hospital, Cheeloo College of Medicine, Shandong University, Jinan, China

Check for updates

OPEN ACCESS

EDITED BY
Zhibin Niu,
Tianjin University, China

REVIEWED BY
Shujun Liang,
Southern Medical University, China
Jian Lian,
Shandong Women's University, China
Jiquan Ma,
Heilongjiang University, China
Aimei Dong,
Qilu University of Technology, China

*CORRESPONDENCE
Wei Xiao,
xweidoc@163.com

SPECIALTY SECTION
This article was submitted to Medical Physics and Imaging,
a section of the journal
Frontiers in Physiology

RECEIVED 17 May 2022
ACCEPTED 04 July 2022
PUBLISHED 10 August 2022

CITATION
Xiao W, Jiang Y, Yao Z, Zhou X, Sui X and Zheng Y (2022), LAD-GCN: Automatic diagnostic framework for quantitative estimation of growth patterns during clinical evaluation of lung adenocarcinoma.
Front. Physiol. 13:946099.
doi: 10.3389/fphys.2022.946099

Quantitative estimation of growth patterns is important for diagnosis of lung adenocarcinoma and prediction of prognosis. However, the growth patterns of lung adenocarcinoma tissue are very dependent on the spatial organization of cells. Deep learning for lung tumor histopathological image analysis often uses convolutional neural networks to automatically extract features, ignoring this spatial relationship. In this paper, a novel fully automated framework is proposed for growth pattern evaluation in lung adenocarcinoma. Specifically, the proposed method uses graph convolutional networks to extract cell structural features; that is, cells are extracted and graph structures are constructed based on histopathological image data without graph structure. A deep neural network is then used to extract the global semantic features of histopathological images to complement the cell structural features obtained in the previous step. Finally, the structural features and semantic features are fused to achieve growth pattern prediction. Experimental studies on several datasets validate our design, demonstrating that methods based on the spatial organization of cells are appropriate for the analysis of growth patterns.

KEYWORDS

lung adenocarcinoma, histopathological, deep learning, polar representation-based model, graph convolutional networks, feature fusion

1 Introduction

Lung cancer is a malignant tumor originating from the bronchial mucosa or glands of the lungs that poses a great threat to human health and life. In recent years, many countries have reported significant increases in rates of lung cancer; in men, lung cancer has the highest morbidity and mortality among all malignant tumors (Ferlay et al., 2020). The 5-year survival rate of patients with lung cancer is relatively low at 19%, mainly owing to the high risk of distant metastasis (Hirsch et al., 2017; Mayekar and Bivona, 2017). Adenocarcinoma is the most common histopathological type of lung cancer and accounts for up to 40% of lung cancer cases (Cheng et al., 2016).

According to the 2011 IASLC/ATS/ERS lung adenocarcinoma classification, lung adenocarcinoma has five predominant growth patterns: lepidic, papillary, acinar, micropapillary, and solid (Travis et al., 2011). Accurate determination of growth patterns and proportions from whole-slide images (WSIs) has crucial clinical implications for diagnosis, subsequent treatment, and prognosis. The World Health Organization recommends that invasive adenocarcinoma should be semi-quantitatively estimated in 5% increments with respect to the various growth patterns on histopathological slides (Travis et al., 2015). Typically, this work involves visual inspection by experienced pathologists through a microscope, which is a very time-consuming and labor-intensive process (Gurcan et al., 2009). In particular, semi-quantitative estimates of growth patterns are required; however, traditional methods only allow estimation and not quantification. WSI technology and computer-aided diagnosis provide an effective strategy for lung adenocarcinoma diagnosis, which can be used as an auxiliary basis for manual evaluation and to alleviate the shortage of pathologists. In this work, we focus on the identification and quantification of lung adenocarcinoma tissue growth patterns from WSIs. This is expected to help pathologists to make rapid diagnoses in practical clinical applications and provide a basis for subsequent treatment. However, identifying growth patterns is challenging because of the high intraclass variation and low interclass distinction among patterns. As shown in, Figure 1, the spatial structure of lung cancer cells has complicated characteristic manifestations; for instance, cells in the lepidic growth pattern grow along alveolar walls in a lepidic fashion, and the acinar growth pattern has well-defined individual tumor glands with well-formed glandular lumina.

Recently, owing to the enormous potential of deep learning, many convolutional neural network (CNN)-based methods have emerged that can automatically extract more beneficial features for classification compared with hand-crafted features (Szegedy

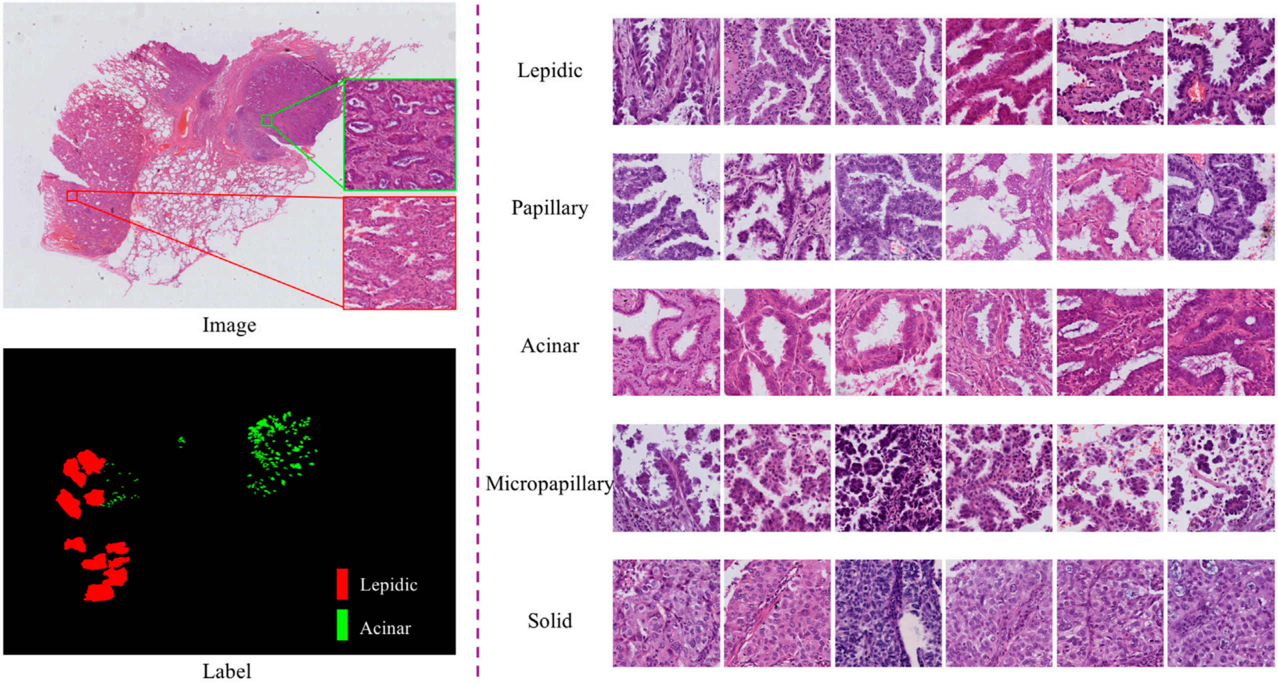


FIGURE 1
Pictures of typical growth patterns. Left: An example of an HandE-stained digital pathology image with manual segmentation of growth patterns, where red is the lepidic growth pattern and green is the acinar growth pattern. Right: Five common growth patterns (lepidic, papillary, acinar, micropapillary, and solid).

et al., 2016; Coudray et al., 2018; Šarić et al., 2019; Noorbakhsh et al., 2020; Yu et al., 2020; Fan et al., 2021). For example, Coudray et al. (2018) used an Inception v3 architecture (Szegedy et al., 2016) to learn a parametric function to automatically classify lung tumor subtypes (adenocarcinoma and squamous cell carcinoma) and predict mutations using a dataset from The Cancer Genome Atlas. Yu et al. (2020) built a CNN model to identify tumor regions from whole-slide histopathology images, achieving an area under the curve value (AUC) > 0.935, and used the proposed model to predict pathologists' diagnoses. Kalra et al. (2020) presented a memory-based exchangeable model that could learn interdependencies among instances through a self-attention mechanism, achieving a competitive accuracy of 84.84% for the classification of lung adenocarcinoma and squamous cell carcinoma. Although CNN can automatically encode rich semantic features contained in captured images, the analysis of histopathological images often focuses on local features, which leads to biased learning. The main reason for this is that the interclass differences among tissue growth patterns are small, that is, cancer cells with different tissue growth patterns show no obvious visual differences. Furthermore, intraclass differences in tissue growth patterns are evident, and models need to cope with altered tumor tissue and cell diversity. Therefore, as well as understanding the semantic features of an image, an algorithm designed for classification of lung tumor histopathological images needs to be able to analyze the spatial structure between cancer cells. Graph convolutional networks (GCN) (Kipf and Welling, 2016), variants of CNN that transform data into spatially structured features, have recently become a popular choice for processing structured data in the field of computer vision. Some studies have proposed the use of GCN to analyze histopathological images. For example, Li et al. (2018) proposed a GCN for WSIs, which obtains graph nodes by sampling representative patches and extracting features for survival prediction. Adnan et al. (2020) introduced attention through graph pooling to infer relations among sampled patches and applied multiple instance learning to classify lung cancer subtypes. Zhou et al. (2019) developed CGC-Net, which converts a large histology image into a graph, where each graph node is a nucleus and the connecting edges of the nodes represent the similarity between nuclei. Wang et al. (2020) also segmented the cell nucleus and extracted topological composition graphs for tumor microenvironment analysis in renal cell carcinoma and patient outcome prediction; however, they did not use GCN.

Inspired by the above methods, in this work we designed a novel deep learning framework, called LAD-GCN (lung adenocarcinoma diagnosis GCN), which aggregates the advantages of GCN and CNN for analyzing histopathology. Specifically, to capture complex tumor microenvironment information and semantic information of entire image patches, we designed a model with two independent feature extraction branches as follows. 1) The GCN module, including a polar representation-based instance segmentation model (Xiao et al., 2021), is used to extract all the cell nuclei contained in the histopathological patch and extract a nuclear feature composition map, which is used as an input to the GCN network to extract cell structural features. 2) The CNN module directly extracts semantic information from the whole patch to supplement the information loss of the GCN module. Then, the cell structural features extracted by the GCN branch and the image patch semantic features extracted by the CNN branch are fused. Compared with the CNN-only models that are widely used in image classification tasks, LAD-GCN could provide complementary semantic and cell structural information during feature extraction. Finally, we quantitatively evaluated the proposed method on a private dataset of lung adenocarcinoma postoperative formalin-fixed, paraffin-embedded (FFPE) tissue slides. The results demonstrate that our method is able to capture features that are beneficial for growth pattern typing. Our major contributions can be summarized as follows.

1. In response to the problem of the small interclass differences in tissue growth patterns that mean there are no obvious visual differences among cancer cells with different growth patterns, we developed a novel GCN-based framework for analysis of the histopathological growth patterns of lung adenocarcinoma. The proposed method adopts a polar representation-based instance segmentation model to segment the nucleus and uses GCN to extract cell spatial structural features.
2. To overcome the limitations of a single feature extraction module, we designed a dual-network joint analysis method: the GCN branch extracts the spatial structural features of cells, while the CNN branch complements these with the extraction of semantic features of patches.
3. We validated the proposed method on a private lung adenocarcinoma WSI dataset, demonstrating the effectiveness of the architecture.

2 Materials and methods

2.1 Materials

Our histopathological image dataset contained data obtained from 243 lung adenocarcinoma patients at Shandong Provincial Hospital; for each patient, there was one FFPE image of the tumor area, stained with hematoxylin and eosin (HandE) and scanned at 20× and 40× magnification with a pixel scale of 0.23 μm × 0.23 μm. All samples represented postoperative pathology, including tumor tissue slides, normal tissue slides, and slides containing the border between normal and tumor tissue. In this dataset, all data were positive samples, that is, slides containing tumor tissue. The tumor/non-tumor area and five histological patterns (lepidic, acinar, papillary, micropapillary,

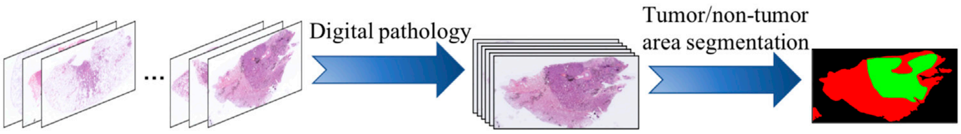


FIGURE 2
Schematic representation of data processing; we used U-Net as the backbone to segment tumor regions.

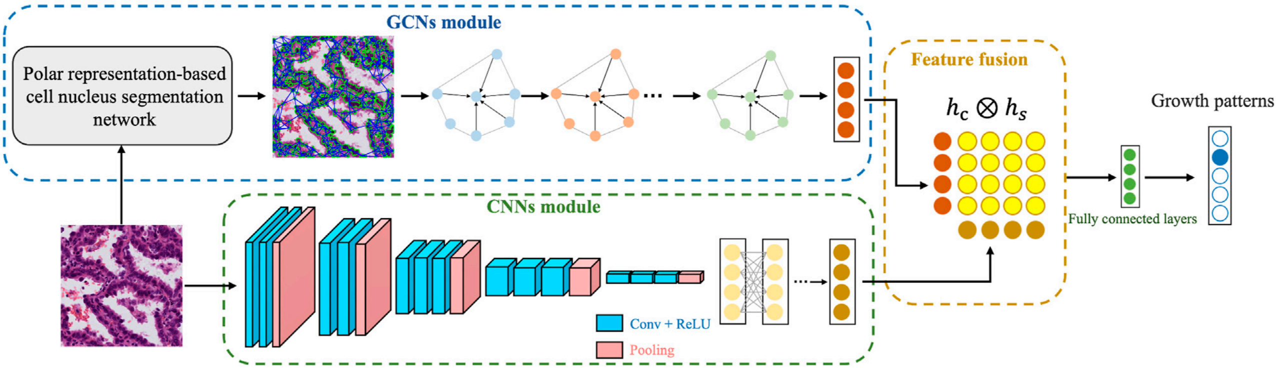


FIGURE 3
Overview of the proposed diagnostic methods for lung adenocarcinoma.

and solid) were manually delineated by an experienced oncology pathologist.

To make the algorithm more effective, we built a segmentation model based on U-Net (Ronneberger et al., 2015) to achieve tumor region extraction; the process is shown in Figure 2. Specifically, for the images in the dataset, both tumor and non-tumor regions were manually annotated by pathologists. We derived 2× magnification WSIs, which were full-coverage images, and trained the U-Net backbone in a traditional fully supervised manner with a cross-entropy loss function (Xie and Tu, 2015) to predict tumor regions in the WSIs. It is worth noting that the size of the original pathological images was non-uniform, and the model was able to achieve region prediction in images of any size.

2.2 Overview of the LAD-GCN architecture

Figure 3 provides an overview of our automatic diagnosis framework. As shown in the figure, instead of directly extracting features using a CNN, we developed both CNN and GCN feature extractors simultaneously. The inputs are patches from digitized postoperative FFPE tissue slides, and the output is the predicted growth pattern type. The whole process consists of three parts. 1) GCN module: a polar representation-based instance segmentation model is used to segment all the cell nuclei contained in the histopathological patch; the nuclear features are extracted to form a composite map that can be used as the input to the GCN; and then GCN are used to extract cell spatial structure features. 2) CNN module: semantic feature extraction is performed using a CNN, VGG16 (Simonyan and Zisserman, 2015). 3) Feature fusion: cell spatial structural features and semantic features are fused for tumor growth pattern prediction.

2.3 Spatial feature encoding with GCN

In histopathology images, each cell has its own characteristic information, and there is structural information between cells. To extract this information, we segment out the nuclei and construct a graph of the tumor microenvironment for graph convolution operations. Specifically, we first extract all the nuclei contained in the patch and calculate the centroids of the nuclei to define the graph node set *V*; then extract the nuclei features, use K-nearest neighbors (KNN) (Muja and Lowe, 2009) to find the connections between adjacent cells to define the edge set *A* (Chen et al., 2020);

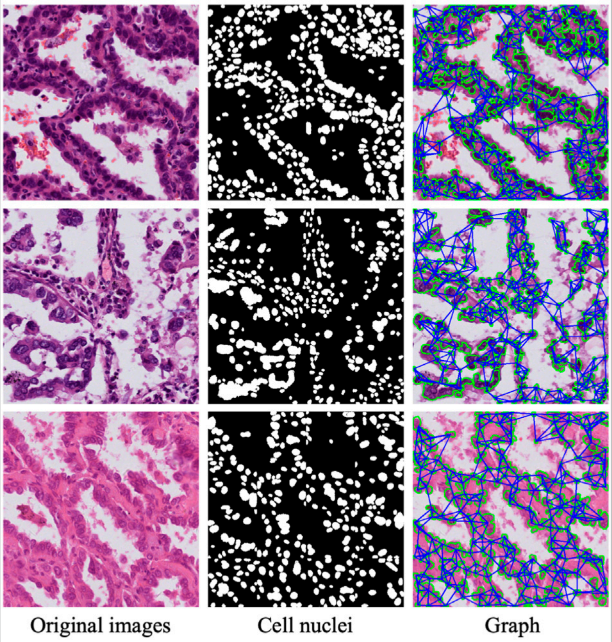


FIGURE 4
Three sample patches of cell nuclei structures. First column: typical patches from lung tumor histopathological images. Second column: nuclei segmentation mask from nuclei segmentation module. Third column: graph nodes and edges.

and, finally, use GCN to learn the graph depth spatial structural features.

2.3.1 Nuclei segmentation module

The major purpose of the nuclei segmentation module is to extract the various nuclei contained in the input image patch, which includes normal cells, tumor cells, and stromal cells. The cell nuclei produced by the nuclei segmentation module are then constructed as a graph and fed into a GCN module for spatial feature extraction. To achieve this aim, we use a polar representation-based instance segmentation model (Xiao et al., 2021) from our previous work to learn the segmentation of nuclei; this model leverages fully convolutional one-stage object detection and consists of a backbone network, feature pyramid network, and task-specific heads. Specifically, when we input an original image via the proposed network, the position of the cell center point and the distance of n ($n = 36$) root rays can be obtained; then, the coordinates of these points on the contour are calculated according to the angle and length, connecting these points starting from 0°; and, finally, the regions within the connected regions are taken as the results of instance segmentation. The nuclei segmentation module models a contour based on the polar coordinate system and transforms the instance segmentation problem into an instance center classification problem and a dense distance regression problem (Xie et al., 2020); thus, the network only needs to return to the length of the fixed angle, which reduces the difficulty of the problem. Through the prediction of the segmentation module, we obtain the mask of the nuclei, and the second column in Figure 4 shows the result of the nuclei segmentation.

2.3.2 Cell feature extraction and construct graph

A feature matrix for graph convolution is generated based on the nuclear segmentation map generated in the previous step. It is used for two main processes: cell feature extraction and graph construction. In the first of these processes, the PyRadiomics package (Van Griethuysen et al., 2017) in Python is used to generate features corresponding to each cell, including eight shape features and four textural features. The shape features include major axis length, minor axis length, angular orientation, eccentricity, roundness, area, and solidity. The textural features, obtained from gray-level co-occurrence matrices, are dissimilarity, homogeneity, angular second moment, and energy. In addition, we use contrastive predictive coding (Henaff, 2020) to encode features of 64 × 64 image patch regions centered on the centroids of cell nuclei.

In the second process, we connect the nuclei into a graph, using the centroid of each nucleus as a graph node, and use the KNN algorithm to build the edge set A of the graph. Specifically, in principle, each nucleus should have contact with the other nuclei, and the nearest neighbor cells are considered to have obvious intercellular interactions. The adjacency matrix is defined as:

$$a_{ij} = \begin{cases} 1, \text{if } j \in \text{KNN}(i) \text{ and } D(i,j) < d \\ 0, \text{otherwise,} \end{cases} \tag{1}$$

where $j \in \text{KNN}(i)$ denotes the K instances closest to instance i. In this work, we set $K = 5$. $D\ (i, j)$ indicates the Euclidean distance between two nucleus instances. Thus, we obtain the input for the GCN, the set of nodes and edges $G = (V, A)$. Figure 4 shows the nuclear segmentation results and the graph structure constructed based on these results for three sample patches.

2.3.3 GCN module

The cell structural information used to construct a graph is very suitable for GCN-based feature extraction. To simplify the operation, we use a spatial-based GCN, where the convolution operation is defined as:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^{(l)}W^{(l)}\right), \tag{2}$$

in which $H^{(l)} \in \mathbb{R}^{m \times k^{(l)}}$ denotes the $k^{(l)}$-channel features at the lth layer, $\sigma(.)$ denotes an activation function, and $W^{(l)}$ is the trainable weight matrix of each layer. $\tilde{A} = A + I_N$ is the adjacency matrix of the undirected graph with added self-connection, and I_N is the identity matrix.

TABLE 1 Effects of each module in our LAD-GCN design. Bold font indicates best result obtained for predictions.

Growth pattern	Methods	CNN	GCN	P (%)	R (%)	F1S (%)
Lepidic	CNN module	✓		**90.64**	87.76	89.18
	GCN module		✓	84.55	78.55	81.44
	LAD-GCN	✓	✓	90.26	**89.83**	**90.04**
Acinar	CNN module	✓		**89.79**	87.24	88.50
	GCN module		✓	83.61	85.98	84.78
	LAD-GCN	✓	✓	89.27	**90.39**	**89.83**
Papillary	CNN module	✓		85.17	79.78	82.39
	GCN module		✓	83.33	81.52	82.42
	LAD-GCN	✓	✓	**87.15**	**85.80**	**86.47**
Micropapillary	CNN module	✓		79.56	**89.44**	84.21
	GCN module		✓	85.63	89.03	**87.30**
	LAD-GCN	✓	✓	**86.53**	86.80	86.67
Solid	CNN module	✓		98.17	98.02	98.10
	GCN module		✓	95.85	97.33	97.07
	LAD-GCN	✓	✓	**98.34**	**98.78**	**98.56**

TABLE 2 Comparison of the performances of each module in terms of accuracy.

Methods	Accuracy (%)
CNN module	88.49
GCN module	86.71
LAD-GCN	**90.35**

2.4 Semantic feature encoding with CNN

The purpose of semantic feature encoding is to encode the semantic features of an entire image patch, which can be used to mine the overall information contained in the image block and as supplementary information regarding the spatial structure of cells. In this work, we apply VGG16 (Simonyan and Zisserman, 2015) as a feature extractor for the CNN, which consist of a convolutional layer, ReLU, max pooling, and fully connected layer. The model structure is shown as the "CNN module" in Figure 3. The size of the input feature map for each pooling operation is 1/2 that of the previous layer. Finally, the semantic feature encoding the CNN module outputs a feature vector with a length of 1,024.

2.5 Multimodal tensor fusion

In many approaches, the features extracted by multiple networks are superimposed onto a set of features through a concatenation operation, followed by convolution operations. However, such approaches are only suitable for extraction of features of the same type. Here, the image semantic features are extracted as a set of feature maps by CNN, and cell structure features are extracted as nodes and edges by GCN. To integrate these multimodal features, we recommend using kronecker product (Zadeh et al., 2017), a feature fusion method for modeling multi-feature interaction. The fusion module combines multimodal tensors through outer product calculation, which can be formulated as:

$$\mathbf{h}_{\text{fusion}} = \mathbf{h}_{\text{c}} \otimes \mathbf{h}_{\text{s}}, \tag{3}$$

where ⊗ is the outer product; $\mathbf{h}_c$ and $\mathbf{h}_s$ denote cell graph and semantic features, respectively, and $\mathbf{h}_{\text{fusion}}$ is a differential multimodal tensor formed in a three-dimensional Cartesian space. After aggregating the multimodal tensors, we use a fully connected layer for the next feature operation. In addition, we adopt a gating-based attention mechanism (Arevalo et al., 2017) to limit the fusion proportion of different modal features. For each modality feature $\mathbf{h}_{\text{m}}$, $\forall m \in \{c, s\}$, we learn a linear transformation $W_{cs\to m}$ for the weight matrix parameters. The importance score of each feature is defined as $\mathbf{z}_m = \sigma(W_{cs\to m} \cdot [\mathbf{h}_c, \mathbf{h}_s])$. Subsequently, the gated representation $\mathbf{h}_{m,\text{ gated}}$ can be calculated as:

$$\mathbf{h}_{m,\text{gated}} = \mathbf{z}_m * \mathbf{h}_m, \forall m \in \{c, s\}, \tag{4}$$

where $\mathbf{h}_m = ReLU(W_m \cdot \mathbf{h}_m)$ denotes the feature after activation by ReLU. Through this gated attention mechanism, the expressive ability of each modality feature can be controlled, and the size of the feature space is also reduced before feature fusion.

2.6 Loss function

The loss function for LAD-GCN is the standard cross-entropy loss:

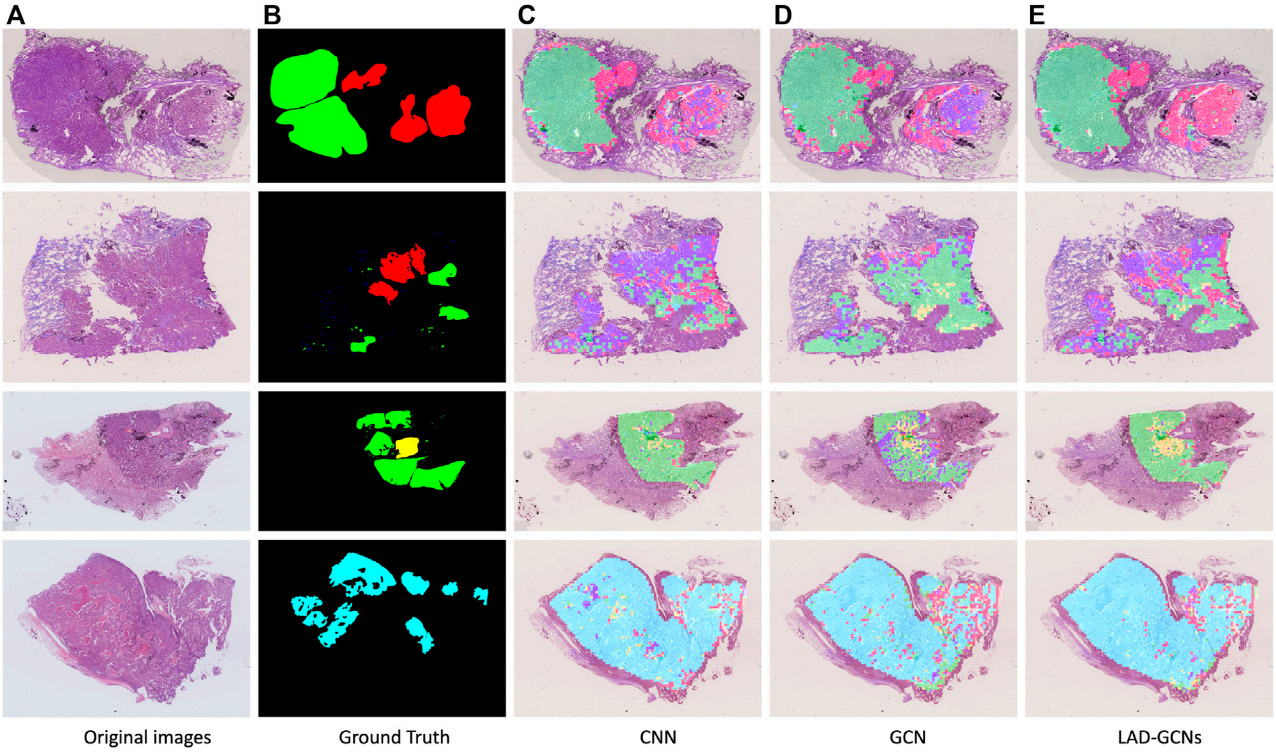


FIGURE 5
Typical histopathological image analysis results obtained with three networks. Red, green, yellow, blue, and cyan masks represent lepidic, acinar, papillary, micropapillary, and solid growth patterns, respectively. **(A)** Example of histopathological images. **(B)** Ground truth by pathologists. **(C,D,E)** Typical histopathological image analysis results obtained with three networks: **(C)** CNN, **(D)** GCN and **(E)** LAD-GCNs. Red, green, yellow, blue, and cyan masks represent lepidic, acinar, papillary, micropapillary, and solid growth patterns, respectively.

$$\min_{P} L = -\sum_{b=1}^{B} p_b \log q_b, \tag{5}$$

where A represents the total number of image patches, and p_b and $q_b = p_{model}\ (y|b)$ indicate the target labels and the predicted class distribution produced by the model for input b, respectively. The whole training process of the network is performed in an end-to-end manner.

3 Experiments and results

3.1 Implementation details

For the GCN module, we first use a pre-trained nuclei segmentation module to extract various nuclei from each patch, and train the GCN model classifiers. We train the CNN module and GCN module using the Adam optimizer with an initial learning rate of 0.001 and a batch size of 64. For the proposed LAD-GCN, the trained GCN module and CNN module are then fine-tuned for 100 epochs with a learning rate of 0.00001. All our modules were implemented with PyTorch and trained on four NVIDIA Tesla A100 GPUs.

3.2 Evaluation metrics

We employed four metrics for performance evaluation of the baseline classification model, GCN module, CNN module, and the proposed LAD-GCN: precision, recall, F1-score, and accuracy. These performance metrics can be understood by considering four terms: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The precision (P) and recall (R) were defined as:

$$P = \frac{TP}{TP + FP}, \tag{6}$$

and

$$R = \frac{TP}{TP + FN}. \tag{7}$$

We further measured the F1 score (F1S), which combines precision and recall, defined as:

$$F1S = \frac{2TP}{2TP + FP + FN}. \tag{8}$$

In addition, we calculated the accuracy of five growth patterns, defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{9}$$

3.3 Ablation study

To verify the effectiveness of the CNN module and GCN module in lung tumor histopathological image analysis, we conducted an ablation study. The results are shown in Table 1 and Table 2. The CNN module demonstrated greater ability in the analysis of lepidic and acinar growth patterns, whereas the GCN module could better capture the micropapillary structure. Both modules worked well for identifying solid growth patterns, possibly because the tumor cells were densely packed and lacked characteristic patterns of adenocarcinoma. The proposed model (LAD-GCN) fuses semantic features and spatial features. Although it could not achieve optimal results in the analysis of every growth mode, its performance was stable. It achieved an accuracy of 90.35%, which was more than 1.8% better than that of the CNN module, and more than 3.6% better than that of the GCN module.

Figure 5 shows the results of histopathological image analysis of four images, which included lepidic, acinar, papillary, micropapillary, and solid growth patterns. The GCN and CNN modules produced very similar masks to the manual ground truth; however, LAD-GCN could still provide a subtle improvement. As shown in the figure, the areas predicted by the deep learning model were often larger than those obtained by manual labeling; this was because manual annotation focused on regions typical of particular growth patterns, whereas the trained deep learning model could predict both typical and atypical growth pattern regions. Pathologists perform semi-quantitative assessments of growth patterns when analyzing histopathological images of lung adenocarcinomas. This process is very dependent on the subjective evaluation of individuals and is difficult to quantify. The trained model could predict the type of growth pattern for each small patch region, enabling quantification of types across the entire histopathological image.

4 Discussion and conclusion

In this study, we proposed the LAD-GCN framework, which consists of a GCN module and a CNN module, for the task of lung adenocarcinoma growth pattern prediction. The GCN module captures the spatial structural features between cells, whereas the CNN module captures semantic features of whole patches; these features can be fused to predict growth patterns. In particular, our proposed model showed significantly enhanced performance in lung adenocarcinoma growth pattern prediction tasks. In the future, our goal is to combine image analysis with patient medical records to predict medication and prognostic status, and to apply our LAD-GCN framework to other histopathological WSI analysis tasks such as images of breast, kidney, and brain.

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 Shandong Provincial Hospital Affiliated to Shandong First Medical University. The patients/participants provided their written informed consent to participate in this study.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work. WX and YZ developed the study concept, and discussed with the other authors. ZY and XZ performed the data analysis. WX, XS and YJ completed the method and experiment, and drafted the manuscript. All authors critically revised the paper for important intellectual content. All authors have read and agreed to the published version of the manuscript.

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

References

Adnan, M., Kalra, S., and Tizhoosh, H. R. (2020). "Representation learning of histopathology images using graph neural networks," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops (Seattle, WA: IEEE), 988–989.

Arevalo, J., Solorio, T., Montes-y Gómez, M., and González, F. A. (2017). *Gated multimodal units for information fusion*. Toulon, France: arXiv preprint arXiv:1702.01992.

Chen, R. J., Lu, M. Y., Wang, J., Williamson, D. F., Rodig, S. J., Lindeman, N. I., et al. (2020). Pathomic fusion: An integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Trans. Med. Imaging* 41, 757–770. doi:10.1109/TMI.2020.3021387

Cheng, T.-Y. D., Cramb, S. M., Baade, P. D., Youlden, D. R., Nwogu, C., Reid, M. E., et al. (2016). The international epidemiology of lung cancer: Latest trends, disparities, and tumor characteristics. *J. Thorac. Oncol.* 11, 1653–1671. doi:10.1016/j.jtho.2016.05.021

Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., et al. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* 24, 1559–1567. doi:10.1038/s41591-018-0177-5

Fan, L., Sowmya, A., Meijering, E., and Song, Y. (2021). "Learning visual features by colorization for slide-consistent survival prediction from whole slide images," in International conference on medical image computing and computer-assisted intervention (Strasbourg, France: Springer), 592–601.

Ferlay, J., Ervik, M., Lam, F., Colombet, M., Mery, L., Piñeros, M., et al. (2020). *Global cancer observatory: Cancer today*. Lyon, France: International agency for research on cancer, 1–6.

Gurcan, M., Boucheron, L., Can, A., Madabhushi, A., Rajpoot, N., Yener, B., et al. (2009). Histopathological image analysis: A review. *IEEE Rev. Biomed. Eng.* 2, 147–171. doi:10.1109/RBME.2009.2034865

Henaff, O. (2020). "Data-efficient image recognition with contrastive predictive coding," in International conference on machine learning (Vienna, Austria: PMLR), 4182–4192.

Hirsch, F. R., Scagliotti, G. V., Mulshine, J. L., Kwon, R., Curran, W. J., Jr, Wu, Y.-L., et al. (2017). Lung cancer: Current therapies and new targeted treatments. *Lancet* 389, 299–311. doi:10.1016/S0140-6736(16)30958-8

Kalra, S., Adnan, M., Taylor, G., and Tizhoosh, H. R. (2020). "Learning permutation invariant representations using memory networks," in European conference on computer vision (Glasgow, United Kingdom: Springer), 677–693.

Kipf, T. N., and Welling, M. (2016). "Semi-supervised classification with graph convolutional networks," in 5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24–26, 2017. Conference Track Proceedings.: arXiv preprint arXiv:1609.02907.

Li, R., Yao, J., Zhu, X., Li, Y., and Huang, J. (2018). "Graph cnn for survival analysis on whole slide pathological images," in International conference on medical image computing and computer-assisted intervention (Granada, Spain: Springer), 174–182.

Mayekar, M. K., and Bivona, T. G. (2017). Current landscape of targeted therapy in lung cancer. *Clin. Pharmacol. Ther.* 102, 757–764. doi:10.1002/cpt.810

Muja, M., and Lowe, D. G. (2009). Fast approximate nearest neighbors with automatic algorithm configuration. *VISAPP* 2 (1), 2.

Noorbakhsh, J., Farahmand, S., Namburi, S., Caruana, D., Rimm, D., Soltanieh-ha, M., et al. (2020). Deep learning-based cross-classifications reveal conserved spatial behaviors within tumor histological images. *Nat. Commun.* 11, 6367. doi:10.1038/s41467-020-20030-5

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation," in International Conference on Medical image computing and computer-assisted intervention (Munich, Germany: Springer), 234–241.

Šarić, M., Russo, M., Stella, M., and Sikora, M. (2019). "Cnn-based method for lung cancer detection in whole slide histopathology images," in 2019 4th international conference on smart and sustainable technologies (SpliTech) (Bol and Split, Croatia: IEEE), 1–4.

Simonyan, K., and Zisserman, A. (2015). "Very deep convolutional networks for large-scale image recognition," in 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, May 7–9, 2015. Conference Track Proceedings.: arXiv preprint arXiv:1409.1556.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). "Rethinking the inception architecture for computer vision," in 2016 IEEE/CVF conference on computer vision and pattern recognition (CVPR) (Las Vegas, NV: IEEE), 2818–2826.

Travis, W. D., Brambilla, E., Nicholson, A. G., Yatabe, Y., Austin, J. H., Beasley, M. B., et al. (2015). The 2015 world health organization classification of lung tumors: Impact of genetic, clinical and radiologic advances since the 2004 classification. *J. Thorac. Oncol.* 10, 1243–1260. doi:10.1097/JTO.0000000000000630

Travis, W. D., Brambilla, E., Noguchi, M., Nicholson, A. G., Geisinger, K. R., Yatabe, Y., et al. (2011). International association for the study of lung cancer/american thoracic society/european respiratory society international multidisciplinary classification of lung adenocarcinoma. *J. Thorac. Oncol.* 6, 244–285. doi:10.1097/JTO.0b013e318206a221

Van Griethuysen, J. J., 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, S., Rong, R., Yang, D. M., Fujimoto, J., Yan, S., Cai, L., et al. (2020). Computational staining of pathology images to study the tumor microenvironment in lung cancer. *Cancer Res.* 80, 2056–2066. doi:10.1158/0008-5472.CAN-19-1629

Xiao, W., Jiang, Y., Yao, Z., Zhou, X., Lian, J., Zheng, Y., et al. (2021). Polar representation-based cell nucleus segmentation in non-small cell lung cancer histopathological images. *Biomed. Signal Process. Control* 70, 103028. doi:10.1016/j.bspc.2021.103028

Xie, E., Sun, P., Song, X., Wang, W., Liu, X., Liang, D., et al. (2020). "Polarmask: Single shot instance segmentation with polar representation," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (Seattle, WA: IEEE), 12193–12202.

Xie, S., and Tu, Z. (2015). "Holistically-nested edge detection," in Proceedings of the IEEE international conference on computer vision (Santiago, Chile: IEEE), 1395–1403.

Yu, K.-H., Wang, F., Berry, G. J., Ré, C., Altman, R. B., Snyder, M., et al. (2020). Classifying non-small cell lung cancer types and transcriptomic subtypes using convolutional neural networks. *J. Am. Med. Inf. Assoc.* 27, 757–769. doi:10.1093/jamia/ocz230

Zadeh, A., Chen, M., Poria, S., Cambria, E., and Morency, L.-P. (2017). "Tensor fusion network for multimodal sentiment analysis," in 2017 conference on empirical methods in natural language processing (Copenhagen, Denmark: Association for Computational Linguistics).

Zhou, Y., Graham, S., Alemi Koohbanani, N., Shaban, M., Heng, P.-A., and Rajpoot, N. (2019). "Cgc-net: Cell graph convolutional network for grading of colorectal cancer histology images," in Proceedings of the IEEE/CVF international conference on computer vision workshops (Seoul, Korea: IEEE).

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 12 August 2022
DOI 10.3389/fphys.2022.928969

OPEN ACCESS

EDITED BY
Yi Xu,
Shanghai Jiao Tong University, China

REVIEWED BY
Hongzan Sun,
Shengjing Hospital of China Medical University, China
Shengping Wang,
Fudan University, China
Shengdong Nie,
University of Shanghai for Science and Technology, China

*CORRESPONDENCE
Zheng Yuan,
yuanzheng0404@163.com
Xiaoxiao Li,
jiayouxiaoxiao@163.com

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

SPECIALTY SECTION
This article was submitted to Medical Physics and Imaging,
a section of the journal
Frontiers in Physiology

RECEIVED 26 April 2022
ACCEPTED 13 July 2022
PUBLISHED 12 August 2022

CITATION
Shi H, Duan Y, Shi J, Zhang W, Liu W, Shen B, Liu F, Mei X, Li X and Yuan Z (2022), Role of preoperative prediction of microvascular invasion in hepatocellular carcinoma based on the texture of FDG PET image: A comparison of quantitative metabolic parameters and MRI.
Front. Physiol. 13:928969.
doi: 10.3389/fphys.2022.928969

Role of preoperative prediction of microvascular invasion in hepatocellular carcinoma based on the texture of FDG PET image: A comparison of quantitative metabolic parameters and MRI

Huazheng Shi[1†], Ying Duan[1†], Jie Shi[2†], Wenrui Zhang[1], Weiran Liu[1], Bixia Shen[1], Fufu Liu[1], Xin Mei[1], Xiaoxiao Li[1]* and Zheng Yuan[3]*

[1]Shanghai Universal Cloud Medical Imaging Diagnostic Center, Shanghai, China, [2]Department of Hepatic Surgery VI, Eastern Hepatobiliary Surgery Hospital, The Second Military Medical University, Shanghai, China, [3]Department of Radiology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Objective: To investigate the role of prediction microvascular invasion (mVI) in hepatocellular carcinoma (HCC) by ^{18}F-FDG PET image texture analysis and hybrid criteria combining PET/CT and multi-parameter MRI.

Materials and methods: Ninety-seven patients with HCC who received the examinations of MRI and ^{18}F-FDG PET/CT were retrospectively included in this study and were randomized into training and testing cohorts. The lesion image texture features of ^{18}F-FDG PET were extracted using MaZda software. The optimal predictive texture features of mVI were selected, and the classification procedure was conducted. The predictive performance of mVI by radiomics classier in training and testing cohorts was respectively recorded. Next, the hybrid model was developed by integrating the ^{18}F-FDG PET image texture, metabolic parameters, and MRI parameters to predict mVI through logistic regression. Furthermore, the diagnostic performance of each time was recorded.

Results: The ^{18}F-FDG PET image radiomics classier showed good predicted performance in both training and testing cohorts to discriminate HCC with/without mVI, with an AUC of 0.917 (95% CI: 0.824–0.970) and 0.771 (95% CI: 0.578, 0.905). The hybrid model, which combines radiomics classier, SUVmax, ADC, hypovascular arterial phase enhancement pattern on contrast-enhanced MRI, and non-smooth tumor margin, also yielded better predictive performance with an AUC of 0.996 (95% CI: 0.939, 1.000) and 0.953 (95% CI: 0.883, 1.000). The differences in AUCs between radiomics classier and hybrid classier were significant in both training and testing cohorts (DeLong test, both $p < 0.05$).

Conclusion: The radiomics classier based on ^{18}F-FDG PET image texture and the hybrid classier incorporating ^{18}F-FDG PET/CT and MRI yielded good predictive performance, which might provide a precise prediction of HCC mVI preoperatively.

KEYWORDS

hepatocellular carcinoma, microvascular invasion, fluorine-18 fluorodeoxyglucose-positron emission tomography (^{18}F-FDG-PET), magnetic resonance imaging, texture feature

1 Introduction

Hepatocellular carcinoma (HCC) is one of the most common liver malignancies and one of the leading causes of cancer death in the world (Sung et al., 2021). The main cause of unsatisfactory HCC prognosis is that it is difficult to detect at the early stage because of poor symptoms. The primary curative treatment modality of HCC is partial hepatectomy. However, upon detection, most HCC patients are unsuitable for hepatectomy because of the underlying liver cirrhosis and hepatic dysfunction. Even though HCC patients receive curative surgical resection, there is still a relatively high recurrence rate, even in HCC patients receiving liver transplantation, as high as 15%–30% (Hoffman and Mehta, 2021; Tampaki et al., 2021). Therefore, early diagnosis and accurate prediction of postoperative HCC recurrence are important. Some established factors such as tumor grade, stage, size, liver function, and treatment acted as predictors of postoperative HCC recurrence (Ochiai et al., 2012).

Vascular invasion of HCC representing invasive tumor behavior is a significant predictor of poor outcomes (Jonas et al., 2001). Macrovascular invasion (MVI) could be readily detected by contrast-enhanced CT/MR imaging before surgical resection (Teefey et al., 2003). However, as a histologic finding, microvascular invasion (mVI) is usually visible only on microscopy by histopathology of the surgical specimen, which is difficult to diagnose before surgical resection. Therefore, it is essential to detect clinical predictors to suggest the presence of mVI preoperatively.

Previous research has reported that the status of mVI can be predicted by key imaging and laboratory tests. Several previous studies have reported that tumor margin, capsule, and peritumoral enhancement on CT/MRI scans were significantly associated with mVI (Nishie et al., 2008; Witjes et al., 2012). However, contradictory results were also reported in some studies (Kim et al., 2009; Chou et al., 2014). In clinical practice, developing a reliable preoperative predictor for mVI is still necessary. As a functional molecular imaging modality, fluorine-18 fluorodeoxyglucose (^{18}F-FDG) positron emission tomography (PET)/computed tomography (CT) is useful for evaluating HCC differentiation grade by estimating the glucose metabolism of tumor cells (Cho et al., 2017). Recently, several studies have reported the role of ^{18}F-FDG PET/CT in defining mVI in patients with HCC. However, there is no consistent conclusion, and the current results showed a wide range of sensitivity, specificity, and accuracy for preoperatively detecting mVI, which indicated that the current PET/CT technique is insufficient alone for establishing a risk factor for mVI (Kornberg et al., 2009; Lee et al., 2009; Lin et al., 2017; Kim and Kim, 2021). Recently, Li et al. conducted radiomics analysis on ^{18}F-FDG PET/CT to preoperatively predict mVI and prognosis in patients with very early and early stages of HCC, which is the importance of precise treatment of patients (Li et al., 2021).

In this retrospective study, we aimed to verify the comprehensive value of ^{18}F-FDG-PET/CT in the prediction of mVI by quantitative uptake measurement and image texture analysis. We also focus on the role of the hybrid model of incorporating ^{18}F-FDG-PET/CT and multi-parameter MRI. We hypothesized if ^{18}F-FDG-PET/CT findings predict mVI and, more importantly, the added value, if any, of PET/CT for the hybrid model in the prediction of mVI.

2 Materials and methods

2.1 Patients

This retrospective study was conducted in accordance with the Declaration of Helsinki proposed in 1975 and revised in 2000 and was approved by the Ethics Committee of the universal medical imaging center, Shanghai University (SHQJ-2019-05). The consecutive HCC patients were confirmed by histopathology after partial hepatectomy from January 2018 to April 2021. The inclusion criteria were as follows: (Sung et al., 2021) age >18 years; (Hoffman and Mehta, 2021) primary HCC confirmed by pathology of surgical specimens; (Tampaki et al., 2021) multi-parameter MR images containing conventional unenhanced MR imaging (including T1WI and T2WI), dynamic contrast-enhanced T1WI (including the arterial phase imaging, portal venous phase imaging, and delayed phase imaging), and diffusion-weighted imaging (DWI) (with *b*-value of 0 and 800 s/mm^2); (Ochiai et al., 2012) PET/CT and multi-parameter MRI examination approximately within 4 weeks before surgery; and (Jonas et al., 2001) no history of preoperative anti-cancer treatment. The exclusion criteria of this study were as follows: (Sung et al., 2021) preoperative images showing macrovascular invasion; (Hoffman and Mehta, 2021) HCC patients who underwent any anti-cancer treatment before partial hepatectomy; (Tampaki et al., 2021) those with time intervals of PET/CT, MRI, and surgery more than 4 weeks; (Ochiai et al., 2012) those with no pathology slides available for review; and (Jonas et al., 2001) images with artifacts affecting evaluation.

The patients enrolled in this study were randomly divided into two cohorts (training and testing) with a ratio of 7:3 using computer-generated random numbers.

2.2 ^{18}F-FDG PET/CT acquisition and image analysis

2.2.1 ^{18}F-FDG PET/CT acquisition

This study examined all ^{18}F-FDG PET/CT acquisitions with the SIEMENS Biograph mCT Flow PET/CT system (Siemens Medical Solutions United States, Inc.). After at least 6 h of fasting, the patient was intravenously administered a standard dose (3.7 MBq/kg) of ^{18}F-FDG, followed by image acquisition 60 min later, from the thigh to the head. Whole-body non-contrast enhancement CT scanning protocols were as follows: 120 kVp, 30–170 mAs adjusted to the patient's body weight and with a section width of 3 mm and collimation of 0.75 mm. An emission scan was performed in a three-dimensional (3D) mode with an acquisition time of 1.7 min per bed position. PET images were reconstructed by a 2-iteration, 21-ordered-subset expectation maximization algorithm using CT images for attenuation correction.

2.2.2 ^{18}F-FDG PET metabolic and volumetric parameters

Standardized uptake values (SUV) were calculated by the region-of-interest (ROI) technique. In order to calculate SUVmax and SUVmean, manually defined circular ROI was drawn on attenuation-corrected emission images selected for the largest axial image of the HCC lesion. On the PET image, an ellipse iso-contour was drawn covering the lesion, and the volume of interest (VOI) in 3D, that is, metabolic tumor volume (MTV), was obtained semi-automatedly with an iso-contour SUV value threshold of 2.5 (Kim et al., 2017). Total lesion glycolysis (TLG) was calculated by multiplying the selected PET volume by the average SUV within that volume: TLG = MTV × (average SUV). If the lesion had a low uptake of ^{18}F-FDG, the VOI was calculated on CT images and was then copied to PET to obtain the VOI on PET. Contrast-enhanced MRI was sometimes used to help delineate lesions.

The parameters of HCC ^{18}F-FDG metabolic avidity SUVmax, SUVmean, the ratio of the maximum standardized uptake value of tumor to the average standardized uptake value of normal liver (TLRmax), and the average tumor-to-normal liver standardized uptake value ratio (TLRmean) were calculated and recorded.

2.2.3 Texture analysis on axial ^{18}F-FDG PET images

2.2.3.1 Data standardization

Before texture analysis of the ^{18}F-FDG PET image, the data standardization procession to minimize the influence of image contrast and brightness variations was performed by adopting a method of normalizing the intensities of greyscale images into the range of mean value ± three standard deviations (SD) (μ – 3SD, μ + 3SD).

2.2.3.2 ^{18}F-FDG PET image texture analysis

Texture feature extraction and selection were performed with the MaZda software package (version 4.6, available at http://www.eletel.p.lodz.pl/mazda/) (Szczypinski et al., 2009). The largest axial ^{18}F-FDG PET image of each HCC lesion was selected for image texture analysis. The ROI was manually circumscribed over the entire HCC lesion possibly on each selected image by an experienced radiologist (Figure 1). MaZda software allows the computation of almost 300 texture features based on the image histogram, co-occurrence matrix (COM), run-length matrix, absolute gradient, auto-regressive model, and wavelet transforms (WAV) (Pinker et al., 2018). These texture features (Supplementary Table S1) were extracted from each ROI.

In order to achieve the highest differentiation power of HCC with or without mVI, avoid the problem of dimensionality, and reduce the bias, the image texture feature selection was performed by a feature selection algorithm combining Fisher's coefficient (Fisher), classification error probability combined with the average correlation coefficients (PA), and mutual information (MI) on module B11. Then, 30 texture features with the highest discriminative power were selected. In the training cohort, the classification procedures were conducted using principal component analysis (PCA), linear discriminant analysis (LDA), and non-linear discriminant analysis (NDA) on module B11 of the MaZda software package, and the sensitivity, specificity, positive predictive value, negative predictive value, and accuracy of diagnosis were calculated. The best classification procedure was selected as the radiomic classifier with the highest accuracy. In the testing cohort, the optimal 30 texture features were selected according to the result of the training cohort, and the predictive performance of the radiomic classifier for differentiation of the status of mVI of HCC was calculated.

2.3 MRI technique and image analysis

All MRI examinations were performed with a 3.0-T MRI scanner (Magnetom Aera, Siemens Healthcare, Erlangen, Germany; or Ingenia, Philips Healthcare, Best, the Netherlands). The imaging protocol was as follows: axial fast spin echo T2WI with fat saturation using a navigator-triggered technique, DWI using a single-shot echoplanar imaging pulse sequence with *b*-values of

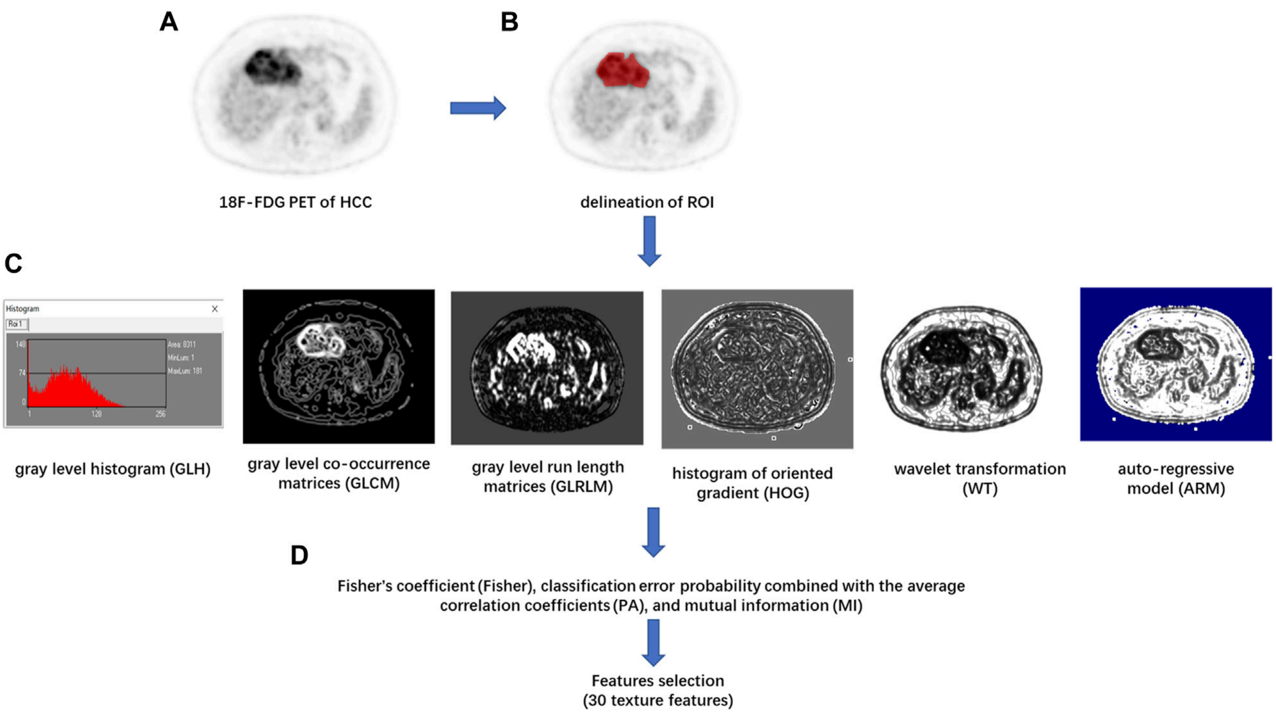


FIGURE 1
Flowchart of texture features (TF) extraction. **(A)** The ^{18}F-FDG PET image of the maximum axial section of hepatocellular carcinoma (HCC) was chosen. **(B)** The region of interest (ROI) of the tumor was drawn in red on MaZda. **(C)** The gray level histogram (GLH), gray level co-occurrence matrices (GLCM), gray level run length matrices (GLRLM), histogram of oriented gradient (HOG), wavelet transformation (WT), and the autoregressive model (ARM) of tumor were calculated, respectively. **(D)** A combination of feature selection algorithms, including Fisher's coefficient (Fisher), classification error probability combined with the average correlation coefficients (PA), and mutual information (MI), was used to determine 30 texture features with the highest discriminative power for differentiation HCC with or without between mVI.

0 and 800 s/mm^2. The apparent diffusion coefficient (ADC) of the HCC lesion was measured on the ADC map by the ROI drawn by two radiologists circumscribing the entire lesion on the largest axial image in consensus. T1WI images were obtained before and after administration of gadolinium injection solution, with a dose of 0.1 mmol per kilogram of body weight and an injection rate of 2 ml/sec. The arterial phase (AP) images were acquired approximately 25 s after contrast material injection. The portal venous phase (PVP) images were acquired approximately 55–65 s after the start of contrast material administration, and delayed phase (DP) images were acquired 90–100 s after contrast material injection. All parameters for the MRI sequences are summarized in Supplementary Table S2.

2.3.1 Image analysis

MR images were retrospectively analyzed on a workstation or a picture archiving and communication system (PACS). Two clinically experienced radiologists evaluated the MR images in consensus to obtain reliable results. Both readers were blinded to the status of mVI.

The two radiologists qualitatively made the following: 1) classified the arterial phase enhancement patterns on dynamic contrast-enhanced MRI into three patterns: hypervascular HCC, isovascular HCC, and hypovascular HCC; (Hoffman and Mehta, 2021) determined the presence or absence of intratumoral artery; 2) classified the patterns of tumor margin into two patterns: smooth margin or non-smooth margin; and 3) determined HCC with or without peritumoral enhancement on the AP images.

Qualitative findings on dynamic contrast-enhanced MR imaging were defined as follows: 1) hypervascular HCC: homogeneously hypervascular, hypervascular with slit-like hypovascular foci, or multifocal hypovascular foci, with a peripheral hypovascular area (Figure 2); 2) hypovascular HCC: with a nodular- or irregular-shaped hypointense portion at an inner area, with irregular rim-like enhancement, with a peripheral hypervascular area, discontinuous rim, or crescent-like, with linear or spot-like hypervascular foci, or completely hypovascular HCC (Figure 2) (Rhee et al., 2021); 3) intratumoral artery: the blood vessels within the tumor in AP images (Figure 2) (Segal et al., 2007); 4) a smooth margin defined as a nodular-shaped tumor without extranodular extension or infiltrative, non-smooth margins defined as a nodule with extranodular extension or an infiltrative margin (Figure 3) (Nakashima et al., 2003; Segal et al., 2007); and 5)

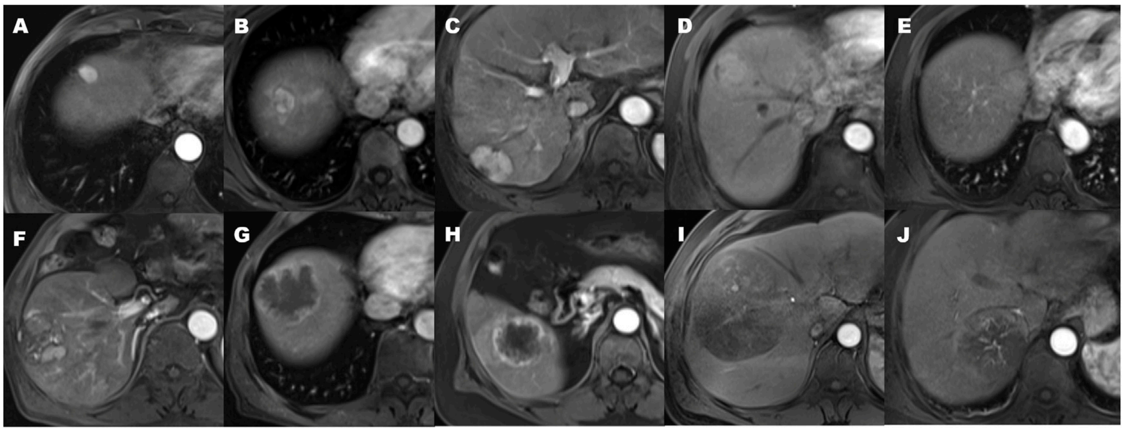


FIGURE 2
The patterns of arterial phase enhancement of hepatocellular carcinoma on contrast-enhanced MRI. **(A)** Homogeneously hypervascular HCC. **(B)** Hypervascular HCC with slit-like hypovascular foci. **(C)** With multifocal hypovascular foci. **(D)** Hypervascular HCC with the peripheral hypovascular area. **(E)** Completely isovascular HCC. **(F)** HCC with a nodular or irregular-shaped hypovascular portion at the inner area. **(G)** Hypovascular HCC with irregular rim-like enhancement. **(H)** Hypovascular HCC with the peripheral hypervascular area, discontinuous rim or crescent-like, and the presence of peritumoral enhancement seen in the arterial phase. **(I)** Hypovascular HCC with linear or spot-like hypervascular foci. **(J)** the presence of an intratumoral artery.

peritumoural enhancement on AP images: defined as a patchy hyperintense area adjacent to the tumor with broad contact to the tumor border on the AP images, presenting isointense to liver parenchyma on the DP images (Figure 2) (Kim et al., 2009).

2.4 Intra-observer and inter-observer agreement

The reproducibility of the intra-observer and inter-observer agreement for texture analysis was measured using 20 randomly chosen samples drawn from axial ^{18}F-FDG PET images and ADC map by two radiologists blinded from patients' characteristics. To evaluate intra-observer reproducibility, the first radiologist delineated an ROI twice within 2 weeks following the same procedure. Meanwhile, the second radiologist also independently delineated the ROI once following the same procedure. Then, the inter-observer agreement was assessed by comparing the results with the texture features extracted calculated from the first ROI delineation by the first radiologist. The intraclass correlation coefficient (ICC) was used to evaluate the intra-observer and inter-observer agreements. An ICC >0.75 indicated satisfactory agreement.

2.5 Microvascular invasion evaluation by histopathology

All surgical specimens and the status of mVI were reviewed and evaluated by an experienced pathologist in liver pathology. mVI was defined as the presence of a tumor in a portal vein, hepatic vein, or a large capsular vessel of the surrounding hepatic tissue lined by the endothelium that was visible only on microscopy (Roayaie et al., 2009; Xu et al., 2019). We categorized the HCC patients with mVI as the mVI-positive (mVI+) group and the HCC patients without mVI as the mVI negative (mVI−) group.

2.6 Statistical analysis

Inter-reader agreement was expressed by Cohen's kappa coefficient. A kappa statistic of 0.8–1.0, 0.6–0.79, 0.40–0.59, 0.2–0.39, and 0–0.19 was considered excellent, good, moderate, fair, and poor agreement, respectively.

For categorical variables, the differences between the mVI (−) and mVI (+) groups were analyzed by the Chi-squared test or Fisher's exact test. For continuous variables with a normal distribution, an independent-samples *t*-test was used to test the significant difference of the mVI (−) and mVI (+) groups; for continuous variables with a skewed distribution, a non-parametric Mann–Whitney *U* test was used. A two-tailed *p*-value less than 0.05 was considered that the difference was statistically significant. SPSS software (SPSS version 24.0; SPSS Inc., Chicago, IL, United States) was used to perform statistical analysis. The predictive value of each factor for mVI was determined by analysis of the area under the ROC curve (AUC). The differences in AUCs were compared by the DeLong test (DeLong et al., 1988) performed using MedCalc software (version 20.023). The calibration of the hybrid model

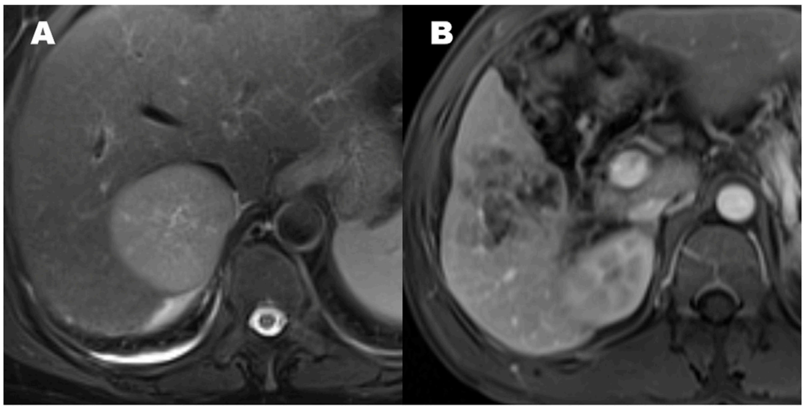


FIGURE 3
The patterns of hepatocellular carcinoma tumor margin. **(A)** Smooth tumor margin and intact tumor capsule. **(B)** Tumor with a non-smooth margin.

was performed by comparing the predicted and actual probability of mVI by the Hosmer–Lemeshow test.

3 Results

For the intra-observer and inter-observer agreement, radiomics features achieved satisfactory consistency. There was no radiomics feature to be eliminated. The mean ICC was 0.956 (range, 0.892 to 1, $p < 0.001$) in the intra-observer agreement and the mean ICC was 0.915 (range, 0.751 to 0.999, $p < 0.001$) in the inter-observer agreement.

In this study, the study flowchart for predicting mVI based on texture features, metabolic and volumetric parameters, and multi-parameter MRI is shown in Figure 4. A total of 97 patients' clinical characteristics are presented in Table 1.

Of the 97 patients with HCC, 58.8% (57/97) had tumors with mVI and 41.2% (40/97) without mVI. In the training cohort, mVI (+) was presented in 55.9% (38/68) of tumors, similar to 65.5% (19/29) seen in the testing cohort ($p > 0.05$). There were also no significant differences in other baseline clinical features between the training and testing groups (all $p > 0.05$, Table 1), which indicated that the distribution of baseline clinical-pathologic characteristics in the training and testing group was balanced.

In the pooled cohorts, the values of SUVmax, SUVmean, TLRmax, TLRmean, and ADC of HCC with mVI ($n = 57$) were higher than HCC without mVI ($n = 40$) (all $p < 0.05$). There was no significant difference in MTV, TLG, and ADC/ADC_liver between HCC with and without mVI (Table 2; Figures 5, 6).

ROC curve analysis showed that the cutoff values of SUVmax, SUVmean, TLRmax, TLRmean, and ADC for predicting HCC with mVI were 5.65, 3.79, 2.53, 1.92, and 1,171.5, respectively, with the largest Youden indexes but moderate diagnostic efficacy (all AUC < 0.70) (Table 3). In all quantitative metabolic parameters, the predictor with the highest diagnostic efficacy is SUVmax with the largest AUC (0.698, 95% CI: 0.593 to 0.803, $p = 0.001$), providing sensitivity and specificity of 59.6% and 80.0% at a cutoff value of 5.65 (Table 4; Supplementary Figure S1).

On dynamic contrast-enhanced MRI review of pooled all cohorts, the HCCs with mVI were significantly associated with enhancement patterns on AP imaging and tumor margin (all $p < 0.05$; Table 2). Most HCCs with mVI (44/57, 77.2%) demonstrated hypovascular lesions (Figure 5). The HCCs with mVI (28/57, 49.1%) manifest as having more frequency of non-smooth tumor margin than the HCCs without mVI (11/40, 27.5%).

Based on the above results, we selected the quantitative parameters of SUVmax and ADC, as well as the qualitative parameters of the hypovascular enhancement pattern on AP MR imaging and non-smooth tumor margin for subsequent analysis.

3.1 Training cohort

Based on the feature selection algorithm combining Fisher, PA, and MI coefficients in features modeling, 30 optimal features are selected and listed in Supplementary Table S3.

Compared to PCA (47.1%, 32/68) and LDA (19.1%, 13/68), NDA had the lowest misclassification rate with 7.4% (5/68) for all these three classification procedures by MaZda software ($p < 0.01$). The NDA classification with an AUC of 0.917 (95% CI: 0.824–0.970) showed a sensitivity of 100%, a specificity of 83.3%, a positive predictive value of 88.4%, a negative predictive value of 100%, and an accuracy of 92.6% (Table 4).

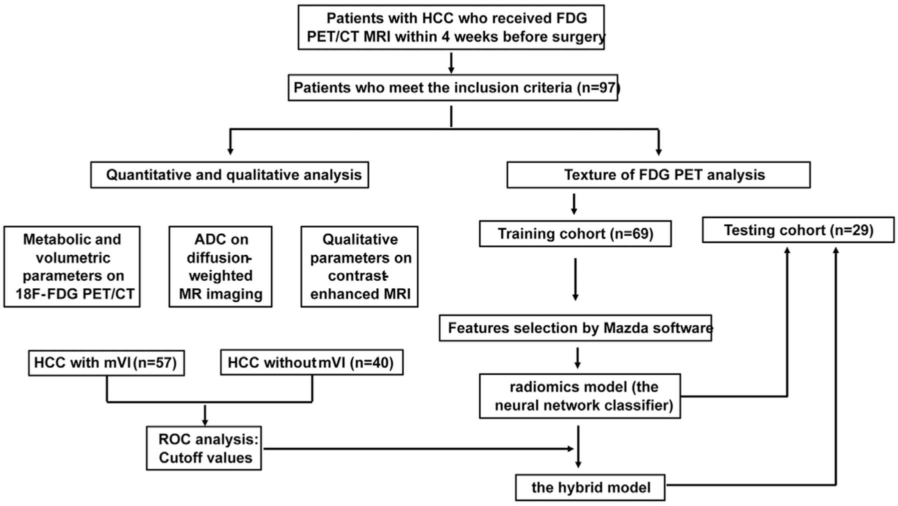


FIGURE 4
Study flowchart for predicting mVI by analyzing PET texture features and the hybrid model incorporated ^{18}F FDG PET/CT and MRI in training and testing set.

Combining radiomics classification results based on the texture of the ^{18}F-FDG-PET image, SUVmax, ADC, and the qualitative parameters of the hypovascular arterial phase enhancement pattern and non-smooth tumor margin, the hybrid model regression equation was as follows:

Logit (P) = −2.077 + 0.203 × SUVmax + 2.825 × ADC − 4.717 × Radiomics − 0.041 × non-smooth − 0.664 × hypovascular.

P is the probability of HCC with mVI. For $p \geq 0.5$, the lesion was expected to be HCC with mVI, whereas the other lesions were categorized as HCC without mVI.

The hybrid model with an AUC of 0.996 (95% CI: 0.939, 1.000; $p < 0.001$) yielded a sensitivity of 94.7%, a specificity of 100%, a positive predictive value of 100%, a negative predictive value of 93.8%, and an accuracy of 97.1% (Table 4). The difference in AUC between the radiomics classification model and the hybrid model was significant ($p = 0.017$). The calibration of the hybrid model was performed by comparing the predicted and actual probability of mVI by the Hosmer–Lemeshow test ($p > 0.05$). (Supplementary Figure S2A) The difference between the predicted and actual probabilities of mVI showed no statistical significance.

3.2 Testing cohort

We selected 30 texture features that were consistent with the training cohort. By using the neural network NDA classifier test included in module B11, the misclassification rate for HCC with *versus* without mVI was 20.7% (6/29). The sensitivity, specificity, positive predictive, negative predictive, and accuracy values were 84.2%, 70.0%, 84.2%, 70.0%, and 79.3%, respectively, with an AUC of 0.771 (95% CI: 0.578, 0.905) (Table 4).

In the testing cohort, the hybrid criteria yielded an AUC of 0.953 (95% CI: 0.883, 1.000; $p < 0.001$), a sensitivity of 95.0%, a specificity of 70%, a positive predictive value of 86.4%, a negative predictive value of 87.5%, and an accuracy of 86.7% (Table 4).

The performances of the radiomics and hybrid models to predict HCC with mVI were also good in the testing cohort, indicating their robustness. The difference in AUC between radiomics criteria and hybrid criteria was also significant ($p = 0.013$), indicating that the hybrid model incorporated ^{18}F-FDG PET/CT and MRI yielded better predictive performance. The calibration of the hybrid model was performed by comparing the predicted and actual probability of mVI by the Hosmer–Lemeshow test ($p > 0.05$) (Supplementary Figure S2B). The difference between the predicted and actual probability of mVI showed no statistical significance.

4 Discussion

To our knowledge, this study has produced the first texture-based radiomics model of ^{18}F FDG PET and a hybrid model that incorporated texture features of ^{18}F-FDG PET, quantitative metabolic parameters, and quantitative and qualitative MRI parameters for predicting the status of mVI in HCC. We found that the radiomics model based on the texture of ^{18}F-FDG PET had a good diagnostic performance with an

TABLE 1 Clinical characteristics of patients with HCC in the training and testing cohort.

		Training cohort (n = 68)			Testing cohort (n = 29)			p-value***
		mVI+ (n = 38) n (%) or median (IQR)	mVI− (n = 30) n (%) or median (IQR)	p-value*	mVI+ (n = 19) n (%) or median (IQR)	mVI− (10) n (%) or median (IQR)	p-value**	
Sex				0.452			0.414	0.519
	Male	26 (38.2)	23 (33.8)		11 (37.9)	8 (27.6)		
	Female	12 (17.6)	7 (10.3)		8 (27.6)	2 (6.9)		
Age (years)		55 (36.5, 59.3)	55 (50, 63.3)	0.158	50 (50, 62)	46 (55, 62)	0.628	0.512
HBV infection				0.965			0.450	0.583
	Absent	15 (22.1)	12 (17.6)		11 (37.9)	4 (13.8)		
	Present	23 (33.8)	18 (26.5)		8 (27.6)	6 (20.7)		
Child–Pugh				0.973			1.000	0.961
	A	29 (42.6)	23 (33.8)		14 (48.3)	7 (24.1)		
	B	9 (13.2)	7 (10.3)		5 (17.2)	3 (10.3)		
Liver cirrhosis				0.833			0.245	0.420
	Absent	13 (19.1)	11 (16.2)		8 (27.6)	7 (24.1)		
	Present	25 (36.8)	19 (27.9)		11 (37.9)	3 (10.3)		
AFP (ng/ml)				0.357			0.270	0.932
	≤200	16 (23.5)	16 (23.5)		12 (41.4)	4 (13.8)		
	>200	22 (32.3)	14 (20.6)		7 (24.1)	6 (20.7)		
Image tumor size (mm)		42 (18.8, 79.3)	40 (17.5, 69.8)	0.923	64.0 (55.0, 85.0)	67.5 (44.0, 96.3)	0.982	0.888

HCC, hepatocellular carcinoma; HBV, hepatitis B virus; AFP, alpha-fetoprotein; mVI, microvascular invasion; IQR, interquartile range.
*The difference in HCC with mVI and without mVI in the training cohort.
**The difference in HCC with mVI and without mVI in the testing cohort.
***The difference in HCC with mVI and without mVI in the pooled cohort.

AUC of 0.917 (95% CI: 0.836–0.998) and 0.771 (95% CI: 0.578 TO 0.966) in the training and testing cohorts, respectively. Furthermore, hybrid criteria combining ^{18}F-FDG PET and MRI could significantly increase diagnostic performance more than the radiomics model ($p < 0.05$) and yield an AUC of 0.996 (95% CI: 0.939, 1.000; $p < 0.001$) and 0.953 (95% CI: 0.883, 1.000) in the training and testing cohorts, respectively. Accordingly, our results may increase the accuracy of preoperative detection of the HCC with or without mVI. It is useful for planning the most appropriate treatment strategy and improving the prognosis of patients with HCC.

Despite progressions in diagnostic and therapeutic modalities, the prognosis of the patient with HCC is still to be improved due to the recurrence rate after treatment remaining high (Mlynarsky et al., 2015). The high heterogeneity of the HCC may have resulted in a varied prognosis (Lu et al., 2016; Lin et al., 2017). As a pathological feature, mVI reflects the invasiveness of the tumor, which usually appears in aggressive HCC while not in low-grade HCC. The patient with HCC presenting mVI has a shorter disease-free survival (DFS) due to a higher risk of tumor recurrence (Miyata et al., 2006). Therefore, mVI is an important prognostic factor of HCC and plays an important role in planning a personalized therapeutic strategy (Erstad and Tanabe, 2019; Xu et al., 2019). It is necessary to detect clinical predictors to suggest the presence of mVI preoperatively in order to establish a personalized therapeutic strategy.

Several previous studies have shown that various imaging modalities, including ultrasound, CT, especially contrast-enhanced CT, ^{18}F-FDG PET/CT, and MRI, have the potential to detect HCC with mVI (Lin et al., 2017; Hyun et al., 2018; Hu et al., 2019; Li et al., 2021; Meng et al., 2021). These studies showed that various imaging modalities might have a comparable predictive performance for mVI, whether the morphologic features, metabolic activity features, radiomics analysis, or combination are used. However, it is still unclear which modality is better and unable to completely meet clinical needs to establish risk factor for mVI only by one modality. In this study, we first extracted texture features from ROI in an ^{18}F-FDG PET image and selected 30 optimal texture features in our study. Meanwhile, the NDA classification procedure and the neural network classifier on module B11 of the MaZda software

TABLE 2 Radiologic findings of primary HCC in the training and testing cohorts.

	Training cohort (*n* = 68)			**Testing cohort (*n* = 29)**			***p*-value*****
	mVI+ (*n* = 38) *n* (%) or median (IQR)	**mVI− (*n* = 30) *n* (%) or median (IQR)**	***p* value***	**mVI+ (*n* = 19) *n* (%) or median (IQR)**	**mVI− (10) *n* (%) or median (IQR)**	***p*-value****	
SUVmax			0.056			0.068	0.011
	6.2 (3.9, 8.3)	4.0 (2.6, 4.8)		7.8 (4.0, 9.3)	4.5 (2.6, 8.2)		
SUVmean			0.022			0.211	0.006
	4.2 (2.6, 5.6)	3.0 (2.1, 3.3)		4.7 (2.9, 6.6)	3.4 (2.4, 5.8)		
TLRmax			0.006[a]			0.153	0.001[a]
	3.1 (2.3, 4.7)	2.1 (1.7, 2.5)		4.0 (2.0, 6.4)	2.6 (1.5, 4.7)		
TLRmean			0.029[a]			0.338	0.017[a]
	2.4 (1.4, 3.2)	1.6 (1.3, 1.8)		2.5 (1.5, 4.3)	1.8 (1.5, 3.4)		
MTV			0.168			0.737	0.303
	33.0 (6.7,141.3)	15.9 (5.8, 104.9)		101.2 (18.5, 180.8)	81.5 (8.3, 241.2)		
TLG			0.155[a]			0.728	0.102
	115.0 (22.6, 940.2)	48.1 (13.2, 350.4)		661.3 (70.4, 1,507.2)	255.5 (31.4, 1728.2)		
ADC (×10^{-3} mm^{2}/s)			0.091			0.066	0.014
	1.29 (1.07, 1.42)	1.12 (0.94, 1.25)		1.19 (0.87, 1.54)	1.03 (0.83, 1.25)		
ADC/ADC_liver			0.365			0.945[a]	0.650[a]
	0.9 (0.8, 1.3)	1.0 (0.9, 1.2)		1.3 (1.0, 1.7)	1.1 (1.0, 1.8)		
Arterial phase enhancement pattern		<0.001			0.698	<0.001	
Hypervascular	7 (10.3)	19 (27.9)		6 (20.7)	4 (13.8)		
Hypovascular	31 (45.6)	11 (16.2)		13 (44.8)	6 (20.7)		
Intratumoral artery		1.000			0.372	0.642	
Presence	5 (7.4)	4 (5.9)		3 (10.3)	3 (10.3)		
Absence	33 (48.5)	26 (38.2)		16 (55.2)	7 (24.1)		
Tumor margin		0.053			0.694	0.033	
Smooth	22 (32.4)	24 (35.3)		7 (24.1)	5 (17.2)		
Non-smooth	16 (23.5)	6 (8.8)		12 (41.4)	5 (17.2)		
Peritumoral enhancement in the arterial phase		0.204			0.449	0.077	
Presence	10 (14.7)	5 (7.4)		9 (31.0)	3 (10.3)		
Absence	28 (41.2)	25 (36.8)		10 (34.5)	7 (24.1)		

HCC, hepatocellular carcinoma; mVI, microvascular invasion; IQR, interquartile range; TLRmax, SUVmax/SUVmen_liver; TLRmean, SUVmean/SUVmen_liver; MTV, metabolic tumor volume; TLG, total lesion glycolysis; ADC, apparent diffusion coefficient.

*The difference in HCC with mVI and without mVI in the training cohort.

**The difference in HCC with mVI and without mVI in the testing cohort.

***The difference in HCC with mVI and without mVI in the pooled cohort.

[a]Non-parametric Mann–Whitney *U* test.

package were adopted to construct a radiomics model for predicting HCC with mVI. Our results showed that the radiomics model derived from the image texture feature of axial ^{18}F-FDG PET achieved a classification accuracy of 92.6% with an AUC of 0.917 in the training cohort and 79.3% with an AUC of 0.771 in the testing cohort. However, the traditional imaging radiologic features or metabolic activity features, such as quantitative metabolic parameters from PET/CT or quantitative and qualitative MRI parameters, yielded an accuracy range from 58.8% to 73.5% in the training cohort and 58.6%–72.4% in the testing cohort. These results showed that the radiomics model may be much better than the traditional morphologic and metabolic activity features. A recent study showed that a radiomics nomogram based on ^{18}F-FDG PET/CT was constructed to

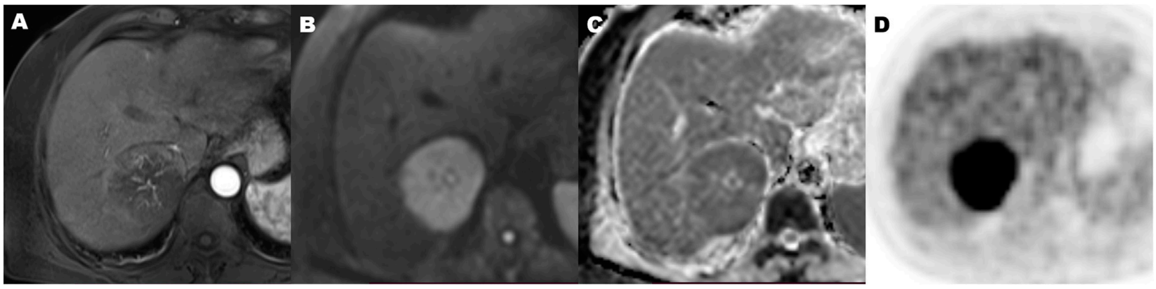


FIGURE 5
A 70-year-old hepatocellular carcinoma in liver segment VII woman with microvascular invasion, demonstrating Hypovascular enhancement pattern, high metabolic of SUVmax 9.2, MTV 134.2 cm^3, TLG 915.2, TLR 4.5, and ADC of 0.86 × 10^{-3} mm^2/s. Contrast-enhanced MR in arterial phase image **(A)**, diffusion-weighted MR image **(B)**, ADC map **(C)**, and ^{18}F-FDG PET **(D)** (MTV, metabolic tumor volume; TLG, total lesion glycolysis; TLR, ratio of tumor-to-liver SUV; ADC, apparent diffusion coefficient).

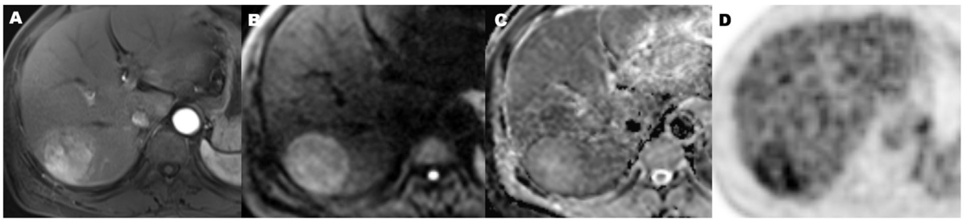


FIGURE 6
A 56-year-old hepatocellular carcinoma in liver segment VII man without microvascular invasion, demonstrating hypervascular enhancement pattern, low metabolic of SUVmax 3.5, MTV 72.2 cm^3, TLG 196.5, TLR 1.9, and ADC value of 1.1 × 10^{-3} mm^2/s. Contrast-enhanced MR in arterial phase image **(A)**, diffusion-weighted MR image **(B)**, ADC map **(C)**, and ^{18}F-FDG PET **(D)** (MTV, metabolic tumor volume; TLG, total lesion glycolysis; TLR, ratio of tumor-to-liver SUV; ADC, apparent diffusion coefficient).

TABLE 3 ROC results of ^{18}F-FDG PET/CT parameters and ADC in predicting HCC with mVI in the pooled cohort.

Characteristic parameter	Area under curve	*p*-values	Sensitivity (%)	Specificity (%)	Cutoff
SUVmax	0.698 (0.593, 0.803)	0.001	59.6	80.0	5.65
SUVmean	0.676 (0.567, 0.786)	0.003	66.7	77.5	3.79
TLRmax	0.693 (0.586, 0.799)	0.001	68.4	70.0	2.53
TLRmean	0.643 (0.531, 0.756)	0.017	64.9	75.0	1.92
ADC($\times 10^{-3} mm^2/s$)	0.651 (0.537, 0.765)	0.012	68.4	75.0	1.17

ROC, receiver operating-characteristic curve; HCC, hepatocellular carcinoma; mVI, microvascular invasion; TLRmax, SUVmax/SUVmen_liver; TLRmean, SUVmean/SUVmen_liver; ADC, apparent diffusion coefficient.

predict the mVI status in patients with very-early- and early-stage HCC with an AUC of 0.891 (95% CI: 0.799–0.984) in the training cohort and an AUC of 0.692 (95% CI: 0.497–0.887) in the testing cohort, which also showed that the radiomics model had a strong predictive power in detecting HCC with mVI (Li et al., 2021).

Of course, there are still several remaining significant challenges for radiomics in detecting HCC with mVI, such as replicability, standardization of images and data, and ethical and regulatory considerations (Forghani et al., 2019). In our study, we adopted a method of normalizing image intensities in the range of mean gray-level value ± three standard deviations (SD) to minimize the influence of contrast and brightness variation. In the testing cohort, our radiomics criteria yielded a classification accuracy of 79.3% with an

TABLE 4 Diagnostic performance of hepatocellular carcinoma with microvascular invasion in the training and testing cohorts.

	True positive	False positive	False negative	True negative	Sensitivity	Specificity	Positive-predictive value	Negative-predictive value	Accuracy
Training cohort ($n = 68$)									
Radiomics model	38	5	0	25	100	83.3	88.4	100	92.6%
SUVmax	20	5	18	25	52.6	83.3	80	58.1	66.2
ADC	26	8	12	22	68.4	73.3	76.5	64.7	70.6
Hypovascular	31	11	7	19	81.6	63.3	73.8	73.1	73.5
Non-smooth margin	16	6	22	24	42.1	80	72.7	52.2	58.8
Hybrid model	36	0	2	30	94.7	100	100	93.8	97.1
Testing cohorts ($n = 29$)									
Radiomics model	16	3	3	7	84.2	70	84.2	70	79.3
SUVmax	14	3	5	7	73.7	70	82.4	58.3	72.4
ADC	13	6	2	8	86.7	57.1	68.4	80	72.4
Hypovascular	13	6	6	4	68.4	40	68.4	40	58.6
Non-smooth margin	12	5	7	5	63.2	50	70.6	41.7	58.6
Hybrid model	19	3	1	7	95	70	86.4	87.5	86.7

AUC of 0.771 (95% CI: 0.578–0.905), which was lower than that classification accuracy of 92.6% with an AUC of 0.917 (95% CI: 0.824–0.970) in the training cohort. This indicated that the robustness of radiomics was to be improved.

The previous studies have also reported that the quantitative and qualitative parameters of SUVmax, ADC values, hypovascular lesion, and non-smooth tumor margin were associated with mVI status in patients with HCC (Mulé et al., 2020; Zhang et al., 2020). Therefore, in this study, we constructed hybrid criteria combining radiomics criteria and quantitative and qualitative parameters derived from ^{18}F-FDG PET/CT and MRI to predict the status of mVI in patients with HCC and achieved better results by significantly increasing AUC, yielded an AUC of 0.996 (95% CI: 0.939 to 1.000; $p < 0.001$) and 0.953 (95% CI: 0.883–1.000) in the training and testing cohorts, respectively. The difference in AUC between hybrid criteria and radiomics criteria showed statistical significance (p-values of 0.0165 and 0.0133 in the training and testing cohorts, respectively). Our findings suggest that the utility of the hybrid model combining ^{18}F-FDG PET/CT and MRI may improve the preoperative prediction of the status of mVI in HCC, compared to the only utility of the radiomics model based on texture features of ^{18}F-FDG PET.

There were several limitations in this study. First, this was a retrospective study with a small sample size, which could influence the robustness and reproducibility of our prediction models. The study samples were only divided into the training and testing cohorts to perform internal validation in this study. There may be a phenomenon of overfitting in the data processing for the AUC significant difference between the radiomics classier training set and the validation set. Therefore, the current study should need further validation with data augmentation and cross-validation in the future. Second, in this study, the HCC ^{18}F-FDG PET image textures were extracted only from the two-dimensional largest axial image of each HCC, which may cause loss of the entire tumor heterogeneity information. Therefore, the 3D structures radiomics of the tumor need further study in the future. Third, there was no significant association between the clinical characteristics and the status of mVI in our study ($p >$ 0.05). It was different from the previous study by Hyun et al. (2018), which showed that some clinical characteristics of clinical stage, AST, and AFP were significant predictors of mVI. The radiological hybrid model incorporating more clinical, pathological, and prognosis characteristics and radiomics should be considered in future studies.

In conclusion, the hybrid radiological model that incorporates the image texture of the ^{18}F-FDG PET signature, quantitative metabolic parameter, and quantitative and qualitative MRI parameters has powerful predictive performance in predicting the status of mVI preoperatively. Thus, such models may facilitate planning clinical treatment and improving survival in selected patients with HCC. Of course, it is warranted to validate the robustness and reproducibility of our prediction models by large-scale multicenter studies in the future.

Data availability statement

The raw data supporting the conclusions 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 the universal medical imaging center, Shanghai University. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

The literature search, analysis, data explanation, and manuscript draft were finished by HS, XL, and ZY. XL and ZY are responsible for the analysis and explanation of the radiomics imaging features data. YD, JS, WZ, WL, BS, FL, and XM acquired the clinical information and imaging. HS and ZY designed the study, explained the data, and made multiple revisions to the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Shanghai Municipal Health Commission (No. 20194Y0148) and the National Natural Science Foundation of China (No. 82073293).

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphys.2022.928969/full#supplementary-material

References

Cho, K. J., Choi, N. K., Shin, M. H., and Chong, A. R. (2017). Clinical usefulness of FDG-PET in patients with hepatocellular carcinoma undergoing surgical resection. *Ann. Hepatobiliary. Pancreat. Surg.* 21, 194–198. doi:10.14701/ahbps.2017.21.4.194

Chou, C. T., Chen, R. C., Lin, W. C., Ko, C. J., Chen, C. B., Chen, Y. L., et al. (2014). Prediction of microvascular invasion of hepatocellular carcinoma: preoperative CT and histopathologic correlation. *AJR. Am. J. Roentgenol.* 203, W253–W259. doi:10.2214/AJR.13.10595

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. doi:10.2307/2531595

Erstad, D. J., and Tanabe, K. K. (2019). Prognostic and therapeutic implications of microvascular invasion in hepatocellular carcinoma. *Ann. Surg. Oncol.* 26, 1474–1493. doi:10.1245/s10434-019-07227-9

Forghani, R., Savadjiev, P., Chatterjee, A., Muthukrishnan, N., Reinhold, C., Forghani, B., et al. (2019). Radiomics and artificial intelligence for biomarker and prediction model development in oncology. *Comput. Struct. Biotechnol. J.* 17, 995–1008. doi:10.1016/j.csbj.2019.07.001

Hoffman, D., and Mehta, N. (2021). Recurrence of hepatocellular carcinoma following liver transplantation. *Expert Rev. Gastroenterol. Hepatol.* 15, 91–102. doi:10.1080/17474124.2021.1823213

Hu, H. T., Wang, Z., Huang, X. W., Chen, S. L., Zheng, X., Ruan, S. M., et al. (2019). Ultrasound-based radiomics score: a potential biomarker for the prediction of microvascular invasion in hepatocellular carcinoma. *Eur. Radiol.* 29, 2890–2901. doi:10.1007/s00330-018-5797-0

Hyun, S. H., Eo, J. S., Song, B. I., Lee, J. W., Na, S. J., Hong, I. K., et al. (2018). Preoperative prediction of microvascular invasion of hepatocellular carcinoma using (18)F-FDG PET/CT: a multicenter retrospective cohort study. *Eur. J. Nucl. Med. Mol. Imaging* 45, 720–726. doi:10.1007/s00259-017-3880-4

Jonas, S., Bechstein, W. O., Steinmuller, T., Herrmann, M., Radke, C., Berg, T., et al. (2001). Vascular invasion and histopathologic grading determine outcome after liver transplantation for hepatocellular carcinoma in cirrhosis. *Hepatology* 33, 1080–1086. doi:10.1053/jhep.2001.23561

Kim, H., Park, M. S., Choi, J. Y., Park, Y. N., Kim, M. J., Kim, K. S., et al. (2009). Can microvessel invasion of hepatocellular carcinoma be predicted by pre-operative MRI? *Eur. Radiol.* 19, 1744–1751. doi:10.1007/s00330-009-1331-8

Kim, K., and Kim, S. J. (2021). Diagnostic test accuracies of F-18 FDG PET/CT for prediction of microvascular invasion of hepatocellular carcinoma: a meta-analysis. *Clin. Imaging* 79, 251–258. doi:10.1016/j.clinimag.2021.06.015

Kim, Y. I., Kim, Y. J., Paeng, J. C., Cheon, G. J., Lee, D. S., Chung, J. K., et al. (2017). Prediction of breast cancer recurrence using lymph node metabolic and volumetric parameters from 18F-FDG PET/CT in operable triple-negative breast cancer. *Eur. J. Nucl. Med. Mol. Imaging* 44 (11), 1787–1795. doi:10.1007/s00259-017-3748-7

Kornberg, A., Freesmeyer, M., Bärthel, E., Jandt, K., Katenkamp, K., Steenbeck, J., et al. (2009). 18F-FDG-uptake of hepatocellular carcinoma on PET predicts microvascular tumor invasion in liver transplant patients. *Am. J. Transpl.* 9, 592–600. doi:10.1111/j.1600-6143.2008.02516.x

Lee, J. W., Paeng, J. C., Kang, K. W., Kwon, H. W., Suh, K. S., Chung, J. K., et al. (2009). Prediction of tumor recurrence by 18F-FDG PET in liver transplantation for hepatocellular carcinoma. *J. Nucl. Med.* 50, 682–687. doi:10.2967/jnumed.108.060574

Li, Y., Zhang, Y., Fang, Q., Zhang, X., Hou, P., Wu, H., et al. (2021). Radiomics analysis of [18F] FDG PET/CT for microvascular invasion and prognosis prediction in very-early- and early-stage hepatocellular carcinoma. *Eur. J. Nucl. Med. Mol. Imaging* 48, 2599–2614. doi:10.1007/s00259-020-05119-9

Lin, C. Y., Liao, C. W., Chu, L. Y., Yen, K. Y., Jeng, L. B., Hsu, C. N., et al. (2017). Predictive value of 18F-FDG PET/CT for vascular invasion in patients with hepatocellular carcinoma before liver transplantation. *Clin. Nucl. Med.* 42, e183. doi:10.1097/RLU.0000000000001545

Lin, D. C., Mayakonda, A., Dinh, H. Q., Huang, P., Lin, L., Liu, X., et al. (2017). Genomic and epigenomic heterogeneity of hepatocellular carcinoma. *Cancer Res.* 77, 2255–2265. doi:10.1158/0008-5472.CAN-16-2822

Lu, L. C., Hsu, C. H., Hsu, C., and Cheng, A. L. (2016). Tumor heterogeneity in hepatocellular carcinoma: facing the challenges. *Liver Cancer* 5, 128–138. doi:10.1159/000367754

Meng, X. P., Wang, Y. C., Zhou, J. Y., Yu, Q., Lu, C. Q., Xia, C., et al. (2021). Comparison of MRI and CT for the prediction of microvascular invasion in solitary hepatocellular carcinoma based on a non-radiomics and radiomics method: which imaging modality is better? *J. Magn. Reson. Imaging.* 54, 526–536. doi:10.1002/jmri.27575

Miyata, R., Tanimoto, A., Wakabayashi, G., Shimazu, M., Nakatsuka, S., Mukai, M., et al. (2006). Accuracy of preoperative prediction of microinvasion of portal vein in hepatocellular carcinoma using superparamagnetic iron oxide-enhanced magnetic resonance imaging and computed tomography during hepatic angiography. *J. Gastroenterol.* 41, 987–995. doi:10.1007/s00535-006-1890-2

Mlynarsky, L., Menachem, Y., and Shibolet, O. (2015). Treatment of hepatocellular carcinoma: steps forward but still a long way to go. *World J. Hepatol.* 7, 566–574. doi:10.4254/wjh.v7.i3.566

Mulé, S., Chalaye, J., Legou, F., Tenenhaus, A., Calderaro, J., Galletto Pregliasco, A., et al. (2020). Hepatobiliary MR contrast agent uptake as a predictive biomarker of aggressive features on pathology and reduced recurrence-free survival in resectable hepatocellular carcinoma: comparison with dual-tracer 18F-FDG and 18F-fch PET/CT. *Eur. Radiol.* 30, 5348–5357. doi:10.1007/s00330-020-06923-5

Nakashima, Y., Nakashima, O., Tanaka, M., Okuda, K., Nakashima, M., Kojiro, M., et al. (2003). Portal vein invasion and intrahepatic micrometastasis in small hepatocellular carcinoma by gross type. *Hepatol. Res.* 26, 142–147. doi:10.1016/s1386-6346(03)00007-x

Nishie, A., Yoshimitsu, K., Asayama, Y., Irie, H., Tajima, T., Hirakawa, M., et al. (2008). Radiologic detectability of minute portal venous invasion in hepatocellular carcinoma. *AJR. Am. J. Roentgenol.* 190, 81–87. doi:10.2214/AJR.07.2810

Ochiai, T., Ikoma, H., Okamoto, K., Kokuba, Y., Sonoyama, T., Otsuji, E., et al. (2012). Clinicopathologic features and risk factors for extrahepatic recurrences of hepatocellular carcinoma after curative resection. *World J. Surg.* 36, 136–143. doi:10.1007/s00268-011-1317-y

Pinker, K., Shitano, F., Sala, E., Do, R. K., Young, R. J., Wibmer, A. G., et al. (2018). Background, current role, and potential applications of radiogenomics. *J. Magn. Reson. Imaging* 47, 604–620. doi:10.1002/jmri.25870

Rhee, H., Cho, E. S., Nahm, J. H., Jang, M., Chung, Y. E., Baek, S., et al. (2021). Gadoxetic acid-enhanced MRI of macrotrabecular-massive hepatocellular carcinoma and its prognostic implications. *J. Hepatol.* 74, 109–121. doi:10.1016/j.jhep.2020.08.013

Roayaie, S., Blume, I. N., Thung, S. N., Park, Y. N., Kim, M. J., Kim, K. S., et al. (2009). A system of classifying microvascular invasion to predict outcome after resection in patients with hepatocellular carcinoma. *Gastroenterology* 137, 850–855. doi:10.1053/j.gastro.2009.06.003

Segal, E., Sirlin, C. B., Ooi, C., Adler, A. S., Gollub, J., Chen, X., et al. (2007). Decoding global gene expression programs in liver cancer by noninvasive imaging. *Nat. Biotechnol.* 25, 675–680. doi:10.1038/nbt1306

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71, 209–249. doi:10.3322/caac.21660

Szczypinski, P. M., Strzelecki, M., Materka, A., and Klepaczko, A. (2009). MaZda–a software package for image texture analysis. *Comput. Methods Programs Biomed.* 94, 66–76. doi:10.1016/j.cmpb.2008.08.005

Tampaki, M., Papatheodoridis, G. V., and Cholongitas, E. (2021). Intrahepatic recurrence of hepatocellular carcinoma after resection: an update. *Clin. J. Gastroenterol.* 14, 699–713. doi:10.1007/s12328-021-01394-7

Teefey, S. A., Hildeboldt, C. C., Dehdashti, F., Siegel, B. A., Peters, M. G., Heiken, J. P., et al. (2003). Detection of primary hepatic malignancy in liver transplant candidates: prospective comparison of CT, MR imaging, US, and PET. *Radiology* 226, 533–542. doi:10.1148/radiol.2262011980

Witjes, C. D., Willemssen, F. E., Verheij, J., van der Veer, S. J., Hansen, B. E., Verhoef, C., et al. (2012). Histological differentiation grade and microvascular invasion of hepatocellular carcinoma predicted by dynamic contrast-enhanced MRI. *J. Magn. Reson. Imaging* 36, 641–647. doi:10.1002/jmri.23681

Xu, X., Zhang, H. L., Liu, Q. P., Sun, S. W., Zhang, J., Zhu, F. P., et al. (2019). Radiomic analysis of contrast-enhanced CT predicts microvascular invasion and outcome in hepatocellular carcinoma. *J. Hepatol.* 70, 1133–1144. doi:10.1016/j.jhep.2019.02.023

Zhang, L., Yu, X., Wei, W., Pan, X., Lu, L., Xia, J., et al. (2020). Prediction of HCC microvascular invasion with gadobenate-enhanced MRI: correlation with pathology. *Eur. Radiol.* 30, 5327–5336. doi:10.1007/s00330-020-06895-6

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 25 August 2022
DOI 10.3389/fphys.2022.948767

OPEN ACCESS

EDITED BY
Yi Xu,
Shanghai Jiao Tong University, China

REVIEWED BY
Alexander E Berezin,
Zaporizhia State Medical University, Ukraine
Xiaohu Li,
First Affiliated Hospital of Anhui Medical University, China

*CORRESPONDENCE
Hongzan Sun,
sunhongzan@126.com

SPECIALTY SECTION
This article was submitted to Medical Physics and Imaging,
a section of the journal
Frontiers in Physiology

RECEIVED 20 May 2022
ACCEPTED 25 July 2022
PUBLISHED 25 August 2022

CITATION
Wang X, Xu C, Grzegorzek M and Sun H (2022), Habitat radiomics analysis of pet/ct imaging in high-grade serous ovarian cancer: Application to Ki-67 status and progression-free survival.
Front. Physiol. 13:948767.
doi: 10.3389/fphys.2022.948767

Habitat radiomics analysis of pet/ ct imaging in high-grade serous ovarian cancer: Application to Ki-67 status and progression-free survival

Xinghao Wang[1], Chen Xu[2], Marcin Grzegorzek[3] and Hongzan Sun[1]*

[1]Department of Radiology, Shengjing Hospital of China Medical University, Shenyang, China, [2]Department of Surgical Oncology, The First Affiliated Hospital of China Medical University, Shenyang, China, [3]Institute of Medical Informatics, University of Luebeck, Luebeck, Germany

Purpose: We aim to develop and validate PET/ CT image-based radiomics to determine the Ki-67 status of high-grade serous ovarian cancer (HGSOC), in which we use the metabolic subregion evolution to improve the prediction ability of the model. At the same time, the stratified effect of the radiomics model on the progression-free survival rate of ovarian cancer patients was illustrated.

Materials and methods: We retrospectively reviewed 161 patients with HGSOC from April 2013 to January 2019. 18F-FDG PET/ CT images before treatment, pathological reports, and follow-up data were analyzed. A randomized grouping method was used to divide ovarian cancer patients into a training group and validation group. PET/ CT images were fused to extract radiomics features of the whole tumor region and radiomics features based on the Habitat method. The feature is dimensionality reduced, and meaningful features are screened to form a signature for predicting the Ki-67 status of ovarian cancer. Meanwhile, survival analysis was conducted to explore the hierarchical guidance significance of radiomics in the prognosis of patients with ovarian cancer.

Results: Compared with texture features extracted from the whole tumor, the texture features generated by the Habitat method can better predict the Ki-67 state ($p < 0.001$). Radiomics based on Habitat can predict the Ki-67 expression accurately and has the potential to become a new marker instead of Ki-67. At the same time, the Habitat model can better stratify the prognosis ($p < 0.05$).

Conclusion: We found a noninvasive imaging predictor that could guide the stratification of prognosis in ovarian cancer patients, which is related to the expression of Ki-67 in tumor tissues. This method is of great significance for the diagnosis and treatment of ovarian cancer.

KEYWORDS

PET/CT, high-grade serous ovarian cancer, Ki-67, radiomics, Habitat, progression-free survival

Introduction

Ovarian cancer is one of the most common gynecological cancers (Torre et al., 2018). In the past few decades, although the survival rate of most tumors has improved, the 5-year survival rate of ovarian cancer has not changed since 1980 (2). Most ovarian tumors belong to high-grade serous ovarian cancer (HGSOC) (Kohn and Ivy, 2017). They usually have extensive peritoneum (III stage) or extraperitoneal (IV stage) spread in the late stage, and the risk of recurrence and death is very high (Chen and Du, 2018). In the early stages of treatment, most ovarian cancer patients respond to surgery and platinum-based chemotherapy but patients often relapse and develop resistance to chemotherapy (Jayson et al., 2014). Therefore, the exploration of prognostic biomarkers for ovarian cancer patients is constantly expanding.

Ki-67 is a kind of nuclear protein, which is expressed in the whole cell cycle of proliferating cells except for G0 cells. It is closely related to cell proliferation and invasion (Schlüter et al., 1993). In ovarian cancer, there is a clear link between Ki-67 and recurrence and prognosis of ovarian cancer (Deng et al., 2015a; Qiu et al., 2019). Positron emission tomography (PET) is a kind of functional imaging method, which can clarify the spatial distribution of the metabolic activity through tracer uptake and accurately locate the malignant lesion area combined with the anatomical information provided by CT (9). Some studies have shown that ovarian cancer PET/CT has higher preoperative staging accuracy than simple CT, and the accuracy of CT and PET/CT staging is between 53%–55% and 55%–89% (Kemppainen et al., 2019; O'Connor et al., 2013; Castellani et al., 2019). At the same time, PET/CT is more accurate in detecting recurrence than other reference standards (such as CA-125, CT, or MRI) (Limei et al., 2013). In other cancers, the radiomics model composed of noninvasive PET/ CT has had a good prediction effect, but it has not been reported in ovarian cancer (Antunovic et al., 2017; Acar et al., 2019; Kong et al., 2019). In this study, we used metabolic subregion evolution (Habitat) to improve the prediction ability of the radiomics model (Mu et al., 2020). In imaging medicine, the Habitat method is often used to divide different tumor subregions (reflecting different functional or material areas of the focus), which is a method with strong application scenarios.

Materials and methods

Patients

This retrospective study was approved by the review committee of our institution and was adherent to the principles and requirements of the Declaration of Helsinki. This retrospective study collected 197 patients with HGSOC in our hospital from April 2013 to January 2019. The exclusion criteria are as follows: 1) have suffered from other tumors, 2) no PET/CT scan performed, 3) any targeted treatment before scanning, and 4) performed within 3 weeks before surgery with negative 18F-FDG uptake. At the same time, the patients were operated on and treated according to NCCN guidelines. After a regular and complete follow-up (imaging data), the patients achieved progression-free survival. Progression-free survival (PFS) refers to the time from randomization to the first occurrence of disease progression or death from any cause. Finally, 161 patients were included in the study. The patients were randomly divided into a training group (n = 112) and test group (n = 49).

18F-FDG PET/CT acquisition

Patients were fasting from food and water for more than 6 h, and their blood sugar level was controlled below 7 mmol/L. One hour after intravenous injection of 18F-FDG (GE MINItrace II; GE Healthcare, Milwaukee, WI) at 0.08–0.16 mci/kg, PET/CT was performed from the head to the middle of the femur (GE Discovery PET/CT Elite; GE Healthcare, Milwaukee, WI). A 3dimensional PET model was used, with a matrix of 192 × 192 and an exposure time of 2 min/bed position. Low-dose spiral CT was performed at 120–140 kV and 80 ma. After CT attenuation correction, PET images were reconstructed using the algorithm of time-of-flight and point-spread-function, including 2 iterations and 24 subsets.

Habitat generation and feature extraction

The workflow of radiomics is shown in Figure 1. For processing images, we have standardized processing, and then in the process of delineation and ROI processing we used LIFEx software (https://www.lifexsoft.org/) and ITK-SNAP (http://www.itksnap.org/pmwiki/pmwiki.php). Based on the metabolic threshold of PET images, we quickly identified the tumor contour. On the python (version 3.8.5) platform, we implemented the Otsu threshold way by self-built code, and we obtained two metabolic subgroups that maximized the variance between groups. Based on the threshold, the corresponding tumors in PET images were divided into the high metabolism region (the red region in Figure 2) and the low metabolism region (the green region in Figure 2), representing different Habitat subgroups. We define the difference in SUV metabolism between these two subregions, and we define them as PET_{hight} and PET_{low}. So far, we have obtained three kinds of ROIs based on PET images, including the whole tumor. Based on each ROI and its effect on the image, we extract 1316 texture features, replace the abnormal value of omics features with the average value, and then separate the feature data according to the average value $\mu = 0$ δ $2 = 1$.

Feature selection and model establishment

For the selection of models, we adopt a variety of methods, such as PCC, PCA, and Lasso, to operate separately or in parallel

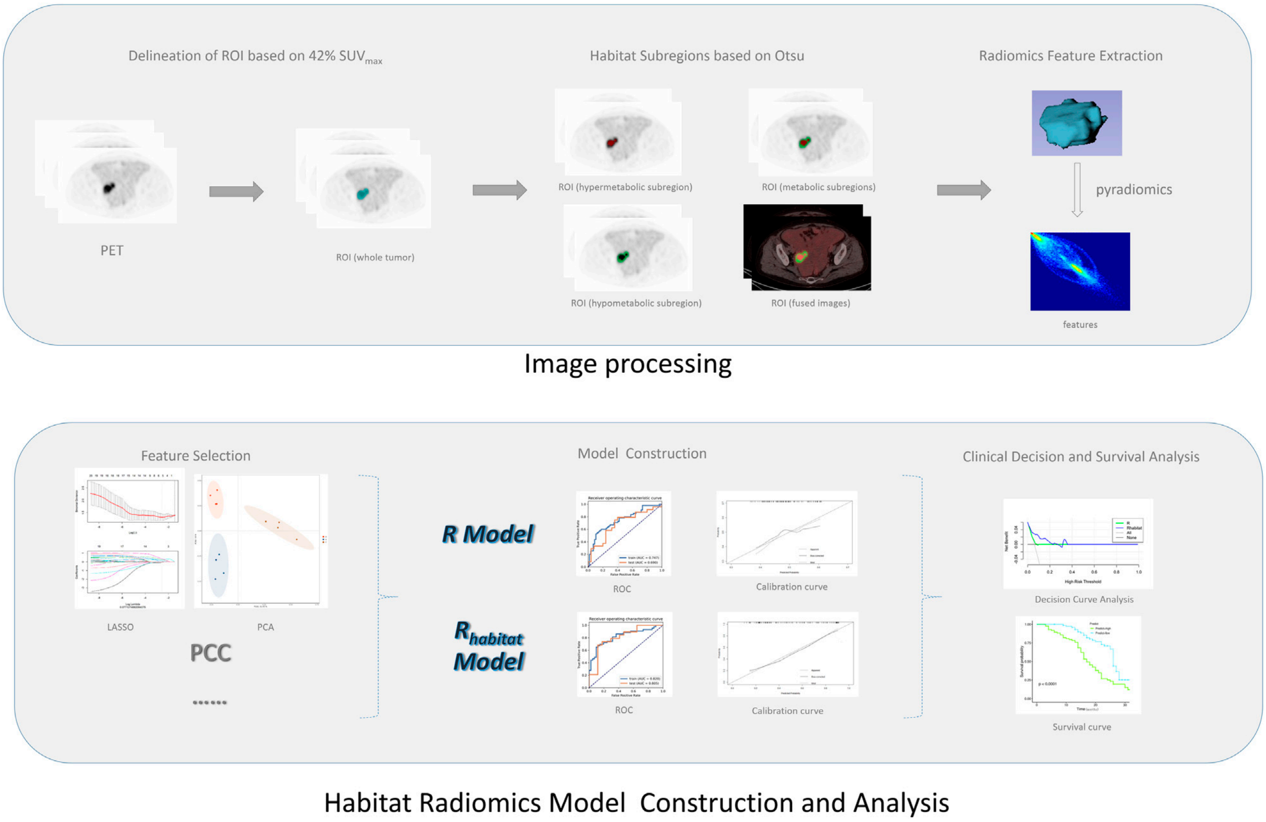


FIGURE 1
Schematic diagram of study design.

many times, in order to remove redundant and strongly collinear or correlated variables, reduce model parameters, make them match the sample size of this study, and avoid overfitting or underfitting phenomenon in the final prediction model.

For modeling, we use common classifiers (SVM, LAD, logistic regression, decision tree, random forest, and the naive Bayesian algorithm). In different models, the model with high AUC and strong generalization ability is selected to be included in the final model selection. According to the different ROI extracted, the feature extracted from the whole tumor region is called R, and the model based on the Habitat region (PET_{hight} and PET_{low}) extraction is $R_{habitat}$.

Statistical analysis

SPSS statistical software (version 24.0; IBM), R (version 3.63), and MedCalc Statistical Software version 15.2.2 (MedCalc Software bvba, Ostend, Belgium; http://www.medcalc.org; 2015) were used for all analyses. The clinical characteristics of the training group and the validation group were statistically tested to test their data distribution. The *t* test was used for data with normal distribution and homogeneity of variance, and the *U* test was used for data without normal distribution. The Delong test was performed on different models in the training group and the test group. Clinical decision curves and survival curves of different group models were compared to explore the significance of PFS.

Results

Clinical features

The clinical characteristics of patients in the training group and the validation group are shown in Table 1. There was no significant difference in clinical characteristics between the two groups.

Description and comparison of prediction models

The model takes the patient's Ki-67 status (>50%) (Qiu et al., 2019) as the label for modeling and analysis. The details of R and $R_{habitat}$ models generated are shown in Table 2. According to the diagnostic efficiency, we found that the $R_{habitat}$ model (the training

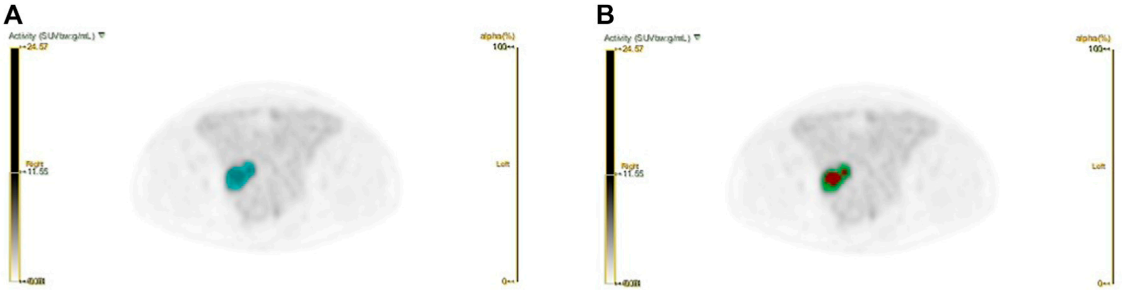


FIGURE 2
Schematic diagram of Habitat method. **(a)** The blue area represents the whole tumor area; **(b)** Based on the whole tumor area, we have implemented the subarea partition by habitat algorithm.

TABLE 1 Clinical characteristics of HGSOC patients in training and test groups.

Characteristic	Training group (n = 112)	Test group (n = 49)	*p* value
Age, mean ± SD, year	53.22±9.31	53.43±10.52	$p > 0.1$
NACT			$p > 0.1$
Yes	43	13	
No	69	36	
LNM			
Yes	59	29	$p > 0.1$
No	53	20	
FIGO stage			$p > 0.1$
Stage III	72	33	
Stage IV	40	16	
Ascites			$p > 0.1$
<200 ml	61	24	
200ml–1000 ml	37	19	
>1000 ml	14	6	

group: the AUC value is 0.835, 95% CIs: [0.7240-0.9460], accuracy: 0.7919, sensitivity: 0.8377; the training group: the AUC value is 0.835, 95% CIs: [0.7240-0.9460], accuracy: 0.7919, sensitivity: 0.8377; the training group: the AUC value is 0; the test group: AUC = 0.8076, 95% CIs: [0.7225-0.9611], accuracy: 0.7557, sensitivity: 0.8192) had the highest diagnostic efficiency (Figure 3), higher than R (the training group: AUC = 0.7670, 95% CIs: [0.6842-0.8498], accuracy: 0.7519, sensitivity: 0.8077); the test group: the AUC value was 0.7488, 95% CIs: [0.6583-0.8393], accuracy: 0.7223, sensitivity: 0.7892). In the Delong test, we found that the efficiency of the $R_{habitat}$ model was higher than that of the R model ($p < 0.05$).

Decision curve analysis and survival analysis

The decision curve analysis (DCA) displays estimates of a series of probability threshold (normalized) net benefits used to classify observations as "high risk." These curves help to assess a treatment policy that recommends that the impact of a risk-based policy on the population be compared with the "treat all" and "no treat" intervention policies, thereby recommending treatment for patients estimated to be "at high risk." DCA of the two models is shown in Figure 4.

Through the aforementioned screening, we obtained the combination model $R_{habitat}$ with strong classification and prediction ability. In order to further explore its prediction ability for the prognosis of patients with HGSOC, we drew the survival curve and found that the model has a strong prediction stratification ability, and the K-M test (Figure 5) has a significant difference ($p < 0.001$).

Discussion

In this study, we developed and validated radiomics from positron emission tomography (PET), computed tomography

TABLE 2 Description of two models.

Model	Standardization method	Feature selection method	Characteristic quantity	Model classifier
R	Z-score	Recursive feature elimination	20	Auto-encoder
$R_{habitat}$	Z-score	Kruskal–Wallis	8	Logistic regression

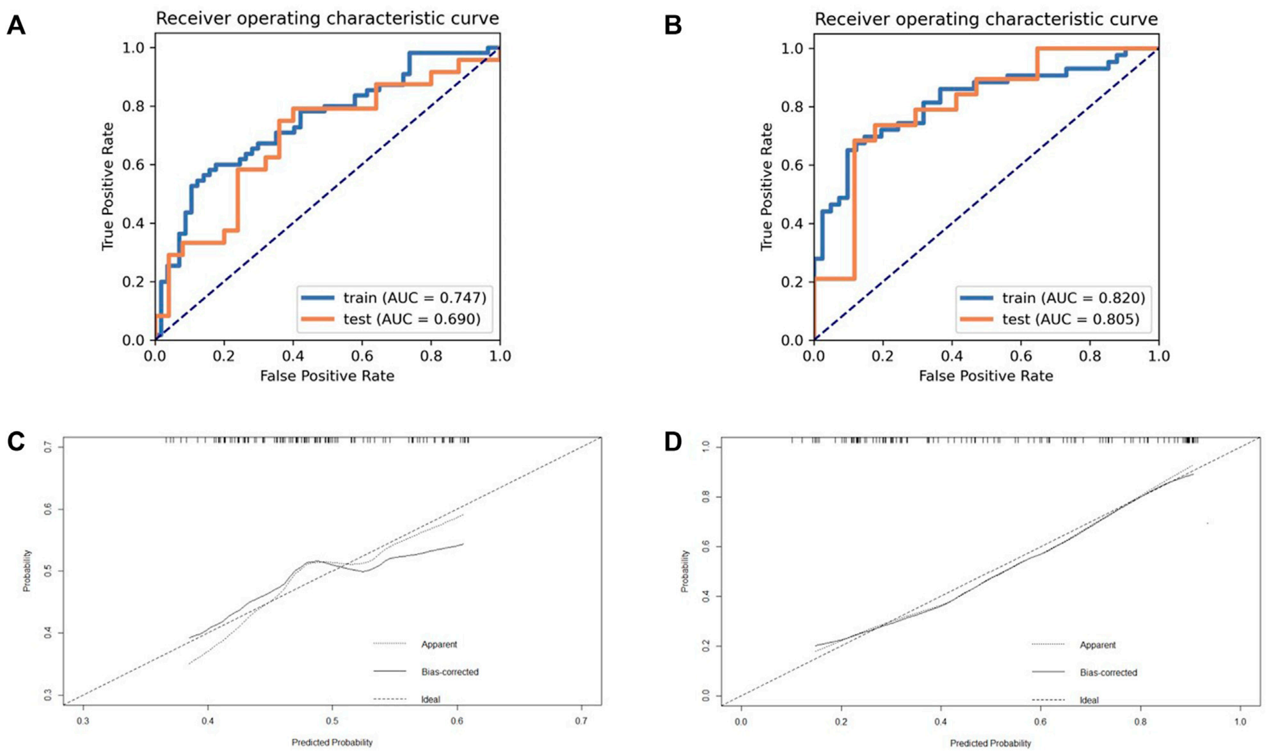


FIGURE 3
The ROC and calibration curve of R and $R_{habitat}$. **(a)** The performance of training set and validation set on model R (based on whole tumor); **(b)** The performance of training set and validation set on model $R_{habitat}$ (based on different metabolic sub-regions). **(c)** Demonstrate the calibration effect of the R model; **(d)** Demonstrate the calibration effect of the $R_{habitat}$ model.

(CT), and the Habitat subregion to predict the Ki-67 status in patients with ovarian cancer and to explore its role in prognostic stratification. Patients can be divided into low-risk and high-risk groups through the establishment of an imaging omics model. There is a significant difference between the model using the Habitat subregion and the traditional model.

In recent years, some scholars have proposed the molecular classification of ovarian cancer related to prognosis, which is marked as differentiation, immunoreactivity, mesenchymal, and proliferative (Tothill et al., 2008; Verhaak et al., 2013). At the same time, studies have shown that the molecular classification based on CT imaging features can effectively distinguish ovarian cancer and can be used as a predictor of prognosis (Vargas et al., 2015). A multicenter study used computed tomography imaging features to assess its association with disease progression time and ovarian cancer transcriptomic characteristics and to develop an image-based risk scoring system (Vargas et al., 2017). Studies have also been conducted to explore the association between proteomics and imaging omics. It is found that four proteins in ovarian cancer patients are related to CT-based imaging. Among them, the correlation between the CRIP2 protein and mesenteric diseases is strongest, and the abundance of other three proteins (STXVP2, ASS1, and CBD) is related to the heterogeneity of tumor location (Beer et al., 2020). There are a lot of research on prognosis but there are some problems of insufficient explanation, such as trying to find the interpretable aspect of gene transcription (Lu et al., 2019). Some studies suggest that SUV_{max} and SUV_{mean} are moderately correlated with the Ki-67

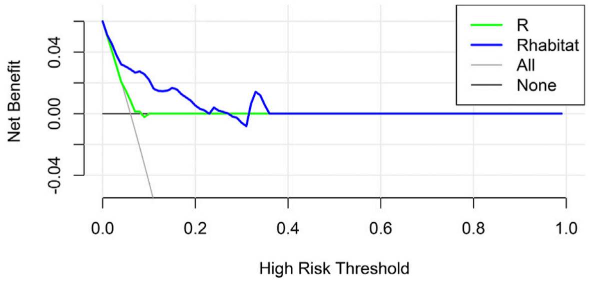


FIGURE 4
Decision curve analysis for R and $R_{habitat}$.

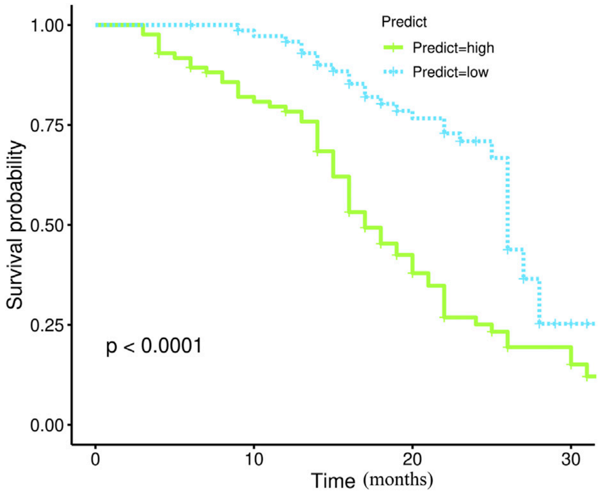


FIGURE 5
The survival curves of $R_{habitat}$.

index, which confirms the value of the PET image for Ki-67 but it is difficult to accurately predict the expression of Ki-67 (Deng et al., 2015b; Mayoral et al., 2018. We use the Habitat method to generate two metabolic subregions with different metabolic characteristics (Mu et al., 2020). In this experiment, the Otsu scheme is used. The principle is to generate two parts so that the overall similarity of each part is the highest and the difference between different parts is the largest. Automatic segmentation is realized by a variance correlation algorithm. At present, the Habitat method has been applied in many medical images. Some studies have shown that it can significantly improve the performance of PFS and OS models in predicting locally advanced cervical cancer patients in PET/ CT (16) and has a significant predictive value for glioma prognosis (Verma et al., 2020; Park et al., 2021), nasopharyngeal carcinoma (Xu et al., 2020), lung cancer (Cherezov et al., 2019), and prostate cancer (Parra et al., 2019). For this research model, the overall characteristics of the tumor were eventually incorporated into the PET mode high metabolic area and low metabolic area, where the high metabolic area represented the strongest part of the tumor activity, which had a significance for the prognosis of patients (Pinho et al., 2020), while the PET low metabolism area was often the edge of the tumor and correlated with immune infiltration (Grove et al., 2021). At the same time, we should note that Habitat subregions generated by PET modality are more heterogeneous and eccentric, and are often associated with poor prognosis (Mu et al., 2020). We used the random grouping method to ensure the independence of data between the training group and test group, and verified by the test group, which showed that the PET/MR mode radiology model based on the Habitat method has some generalization ability. We have also noted that some scholars used enhanced CT scanning images to process images with the Gaussian mixture model (another common analysis method in the field of habitat analysis) to identify cystic and solid tumor subregions and help ovarian cancer patients with accurate puncture (Beer et al., 2021). The Gaussian mixture model could aim to distinguish the heterogeneity of mixture since its birth, which was undoubtedly in line with the application background of the aforementioned research. Just as the Gaussian mixture model could better distinguish the cystic and solid parts of ovarian cancer, our research used the Otsu model with maximized variance between groups to identify areas with more active metabolism to represent the tumor proliferation activity (Ki-67) and prognosis. Our outlook for this technology provides a noninvasive method to evaluate the activity of local lesions in chemotherapy for patients with high-grade serous ovarian cancer, which is also related to the overall survival time of patients.

At the same time, this study has the following shortcomings (Torre et al., 2018): a single center, that is, the lack of effective external verification, will undoubtedly greatly affect the generalization and application ability of the model (Lisio et al., 2019); the PET modal images processed by Habitat have been verified by the literature, but CT images were not included in this study due to cautious attitude and early development (Kohn and Ivy, 2017). For the study of PFS, there is a lack of further involvement of clinical factors, which needs further multicenter and large-scale data research.

Conclusion

Noninvasive imaging prediction indicators based on PET images can guide the prognosis stratification of ovarian cancer, which is related to the expression of Ki-67 in tumor tissues, and the accuracy of Habitat is improved. In the diagnosis and treatment of ovarian cancer, it is important to use a variety of technical means to guide the prognosis and molecular typing, especially for noninvasive means.

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 Shengjing Hospital Ethics Committee. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

XW: data curation; formal analysis; investigation; project administration; resources; supervision; validation; visualization; roles/writing—original draft; and writing—review and editing. CX: software and writing—review and editing. MG: methodology and writing—review and editing. HS: conceptualization; funding acquisition; and writing—review and editing.

Funding

This work was sponsored in part by the National Natural Science Foundation of China (No. 82071967), LIAONING Science Natural Science Foundation (No. 2019-MS-373), Shenyang High Level Innovative Talents Support Program (No. RC210138), 345 Talent Project (No. M0377), and Outstanding Scientific Fund of Shengjing Hospital (No. M1149).

Conflict of interest

Publisher's note

References

Acar, E., Turgut, B., Yiğit, S., and Kaya, G. (2019). Comparison of the volumetric and radiomics findings of 18F-FDG PET/CT images with immunohistochemical prognostic factors in local/locally advanced breast cancer. *Nucl. Med. Commun.* 40 (7), 764–772. doi:10.1097/mnm.0000000000001019

Antunovic, L., Gallivanone, F., Sollini, M., Sagona, A., Invento, A., Manfrinato, G., et al. (2017). [(18)F]FDG PET/CT features for the molecular characterization of primary breast tumors. *Eur. J. Nucl. Med. Mol. Imaging* 44 (12), 1945–1954. doi:10.1007/s00259-017-3770-9

Beer, L., Martin-Gonzalez, P., Delgado-Ortet, M., Reinius, M., Rundo, L., Woitek, R., et al. (2021). Ultrasound-guided targeted biopsies of CT-based radiomic tumour habitats: Technical development and initial experience in metastatic ovarian cancer. *Eur. Radiol.* 31 (6), 3765–3772. doi:10.1007/s00330-020-07560-8

Beer, L., Sahin, H., Bateman, N. W., Blazic, I., Vargas, H. A., Veeraraghavan, H., et al. (2020). Integration of proteomics with CT-based qualitative and radiomic features in high-grade serous ovarian cancer patients: An exploratory analysis. *Eur. Radiol.* 30 (8), 4306–4316. doi:10.1007/s00330-020-06755-3

Castellani, F., Nganga, E. C., Dumas, L., Banerjee, S., and Rockall, A. G. (2019). Imaging in the pre-operative staging of ovarian cancer. *Abdom. Radiol.* 44 (2), 685–696. doi:10.1007/s00261-018-1779-6

Chen, Y., and Du, H. (2018). The promising PARP inhibitors in ovarian cancer therapy: From Olaparib to others. *Biomed. Pharmacother.* 99, 552–560. doi:10.1016/j.biopha.2018.01.094

Cherezov, D., Goldgof, D., Hall, L., Gillies, R., Schabath, M., Müller, H., et al. (2019). Revealing tumor habitats from texture heterogeneity analysis for classification of lung cancer malignancy and aggressiveness. *Sci. Rep.* 9 (1), 4500. doi:10.1038/s41598-019-38831-0

Deng, S. M., Zhang, W., Zhang, B., Chen, Y. Y., Li, J. H., and Wu, Y. W. (2015). Correlation between the uptake of 18F-fluorodeoxyglucose (18F-fdg) and the expression of proliferation-associated antigen ki-67 in cancer patients: A meta-analysis. *PLoS One* 10 (6), e0129028. doi:10.1371/journal.pone.0129028

Deng, Y., Chen, C., Hua, M., Xi, Q., Liu, R., Yang, S., et al. (2015). Annexin A2 plays a critical role in epithelial ovarian cancer. *Arch. Gynecol. Obstet.* 292 (1), 175–182. doi:10.1007/s00404-014-3598-5

Grove, O., Berglund, A. E., Schabath, M. B., Aerts, H., Dekker, A., Wang, H., et al. (2021). Correction: Quantitative computed tomographic descriptors associate tumor shape complexity and intratumor heterogeneity with prognosis in lung adenocarcinoma. *PLoS One* 16 (3), e0248541. doi:10.1371/journal.pone.0248541

Jayson, G. C., Kohn, E. C., Kitchener, H. C., and Ledermann, J. A. Ovarian cancer. *Lancet* (2014) 384(9951):1376–1388. doi:10.1016/s0140-6736(13)62146-

Kemppainen, J., Hynninen, J., Virtanen, J., and Seppänen, M. (2019). PET/CT for evaluation of ovarian cancer. *Semin. Nucl. Med.* 49 (6), 484–492. doi:10.1053/j.semnuclmed.2019.06.010

Kohn, E. C., and Ivy, S. P. (2017). Whence high-grade serous ovarian cancer. *Am. Soc. Clin. Oncol. Educ. Book.* 37, 443–448. doi:10.1200/edbk_174718

Kong, Z., Li, J., Liu, Z., Liu, Z., Zhao, D., Cheng, X., et al. (2019). Radiomics signature based on FDG-PET predicts proliferative activity in primary glioma. *Clin. Radiol.* 74 (10), 815. doi:10.1016/j.crad.2019.06.019

Limei, Z., Yong, C., Yan, X., Shuai, T., Jiangyan, X., and Zhiqing, L. (2013). Accuracy of positron emission tomography/computed tomography in the diagnosis and restaging for recurrent ovarian cancer: A meta-analysis. *Int. J. Gynecol. Cancer* 23 (4), 598–607. doi:10.1097/IGC.0b013e31828a183c

Lisio, M. A., Fu, L., Goyeneche, A., Gao, Z. H., and Telleria, C. (2019). High-grade serous ovarian cancer: Basic sciences, clinical and therapeutic standpoints. *Int. J. Mol. Sci.* 20 (4), E952. doi:10.3390/ijms20040952

Lu, H., Arshad, M., Thornton, A., Avesani, G., Cunnea, P., Curry, E., et al. (2019). A mathematical-descriptor of tumor-mesoscopic-structure from computed-tomography images annotates prognostic- and molecular-phenotypes of epithelial ovarian cancer. *Nat. Commun.* 10 (1), 764. doi:10.1038/s41467-019-08718-9

Mayoral, M., Paredes, P., Saco, A., Fusté, P., Perlaza, P., Tapias, A., et al. (2018). Correlation of (18)F-FDG uptake on PET/CT with Ki67 immunohistochemistry in pre-treatment epithelial ovarian cancer. *Rev. Esp. Med. Nucl. Imagen Mol.* 37 (2), 80–86. doi:10.1016/j.remn.2017.07.005

Mu, W., Liang, Y., Hall, L. O., Tan, Y., Balagurunathan, Y., Wenham, R., et al. (2020). F-FDG PET/CT habitat radiomics predicts outcome of patients with cervical cancer treated with chemoradiotherapy. *Radiol. Artif. Intell.* 2 (6), e190218. doi:10.1148/ryai.2020190218

O'Connor, O. J., Prakash, P., Cronin, C. G., McDermott, S., and Blake, M. A. (2013). PET/CT in the imaging of ovarian Cancer. *Front. Biosci.* 5, 141–153. doi:10.2741/e603

Park, J. E., Kim, H. S., Kim, N., Park, S. Y., Kim, Y. H., and Kim, J. H. (2021). Spatiotemporal heterogeneity in multiparametric physiologic MRI is associated with patient outcomes in IDH-wildtype glioblastoma. *Clin. Cancer Res.* 27 (1), 237–245. doi:10.1158/1078-0432.Ccr-20-2156

Parra, N. A., Lu, H., Choi, J., Gage, K., Pow-Sang, J., Gillies, R. J., et al. (2019). Habitats in DCE-MRI to predict clinically significant prostate cancers. *Tomography* 5 (1), 68–76. doi:10.18383/j.tom.2018.00037

Pinho, D. F., King, B., Xi, Y., Albuquerque, K., Lea, J., and Subramaniam, R. M. (2020). Value of intratumoral metabolic heterogeneity and quantitative (18)F-fdg PET/CT parameters in predicting prognosis for patients with cervical cancer. *AJR. Am. J. Roentgenol.* 214 (4), 908–916. doi:10.2214/ajr.19.21604

Qiu, D., Cai, W., Zhang, Z., Li, H., and Zhou, D. (2019). High ki-67 expression is significantly associated with poor prognosis of ovarian cancer patients: Evidence from a meta-analysis. *Arch. Gynecol. Obstet.* 299 (5), 1415–1427. doi:10.1007/s00404-019-05082-3

Schlüter, C., Duchrow, M., Wohlenberg, C., Becker, M. H., Key, G., Flad, H. D., et al. (1993). The cell proliferation-associated antigen of antibody ki-67: A very large, ubiquitous nuclear protein with numerous repeated elements, representing a new kind of cell cycle-maintaining proteins. *J. Cell. Biol.* 123 (3), 513–522. doi:10.1083/jcb.123.3.513

Torre, L. A., Trabert, B., DeSantis, C. E., Miller, K. D., Samimi, G., Runowicz, C. D., et al. (2018). Ovarian cancer statistics, 2018. *Ca. Cancer J. Clin.* 68 (4), 284–296. doi:10.3322/caac.21456

Tothill, R. W., Tinker, A. V., George, J., Brown, R., Fox, S. B., Lade, S., et al. (2008). Novel molecular subtypes of serous and endometrioid ovarian cancer linked to clinical outcome. *Clin. Cancer Res.* 14 (16), 5198–5208. doi:10.1158/1078-0432.Ccr-08-0196

Vargas, H. A., Huang, E. P., Lakhman, Y., Ippolito, J. E., Bhosale, P., Mellnick, V., et al. (2017). Radiogenomics of high-grade serous ovarian cancer: Multireader multi-institutional study from the cancer genome atlas ovarian cancer imaging research group. *Radiology* 285 (2), 482–492. doi:10.1148/radiol.2017161870

Vargas, H. A., Miccò, M., Hong, S. I., Goldman, D. A., Dao, F., Weigelt, B., et al. (2015). Association between morphologic CT imaging traits and prognostically relevant gene signatures in women with high-grade serous ovarian cancer: A hypothesis-generating study. *Radiology* 274 (3), 742–751. doi:10.1148/radiol.14141477

Verhaak, R. G., Tamayo, P., Yang, J. Y., Hubbard, D., Zhang, H., Creighton, C. J., et al. (2013). Prognostically relevant gene signatures of high-grade serous ovarian carcinoma. *J. Clin. Invest.* 123 (1), 517–525. doi:10.1172/jci65833

Verma, R., Correa, R., Hill, V. B., Statsevych, V., Bera, K., Beig, N., et al. (2020). Tumor habitat-derived radiomic features at pretreatment mri that are prognostic for progression-free survival in glioblastoma are associated with key morphologic attributes at histopathologic examination: A feasibility study. *Radiol. Artif. Intell.* 2 (6), e190168. doi:10.1148/ryai.2020190168

Xu, H., Lv, W., Feng, H., Du, D., Yuan, Q., Wang, Q., et al. (2020). Subregional radiomics analysis of PET/CT imaging with intratumor partitioning: Application to prognosis for nasopharyngeal carcinoma. *Mol. Imaging Biol.* 22 (5), 1414–1426. doi:10.1007/s11307-019-01439-x

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 26 September 2022
DOI 10.3389/fphys.2022.980996

Machine-learning-derived radiomics signature of pericoronary tissue in coronary CT angiography associates with functional ischemia

Yan Feng[1,2], Zhihan Xu[3], Lin Zhang[2], Yaping Zhang[2], Hao Xu[4], Xiaozhong Zhuang[4], Hao Zhang[1,2]* and Xueqian Xie[1,2]*

[1]Department of Radiology, Shanghai General Hospital of Nanjing Medical University, Shanghai, China, [2]Department of Radiology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China, [3]DI CT Collaboration, Siemens Healthineers Ltd, Shanghai, China, [4]Department of Cardiology, Shanghai General Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Objectives: To determine the association between radiomics signature (Rad-signature) of pericoronary tissue (PCT) in coronary computed tomography angiography (CCTA) and CT-derived fractional flow reserve (CT-FFR), and explore the influential factors of functional ischemia.

Methods: We retrospectively included 350 patients who underwent CCTA from 2 centers, consisting of the training (n = 134), validation (n = 66), and testing (with CCTA and invasive coronary angiography, n = 150) groups. After evaluating coronary stenosis level in CCTA (anatomical CT), pericoronary fat attenuation index (FAI), and CT-FFR, we extracted 1,691 radiomic features from PCT. By accumulating and weighting the most contributive features to functional ischemia (CT-FFR $\leq$ 0.8) the Rad-signature was established using Boruta integrating with a random forest algorithm. Another 45 patients who underwent CCTA and invasive FFR were included to assure the performance of Rad-signature.

Results: A total of 1046 vessels in 350 patients were analyzed, and functional ischemia was identified in 241/1046 (23.0%) vessels and 179/350 (51.1%) patients. From the 47 features highly relevant to functional ischemia, the top-8 contributive features were selected to establish Rad-signature. At the vessel level, the area under the curve (AUC) of Rad-signature to discriminate functional ischemia was 0.83, 0.82, and 0.82 in the training, validation, and testing groups, higher than 0.55, 0.55, and 0.52 of FAI ($p < 0.001$), respectively, and was higher than 0.72 of anatomical CT in the testing group ($p = 0.017$). The AUC of the combined model (Rad-signature + anatomical CT) was 0.86, 0.85, and 0.83, respectively, significantly higher than that of anatomical CT and FAI ($p < 0.05$). In the CCTA-invasive FFR group, using invasive FFR as the standard, the mean AUC of Rad-signature was 0.83 ± 0.02. At the patient level, multivariate logistic regression analysis showed that Rad-signature of left anterior descending (LAD) [odds ratio (OR) = 1.72; $p = 0.012$] and anatomical CT (OR = 3.53; $p < 0.001$) were independent influential factors of functional

OPEN ACCESS

EDITED BY
Tao Tan,
Eindhoven University of Technology,
Netherlands

REVIEWED BY
Junjie Zhang,
Nanjing No. 1 Hospital, China
Hiroyuki Daida,
Juntendo University, Japan
Seokhun Yang,
Seoul National University Hospital,
South Korea
Bin Lu,
Chinese Academy of Medical Sciences
and Peking Union Medical College,
China

*CORRESPONDENCE
Xueqian Xie,
xiexueqian@hotmail.com
Hao Zhang,
zhanghao021@126.com

SPECIALTY SECTION
This article was submitted to Medical Physics and Imaging,
a section of the journal
Frontiers in Physiology

RECEIVED 29 June 2022
ACCEPTED 12 September 2022
PUBLISHED 26 September 2022

CITATION
Feng Y, Xu Z, Zhang L, Zhang Y, Xu H, Zhuang X, Zhang H and Xie X (2022), Machine-learning-derived radiomics signature of pericoronary tissue in coronary CT angiography associates with functional ischemia.
Front. Physiol. 13:980996.
doi: 10.3389/fphys.2022.980996

ischemia ($p < 0.05$). In the subgroup of nonobstructive (stenosis <50% in invasive coronary angiography) and obstructive (≥50%) cases of the testing group, the independent factor of functional ischemia was FAI of LAD (OR = 1.10; $p = 0.041$) and Rad-signature of LAD (OR = 2.45; $p = 0.042$), respectively.

Conclusion: The machine-learning-derived Rad-signature of PCT in CCTA demonstrates significant association with functional ischemia.

KEYWORDS

machine learning, fractional flow reserve, radiomics, functional ischemia, coronary computed tomography angiography

Introduction

Functional ischemia is a state in which the blood flow cannot meet the metabolic needs of tissues even in the absence of vascular obstruction (Moroni et al., 2021). Coronary computed tomography angiography (CCTA) is widely used for demonstrating the degree of coronary stenosis, but lacks the information of blood flow function. Invasive fractional flow reserve (FFR) is the gold standard for assessing coronary blood flow, and used for clinical decision-making in the treatment of coronary artery disease (CAD) (Pijls et al., 2007; Tonino et al., 2009; De Bruyne et al., 2012; Windecker et al., 2014). However, its clinical application is limited due to the invasive and high-cost pressure guide wire. Recently, CT-derived FFR (CT-FFR) based on hydrodynamics or deep learning has been developed to noninvasively measure lumen blood flow, avoiding additional radiation exposure and invasive procedure. A CT-FFR value ≤0.8 is considered coronary functional ischemia, and studies have proved that CT-FFR is highly consistent with invasive FFR (Tesche et al., 2017; Alex et al., 2020). In addition to assessing blood flow from the perspective of lumen, exploring the association between functional ischemia and pericoronary tissue (PCT) may provide more evidence for the diagnosis and treatment of CAD.

The bidirectional interaction between pericoronary adipose tissue (PCAT) and the adjacent coronary wall leads to coronary artery inflammation and plaque formation (Margaritis et al., 2013; Antonopoulos et al., 2014; Antonopoulos et al., 2015; Antonopoulos et al., 2017). Then, inflammatory cell infiltration and edema in PCAT result in increased CT attenuation, realizing the visualization and quantitative evaluation of vascular inflammation. A PCAT imaging biomarker, fat attenuation index (FAI), has been introduced as a strong and independent predictor of major adverse cardiovascular events (Dai et al., 2022). Ma et al. (2021) found that overall FAI was not significantly associated with abnormal FFR, but lesion-specific PCAT was independently related to abnormal FFR. FAI only incorporates PCAT density, but does not reflect the complex tissue structures around the coronary artery. Although the widely used CT attenuation range of PCAT is from −190 to −30 Hounsfield unit (HU), the PCAT attenuation of high-risk plaque with a “fat stranding” sign can reach 31 HU, because of the complex plaque components (Hedgire et al., 2018). In order to analyze PCT, machine learning-based radiomics allows to extract and analyze numerous quantitative features inside the medical images (Gillies et al., 2016).

Therefore, we hypothesize that the radiomic features of PCT are associated with functional coronary ischemia. Considering that some information may be omitted when solely determining the adipose tissue by CT attenuation, we aim to comprehensively analyze PCT by extracting the radiomic features of adipose and other tissues around the coronary artery. We established a machine-leaning-derived radiomics signature (Rad-signature) based on PCT to discriminate functional ischemia, compared with the conventional stenosis grading on CCTA (anatomical CT) and FAI, and then analyzed the influential factors of functional ischemia.

Materials and methods

Study sample

The patients were retrospectively included in two medical centers [Hospital-1: Shanghai General Hospital-North (city center); Hospital-2: Shanghai General Hospital-South (Songjiang new city)]. The inclusion criteria of subjects with CCTA were as follows: 1) patients with suspected or diagnosed CAD, defined by the guidelines (Fox et al., 2006; Fihn et al., 2012; Montalescot et al., 2013); 2) patients with CCTA from January to December 2020 to establish the model and validate its performance; 3) patients who underwent CCTA and invasive coronary angiography (ICA) from January 2014 to December 2019 to test the model and analyze the influential factors of functional ischemia by subgrouping the patients into obstructive and nonobstructive CAD. Additionally, patients who underwent CCTA and invasive FFR from January 2020 to December 2021 were included to assure the performance of Rad-signature in discriminating standard functional ischemia.

The exclusion criteria were: 1) history of coronary stenting or bypass surgery; 2) poor image quality and insufficient for diagnosis; 3) coronary artery variation; 4) total occlusion of

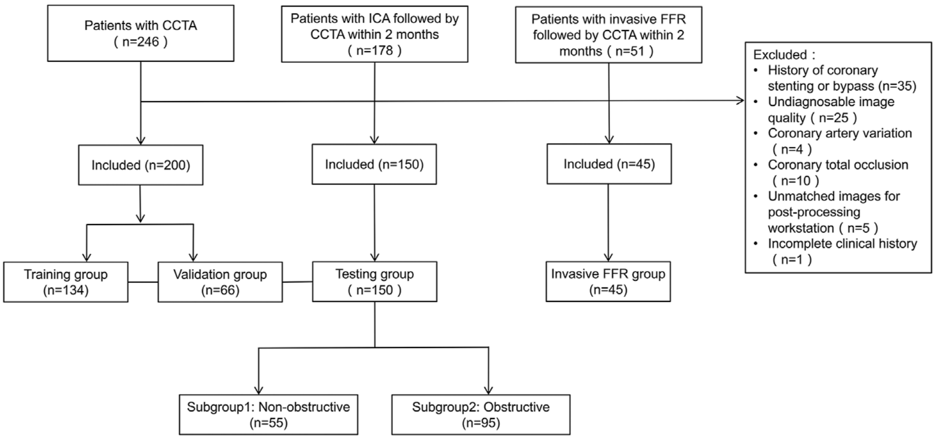


FIGURE 1
Flowchart of patient selection and grouping. CCTA, coronary computed tomography angiography; CAD, coronary artery disease; ICA, invasive coronary angiography.

coronary artery limiting the calculation of CT-FFR; 5) images that cannot be processed by the post-processing workstation, resulting in the failure of CT-FFR calculation or PCT segmentation; 6) time interval between CCTA and ICA (or invasive FFR) >2 months.

The patients were divided into four groups, including the training, validation (patients with CCTA), testing (with CCTA and ICA), and CCTA-invasive FFR groups. Figure 1 displays patient selection and grouping. The institutional review board approved this retrospective study and exempted the patient informed consent.

In the training, validation, and testing groups, a CT-FFR value ≤0.8 was defined as coronary functional ischemia (Tesche et al., 2017; Alex et al., 2020). In the invasive FFR group, an invasive FFR value ≤ 0.8 was defined as standard coronary functional ischemia. Basic characteristics and medical history were collected from the electronic medical record system, including age, sex, body mass index (BMI), smoking history, family history of CAD, diabetes, hypertension, and hyperlipidemia.

Coronary computed tomography angiography acquisition and anatomic evaluation

CCTA examination was performed on different CT equipment (Somatom Definition Flash and Somatom Force, Siemens Healthineers; Revolution CT and Discovery750 HD, GE Healthcare; Aquilion ONE, Canon Medical Systems) in the two centers (Supplementary Table S1). Beta-blockade was administrated for patients with a heart rate >90 beats per minute. Prospective gated CT scanning was performed. The scan covering range was from 1 cm below the tracheal carina to 2 cm below the left diaphragm. The tube voltage and current were automatically set using smart mode. The volume of contrast medium (Iopamidol 370 mg I/mL; Bracco) was calculated as 0.8 ml/kg, and the flow rate was 4–5 ml/sec. Then 20 ml of normal saline was injected at the same flow rate.

According to the standard segments of coronary artery recommended by the Society of Cardiovascular Computed Tomography of America (Leipsic et al., 2014), two experienced radiologists with 5- and 15-years experience in cardiovascular imaging independently evaluated the diameter stenosis level of the left anterior descending (LAD), left circumflex (LCx), and right coronary artery (RCA), using a dedicated image processing workstation (Advanced Workstation 4.6, GE Healthcare) with curved planar reformation, multiplanar reformation and volume rendering, and resolved the disagreement by mutual consultation. The vessels of <1.5 mm in diameter were excluded from anatomical evaluation.

Invasive coronary angiography and invasive fractional flow reserve methods

Invasive coronary angiography was performed by experienced interventional cardiologists according to local clinic standards, who were blinded to the results of CCTA. After radial artery or femoral artery puncture, left and right coronary angiography was performed. Based on the 5 projection angles of the left coronary artery and 2 of the right coronary artery, the degree of stenosis was quantitatively determined. A stenosis level ≥50% was considered obstructive CAD. An interventional cardiologist performed invasive FFR measurements based on coronary stenosis to determine the hemodynamic significance and the need for revascularization. The FFR pressure wire (Pressure Wire Abbott, St. Jude Medical, United States) was placed distal to the stenosis position. An FFR ≤ 0.80 was considered functional ischemia. The dose of contrast agent for each view was approximately 5 ml. The X-ray dose pattern was electrophysiological. The image acquisition rate was 7.5 frames per second.

Functional CT-derived fractional flow reserve

An on-site research prototype application (cFFR v3.5.0, Siemens Healthineers, not currently commercially available) was implemented for CT-FFR computation, which has been previously described and validated (Itu et al., 2016; Coenen et al., 2018; Tesche et al., 2018). This application calculates the CT-FFR value using a deep learning-based framework, which integrates the complex nonlinear relationship between the various features extracted from the coronary tree geometry and computes the blood flow of a coronary position (Itu et al., 2016). This model calculates CT-FFR based on reduced-order hydrodynamics. This application program can segment the coronary artery lumen on CT images to generate a coronary tree, and semiautomatically represent the CT-FFR value of any point in this tree. Before calculating CT-FFR, the observer needs to confirm or manually edit the automatically recognized coronary centerlines, and then the software generates a 3D pseudocolor map to comprehensively visualize the CT-FFR values of the coronary artery tree.

A radiologist with 10 years of cardiac imaging experience measured CT-FFR values, blinded to the results of medical history and other examinations. In this study, the CT-FFR values of LAD, LCx, and RCA were recorded. For vessels with stenosis, the measuring position was 2–3 cm distal from the stenosis. In the case of multiple stenoses in a single vessel, the distal end of the farthest lesion was measured. For normal vessels, the measuring position was the farthest end (about 1.5 mm in diameter). In order to match the positions of the invasive FFR and CT-FFR measurements, an independent radiologist, blinded to the functional results, marked the corresponding location on the CT-FFR image after identifying the location of the invasive FFR on the fluoroscopic image.

Pericoronary tissue segmentation and radiomic feature extraction

The same radiologist segmented PCT using dedicated software (Coronary plaque analysis v5.0.2, Frontier, Syngo. *via*, Siemens Healthineers). The software automatically segmented the image, and the radiologist manually modified them in case of inaccuracy. PCT is defined as all voxels extending outward from the outer wall of the vessel with a radius equal to the vessel diameter (Goeller et al., 2019). For LAD and LCx, the analyzed PCT was 4 cm long in the proximal segment of the vessels. For RCA, the analyzed tissue was 4 cm long in the proximal segment (1–5 cm from RCA ostium). The software then calculated FAI based on the segmented volume.

After importing the PCT mask, the dedicated software (Radiomics 13.0, Frontier, Syngo. *via*, Siemens Healthineers) automatically extracted and calculated 1,691 radiomic features of each vessel in about 10 s, including three major categories: 18 first-order, 75 texture, and 17 size and shape features.

Feature selection and rad-signature construction

To select stable and repeatable features, a radiologist with 10 years of experience in cardiac imaging randomly selected 30 vessels from the training group, segmented PCT and extracted radiomic features, and repeated the same procedure 1 month later. In these two measurements, the features with an intraclass correlation coefficient >0.8 were considered stable. Boruta algorithm integrated with random forest selected highly associated features by iteratively deleting features. Then the Boruta algorithm-selected features were converged by hierarchical clustering, and the most important features were selected as candidate features from each cluster. In this study, a random forest algorithm constructed a Rad-signature incorporating multiple features into one value (Dercle et al., 2020; Wu et al., 2020). The parameters in the training group were estimated by grid search with 10-fold cross-validation to avoid overfitting. The feature importance was assessed by the Gini impurity decreased overall decision trees. Each coronary vessel has a Rad-signature with a rad-score range of 0–1, indicating the probability of functional ischemia. The greater the Rad-signature, the more likely functional ischemia happens. In the CCTA-invasive FFR group, 4-fold cross-validation was used to estimate the performance of Rad-signature in discriminating standard functional ischemia. Figure 2 shows the process of establishing Rad-signature and performance evaluation. Supplementary Figure S1 demonstrates a representative case of Rad-signature establishment.

Statistics

Continuous variables were represented by median (25% and 75% quartile), and the difference between groups was tested by the Mann-Whitney *U* test. Categorical variables were expressed by frequency (percentages) and tested by Chi-Square test. At the vessel level, the performance of Rad-signature, anatomical CT, and FAI in identifying functional ischemia (CT-FFR ≤ 0.8) was

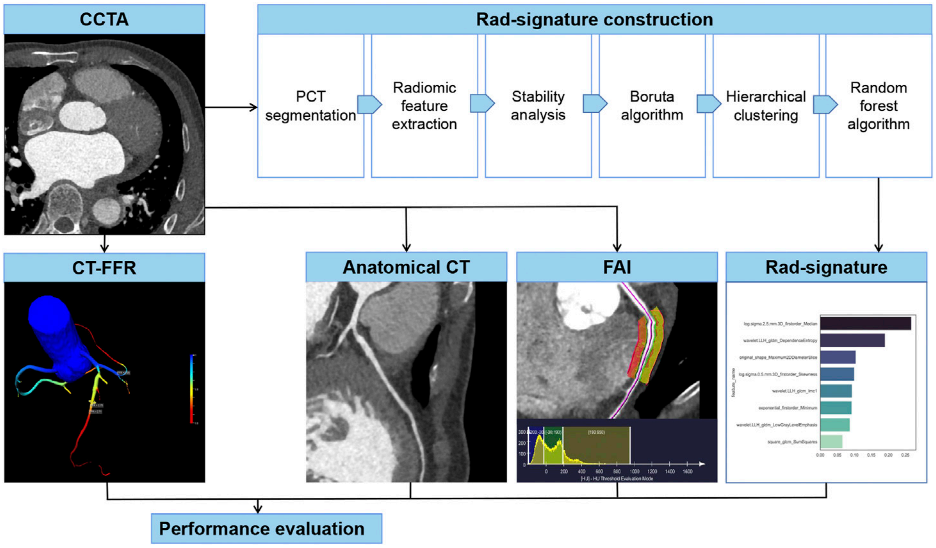


FIGURE 2
Establishment and evaluation of anatomical CT, FAI, and Rad-signature. CCTA, coronary computed tomography angiography. PCT, pericoronary tissue. CT-FFR, CT-derived fraction flow reserve; FAI, fat attenuation index.

evaluated by receiver operating characteristic (ROC) curve, area under the curve (AUC), accuracy, sensitivity, and specificity. The optimal cut-off value was indicated by Youden's index based on the training group. The difference between AUCs was evaluated by Delong's test. The incremental value of Rad-signature to anatomical CT was further evaluated by net reclassification index. At the patient level, the vessel with the lowest CT-FFR was used to define functional ischemia. The potential risk factors of functional ischemia were explored in the entire study sample by univariate and multivariate logistic regression analysis. Additionally, in the testing group (with CCTA and ICA), subgroup analysis was implemented to further assess the correlation between Rad-signature and functional ischemia in different risk levels (obstructive or nonobstructive CAD). The statistics was performed by open-source packages (R v3.6.0, http://www.Rproject.org; Python v3.7 with Scikit-survival library v0.13.2, https://scikit-survival.readthedocs.io/en/latest/). Supplementary Table S2 lists the details of software packages and functions. A $p < 0.05$ was considered statistically significant.

Results

Patient and vessel characteristics

In the training, validation, and testing groups, a total of 350 patients (66 years; 61–71 years) were eligible for this study from 442 candidates, including 134, 66, and 150 in the training, validation, and testing groups, respectively. In the three groups, 1046 vessels were analyzed including 350 LAD, 346 LCx, and 350 RCA, and functional ischemia (CT-FFR ≤ 0.8) was identified in 241/1046 (23.0%) vessels and 179/350 (51.1%) patients. In the three groups, anatomical CT detected coronary stenosis ≥50% in 377/1046 (36.0%) vessels and 236/350 (67.4%) patients. At the vessel level, the median FAI was −83.5 HU (−89.7 to −77.6 HU). In the testing group, ICA diagnosed obstructive CAD (stenosis ≥50%) in 148/447 (33.1%) vessels and in 95/150 (63.0%) patients. Tables 1, 2 summarize the clinical and CT characteristics of the three groups, respectively.

The CCTA-invasive FFR group included 55 vessels in another 45 patients (65 years; 59–71 years), including 7 RCA, 37 LAD, and 11 LCx. Functional ischemia (FFR ≤ 0.8) was detected in 13/55 (23.6%) vessels and 11/45 (24.4%) patients.

Radiomics feature selection

The stability analysis revealed 429 radiomic features with an intraclass correlation coefficient >0.8 out of 1691 features extracted from the training group. Boruta algorithm identified 47 candidate features which were highly associated with functional ischemia (Figure 3 and Supplementary Table S3). Hierarchical clustering demonstrated 8 distinct clusters of highly correlated radiomic features to functional ischemia (Figure 4). According to the feature importance, the most contributive feature from each of the 8 clusters was selected to establish the Rad-signature (Figure 5). The eight most contributive features included 4 texture features, 3 gray-level

TABLE 1 Clinical characteristics.

Characteristics	All	Training group	Validation group	Testing group	p-Value
	n = 350	n = 134	n = 66	n = 150	
Male	208 (59.4%)	79 (59.0%)	41 (62.1%)	88 (58.7%)	0.884
Age, years	66.0 [61.0; 71.0]	66.5 [62.0; 71.0]	65.0 [59.5; 71.0]	65.0 [60.0; 71.0]	0.599
Hypertension	197 (56.3%)	76 (56.7%)	37 (56.1%)	84 (56.0%)	0.992
Diabetes	124 (35.4%)	64 (47.8%)	23 (34.8%)	37 (24.7%)	**<0.001**
Hyperlipemia	124 (35.4%)	51 (38.1%)	28 (42.4%)	45 (30.0%)	0.153
Smoking	94 (26.9%)	23 (17.2%)	23 (34.8%)	48 (32.0%)	**0.005**
Family history of CAD	29 (8.29%)	10 (7.46%)	11 (16.7%)	8 (5.33%)	**0.019**
Body mass index, kg/m^2	24.3 [22.0; 26.4]	24.2 [21.5; 26.3]	24.3 [22.0; 26.8]	24.2 [22.4; 26.4]	0.678

Data are represented by median [25% and 75% quartile] or frequency (percentage).
p-Value represents the difference among the train, validation, and testing groups.
Bold values signify statistical significance.
CAD, coronary artery disease.

features, and 1 geometric feature. The correlation diagram and paired plots show minimal association among these top-8 contributive features (Supplementary Figure S2). The swarm plots show the distribution of radiomic features in the ischemia (CT-FFR ≤ 0.8) and nonischemic (CT-FFR > 0.8) groups (Supplementary Figure S3). The optimal cut-off value of Rad-signature to define functional ischemia was 0.202 according to the Youden index in the training group.

Identifying functional ischemia of rad-signature, fat attenuation index, and anatomical CT

The performance of Rad-signature, FAI, and anatomical CT in identifying functional ischemia was evaluated at the vessel level. In the training group, the AUC of Rad-signature was 0.83 [95% confidence interval (CI): 0.78–0.87], significantly higher than 0.55 (0.49–0.62) of FAI (DeLong's $p < 0.001$), but not higher than 0.79 (0.74–0.83) of anatomical CT ($p = 0.175$). The AUC of the combined model of Rad-signature and anatomical CT increased to 0.86 (0.82–0.90), higher than 0.83 (0.78–0.87) of only Rad-signature ($p = 0.054$). After adding Rad-signature to the traditional anatomical CT model, the net reclassification index was 0.25 (0.13–0.38, $p < 0.001$), which indicated the incremental value of Rad-signature in discriminating functional ischemia.

In the validation group, the AUC of Rad-signature was 0.82 (0.74–0.89), significantly higher than 0.55 (0.46–0.64) of FAI ($p < 0.001$), but not higher than 0.78 (0.71–0.85) of anatomical CT ($p = 0.443$). The AUC of the combined Rad-signature and anatomical CT model was 0.85 (0.79–0.91), significantly higher than that of FAI ($p < 0.001$) and anatomical CT ($p = 0.012$).

In the testing group, the AUC of Rad-signature was 0.82 (0.77–0.86), significantly higher than 0.52 (0.45–0.59) of FAI ($p < 0.001$) and 0.72 (0.65–0.79) of anatomical CT ($p = 0.017$). The AUC of the combined Rad-signature and anatomical CT model was 0.83 (0.77–0.91), significantly higher than that of FAI or anatomical CT ($p < 0.001$). Table 3 and Figure 6 show the comparison of AUCs among the three groups**.** In addition, in the subgroup of obstructive CAD (stenosis ≥ 50%) identified by ICA (284 vessels in 95 patients), the AUC of Rad-signature was 0.76 (0.68–0.84), significantly higher than 0.53 (0.42–0.63) of FAI ($p < 0.001$) (Supplementary Table S4). The AUC of the combined Rad-signature and anatomical CT model was 0.80 (0.73–0.88), significantly higher than that of other models ($p < 0.05$). In the analysis of subgroups divided by vessels, the RCA, LAD, and LCx groups, the AUC of the combined model was significantly higher than that of FAI and anatomical CT in the three subgroups (all $p < 0.05$) and showed an incremental value on the basis of Rad-signature (Supplementary Table S5).

In the CCTA-invasive FFR group, at the vessel level, using invasive FFR as the gold standard, the accuracy of CT-FFR was 85.4%, and the AUC of CT-FFR was 0.928 (0.85–1.00). The AUC of Rad-signature in the 4-fold cross-validation was 0.81, 0.86, 0.84, and 0.83, and the mean AUC was 0.83 ± 0.02. Supplementary Figure S4 shows the AUCs of Rad-signature in the cross-validation.

Influential factors of functional ischemia

For all 350 patients in this study, at the patient level, univariate logistic regression showed that sex, left ventricular mass derived from CCTA images (LVM-CT), Rad-signature of LAD (Rad-LAD), Rad-signature of LCx (Rad-LCx), the mean Rad-signature of RCA, LAD, and LCx (Rad-mean), and anatomical CT were significantly associated with functional ischemia ($p < 0.05$). Multivariate logistic regression showed that Rad-LAD [OR = 1.7 (95% CI: 1.1–2.6), $p = 0.018$] and

TABLE 2 CT characteristics.

Variables	Training group	Validation group	Testing group	*p*-value
	n = 134	*n* = 66	*n* = 150	
LVM-CT	127 [108; 151]	138 [113; 163]	134 [115; 152]	0.221
CT-FFR_RCA	89.0 [84.0; 93.0]	91.0 [86.0; 94.0]	92.0 [86.0; 94.0]	**0.011**
≤0.8	26 (19.4%)	8 (12.1%)	13 (8.7%)	
>0.8	108 (80.6%)	58 (87.9%)	137 (91.3%)	
CT-FFR_LAD	80.5 [67.0; 86.0]	80.5 [66.2; 87.8]	85.0 [76.0; 90.0]	**0.001**
≤0.8	67 (50%)	33 (50%)	52 (34.7%)	
>0.8	67 (50%)	33 (50%)	98 (65.3%)	
CT-FFR_LCx	91.0 [85.0; 94.8]	92.0 [87.0; 95.0]	92.0 [86.0; 95.0]	0.816
≤0.8	18 (13.4%)	7 (10.8%)	17 (11.6%)	
>0.8	116 (86.6%)	58 (89.2%)	130 (88.4%)	
CT-FFR_patient	76.0 [63.5; 84.0]	77.5 [65.0; 85.0]	82.5 [71.0; 88.0]	**0.001**
≤0.8	78 (58.2%)	37 (56.1%)	64 (42.7%)	
>0.8	56 (41.8%)	29 (43.9%)	86 (57.3%)	
FAI_LAD	−87.18 [−92.52; −82.17]	−84.08 [−92.75; −78.84]	−84.97 [−91.15; −78.18]	0.082
FAI_RCA	−86.29 [−91.90; −79.47]	−84.04 [−91.46; −79.78]	−84.16 [−90.22; −78.21]	0.229
FAI_LCx	−79.86 [−85.44; −75.03]	−79.45 [−83.70; −74.01]	−79.83 [−87.09; −74.06]	0.745
Anatomical CT_RCA				0.138
<50%	86 (64.2%)	48 (72.7%)	112 (74.7%)	
≥50%	48 (35.8%)	18 (27.3%)	38 (25.3%)	
Anatomical CT_LAD				0.125
<50%	50 (37.3%)	29 (43.9%)	74 (49.3%)	
≥50%	84 (62.7%)	37 (56.1%)	76 (50.7%)	
Anatomical CT_LCx				0.991
<50%	105 (78.4%)	52 (78.8%)	117 (78.0%)	
≥50%	29 (21.6%)	14 (21.2%)	33 (22.0%)	
Anatomical CT_patient				0.404
<50%	38 (28.4%)	24 (36.4%)	52 (34.7%)	
≥50%	96 (71.6%)	42 (63.6%)	98 (65.3%)	

LVM-CT, left ventricular mass on CT; CT-FFR, CT-derived fraction flow reserve; FAI, fat attenuation index.
Data are represented by median [25% and 75% quartile] or frequency (percentage).
Bold values signify statistical significance.

anatomical CT [OR = 3.5 (2.1–6.0), $p < 0.01$] were independent influential factors of functional ischemia.

The patients in the testing group ($n = 150$) were divided into two subgroups according to the ICA results, i.e., 95 (63.3%) patients with obstructive CAD (stenosis ≥ 50%) and 55 (36.7%) non-obstructive (<50%). In the obstructive subgroup, univariate logistic regression showed that age, LVM-CT, Rad-LAD, Rad-LCx, and Rad-mean were significantly associated with functional ischemia (all $p < 0.05$). Multivariate logistic regression showed that Rad-LAD [OR = 2.45 (95% CI: 1.08–6.27), $p = 0.042$] independently associated with functional ischemia. In the nonobstructive subgroup, univariate logistic regression showed that FAI-LAD ($p = 0.026$) and Rad-LCx ($p = 0.067$) were strongly associated with functional ischemia. Multivariate logistic regression showed that FAI-LAD [OR = 1.1 (1.0–1.2), $p = 0.041$] independently associated with functional ischemia. Tables 4, 5 show the logistic regression results.

Additionally, in the subgroup of patients with single-vessel disease, Rad-LAD, Rad-mean, and FAI-RCA were independently associated with functional ischemia. In the subgroup of patients with multiple vessel disease, Rad-LAD, Rad-LCx, and Rad-mean were independently associated with functional ischemia (Supplementary Table S6).

Discussion

Based on the comprehensive imaging features of PCT, we successfully established a radiomics signature (Rad-signature) to discriminate coronary functional ischemia. The AUC of the Rad-

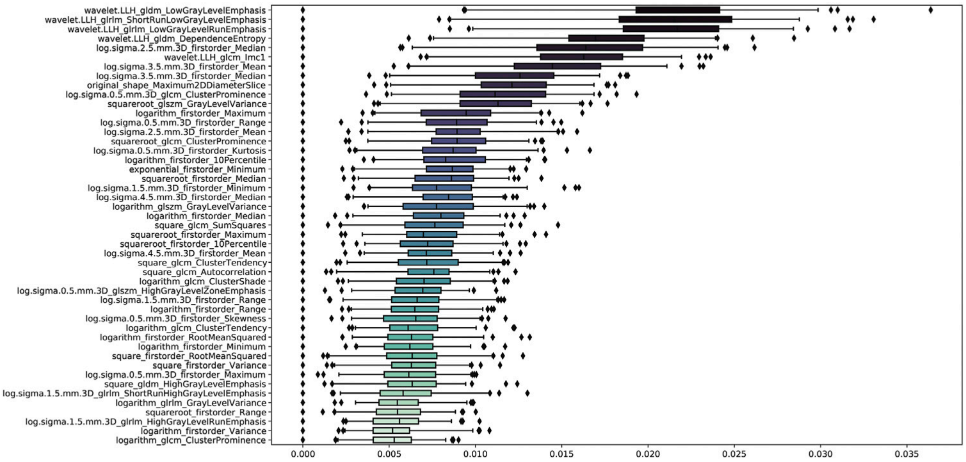


FIGURE 3
Importance ranking of 47 relevant radiomic features identified by Boruta algorithm in the training group. GLDM, gray level dependence matrix; GLRLM, gray level run length matrix; GLCM, gray level co-occurrence matrix; GLSZM, gray level size zone matrix.

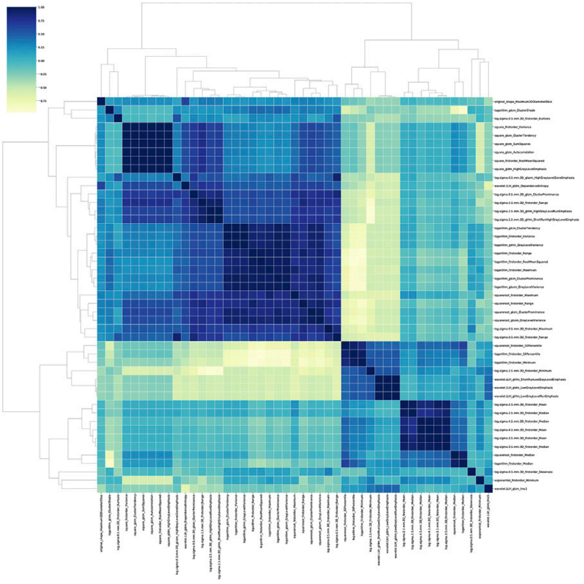


FIGURE 4
Correlation heatmap and dendrogram derived from hierarchical clustering in the training group. Boruta algorithm identified 8 different clusters derived from 47 important features. The darker color indicates higher correlation coefficient. GLDM, gray level dependence matrix; GLCM, gray level co-occurrence matrix; GLSZM, gray level size zone matrix; GLRLM, gray level run length matrix.

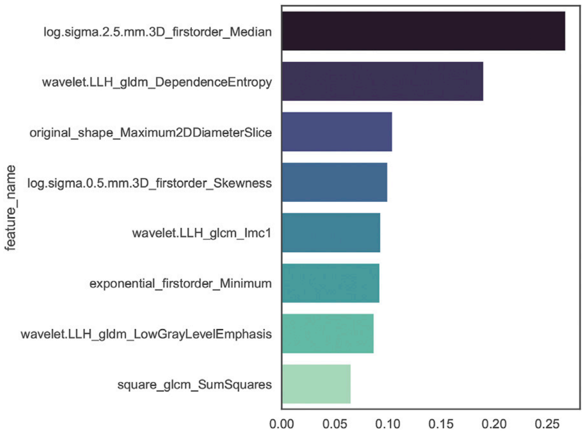


FIGURE 5
Importance ranking of the top-8 contributive features established by random forest algorithm of discriminating functional ischemia in the training group. The darker color indicates higher correlation. GLDM, gray level dependence matrix; GLCM, gray level co-occurrence matrix.

signature reached 0.82 in the validation and testing groups, which was significantly higher than that of FAI and numerically higher than anatomical CT but not statistically significant. The combined model of Rad-signature and anatomical CT increased the AUC compared with only anatomical CT, so adding Rad-signature has incremental value in discriminating functional ischemia. We also ensured that Rad-signature provides a good discrimination ability of standard functional ischemia (invasive FFR ≤ 0.8).

Coronary stenosis degree on CCTA is a widely-used indicator in diagnosing CAD with high accuracy and negative predictive value (Meijboom et al., 2008; Garg et al., 2016). However, conventional CCTA only provides morphological information but does not provides

TABLE 3 Performance metrics of all models at the vessel level.

Model	AUC	95% CI	Accuracy	95%CI	Sensitivity	Specificity	*p*-Value
Training group							
Rad-signature	0.83	0.78–0.87	0.72	0.67–0.76	0.80	0.69	0.020
Anatomical CT	0.79	0.74–0.83	0.71	0.67–0.76	0.73	0.71	**<0.001**
FAI	0.55	0.49–0.62	0.46	0.41–0.51	0.36	0.50	**<0.001**
Combined model (Rad-signature and anatomical CT)	0.86	0.82–0.90	0.81	0.77–0.85	0.79	0.81	N/A
Validation group							
Rad-signature	0.82	0.74–0.89	0.71	0.64–0.77	0.81	0.67	0.098
Anatomical CT	0.78	0.71–0.85	0.72	0.65–0.78	0.64	0.74	**0.012**
FAI	0.55	0.46–0.64	0.45	0.38–0.52	0.51	0.43	**<0.001**
Combined model (Rad-signature and anatomical CT)	0.85	0.79–0.91	0.82	0.76–0.87	0.68	0.86	N/A
Testing group							
Rad-signature	0.82	0.77–0.86	0.69	0.64–0.73	0.82	0.66	0.531
Anatomical CT	0.72	0.65–0.79	0.71	0.66–0.75	0.61	0.73	**<0.001**
FAI	0.52	0.45–0.59	0.51	0.46–0.56	0.54	0.50	**<0.001**
Combined model (Rad-signature and anatomical CT)	0.83	0.77–0.91	0.78	0.74–0.82	0.66	0.80	N/A

p-Values represents the difference between the AUC of the model and combined model (Rad-signature and anatomic CT).
Bold values signify statistical significance.
AUC, area under the ROC curve; CI, confidence interval; Anatomical CT, coronary stenosis grade on CCTA; FAI, fat attenuation index.

hemodynamic significance and lesion-specific ischemia, the functional ischemia (Park et al., 2012). In recent years, researchers have realized the inconsistency between morphologic stenosis and functional ischemia. Tonino et al. reported that 65% of the patients with moderate coronary stenosis and 20% with severe stenosis were not functionally significant (Tonino et al., 2010). Park et al. conducted a prospective study and found that the frequency of visual-functional mismatch was 40% between coronary angiography and FFR (Park et al., 2012). Furthermore, Pijls et al. found that the patients with CAD benefited more from FFR-guided revascularization strategies than morphological assessment by standard angiography, and 2-years mortality and incidence of myocardial infarction were significantly reduced (Pijls et al., 2010). CT-FFR noninvasively measures lumen blood flow, but it often depends on dedicated software, and sometimes the remote processing results are not timely, which limits the wide application of CT-FFR. The establishment of Rad-signature can be performed with an on-site workstation or personal computer, thus Rad-signature may be a practical and economical indicator. Since clinical decision-making should depend on the coronary functional significance, our study provides new evidence for the diagnosis and treatment of CAD by mining the functional correlation from anatomical CT.

Radiomics can extract a large number of imaging features from CCTA from a computational point of view. Meanwhile, machine learning can effectively select valuable information from numerous features and establish predictive models (Zhang et al., 2021). In our study, the Rad-signature was a powerful predictor and independent influential factor of functional ischemia. It derived from the eight most contributive radiomic features extracted from PCT on CCTA images, including 4 texture, 3 gray-level, and 1 geometric feature. Half of these contributive features were texture features, which were wavelet and log transformation based on the Gray-Level Co-occurrence Matrix which describes the spatial relationships of pixel pairs or voxel pairs with predefined gray intensity, and Gray Level Dependence Matrix which describes the grayscale relationship between the central pixel or the voxel and its neighborhood. The texture features may reflect the heterogeneity of PCT. The gray level features may reflect the intensity of PCT. Therefore, the main components of Rad-signature are not only the intensity information similar to FAI, but also the image heterogeneity information beyond the traditional image analysis standards. Similarly, Oikonomou et al. discussed that the radiomic features of PCAT derived from CCTA were highly associated with pathologically confirmed fibrosis and microvascular remodeling, and can differentiate patients with acute myocardial infarction and stable CAD, because they capture the spatial shifts in composition and lipid content of PCAT (Oikonomou et al., 2019). Vascular inflammation leads not only to plaque formation and lumen stenosis, but also to endothelial dysfunction and impaired vasodilation (Margaritis et al., 2013;

Antonopoulos et al., 2017), which may decrease distal flow reserve and cause functional ischemia.

We also conducted a subgroup analysis and evaluated the influential factors of functional ischemia. In the subgroup of obstructive CAD, Rad-LAD was an independent influential factor of functional ischemia. It suggests that the radiomic phenotype may be associated with pericoronary inflammation and easier to be captured in patients with obstructive CAD. In the nonobstructive subgroup, FAI-LAD rather than Rad-signature was an independent factor of functional ischemia. With the development of coronary atherosclerosis, the histological structure of PCT changes accordingly. Therefore, Rad-signature may be suitable for patients with moderate and severe coronary stenosis and FAI may be applicable for those with minimal and mild stenosis. Similarly, Antonopoulos et al. (2017) reported that FAI can be used to detect vascular inflammation at an early stage and change dynamically with the status of inflammation.

here are limitations. First, our study used a time-independent testing group instead of an external testing group. Multiple CT equipment enhanced the robustness of the model, but if external testing is adopted, the difference in imaging features caused by different CT devices between the two centers may interfere with the exploration of the relationship between PCT and functional ischemia. Second, we evaluated the ability of Rad-signature in discriminating functional ischemia defined as CT-FFR ≤ 0.8. Although CT-FFR ≤ 0.8 is widely considered coronary functional significance, noninvasive CT-FFR does not directly measure blood flow. Different reference standards may lead to a different selection of radiomic features, which means that further study is necessary to refine and calibrate the Rad-signature model. Furthermore, preliminary exploration in the CCTA-invasive FFR group suggested that Rad-signature may have a good ability in discriminating standard functional ischemia. However, more cases were needed to further validate the results.

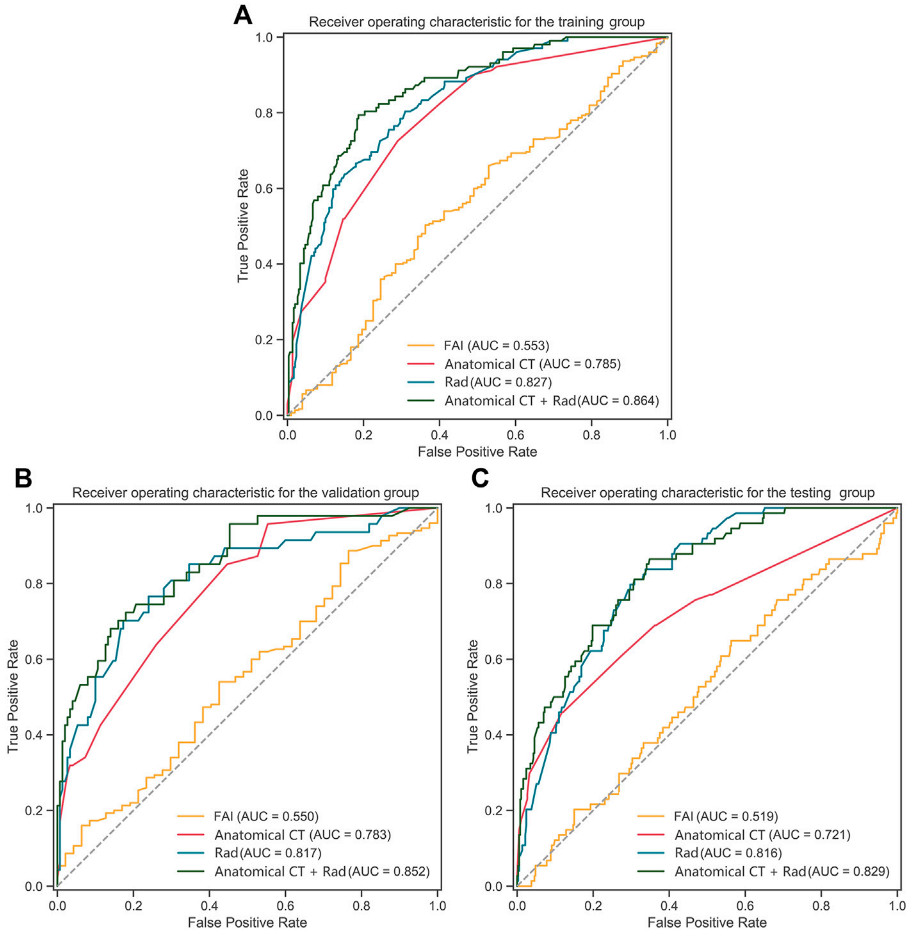


FIGURE 6
ROC curves of Rad-signature, FAI, anatomical CT and combined model (Rad-signature and anatomical CT) in **(A)** the training, **(B)** validation, and **(C)** testing groups. ROC, receiver operating characteristic; AUC, area under the ROC curve; Rad, Rad-signature; FAI, fat attenuation index.

TABLE 4 Logistic regression analysis of discriminating coronary functional ischemia in the whole study population at the patient level.

Variables	Univariate analysis		Multivariate analysis	
	OR (95% CI)	*p*-Value	OR (95% CI)	*p*-Value
Sex	0.49 (0.32–0.76)	**0.001**	0.78 (0.38–1.62)	0.506
Age	0.98 (0.96–1.01)	0.181		
Hypertension	0.96 (0.63–1.47)	0.861		
Diabetes	1.25 (0.8–1.94)	0.326		
Hyperlipemia	1.42 (0.92–2.22)	0.115		
Smoking	1.25 (0.78–2.01)	0.359		
Family history of CAD	0.83 (0.38–1.77)	0.628		
BMI	0.97 (0.91–1.03)	0.282		
LVM-CT	1.01 (1.00–1.01)	**0.023**	1.01 (1.00–1.01)	0.106
Rad-LAD	1.99 (1.57–2.56)	**<0.001**	1.72 (1.13–2.64)	**0.018**
Rad-RCA	3.17 (0.28–37.65)	0.352		
Rad-LCx	1.49 (1.19–1.88)	**<0.001**	1.22 (0.85–1.74)	0.274
Rad-mean	2.02 (1.59–2.61)	**<0.001**	1.25 (0.76–2.09)	
FAI-LAD	1.01 (0.98–1.03)	0.652		
FAI-RCA	1.02 (1.00–1.04)	0.097	1.00 (0.94–1.03)	0.980
FAI-LCx	1.00 (0.98–1.02)	0.878		
Anatomic CT	3.74 (2.32–6.13)	**<0.001**	3.52 (2.10–6.04)	**<0.001**

OR, odds ratio; CI, confidence interval; CAD, coronary artery disease; BMI, body mass index; LVM-CT, left ventricular mass on CT; Rad-LAD, rad-score of left anterior descending; Rad-RCA, rad-score of right coronary artery; Rad-LCx, rad-score of left circumflex; Rad-mean, the mean rad-score of LAD, RCA and LCx; FAI-LAD, fat attenuation index of LAD; FAI-RCA, FAI of RCA; FAI-LCx, FAI of LCx; Anatomic CT, coronary stenosis grade on CCTA.
Bold values signify statistical significance.

TABLE 5 Logistic regression analysis of discriminating coronary functional ischemia in subgroups divided by invasive coronary angiography in the testing group at the patient level.

Variables	Non-obstructive (stenosis < 50%) (*n* = 55)				Obstructive (stenosis ≥ 50%) (*n* = 95)			
	Univariate analysis		Multivariate analysis		Univariate analysis		Multivariate analysis	
	OR (95% CI)	*p*-Value	OR (95% CI)	*p*-Value	OR (95% CI)	*p*-Value	OR (95% CI)	*p*-Value
Sex	0.37 (0.10–1.24)	0.117			0.31 (0.12–0.74)	**0.01**	1.35 (0.27–7.26)	0.717
Age	0.98 (0.91–1.06)	0.652			0.95 (0.9–0.99)	**0.021**	0.94 (0.88–1.00)	0.054
Hypertension	0.84 (0.25–2.84)	0.782			0.83 (0.37–1.86)	0.648		
Diabetes	0.66 (0.13–2.58)	0.571			0.81 (0.31–2.09)	0.668		
Hyperlipemia	1.31 (0.39–4.38)	0.656			2.37 (0.9–6.6)	0.086	2.09 (0.64–7.22)	0.227
Smoking	1.5 (0.39–5.38)	0.538			1.88 (0.8–4.48)	0.148		
Family history of CAD	1.36 (0.06–15.27)	0.809			0.26 (0.01–1.85)	0.238		
BMI	0.92 (0.75–1.1)	0.37			1.03 (0.91–1.16)	0.671		
LVMCT.	1.01 (0.99–1.03)	0.241			1.01 (1–1.03)	**0.031**	1.01 (0.99–1.03)	0.267
Rad_LAD	1.60 (0.86–3.17)	0.15			2.07 (1.32–3.46)	**0.003**	2.45 (1.08–6.27)	**0.042**
Rad_RCA	0.78 (0.39–1.43)	0.447			0.88 (0.56–1.33)	0.566		
Rad_LCx	1.74 (0.97–3.26)	0.067	1.65 (0.89–3.21)	0.117	1.72 (1.10–2.87)	**0.025**	1.71 (0.80–4.30)	0.2
Rad_mean	1.54 (0.85–2.94)	0.166			1.92 (1.23–3.17)	**0.006**	0.84 (0.25–2.24)	0.739
FAI_LAD	1.11 (1.02–1.22)	**0.026**	1.10 (1.01–1.22)	**0.041**	1 (0.96–1.04)	0.894		
FAI_RCA	1.04 (0.98–1.12)	0.236			1.02 (0.98–1.07)	0.359		
FAI_LCx	1.02 (0.93–1.11)	0.711			1.01 (0.97–1.04)	0.745		

OR, odds ratio; CI, confidence interval; CAD, coronary artery disease; BMI, body mass index; LVM-CT, left ventricular mass on CT; Rad-LAD, rad-score of left anterior descending; Rad-RCA, rad-score of right coronary artery; Rad-LCx, rad-score of left circumflex; Rad-mean, the mean rad-score of LAD, RCA, and LCx; FAI, fat attenuation index.
Bold values signify statistical significance.

Conclusion

The machine-learning-derived radiomics model of Rad-signature of PCT showed a good ability in discriminating coronary functional ischemia. It may potentially become a noninvasive, fast, and economical indicator to screen functional ischemia before expensive invasive examinations. The combined model demonstrated the incremental value of Rad-signature to anatomical CT, rather than the superiority of Rad-signature alone in discriminating functional ischemia, which may help identify high-risk patients.

Data availability statement

The raw data supporting the conclusions 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 IRB of Shanghai General Hospital. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

Conception and design: XX, HZ, and YF. Administrative support: XX and HZ. Provision of study materials or patients: YF, XX, HX, and ZX. Collection and assembly of data: YF, YZ, LZ, XZ, and XX. Data analysis and interpretation: YF, ZX, and XX. All authors contributed to the article and approved the submitted version.

Funding

This study was sponsored by the Ministry of Science and Technology of China (2016YFE0103000), the National Natural Science Foundation of China (project no. 81971612 and 81471662), and Shanghai Jiao Tong University (ZH2018ZDB10).

Conflict of interest

Author ZX was employed by the company Siemens Healthineers Ltd.

The remaining 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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphys.2022.980996/full#supplementary-material

References

Alex, L. H., Paul, L. M., Richard, A. B., Mansi, T., Anna, B. R., Jacob, M., et al. (2020). CT-derived fractional flow reserve (FFRCT): From gatekeeping to roadmapping. *Can. Assoc. Radiol. J.* 71 (2), 201–207. doi:10.1177/0846537119893752

Antonopoulos, A. S., Margaritis, M., Coutinho, P., Digby, J., Patel, R., Psarros, C., et al. (2014). Reciprocal effects of systemic inflammation and brain natriuretic peptide on adiponectin biosynthesis in adipose tissue of patients with ischemic heart disease. *Arterioscler. Thromb. Vasc. Biol.* 34, 2151–2159. doi:10.1161/ATVBAHA.114.303828

Antonopoulos, A. S., Margaritis, M., Coutinho, P., Shirodaria, C., Psarros, C., Herdman, L., et al. (2015). Adiponectin as a link between type 2 diabetes and vascular NADPH oxidase activity in the human arterial wall: The regulatory role of perivascular adipose tissue. *Diabetes* 64, 2207–2219. doi:10.2337/db14-1011

Antonopoulos, A. S., Sanna, F., Sabharwal, N., Thomas, S., Oikonomou, E. K., Herdman, L., et al. (2017). Detecting human coronary inflammation by imaging perivascular fat. *Sci. Transl. Med.* 9 (398), eaal2658. doi:10.1126/scitranslmed.aal2658

Coenen, A., Kim, Y. H., Kruk, M., Tesche, C., De Geer, J., Kurata, A., et al. (2018). Diagnostic accuracy of a machine-learning approach to coronary computed tomographic angiography-based fractional flow reserve: Result from the MACHINE Consortium. *Circ. Cardiovasc. Imaging* 11 (6), e007217. doi:10.1161/CIRCIMAGING.117.007217

Dai, X., Hou, Y., Tang, C., Lu, Z., Shen, C., Zhang, L., et al. (2022). Long-term prognostic value of the serial changes of CT-derived fractional flow reserve and perivascular fat attenuation index. *Quant. Imaging Med. Surg.* 12 (1), 752–765. doi:10.21037/qims-21-424

De Bruyne, B., Pijls, N. H., Kalesan, B., Barbato, E., Tonino, P. A., Piroth, Z., et al. (2012). Fractional flow reserve guided PCI versus medical therapy in stable coronary disease. *N. Engl. J. Med.* 367 (11), 991–1001. doi:10.1056/NEJMoa1205361

Dercle, L., Fronheiser, M., Lu, L., Du, S., Hayes, W., Leung, D. K., et al. (2020). Identification of non–small cell lung cancer sensitive to systemic cancer therapies using radiomics. *Clin. Cancer Res.* 26 (9), 2151–2162. doi:10.1158/1078-0432.CCR-19-2942

Fihn, S. D., Gardin, J. M., Abrams, J., Berra, K., Blankenship, J. C., Dallas, A. P., et al. (2012). 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: A report of the American College of cardiology foundation/American heart association task

Force on practice guidelines, and the American College of physicians, American association for thoracic surgery, preventive cardiovascular nurses association, society for cardiovascular angiography and interventions, and society of thoracic surgeons. *J. Am. Coll. Cardiol.* 126 (25), e44–e164. doi:10.1016/j.jacc.2012.07.013

Fox, K., Garcia, M. A., Ardissino, D., Buszman, P., Camici, P. G., Crea, F., et al. (2006). Guidelines on the management of stable angina pectoris: Executive summary: The task Force on the management of stable Angina pectoris of the European society of cardiology. *Eur. Heart J.* 27 (11), 1341–1381. doi:10.1093/eurheartj/ehl001

Garg, P., Underwood, S. R., Senior, R., Greenwood, J. P., and Plein, S. (2016). Noninvasive cardiac imaging in suspected acute coronary syndrome. *Nat. Rev. Cardiol.* 13 (5), 266–275. doi:10.1038/nrcardio.2016.18

Gillies, R. J., Kinahan, P. E., and Hricak, H. (2016). Radiomics: Images are more than pictures, they are data. *Radiology* 278, 563–577. doi:10.1148/radiol.2015151169

Goeller, M., Tamarappoo, B. K., Kwan, A. C., Cadet, S., Commandeur, F., Razipour, A., et al. (2019). Relationship between changes in pericoronary adipose tissue attenuation and coronary plaque burden quantified from coronary computed tomography angiography. *Eur. Heart J. Cardiovasc. Imaging* 20, 636–643. doi:10.1093/ehjci/jez013

Hedgire, S., Baliyan, V., Zucker, E. J., Bittner, D. O., Staziaki, P. V., Takx, R. A., et al. (2018). Perivascular epicardial fat strand- ing at coronary CT angiography: A marker of acute plaque rupture and spontaneous coronary artery dissection. *Radiology* 287 (3), 808–815. doi:10.1148/radiol.2017171568

Itu, L., Rapaka, S., Passerini, T., Georgescu, B., Schwemmer, C., Schoebinger, M., et al. (2016). A machine-learning approach for computation of fractional flow reserve from coronary computed tomography. *J. Appl. Physiol.* 121 (1), 42–52. doi:10.1152/japplphysiol.00752.2015

Leipsic, J., Abbara, S., Achenbach, S., Cury, R., Earls, J. P., Mancini, G. J., et al. (2014). SCCT guidelines for the interpretation and reporting of coronary CT angiography: A report of the society of cardiovascular computed tomography guidelines committee. *J. Cardiovasc. Comput. Tomogr.* 8, 342–358. doi:10.1016/j.jcct.2014.07.003

Ma, S., Chen, X., Ma, Y., Liu, H., Zhang, J., Xu, L., et al. (2021). Lesion-specific peri-coronary fat attenuation index is associated with functional myocardial ischemia defined by abnormal fractional flow reserve. *Front. Cardiovasc. Med.* 8, 755295. doi:10.3389/fcvm.2021.755295

Margaritis, M., Antonopoulos, A. S., Digby, J., Lee, R., Reilly, S., Coutinho, P., et al. (2013). Interactions between vascular wall and perivascular adipose tissue reveal novel roles for adiponectin in the regulation of endothelial nitric oxide synthase function in human vessels. *Circulation* 127, 2209–2221. doi:10.1161/CIRCULATIONAHA.112.001133

Meijboom, W. B., Meijs, M. F., Schuijf, J. D., Cramer, M. J., Mollet, N. R., van Mieghemet, C. A. G., et al. (2008). Diagnostic accuracy of 64-slice computed tomography coronary angiography: A prospective, multicenter, multivendor study. *J. Am. Coll. Cardiol.* 52 (25), 2135–2144. doi:10.1016/j.jacc.2008.08.058

Montalescot, G., Sechtem, U., Achenbach, S., Andreotti, F., Arden, C., Budaj, A., et al. (2013). 2013 ESC guidelines on the management of stable coronary artery disease: The task Force on the management of stable coronary artery disease of the European society of cardiology. *Eur. Heart J.* 34 (38), 2949–3003. doi:10.1093/eurheartj/eht296

Moroni, F., Gertz, Z., and Azzalini, L. (2021). Relief of ischemia in ischemic cardiomyopathy. *Curr. Cardiol. Rep.* 23 (7), 80. doi:10.1007/s11886-021-01520-4

Oikonomou, E. K., Williams, M. C., Kotanidis, C. P., Desai, M. Y., Marwan, M., Antonopoulos, A. S., et al. (2019). A novel machine learning-derived radiotranscriptomic signature of perivascular fat improves cardiac risk prediction using coronary CT angiography. *Eur. Heart J.* 40, 3529–3543. doi:10.1093/eurheartj/ehz592

Park, S. J., Kang, S. J., Ahn, J. M., Shim, E. B., Kim, Y. T., Yun, S. C., et al. (2012). Visual-functional mismatch between coronary angiography and fractional flow reserve. *JACC. Cardiovasc. Interv.* 5 (10), 1029–1036. doi:10.1016/j.jcin.2012.07.007

Pijls, N. H., Fearon, W. F., Tonino, P. A., Siebert, U., Ikeno, F., Bornschein, B., et al. (2010). Fractional flow reserve versus angiography for guiding percutaneous coronary intervention in patients with multivessel coronary artery disease: 2-year follow-up of the FAME (fractional flow reserve versus angiography for multivessel evaluation) study. *J. Am. Coll. Cardiol.* 56 (3), 177–184. doi:10.1016/j.jacc.2010.04.012

Pijls, N. H., van Schaardenburgh, P., Manoharan, G., Boersma, E., Bech, J. W., van't Veer, M., et al. (2007). Percutaneous coronary intervention of functionally nonsignificant stenosis: 5-year follow-up of the DEFER study. *J. Am. Coll. Cardiol.* 49, 2105–2111. doi:10.1016/j.jacc.2007.01.087

Tesche, C., De Cecco, C. N., Albrecht, M. H., Duguay, T. M., Bayer, R. R., 2nd, Litwin, S. E., et al. (2017). Coronary CT angiography-derived fractional flow reserve. *Radiology* 285 (1), 17–33. doi:10.1148/radiol.2017162641

Tesche, C., De Cecco, C. N., Baumann, S., Renker, M., McLaurin, T. W., Duguay, T. M., et al. (2018). Coronary CT angiography-derived fractional flow reserve: Machine learning algorithm versus computational fluid dynamics modeling. *Radiology* 288 (1), 64–72. doi:10.1148/radiol.2018171291

Tonino, P. A., De Bruyne, B., Pijls, N. H., Siebert, U., Ikeno, F., van' t Veer, M., et al. (2009). Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. *N. Engl. J. Med.* 360 (3), 213–224. doi:10.1056/NEJMoa0807611

Tonino, P. A., Fearon, W. F., De Bruyne, B., Oldroyd, K. G., Leesar, M. A., Ver Lee, P. N., et al. (2010). Angiographic versus functional severity of coronary artery stenoses in the FAME study fractional flow reserve versus angiography in multivessel evaluation. *J. Am. Coll. Cardiol.* 55 (25), 2816–2821. doi:10.1016/j.jacc.2009.11.096

Windecker, S., Kolh, P., Alfonso, F., Collet, J. P., Cremer, J., Falk, V., et al. (2014). 2014 ESC/EACTS guidelines on myocardial revascularization: The task Force on myocardial revascularization of the European society of cardiology (ESC) and the European association for cardio-thoracic surgery (EACTS)developed with the special contribution of the European association of percutaneous cardiovascular interventions (EAPCI). *Eur. Heart J.* 35 (37), 2541–2619. doi:10.1093/eurheartj/ehu278

Wu, G., Woodruff, H. C., Shen, J., Refaee, T., Sanduleanu, S., Ibrahim, A., et al. (2020). Diagnosis of invasive lung adenocarcinoma based on chest CT radiomic features of part-solid pulmonary nodules: A multicenter study. *Radiology* 297 (2), 451–458. doi:10.1148/radiol.2020192431

Zhang, L., Sun, J., Jiang, B., Wang, L., Zhang, Y., and Xie, X. (2021). Development of artificial intelligence in epicardial and pericoronary adipose tissue imaging: A systematic review. *Eur. J. Hybrid. Imaging* 5 (1), 14. doi:10.1186/s41824-021-00107-0

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 14 October 2022
DOI 10.3389/fphys.2022.994304

OPEN ACCESS

EDITED BY
Yi Xu,
Shanghai Jiao Tong University, China

REVIEWED BY
Han Lv,
Beijing Friendship Hospital, Capital Medical University, China
Haiming Li,
Fudan University, China
Zhongxiang Ding,
Zhejiang University, China

*CORRESPONDENCE
Hongzan Sun,
sunhongzan@126.com

SPECIALTY SECTION
This article was submitted to Medical Physics and Imaging,
a section of the journal
Frontiers in Physiology

RECEIVED 14 July 2022
ACCEPTED 03 October 2022
PUBLISHED 14 October 2022

CITATION
Zhang Z, Li X and Sun H (2022), Development of machine learning models integrating PET/CT radiomic and immunohistochemical pathomic features for treatment strategy choice of cervical cancer with negative pelvic lymph node by mediating COX-2 expression.
Front. Physiol. 13:994304.
doi: 10.3389/fphys.2022.994304

Development of machine learning models integrating PET/CT radiomic and immunohistochemical pathomic features for treatment strategy choice of cervical cancer with negative pelvic lymph node by mediating COX-2 expression

Zhe Zhang, Xiaoran Li and Hongzan Sun*

Department of Radiology, Shengjing Hospital of China Medical University, Shenyang, China

Objectives: We aimed to establish machine learning models based on texture analysis predicting pelvic lymph node metastasis (PLNM) and expression of cyclooxygenase-2 (COX-2) in cervical cancer with PET/CT negative pelvic lymph node (PLN).

Methods: Eight hundred and thirty-seven texture features were extracted from PET/CT images of 148 early-stage cervical cancer patients with negative PLN. The machine learning models were established by logistic regression from selected features and evaluated by the area under the curve (AUC). The correlation of selected PET/CT texture features predicting PLNM or COX-2 expression and the corresponding immunohistochemical (IHC) texture features was analyzed by the Spearman test.

Results: Fourteen texture features were reserved to calculate the Rad-score for PLNM and COX-2. The PLNM model predicting PLNM showed good prediction accuracy in the training and testing dataset (AUC = 0.817, $p < 0.001$; AUC = 0.786, $p < 0.001$, respectively). The COX-2 model also behaved well for predicting COX-2 expression levels in the training and testing dataset (AUC = 0.814, $p < 0.001$; AUC = 0.748, $p = 0.001$). The wavelet-LHH-GLCM ClusterShade of the PET image selected to predict PLNM was slightly correlated

---

Abbreviations: AUC, area under the curve; COX-2, cyclooxygenase-2; CT, computed tomography; FIGO, federation international of gynecology and obstetrics; GLCM, grey-level co-occurrence matrix; GLDM, gray level dependence matrix; GLRLM, gray level run length matrix; GLSZM, gray level size zone matrix; IHC, immunohistochemical; IRS, immunoreactive scoring system; LASSO, least absolute shrinkage and selection operator; LNM, lymph node metastasis; MTV, metabolic tumor volume; NGTDM, neigbouring gray tone difference matrix; PET, positron emission tomography; PLN, pelvic lymph node; PLNM, pelvic lymph node metastasis; Rad-score, radiomics score; ROC, receiver operating characteristic; ROI, region of interest; SUV, standardized uptake value; TLG, total pathological glycolysis; 18F-FDG, 18F-fluorodeoxyglucose.

with the corresponding feature of the IHC image (r = −0.165, $p < 0.05$). There was a weak correlation of wavelet-LLL-GLRLM LongRunEmphasis of the PET image selected to predict COX-2 correlated with the corresponding feature of the IHC image (r = 0.238, $p < 0.05$). The correlation between PET image selected to predict COX-2 and the corresponding feature of the IHC image based on wavelet-LLL-GLRLM LongRunEmphasis is considered weak positive (r = 0.238, *p*=<0.05).

Conclusion: This study underlined the significant application of the machine learning models based on PET/CT texture analysis for predicting PLNM and COX-2 expression, which could be a novel tool to assist the clinical management of cervical cancer with negative PLN on PET/CT images.

KEYWORDS

PET/CT, lymphatic metastasis, radiomics, cyclooxygenase 2, machine learning

Introduction

Cervical cancer is the fourth most prevalent cancer and the fourth leading cause of female cancer, with more than 6,00,000 incidences and 3,00,000 death cases reported in 2020 (Sung et al., 2021). The most commonly used clinical treatments are radical hysterectomy with pelvic lymph node dissection. However, the prognosis of post-operative patients varied significantly due to tumor heterogeneity (Bhatla et al., 2021). Several studies have indicated the most relevant prognostic factor in early cervical cancer is pelvic lymph node metastasis (PLNM) (Rudtanasudjatum et al., 2011; Horn et al., 2014; Zyla et al., 2020; Federico et al., 2022; Wenzel et al., 2022). Kulisara et al. have proved that patients with lymph node metastasis (LNM) had poorer 5-year overall survival than the patients without LNM ($p < 0.05$) (Nanthamongkolkul and Hanprasertpong, 2018). In addition, lymphadenectomy may increase the probability of some complications including lower limb lymphedema, ileus, and chylous ascites (Yost et al., 2014; Kuroda et al., 2017; Nica et al., 2020; Umbreit et al., 2020). The accuracy of predicting LNM in cervical cancer patients is crucial for treatment decision-making.

Previous studies have indicated that PET/CT could be used for the evaluation of LNM as a preoperative imaging test, which is of vital importance for clinical strategies and individualized treatment (Sironi et al., 2006; Lv et al., 2014; Fasmer et al., 2020). Nevertheless, these studies usually used lymph node metabolism and diameter to assess LNM. Few studies have evaluated the metastasis of PLN with slightly higher FDG metabolism and diameter less than 1 cm on PET images which has limitations in detecting micrometastasis. Radiomics is rapidly gaining momentum and this technique is characterized by quantifying tumor heterogeneity through extraction of computational features using advanced computational algorithms. Texture parameters of radiomics features of PET images and IHC pathomic features could potentially be adopted to predict the PLNM for strategy choice of cervical cancer patients.

Substantial evidence suggests that cyclooxygenase-2 (COX-2), a key protein in prostaglandin metabolism, has a critical role in PLNM in cervical cancer (Ryu et al., 2000; Hoellen et al., 2016). Previous studies have indicated elevated COX-2 was strongly related to LNM in stage IB cervical cancer (Kang et al., 2006), the high COX expression has been revealed positive correlation with malignancy in the parametrial tumor tissue or LNM (Ryu et al., 2000). Other studies also found that high-level expression of COX-2 was correlated with a poorer prognosis, recurrence, low sensitivity of nedaplatin, and radiosensitivity (Kim et al., 2003; Chen et al., 2005; Ishikawa et al., 2006; Manchana et al., 2006; Jo et al., 2007; Huang et al., 2013; Stasinopoulos et al., 2013; Kato et al., 2015). In neoplasia, COX-2 stimulates cell proliferation which promotes angiogenesis through pathways involving an increase in VEGF production (Huang et al., 2013; Xu et al., 2014). It has been suggested that COX-2 expression may enhance LNM after the onset of lymphovascular space invasion (Khunamornpong et al., 2009; Hoellen et al., 2016). The heterogenic ^{18}F-FDG uptake was strongly related to the histopathological appearance in the tumor region. ^{18}F-FDG heterogenic uptake within the tumor was correlated with the heterogeneity of tumor histopathological tissues (Zhao et al., 2005; Henriksson et al., 2007). IHC assay demonstrated that tumor angiogenesis and cancer cell proliferation were significantly related to the enhancement of tumor heterogeneity. Therefore, the high expression of COX-2 played a connecting role between the increase of tumor heterogeneity and PLNM (Liu et al., 2011).

Based on texture parameters of radiomics features of PET images, the global and local-regional heterogeneities of ^{18}F-FDG distribution could be potentially assessed. Moreover, some mathematical methods were obtained to describe the relationships between their position in PET images and the gray-level intensity of pixels or voxels (Chicklore et al., 2013; O'Connor, 2017). In this study, we hypothesized that the overexpression of COX-2 promoted the increase of tumor heterogeneity and then caused the change of texture features

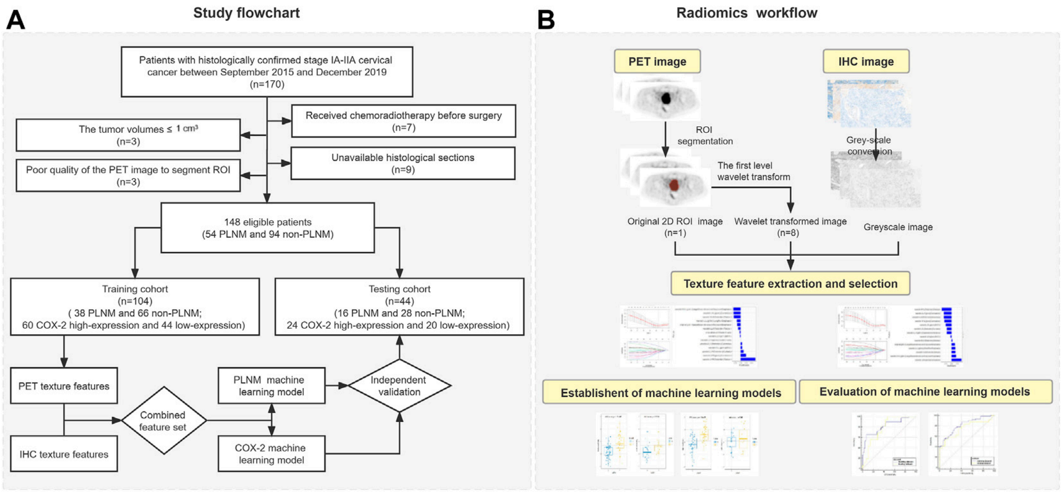


FIGURE 1
Study flowchart and radiomics workflow. **(A)** Study flowchart. **(B)** Radiomics workflow.

of radiomics derived from IHC and PET/CT imaging. The texture features of the primary tumor lesion may be correlated with PLNM in patients with early-stage cervical cancer. Therefore, we aimed to establish machine learning models of texture analysis that could predict PLNM and COX-2 expression based on PET/CT imaging to assist the clinical management of PLNM therapy in cervical cancer with PET/CT negative PLN.

Materials and methods

Radiomics workflow

The study flowchart and radiomics workflow are shown in Figure 1 and includes the collection and exclusion of patients, image acquisition, ROI segmentation, feature extraction and selection, establishment and evaluation of machine learning models, and correlation analysis between PET and IHC images with the same texture features.

Patients

This retrospective study consisted of 170 patients with histologically confirmed stage IA-IIA cervical cancer who underwent radical hysterectomy with pelvic node dissection between September 2015 and December 2019 in Shengjing Hospital of China Medical University. All patients underwent ^{18}F-FDG PET/CT scan within 1 week before treatment. The Hospital Institutional Review Board approved this study and informed consents were nor required due to retrospective nature. Inclusion criteria for cases: (Sung et al., 2021) Cervical squamous cell carcinoma confirmed by pathology and lymph node dissection performed in the patient; (Bhatla et al., 2021) Ia-IIa stage identified by 2021 Federation of Gynecology and Obstetrics (FIGO) staging (Bhatla et al., 2021); (Horn et al., 2014) The SUV$_{max}$ of PLNM was less than 2.5 and the diameter was less than 1 cm; (Zyla et al., 2020) Normal serum glucose level before PET/CT scanning; (Rudtanasudjatum et al., 2011) No other tumor or metabolic disease. A total of twenty-two patients were excluded from the sample. Seven of them were excluded because they had received chemoradiotherapy before surgery. Nine histological sections could not be obtained. The tumor volumes of three patients were less than 1 cm^3 to be unable to extract texture parameters. The PET image quality of the three cases was too bad to segment regions of interest (ROI). Eventually, 148 cases (54 PLNM and 94 non-PLNM) were enrolled in the study and randomly divided into a training dataset and a testing dataset according to the 7:3 ratio.

^{18}F-FDG PET/CT technique

The patients were all performed with the PET-CT (Discovery PET/CT 690; GE Healthcare, Chicago, Illinois, United States) and received an injection of 3.7 MBq/kg ^{18}F-FDG intravenous. The CT parameters were 3.27 mm slice thickness, 120 kV tube voltage, and 30–210 Ma. Then, with a three-dimensional acquisition mode and a matrix size of 192 × 192, PET data were captured at a speed of 1.5 min/bed (total of seven to eight beds). Using an iterative reconstruction algorithm of order subset expectation maximization, the PET image was reconstructed with twice iteration, 24 subsets and 6.4 mm Gaussian filter. In

TABLE 1 Patient characteristics.

	Training dataset (N = 104)	Testing dataset (N = 44)	p value
Age	51 (33–64)	53 (35–74)	0.266
Stage			
IA	27	11	0.743
IB	43	21	
IIA	34	12	
Differentiation			0.482
Well	24	12	
Moderate	65	23	
Poor	15	9	
PLNM			1.000
No	66	28	
Yes	38	16	
COX-2			0.856
Low expression	44	20	
High expression	60	24	
WBC ($\times 10^9$/L)	6.106	5.784	0.272
NEU (%)	60.622	59.048	0.537

the AW4.5 workstation (GE Healthcare), and all PET images were transferred. The conventional metabolic parameters of tumors in ^{18}F-FDG PET images for all patients consisted of total pathological glycolysis (TLG), the metabolic tumor volume (MTV), SUV_{max}, SUV_{peak}, and SUV_{mean}. SUV corrected for body weight and was measured automatically using a threshold of 42% SUV_{max} from the ROIs.

Immunohistochemical analysis

Department of Pathology in our hospital prepared all paraffin sections for cervical cancer. IHC staining was performed by Leica BOND MAX™ (Leica Biosystems). Goat anti-human COX-2 (1:400 dilution) polyclonal primary antibodies (Abcam) were used to incubate these sections, following species-appropriate secondary antibodies and the standard procedures were performed as in the previous report (Li et al., 2021a). Tumor sections were scanned using the Pannoramic MIDI slice scanner (3DHISTECH Ltd.) forming a digital image (×400) and analyzed by QuantCenter software with the Pannoramic viewer. The whole images were scanned by the DensitoQuant software and the analysis procedure was performed as in previous reports (Yeo et al., 2015; Li et al., 2021b). The immunoreactive scoring system (IRS) was utilized to assess the expression level of COX-2 (Kim et al., 2008). The IRS was derived from the addition of staining intensity (scored on a 0–3 scale: 0, negative; 1, weakly positive; 2, moderately positive; and 3, strongly positive) and staining extent (scored on a 0–4 scale: 0, no staining; 1, 1%–25% positive 2, 26%–50% positive; 3, 51%–75% positive; and 4, 76%–100% positive tumor cells). The level of COX-2 expression was classified as a dichotomous variable for high (IRS, 4–7) or low (IRS, 0–3) expression.

Extracting texture features of PET and immunohistochemical images

All PET images were loaded to 3D slicer (https://www.slicer.org) software 4.10.2 version. Two nuclear medicine physicians manually segmented independently the largest slice of all tumors in PET images to form 2D ROI, blinded to patient clinical information. Then, the texture features of ROI in PET images were extracted by the pyradiomics package (van Griethuysen JJMFedorov et al., 2017). The resampled voxel size was set to 1 mm × 1 mm × 1 mm to be isotropic of the image. The discretization of the grayscale was set to 25 bin width. The PET original images were transformed into eight images by the first level wavelet transform. Then the texture features were extracted from ROI based on PET original images and wavelet transformed images.

A pathologist randomly captured the cancer tissue area of the cervix on the digital IHC image (×20). The captured images that were native red/green/blue (RGB) images were converted to greyscale before computing the texture features (Kather et al.,

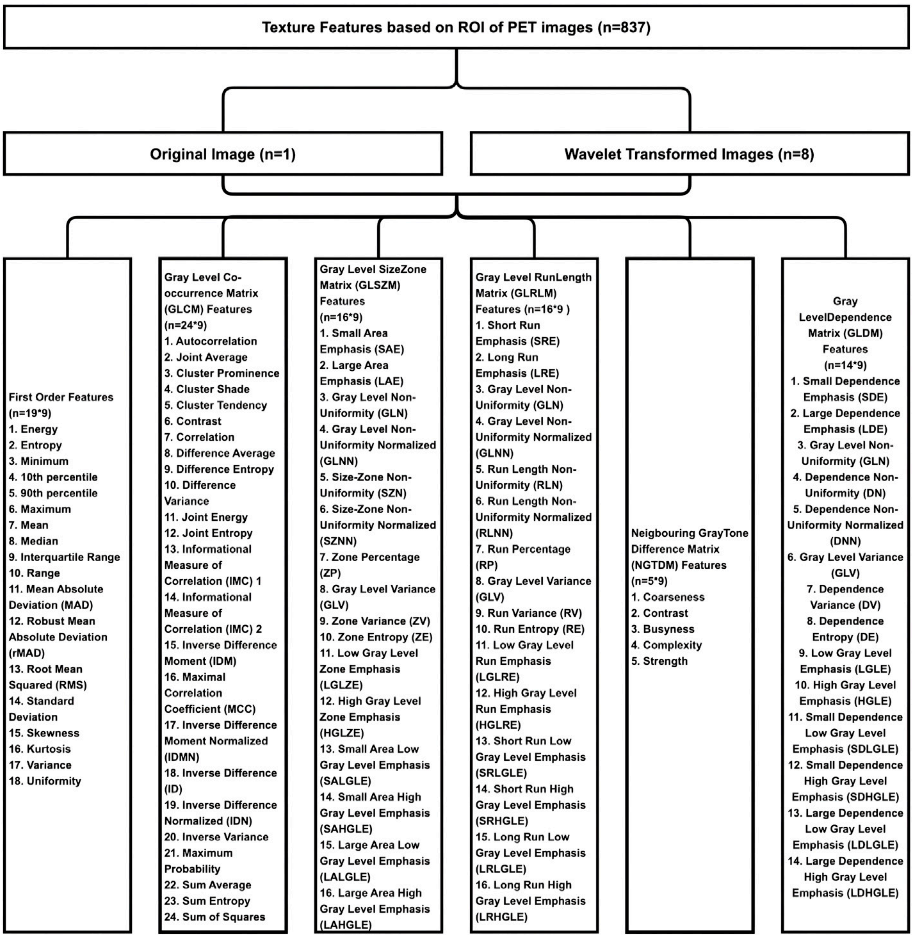


FIGURE 2
The result of extracting texture feature parameters of PET images. Eight hundred and thirty-seven texture features were extracted from the ROI in the PET image, including the first order features (n = 18), gray level co-occurrence matrix (GLCM) features (n = 24), gray level size zone matrix (GLSZM) features (n = 16), gray level run length matrix (GLRLM) features (n = 16), neigbouring gray tone difference matrix (NGTDM) features (n = 5), gray level dependence matrix (GLDM) features (n = 14), and wavelet features derived from one level of wavelet decompositions yielding eight derived images (n = 93*8).

2016). Then the same texture features as PET images were extracted with 3D slicer.

Dimensionality reduction of texture features

All texture feature parameters were standardized using the Z-score method. In the training dataset, with 10-fold cross-validation, the least absolute shrinkage and selection operator (LASSO) algorithm was used to filter clinical features, the conventional metabolic parameters, and texture features derived from PET images that could be used to predict PLNM and COX-2 expression. A classical metabolic parameter was generated using a linear combination of selected texture features of PET images of non-zero coefficient after dimensionality reduction. Afterward, both of them were weighted by their respective coefficients to establish the radiomics score (Rad-score) (Huang et al., 2016). The Rad-score was utilized to construct machine learning models to prognosticate PLNM and COX-2 expression.

Establishing and testing machine learning model

The Rad-score (PLNM) in the training dataset was utilized to establish the PLNM model for predicting PLNM with logistic regression algorithm. And the Rad-score (COX-2) was the

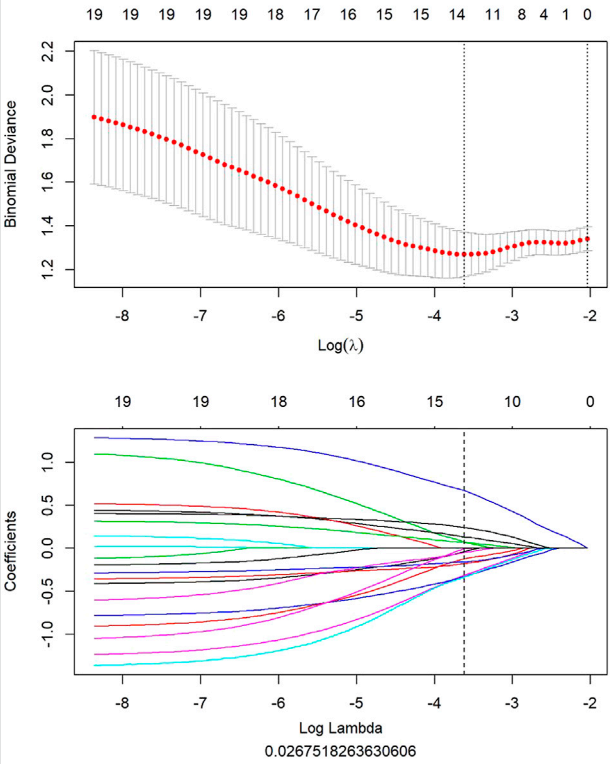


FIGURE 3
All features of PET image selected to predict PLNM by LASSO algorithm. When lambda was 0.027 in predicting PLNM with the LASSO algorithm, 14 most informative features were reserved to predict PLNM in the training dataset.

parameter for establishing the COX-2 model. The PLNM and COX-2 models in the testing dataset were tested independently.

Statistical analysis

The Mann-Whitney U test (continuous variables) or the Pearson chi-square test (rank variables) was used to evaluate the distribution of the clinical feature between the training and testing dataset. The correlation of selected PET texture features and the corresponding IHC texture images was analyzed with the Spearman correlation method. The differences in Rad-score (PLNM) between the PLNM group and non-PLNM group in all datasets and Rad-score (COX-2) between the COX-2 high expression group and COX-2 low expression group were analyzed with the Wilcoxon test. To evaluate the results of PLNM and COX-2 models, the ROC curve was used. All data processing, establishing machine learning models, and statistical analysis were performed with R software version 3.5.1 or SPSS software version 25.0 (IBM Corp., Armonk, NY, United States). A two-tailed $p < 0.05$ was considered statistically significant in all statistical analysis.

Results

The distribution of clinical characteristics of patients

The basic clinical characteristics of patients were summarized in Table 1. With a median age of 51 years, 104 patients were randomly assigned to the training dataset, containing 38 patients with PLNM. COX-2 was highly expressed in 60 patients in the training dataset. There were 44 patients with a median age of 53 years old in the testing dataset. Sixteen of them were PLNM positive. In the testing dataset, there were 24 patients with COX-2 high expression. Statistical analysis showed that there was no statistically significant difference in the distribution of all clinical features between the training and the testing dataset ($p > 0.05$).

Filtering and integration of features

A total of 837 texture features were extracted from the ROI in the PET image, including the First Order Features ($n = 18$), Gray Level Co-occurrence Matrix (GLCM) Features ($n = 24$), Gray Level Size Zone Matrix (GLSZM) Features ($n = 16$), Gray Level Run Length Matrix (GLRLM) Features ($n = 16$), Neigbouring Gray Tone Difference Matrix (NGTDM) Features ($n = 5$), Gray Level Dependence Matrix (GLDM) Features ($n = 14$) and Wavelet Features derived from one level of Wavelet decomposistions yielding eight derived images ($n = 93*8$). The detailed texture feature parameters of PET images were shown in Figure 2. The same 837 texture features were extracted based on the IHC images.

All clinical features, conventional metabolic parameters, and texture features derived from PET images were selected to predict PLNM. When Lambda was 0.027 in predicting PLNM with the LASSO algorithm, the 14 most informative features were reserved in the training dataset (Figure 3). And minimal binomial deviation for predicting PLNM was acquired with the 14 reserved features. Figure 4 showed that the coefficients of the reserved texture feature were used to predict PLNM with logistic regression algorithm. Partial regression coefficients were negative for nine of the reserved features and positive for five of the features. Then the reserved features were multiplied by their partial regression coefficients and linearly integrated into Rad-score (PLNM). The Rad-score (PLNM) was used to establish the PLNM model with logistic regression algorithm.

Rad-score (PLNM) = − 0.038 * wavelet-LLH-GLCM MCC
−0.177 * wavelet-LLL-GLRLM LongRunEmphasis
+0.076 * wavelet-HLL-GLCM MCC
−0.334 * wavelet-HHL-GLDM LargeDependenceLowGrayLevelEmphasis
+0.671 * wavelet-LHH-Firstorder Median
−0.311 * wavelet-HLH-Firstorder Median

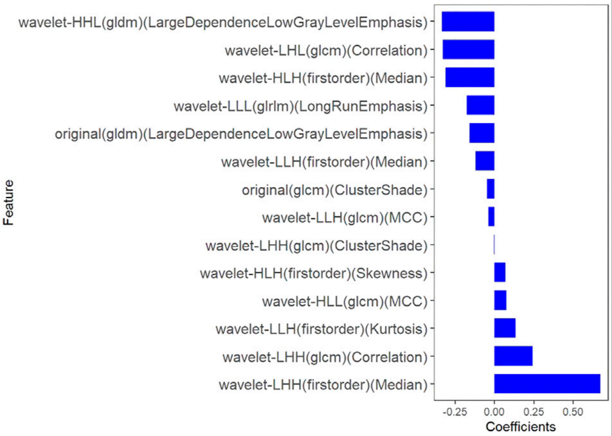


FIGURE 4
The coefficients histogram of texture features was selected to predict PLNM. Using the LASSO model, 14 corresponding texture features were selected to predict PLNM.

+0.133 * wavelet-LLH-Firstorder Kurtosis
−0.327 * wavelet-LHL-GLCM Correlation
−0.048 * original-GLCM ClusterShade
+0.242 * wavelet-LHH-GLCM Correlation
−0.121 * wavelet-LLH-Firstorder Median
+0.069 * wavelet-HLH-Firstorder Skewness
−0.158 * original-GLDM LargeDependenceLowGrayLevel Emphasis
−0.004 * wavelet-LHH-GLCM ClusterShade − 0.641.

Fourteen features were selected to predict COX-2 by LASSO 10-fold cross-validation in the training dataset (Figure 5). The histogram showed the coefficients of the reserved features in Figure 6. Partial regression coefficients were negative for eight of the reserved features and positive for six of the features. Rad-score (COX-2) was integrated according to the calculation formula below. The Rad-score (COX-2) was used to establish the COX-2 model with logistic regression algorithm.

Rad-score (COX-2) = − 0.044 * wavelet-LLH-GLCM MCC
−0.326 * wavelet-LLH-GLCM Correlation
−0.245 * wavelet-LHL-GLCM MCC
+0.165 * wavelet-LLL-GLRLM ShortRunEmphasis
+0.144 * original-GLDM LargeDependenceLowGrayLevel Emphasis
+0.198 * wavelet-HLL-GLDM LargeDependenceLowGray LevelEmphasis
−0.206 * wavelet-LHL-Firstorder Kurtosis
+0.17 * wavelet-HLH-Firstorder Mean
+0.042 * wavelet-HLH-Firstorder Median
−0.31 * wavelet-HHL-GLCM Correlation
−0.205 * wavelet-LLH-GLDM DependenceVariance

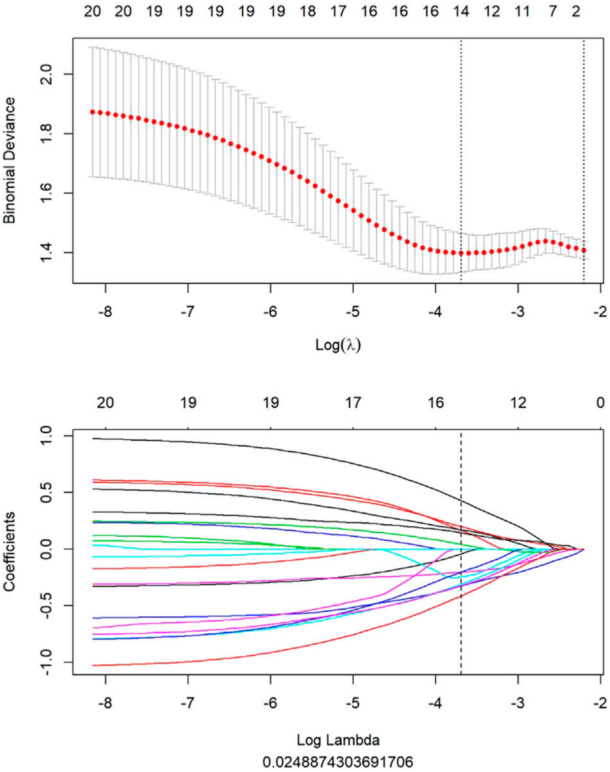


FIGURE 5
All features of PET image selected to predict COX-2 by LASSO algorithm. When lambda was 0.025 in predicting PLNM with the LASSO algorithm, 14 most informative features were reserved to predict PLNM in the training dataset.

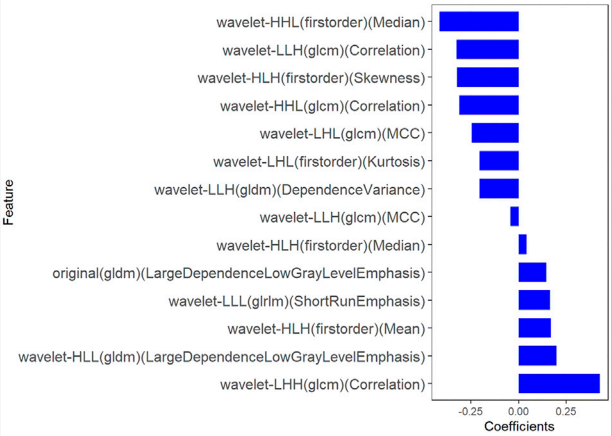


FIGURE 6
The coefficients histogram of texture features was selected to predict COX-2 expression. Using the LASSO model, 14 corresponding texture features were selected to predict COX-2 expression.

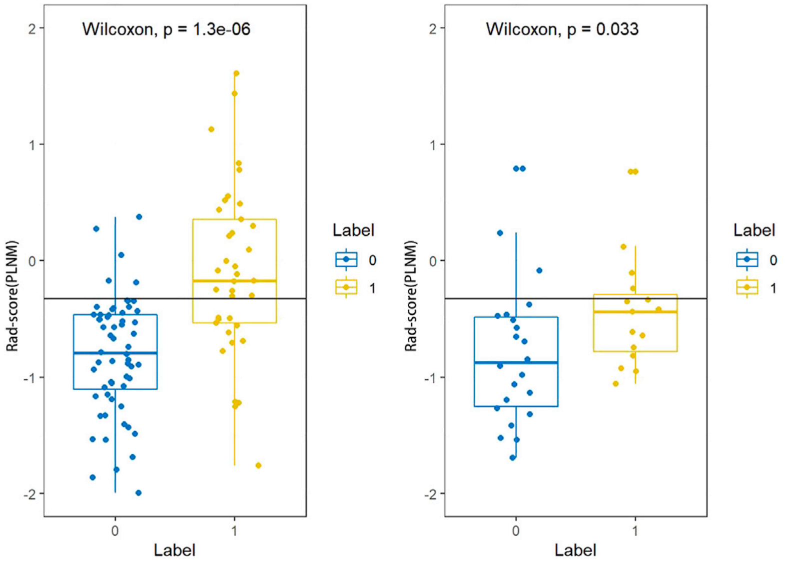


FIGURE 7
Rad-score (PLNM) distribution difference between PLNM negative subgroup and positive subgroup in the training dataset and testing dataset. The rad-score (PLNM) of patients with PLNM was higher than the rad-score (PLNM) of patients with non-PLNM in the training dataset and testing dataset ($p < 0.001$, $p < 0.05$, respectively).

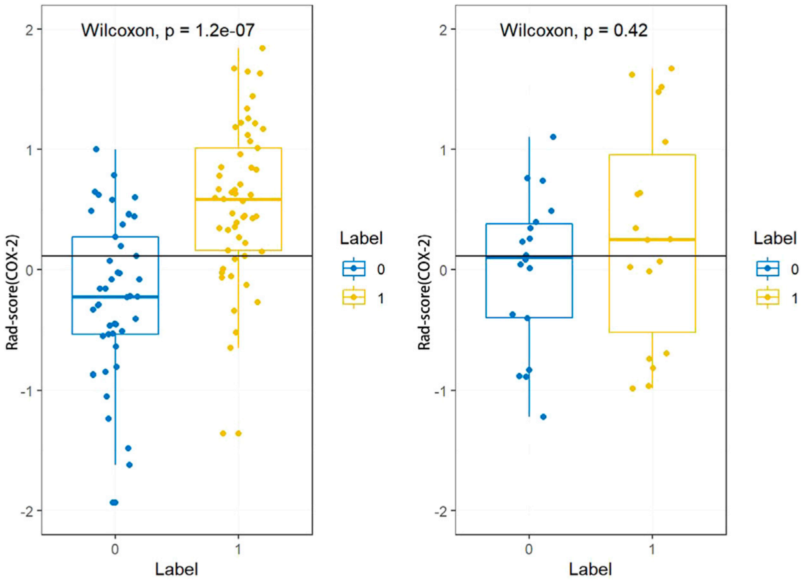


FIGURE 8
Rad-score (COX-2) distribution difference between COX-2 low expression subgroup and high expression subgroup in the training dataset and testing dataset. Rad-score (COX-2) of patients with high expression of COX-2 was higher than that with low expression of COX-2 in the training dataset ($p < 0.001$). In the testing dataset, there was no significant difference in Rad-score (COX-2) between patients with high COX-2 expression and patients with low COX-2 expression ($p < 0.05$).

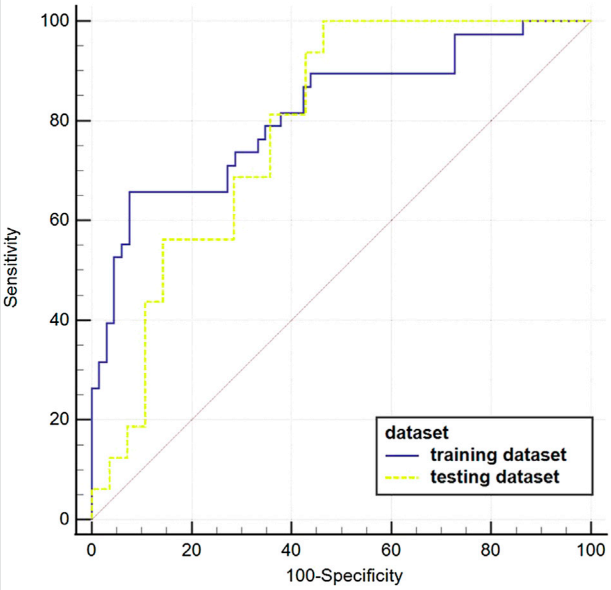


FIGURE 9
The ROC curves of the PLNM model in the training dataset and testing dataset. The blue ROC curve represents the training dataset; the yellow ROC curve represents the testing set. The AUC value of each model that was aimed to predict PLNM was 0.817 ($p < 0.001$) in the training dataset, and 0.786($p < 0.001$) in the testing dataset.

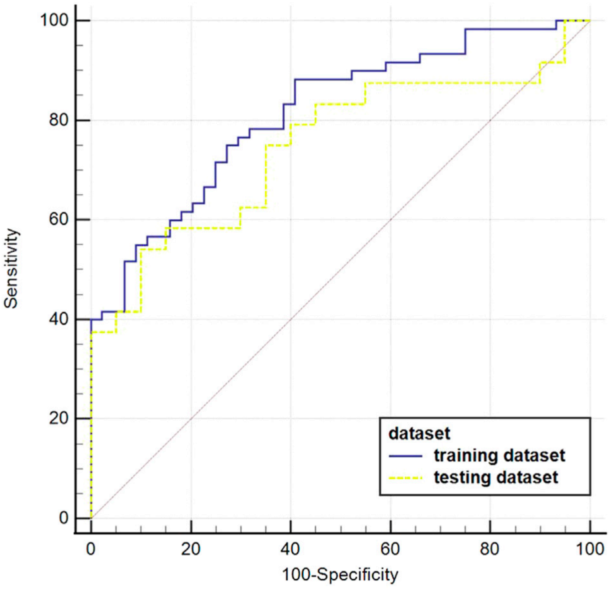


FIGURE 10
The ROC curves of the COX-2 model in the testing dataset. The blue ROC curve represents the training dataset, the yellow ROC curve represents the testing set. Both of them behaved well in predicting COX-2 expression levels in the training dataset and testing dataset (AUC = 0.814, $p < 0.001$; AUC = 0.748, $p = 0.001$, respectively).

-0.324 * wavelet-HLH-Firstorder Skewness
$+0.427$ * wavelet-LHH-GLCM Correlation
-0.415 * wavelet-HHL-Firstorder Median + 0.324.

Distribution differences of rad-scores in all datasets

The Rad-score (PLNM) of patients with PLNM was higher than the Rad-score (PLNM) of patients with Non-PLNM in the training dataset and testing dataset ($p < 0.001$, $p < 0.05$, respectively) (Figure 7). And the Rad-score (COX-2) of patients with high expression of COX-2 was higher than that with low expression of COX-2 in the training dataset ($p < 0.001$). Whereas, in the testing dataset, there was no statistically significant difference in Rad-score (COX-2) between patients with high COX-2 expression and patients with low COX-2 expression ($p < 0.05$) (Figure 8).

Evaluation of machine learning models

The AUC value of the machine learning model was shown in Figure 9 which was aimed to predict PLNM was 0.817 ($p < 0.001$) in the training dataset, and 0.786 ($p < 0.001$) in the testing dataset. And the COX-2 model also behaved well for predicting COX-2 expression levels in the training and testing dataset (AUC = 0.814, $p < 0.001$; AUC = 0.748, $p = 0.001$) (Figure 10). The sensitivity (Sen) value of the PLNM model was 65.8% in the training dataset, and 100.0% in the testing dataset. The specificity (Spe) value of the COX-2 model was 72.7% in the training dataset, and 90.0% in the testing dataset (Table 2).

The correlation of texture features derived from PET image with same texture features of immunohistochemical image

Table 3 showed the correlation between the texture features derived from ROI in the PET image selected to predict PLNM and the same texture features of the IHC image by Spearman correlation analysis. Only the wavelet-LHH (glcm) (ClusterShade) derived from the PET image was slightly correlated with the same feature of the IHC image ($r = -0.165$, $p < 0.05$). The correlation of texture features of the PET image selected to predict the COX-2 expression level with the IHC image's same texture features as illustrated in Table 4. There was a weak correlation that wavelet-LLL (glrlm) (LongRunEmphasis) derived from the ROI of the PET image correlated with the same feature of the IHC image ($r = 0.238$, $p < 0.05$).

TABLE 2 Evaluation of machine learning model prediction in the training and testing dataset.

Model	Training dataset				Testing dataset			
	AUC (95% CI)	*p*	Sen	Spe	AUC (95% CI)	*p*	Sen	Spe
PLNM model	0.817 (0.730–0.886)	<0.001	0.658	0.924	0.786 (0.636–0.895)	<0.001	1.000	0.538
COX-2 model	0.814 (0.726–0.884)	<0.001	0.750	0.727	0.748 (0.594–0.866)	0.001	0.542	0.900

AUC, area under the curve; Sen, sensitivity; Spe, specificity; CI, confidence interval.

TABLE 3 Correlation of PET texture parameters selected to predict the PLNM with same texture parameters of IHC image.

	Feature		Spearman test	
Filter	Class	Name	r	*p*
Wavelet-LLH	GLCM	MCC	−0.157	0.057
Wavelet-LLL	GLRLM	Long Run Emphasis	−0.002	0.980
Wavelet-HLL	GLCM	MCC	0.024	0.772
Wavelet-LHH	First order	Median	−0.025	0.767
Wavelet-HLH	First order	Median	−0.006	0.946
Wavelet-LLH	First order	Kurtosis	−0.070	0.400
Wavelet-LHL	GLCM	Correlation	0.026	0.754
Wavelet-HHL	GLDM	LDLGLE	0.085	0.302
Original	GLCM	Cluster Shade	0.115	0.163
Wavelet-LHH	GLCM	Correlation	−0.016	0.849
Wavelet-LLH	First order	Median	−0.008	0.920
Wavelet-HLH	First order	Skewness	−0.144	0.080
Original	GLDM	LDLGLE	−0.155	0.060
Wavelet-LHH	GLCM	Cluster Shade	−0.165*	0.045

GLCM, gray level co-occurrence matrix; MCC, maximum correlation coefficient; GLRLM, gray level run length matrix; GLDM, gray level dependence matrix; LDLGLE, large dependence low gray level emphasis. * indicating significant correlation.

TABLE 4 Correlation of PET texture parameters selected to predict the COX-2 expression level with same texture parameters of IHC image.

	Feature		Spearman test	
Filter	Class	Name	r	*p*
Wavelet-LLH	GLCM	MCC	−0.157	0.057
Wavelet-LLL	GLRLM	Short Run Emphasis	−0.042	0.614
Wavelet-LLH	GLCM	Correlation	**0.238***	**0.004**
Wavelet-LHL	GLCM	MCC	−0.098	0.238
Wavelet-LHL	First order	Kurtosis	−0.043	0.606
Wavelet-HLH	First order	Mean	0.123	0.135
Wavelet-HLL	GLDM	LDLGLE	0.079	0.338
Wavelet-HLH	First order	Median	−0.005	0.947
Wavelet-HHL	GLCM	Correlation	0.109	0.187
Original	GLDM	LDLGLE	0.050	0.545
Wavelet-LLH	GLDM	Dependence Variance	0.047	0.568
Wavelet-HLH	First order	Skewness	−0.144	0.080
Wavelet-LHH	GLCM	Correlation	−0.016	0.849
Wavelet-HHL	First order	Median	0.114	0.168

GLCM, gray level co-occurrence matrix; MCC, maximum correlation coefficient; GLRLM, gray level run length matrix; GLDM, gray level dependence matrix; LDLGLE, large dependence low gray level emphasis. * indicating significant correlation.
The bold values represents the wavelet-LLL (glrlm) (LongRunEmphasis) derived from the ROI of PET images correlated better with the same feature of IHC images relative to other texture features (r = 0.238, $p<0.05$).

Discussion

We provided machine learning models to study the diagnostic **value** of the textural features in PET images for predicting PLNM and performed well with good accuracy, sensitivity and specificity. Based on PET texture analysis predicting PLNM and COX-2 expression levels, this study revealed that machine learning models could assist clinical treatment of PLN in patients with early-stage cervical cancer. The rate of PLNM among patients with cervical squamous cell carcinoma stages IA-IIA was 36.54% in the training dataset and 36.36% in the testing dataset. The rate of COX-2 high expression among the patients was 57.69% in the training dataset and 54.55% in the testing dataset. The high expression of COX-2 were characteristics to predict PLNM associated with PET texture analysis and enriched level of COX-2 in the IHC images located in tumor, respectively. The Chi-square test or M-U analysis confirmed that the distribution of all clinical features was balanced between the training and testing dataset avoiding the inaccuracy and overfitting of the imbalanced feature distribution for building machine learning models.

The correlation between COX-2 high expression of primary tumor lesions and PLNM has been widely reported in cervical cancer (Ryu et al., 2000; Kim et al., 2003; Liu et al., 2011). The COX-2 model in this study performed well in the training and testing dataset (AUC = 0.814/0.748, $p < 0.001$/$p = 0.001$, respectively). And the specificity was 0.727 in the training dataset, and 0.900 in the testing dataset. The correlation feature of GLCM based on original and wavelet transformed images was also selected to calculate the Rad-score (COX-2). This

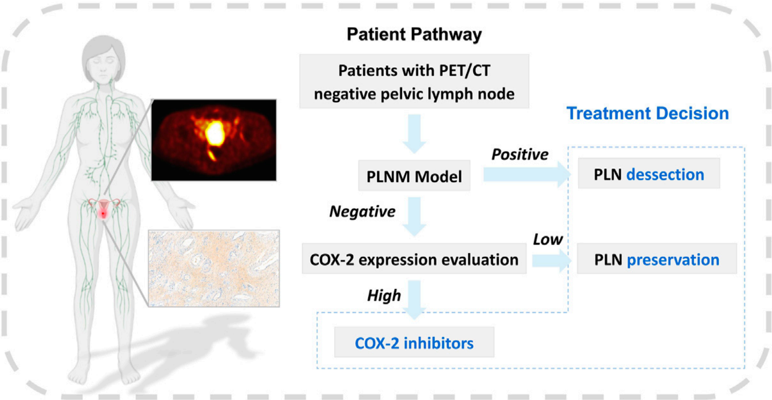


FIGURE 11
The workflow of machine learning models. The workflow of machine learning models in assisting clinical management of early-stage cervical cancer with negative PLN on PET/CT image.

indirectly confirmed that PET texture analysis of the primary tumor can predict that PLNM may be partly due to the high expression of COX-2. At present, the clinical evaluation of COX-2 mainly relies on IHC analysis, and the study ulitized PET/CT texture features for predicting COX-2 expression level before treatment is evaluated. We made a workflow based on the prediction of PLNM and COX-2 expression to assist the clinical management of PLN in early-stage cervical cancer (Figure 11). Kim et al. (2008) have confirmed that images with para-aortic lymph node recurrence possessed valuable expression of COX-2 attributes in cervical cancer across different patients. The prediction of the PLNM model was positive, the PLN dissection may be necessary whereas the prediction of the PLNM model was negative and the predicting COX-2 model was positive, COX-2 inhibitors were helpful for patients to control micrometastasis or recurrence of lymph nodes. Moreover, our results demonstrated the two of selected PET texture to predict PLNM and COX-2 expression that were slightly correlated with corresponding texture features from IHC images.

There are several limitations in the current study. Firstly, it was retrospective and performed at a single institution based on a small sample size. Prospective multicenter studies on automatic image acquisition and reconstruction are required to improve the process. Secondly, although we chose 2D ROI to be consistent with the 2D IHC image feature extraction method, the cross-sections of the IHC images in this study may not correspond exactly to the cross-sections of PET/CT. Further clinical studies on large-scale data sets based on the 3D printing technology are needed to achieve more accurate matching of PET images and pathological images to fully address this question. Thirdly, further test-retest studies and more standardized workflow are needed to assess feature robustness of PLNM for better generalization.

Conclusion

In conclusion, combining PET/CT texture analysis to predict PLNM and COX-2 expression can improve the predictive ability of machine learning models for PLNM trends in PLN-negative patients. In addition, the correlation between the texture features of PET images and the corresponding texture features of IHC images provides a reasonable explanation that the texture features of the primary tumor on PET images can predict PLNM. Based on this machine learning model integrating PET/CT radiomic and IHC pathomic features, it is expected to provide guidance for the treatment strategy of negative pelvic lymph node cervical cancer in the near future.

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 Shengjing Hospital Ethics Committee.

Author contributions

ZZ: Conceptualization, methodology, investigation, formal analysis, data curation, writing—original draft; XL: Formal analysis, data curation, validation, writing—review and editing; HS: Conceptualization, funding acquisition, writing—review and editing, supervision. All authors contributed to the article and approved the submitted version.

Funding

This study has received funding by the National Natural Science Foundation of China (No. 82071967), Liaoning Science Natural Science Foundation (No. 2019-MS-373), Shenyang High Level Innovative Talents Support Program (No. RC210138) and 345 Talent Project (No. M0377).

Acknowledgments

We would like to express my gratitude to all those who helped us for this paper. Thanks to Xinghao Wang and Chen Xu for their contributions to the paper.

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

References

Bhatla, N., Aoki, D., Sharma, D. N., and Sankaranarayanan, R. (2021). Cancer of the cervix uteri: 2021 update. *Int. J. Gynaecol. Obstet.* 155 , 28–44. doi:10.1002/ijgo.13865

Chen, H. H., Su, W. C., Chou, C. Y., Guo, H. R., Ho, S. Y., Que, J., et al. (2005). Increased expression of nitric oxide synthase and cyclooxygenase-2 is associated with poor survival in cervical cancer treated with radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 63 (4), 1093–1100. doi:10.1016/j.ijrobp.2005.03.062

Chicklore, S., Goh, V., Siddique, M., Roy, A., Marsden, P. K., and Cook, G. J. (2013). Quantifying tumour heterogeneity in 18f-fdg pet/ct imaging by texture analysis. *Eur. J. Nucl. Med. Mol. Imaging* 40 (1), 133–140. doi:10.1007/s00259-012-2247-0

Fasmer, K. E., Gulati, A., Dybvik, J. A., Ytre-Hauge, S., Salvesen, O., Trovik, J., et al. (2020). Preoperative 18f-fdg pet/ct tumor markers outperform mri-based markers for the prediction of lymph node metastases in primary endometrial cancer. *Eur. Radiol.* 30 (5), 2443–2453. doi:10.1007/s00330-019-06622-w

Federico, A., Anchora, L. P., Gallotta, V., Fanfani, F., Cosentino, F., Turco, L. C., et al. (2022). Clinical impact of pathologic residual tumor in locally advanced cervical cancer patients managed by chemoradiotherapy followed by radical surgery: A large, multicenter, retrospective study. *Ann. Surg. Oncol.* 29, 4806–4814. doi:10.1245/s10434-022-11583-4

Henriksson, E., Kjellen, E., Wahlberg, P., Ohlsson, T., Wennerberg, J., and Brun, E. (2007). 2-Deoxy-2-[18f] fluoro-D-glucose uptake and correlation to intratumoral heterogeneity. *Anticancer Res.* 27 (4B), 2155–2159.

Hoellen, F., Waldmann, A., Banz-Jansen, C., Rody, A., Heide, M., Koster, F., et al. (2016). Expression of cyclooxygenase-2 in cervical cancer is associated with lymphovascular invasion. *Oncol. Lett.* 12 (4), 2351–2356. doi:10.3892/ol.2016.4925

Horn, L. C., Bilek, K., Fischer, U., Einenkel, J., and Hentschel, B. (2014). A cut-off value of 2 Cm in tumor size is of prognostic value in surgically treated figo stage ib cervical cancer. *Gynecol. Oncol.* 134 (1), 42–46. doi:10.1016/j.ygyno.2014.04.011

Huang, M., Chen, Q., Xiao, J., Liu, C., and Zhao, X. (2013). Prognostic significance of cyclooxygenase-2 in cervical cancer: A meta-analysis. *Int. J. Cancer* 132 (2), 363–373. doi:10.1002/ijc.27686

Huang, Y., Liu, Z., He, L., Chen, X., Pan, D., Ma, Z., et al. (2016). Radiomics signature: A potential biomarker for the prediction of disease-free survival in early-stage (I or ii) non-small cell lung cancer. *Radiology* 281 (3), 947–957. doi:10.1148/radiol.2016152234

Ishikawa, H., Ohno, T., Kato, S., Wakatsuki, M., Iwakawa, M., Ohta, T., et al. (2006). Cyclooxygenase-2 impairs treatment effects of radiotherapy for cervical cancer by inhibition of radiation-induced apoptosis. *Int. J. Radiat. Oncol. Biol. Phys.* 66 (5), 1347–1355. doi:10.1016/j.ijrobp.2006.07.007

Jo, H., Kang, S., Kim, J. W., Kang, G. H., Park, N. H., Song, Y. S., et al. (2007). Hypermethylation of the cox-2 gene is a potential prognostic marker for cervical cancer. *J. Obstet. Gynaecol. Res.* 33 (3), 236–241. doi:10.1111/j.1447-0756.2007.00517.x

Kang, S., Kim, M. H., Park, I. A., Kim, J. W., Park, N. H., Kang, D., et al. (2006). Elevation of cyclooxygenase-2 is related to lymph node metastasis in adenocarcinoma of uterine cervix. *Cancer Lett.* 237 (2), 305–311. doi:10.1016/j.canlet.2005.06.027

Kather, J. N., Weis, C. A., Bianconi, F., Melchers, S. M., Schad, L. R., Gaiser, T., et al. (2016). Multi-class texture analysis in colorectal cancer histology. *Sci. Rep.* 6, 27988. doi:10.1038/srep27988

Kato, R., Hasegawa, K., Torii, Y., Udagawa, Y., and Fukasawa, I. (2015). Factors affecting platinum sensitivity in cervical cancer. *Oncol. Lett.* 10 (6), 3591–3598. doi:10.3892/ol.2015.3755

Khunamornpong, S., Settakorn, J., Sukpan, K., Srisomboon, J., Ruangvejvorachai, P., Thorner, P. S., et al. (2009). Cyclooxygenase-2 expression in squamous cell carcinoma of the uterine cervix is associated with lymph node metastasis. *Gynecol. Oncol.* 112 (1), 241–247. doi:10.1016/j.ygyno.2008.09.017

Kim, H. J. W. H., Park, I. A., and Ha, S. W. (2003). High cyclooxygenase-2 expression is related with distant metastasis in cervical cancer treated with radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.* 55 (1), 16–20. doi:10.1016/s0360-3016(02)03821-x

Kim, J. S., Li, S., Kim, J. M., Yeo, S. G., Kim, K. H., and Cho, M. J. (2008). Cyclooxygenase-2 expression as a predictor of para-aortic lymph node recurrence in uterine cervical cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 70 (5), 1516–1521. doi:10.1016/j.ijrobp.2007.08.048

Kuroda, K., Yamamoto, Y., Yanagisawa, M., Kawata, A., Akiba, N., Suzuki, K., et al. (2017). Risk factors and a prediction model for lower limb lymphedema following lymphadenectomy in gynecologic cancer: A hospital-based retrospective cohort study. *BMC Womens Health* 17 (1), 50. doi:10.1186/s12905-017-0403-1

Li, X., Xu, C., Yu, Y., Guo, Y., and Sun, H. (2021). Prediction of lymphovascular space invasion using a combination of tenascin-C, cox-2, and pet/ct radiomics in patients with early-stage cervical squamous cell carcinoma. *BMC Cancer* 21 (1), 866. doi:10.1186/s12885-021-08596-9

Li, X. R., Jin, J. J., Yu, Y., Wang, X. H., Guo, Y., and Sun, H. Z. (2021). Pet-ct radiomics by integrating primary tumor and peritumoral areas predicts E-cadherin expression and correlates with pelvic lymph node metastasis in early-stage cervical cancer. *Eur. Radiol.* 31 (8), 5967–5979. doi:10.1007/s00330-021-07690-7

Liu, H., Xiao, J., Yang, Y., Liu, Y., Ma, R., Li, Y., et al. (2011). Cox-2 expression is correlated with vegf-C, lymphangiogenesis and lymph node metastasis in human cervical cancer. *Microvasc. Res.* 82 (2), 131–140. doi:10.1016/j.mvr.2011.04.011

Lv, K., Guo, H. M., Lu, Y. J., Wu, Z. X., Zhang, K., and Han, J. K. (2014). Role of 18f-fdg pet/ct in detecting pelvic lymph-node metastases in patients with early-stage uterine cervical cancer: Comparison with mri findings. *Nucl. Med. Commun.* 35 (12), 1204–1211. doi:10.1097/MNM.0000000000000198

Manchana, T., Triratanachat, S., Sirisabya, N., Vasuratna, A., Termrungruanglert, W., and Tresukosol, D. (2006). Prevalence and prognostic significance of cox-2 expression in stage ib cervical cancer. *Gynecol. Oncol.* 100 (3), 556–560. doi:10.1016/j.ygyno.2005.09.014

Nanthamongkolkul, K., and Hanprasertpong, J. (2018). Predictive factors of pelvic lymph node metastasis in early-stage cervical cancer. *Oncol. Res. Treat.* 41 (4), 194–198. doi:10.1159/000485840

Nica, A., Gien, L. T., Ferguson, S. E., and Covens, A. (2020). Does small volume metastatic lymph node disease affect long-term prognosis in early cervical cancer? *Int. J. Gynecol. Cancer* 30 (3), 285–290. doi:10.1136/ijgc-2019-000928

O'Connor, J. P. B. (2017). Cancer heterogeneity and imaging. *Semin. Cell Dev. Biol.* 64, 48–57. doi:10.1016/j.semcdb.2016.10.001

Rudtanasudjatum, K., Charoenkwan, K., Khunamornpong, S., and Siriaunkgul, S. (2011). Impact of histology on prognosis of patients with early-stage cervical cancer treated with radical surgery. *Int. J. Gynaecol. Obstet.* 115 (2), 183–187. doi:10.1016/j.ijgo.2011.06.011

Ryu, H. S., Chang, K. H., Yang, H. W., Kim, M. S., Kwon, H. C., and Oh, K. S. (2000). High cyclooxygenase-2 expression in stage ib cervical cancer with lymph node metastasis or parametrial invasion. *Gynecol. Oncol.* 76 (3), 320–325. doi:10.1006/gyno.1999.5690

Sironi, S., Buda, A., Picchio, M., Perego, P., Moreni, R., Pellegrino, A., et al. (2006). Lymph node metastasis in patients with clinical early-stage cervical cancer: Detection with integrated fdg pet/ct. *Radiology* 238 (1), 272–279. doi:10.1148/radiol.2381041799

Stasinopoulos, I., Shah, T., Penet, M. F., Krishnamachary, B., and Bhujwalla, Z. M. (2013). Cox-2 in cancer: Gordian knot or achilles heel? *Front. Pharmacol.* 4, 34. doi:10.3389/fphar.2013.00034

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *Ca. Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660

Umbreit, E. C., McIntosh, A. G., Suk-Ouichai, C., Segarra, L. A., Holland, L. C., Fellman, B. M., et al. (2020). Intraoperative and early postoperative complications in postchemotherapy retroperitoneal lymphadenectomy among patients with germ cell tumors using validated grading classifications. *Cancer* 126 (22), 4878–4885. doi:10.1002/cncr.33051

van Griethuysen JjmFedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., et al. (2017). Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* 77 (21), e104–e107. doi:10.1158/0008-5472.CAN-17-0339

Wenzel, H. H. B., Olthof, E. P., Bekkers, R. L. M., Boere, I. A., Lemmens, V., Nijman, H. W., et al. (2022). Primary or adjuvant chemoradiotherapy for cervical cancer with intraoperative lymph node metastasis - a review. *Cancer Treat. Rev.* 102, 102311. doi:10.1016/j.ctrv.2021.102311

Xu, L., Stevens, J., Hilton, M. B., Seaman, S., Conrads, T. P., Veenstra, T. D., et al. (2014). Cox-2 inhibition potentiates antiangiogenic cancer therapy and prevents metastasis in preclinical models. *Sci. Transl. Med.* 6 (242), 242ra84. doi:10.1126/scitranslmed.3008455

Yeo, W., Chan, S. L., Mo, F. K., Chu, C. M., Hui, J. W., Tong, J. H., et al. (2015). Phase I/ii study of temsirolimus for patients with unresectable hepatocellular carcinoma (hcc)- a correlative study to explore potential biomarkers for response. *BMC Cancer* 15, 395. doi:10.1186/s12885-015-1334-6

Yost, K. J., Cheville, A. L., Al-Hilli, M. M., Mariani, A., Barrette, B. A., McGree, M. E., et al. (2014). Lymphedema after surgery for endometrial cancer: Prevalence, risk factors, and quality of life. *Obstet. Gynecol.* 124, 307–315. doi:10.1097/AOG.0000000000000372

Zhao, S., Kuge, Y., Mochizuki, T., Takahashi, T., Nakada, K., Sato, M., et al. (2005). Biologic correlates of intratumoral heterogeneity in 18f-fdg distribution with regional expression of glucose transporters and hexokinase-ii in experimental tumor. *J. Nucl. Med.* 46 (4), 675–682.

Zyla, R. E., Gien, L. T., Vicus, D., Olkhov-Mitsel, E., Mirkovic, J., Nofech-Mozes, S., et al. (2020). The prognostic role of horizontal and circumferential tumor extent in cervical cancer: Implications for the 2019 figo staging system. *Gynecol. Oncol.* 158 (2), 266–272. doi:10.1016/j.ygyno.2020.05.016

Frontiers in Physics

TYPE Original Research
PUBLISHED 14 March 2023
DOI 10.3389/fphy.2023.1003874

EDITED BY
Yukun Lai,
Cardiff University, United Kingdom

REVIEWED BY
Haoran Dou,
University of Leeds, United Kingdom
Zhenyuan Ning,
Southern Medical University, China
Xuzhe Zhang,
Columbia University, United States

*CORRESPONDENCE
Chun Wang,
✉ chunwang@sdu.edu.cn
Kun Zhao,
✉ zhaokunbj@inspur.com
Zhi Liu,
✉ liuzhi@sdu.edu.cn

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

SPECIALTY SECTION
This article was submitted to Medical Physics and Imaging,
a section of the journal
Frontiers in Physics

RECEIVED 01 August 2022
ACCEPTED 22 February 2023
PUBLISHED 14 March 2023

CITATION
An M, Chen J, Cao Y, Tao K, Wang J, Wang C, Zhao K and Liu Z (2023), ResAttn-recon: Residual self-attention based cortical surface reconstruction.
Front. Phys. 11:1003874.
doi: 10.3389/fphy.2023.1003874

ResAttn-recon: Residual self-attention based cortical surface reconstruction

Mujun An[1†], Jianzhang Chen[2†], Yankun Cao[3], Kemeng Tao[1], Jianlei Wang[4], Chun Wang[4*], Kun Zhao[5*] and Zhi Liu[1*]

[1]The Research Center of Intelligent Medical Information Processing, School of Information Science and Engineering, Shandong University, Qingdao, China, [2]Department of Clinical Psychology, The 960th Hospital of the PLA Joint Logistics Support Force, Jinan, China, [3]School of Software, Shandong University, Jinan, China, [4]Optical Advanced Research Center, Shandong University, Qingdao, China, [5]Inspur Electronic Information Industry Co., Ltd., Qingdao, China

Introduction: The accurate cerebral cortex surface reconstruction is crucial for the study of neurodegenerative diseases. Existing voxelwise segmentation-based approaches like FreeSurfer and FastSurfer are limited by the partial volume effect, meaning that reconstruction details highly rely on the resolution of the input volume. In the computer version area, the signed distance function has become an efficient method for 3D shape representation, the inherent continuous nature makes it easy to capture the fine details of the target object at an arbitrary resolution. Additionally, as one of the most valuable breakthroughs in deep learning research, attention is a powerful mechanism developed to enhance the performance of the encoder-decoder architecture.

Methods: To further improve the reconstruction accuracy of the cortical surface, we proposed ResAttn-Recon, a residual self-attention based encoder-decoder framework. In this framework, we also developed a lightweight decoder network with skip connections. Furthermore, a truncated and weighted L1 loss function are proposed to accelerate network convergence, compared to simply applying the L1 loss function.

Results: The intersection over union curve in the training process achieved a steeper slope and a higher peak (0.948 vs. 0.920) with a truncated L1 loss. Thus, the average symmetric surface distance (AD) for the inner and outer surfaces is 0.253 ± 0.051 and the average Hausdorff distance (HD) is 0.629 ± 0.186, which is lower than that of DeepCSR, whose absolute distance equals 0.283 ± 0.059 and Hausdorff distance equals 0.746 ± 0.245.

Discussion: In conclusion, the proposed residual self-attention-based framework can be a promising approach for improving the cortical surface reconstruction performance.

KEYWORDS

cortical surface reconstruction, residual self-attention mechanism, signed distance function, deep learning, MRI volume

1 Introduction

In neural image processing, the brain cortical surface reconstruction plays an essential role in the study of neurodegenerative diseases [1] and psychological disorders [2]. Specifically, the cortical surface reconstruction aims to extract two surface meshes from brain magnetic resonance imaging (MRI). The inner white matter surface separates the white matter and the gray matter tissues, and the outer pial surface separates the gray matter tissue and the cerebrospinal fluid [3]. Considering the highly curved and folded intrinsic folding pattern of the cortical surface [4], it is challenging to extract anatomically plausible and topologically correct cortical surfaces in practice.

To address this, traditional approaches use a series of lengthy and computationally intensive processing algorithms, with manual intervention for hyperparameter fine-turning [5–11]. For instance, the widely used and reliable [9] toolkit usually takes hours to process a MRI volume data. In recent years, several deep learning approaches have emerged to overcome this shortage, and according to the data format being processed, these approaches can be categorized as voxel-based, mesh-based, and implicit surface representation-based. Voxel-based approaches first obtain the brain white matter tissue segmentation based on 3D-CNN [12] or 3D-Unet-like [13] architecture. Then the triangular mesh of the inner surface is extracted by applying mesh tessellation to the segmentation masks, with surface mesh smoothing and topology correction [14]. The outer pial surface mesh can then be derived from inflating the white matter surface mesh [15]. Leonie et al. [15] proposed FastSurfer to accelerate the FreeSurfer pipeline by replacing the traditional white matter segmentation algorithm with a 3D-CNN network. [16] proposed the SegRecon framework for cortical surface reconstruction and segmentation. Due to the partial volume effect (PVE) [17], voxel-based approaches have inherent limitations in capturing fine details at high resolution. Mesh-based approaches are mainly implemented by deforming the initial surface mesh to the target surface mesh, with a geometric deep-learning model. For instance, [18] proposed PialNN to reconstruct the pial surface from the white matter surface handled by the FreeSurfer pipeline. [19] proposed Voxel2Mesh to deform predefined sphere template meshes to cortical surfaces. [20] Proposed Vox2Cortex that leverages convolutional and graph convolutional neural networks to deform the template mesh to densely folded target cortical surface. Despite fast processing, theoretical guarantees are to be further developed to prevent self-intersections of the surface mesh. Implicit surface representation-based methods reformulate cortical surface reconstruction as the prediction of the implicit surface representation [21]. Typically, [3] proposed the DeepCSR network to learn an implicit surface function in a continuous coordinate system, with topology correction algorithm to ensure the geometric accuracy of the target surface.

Given brain MRI volume, existing deep learning approaches spend less time reconstructing cortical surface compared to traditional pipeline, with high reliability. Most of these approaches require voxelwise or vertexwise features extracted from the input MRI volume, however, none of them considered the long range feature dependencies, which plays an important role in model performance improvement. For instance, DeepCSR [3] directly concatenates the local and global features from the encoder feature maps, combined with the location coordinates of the query point as the input of the decoder network, PialNN [18] combined the norm and the location coordinate of the initial mesh vertex with the volumetric features extracted from local convolution, to predict the deformation displacement in the inflating process. Both approaches ignored the relationship between query points or mesh vertices. In this work, to efficiently model the long range feature dependencies in cortical surface reconstruction, the concept of the self-attention mechanism is introduced from neural language processing [22–24] and computer version [25–27] area. For pioneer works that apply self-attention to vision tasks [28], proposed Vision Transformer (ViT) for image recognition [29], proposed Tokens-To-Token Vision Transformer (T2T-ViT) to improve classification accuracy [30], proposed Swin Transformer, a general framework for image classification and segmentation, all the above works are discussed around 2D images. In this work, firstly, the input MRI volume is registered to a standard brain space, such as MNI105. Secondly, after the 3D Convolution block, the residual connected multi-head self-attention block and global flatten block, multi-scale feature maps and global feature vector are prepared to obtain volumetric features of sampling points at any given resolution. Then combined with location coordinates in standard space, the signed distance values of sampling points toward four cortical surfaces are predicted, Thirdly, after topology correction and iso-surface extraction operation, the inner and outer surface of the left and right hemispheres are reconstructed in parallel.

In this paper, we proposed ResAttn-Recon, a novel implicit surface representation approach based framework for inner and outer cortical surface reconstruction. In this work, we propose employing the concept of the self-attention mechanism and residual connection trick to the 3D convolutional neural network (3D CNN) encoder, 1 × 1 convolution is embedded into a multi-head self-attention block to fit the 3D feature map input. The proposed framework is able to reconstruct the cortical surface at an arbitrary resolution and benefit from the theoretical support of the implicit surface representation approach. The experimental performance has been substantially improved compared to the DeepCSR and the simple encoder-decoder framework without the attention block. The main contributions of this paper are as follows.

1) To the best of our knowledge, this is the first exploration in employing the residual self-attention mechanism in 3D cortical surface reconstruction.
2) A Commit2 from Review4 with skip connections is developed as an improvement over the DeepCSR decoder network, to simplify the network structure without losing performance.
3) The prior constraints are imposed on the network training with the proposed truncated L1 loss and Gaussian decay weighted L1 loss, as a new strategy for model performance improvement.

The rest of this paper is organized as follows. Section 2 introduces the basic theories and the proposed framework, as well as the dataset enrolled in this work, In Section 3, the details of the experimental evaluation and analysis are given. Section 4 and Section 5 provide a discussion and conclude this paper.

2 Materials and methods

In this section, we introduced ResAttn-Recon, a residual self-attention-based cortical surface reconstruction network. Simultaneously, the lightweight decoder networks and the loss function with prior constrains are also explored to improve the reconstruction performance. As shown in Figure 1, the proposed framework consists of four main parts: 1) data preprocessing, including data acquisition from FreeSurfer toolkit and MRI

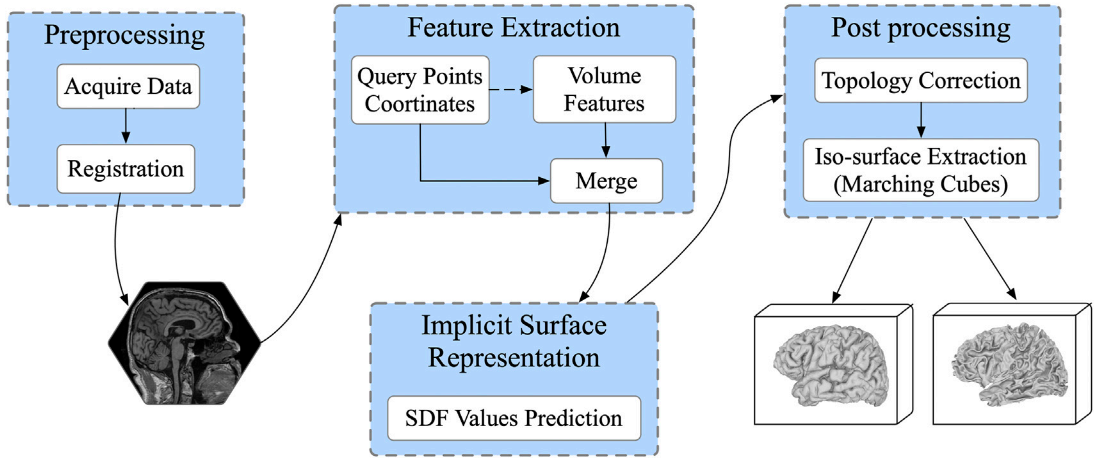


FIGURE 1
The workflow for cortical surface reconstruction. Input the 3D MRI volume, given enough sampling points, it predicts 4 mesh surfaces of arbitrary resolution in parallel.

volume registration; 2) feature extraction; 3) implicit surface representation performed with the proposed ResAttn-Recon network; and 4) post-processing to extract the target inner and outer cortical surfaces from the predicted signed distance function representation. The residual self-attention block is embedded in the encoder network following the intermediate feature map.

2.1 Data acquisition

In this paper, we used the publicly available dataset from Alzheimer's Disease Neuroimaging Initiative (ADNI) [31], MRI scans of 560 T1 original images are enrolled in this work, 470 images for training, 30 images for validation and the remaining 60 images for testing. Ground-truth of inner and outer surfaces from the left and right hemispheres are extracted by the FreeSurfer pipeline. Images are normalized to the size of 182 × 218 × 182, with voxel spacing to [1, 1, 1]. To unify the coordinate systems of the input MRI scans, affine registration is first performed to the MNI105 brain template [32].

2.2 Surface representation

The signed distance function (SDF) is a continuous function to represent the surface distribution, and has been widely employed in 3D shape representation.

The SDF function can be defined as follows:

$$SDF_{\text{surface}}(\boldsymbol{x}_i) = s_i, \text{where } \boldsymbol{x}_i \in \mathcal{R}^3, s_i \in \mathcal{R} \tag{1}$$

Here, x_i stands for any point in Euclidean space represented by its 3D location coordinate, and s_i is the shortest Euclidean distance from the point x_i to the surface, with a positive sign if the point is inside the watertight surface or negative sign if the point is outside the surface.

With the SDF values of given spatial points, the target surface can be expressed as a set consisting of all points satisfying the following:

$$SDF_{\text{surface}}(\cdot) = 0 \tag{2}$$

Then, after Gaussian smoothing and topology correction processing, the target surface is extracted with a zero iso-surface extraction algorithm, such as marching cubes [33].

In this work, the continuous SDF in Ω space is approximated by the deep learning model. Given query point, the well-trained network predicts its SDF value to the target surface directly. This provides theoretical support for reconstructing surfaces of arbitrary resolutions. In detail, the approximator is implemented by a decoder network parameterized by θ, which is further described below.

2.3 ResAttn-recon framework

2.3.1 Feature extraction encoder module

The network architecture is illustrated in Figure 2, and the raw brain MRI need to be firstly registered to MNI105 space before being sent to the network. The feature extraction encoder module consists of three subblocks, the 3D Convolution (Conv3D) block, the Residual Self-Attention block, and the Global Flatten block. The Conv3D block consists of five Conv3D layers, each of which is followed by a Rectified Linear Units (ReLU) activation, with the 3D max pooling operation before the fourth convolution. The number of convolution output channels is sequentially increased to 2^3; 2^4; 2^5; 2^6; 2^7 after each convolution. The convolutional kernel is set to 3 × 3 × 3, with the stride equal to two and padding equal to one. The fourth output feature map is then input into the residual self-attention block, followed by the global flatten block to generate the global feature map. The image features represented by the encoder intermediate feature maps and outputs are firstly extracted by the proposed feature extraction encoder module from the registered MRI, then we construct a bounding box grid with evenly spaced points at a predefined desired resolution (e.g;,

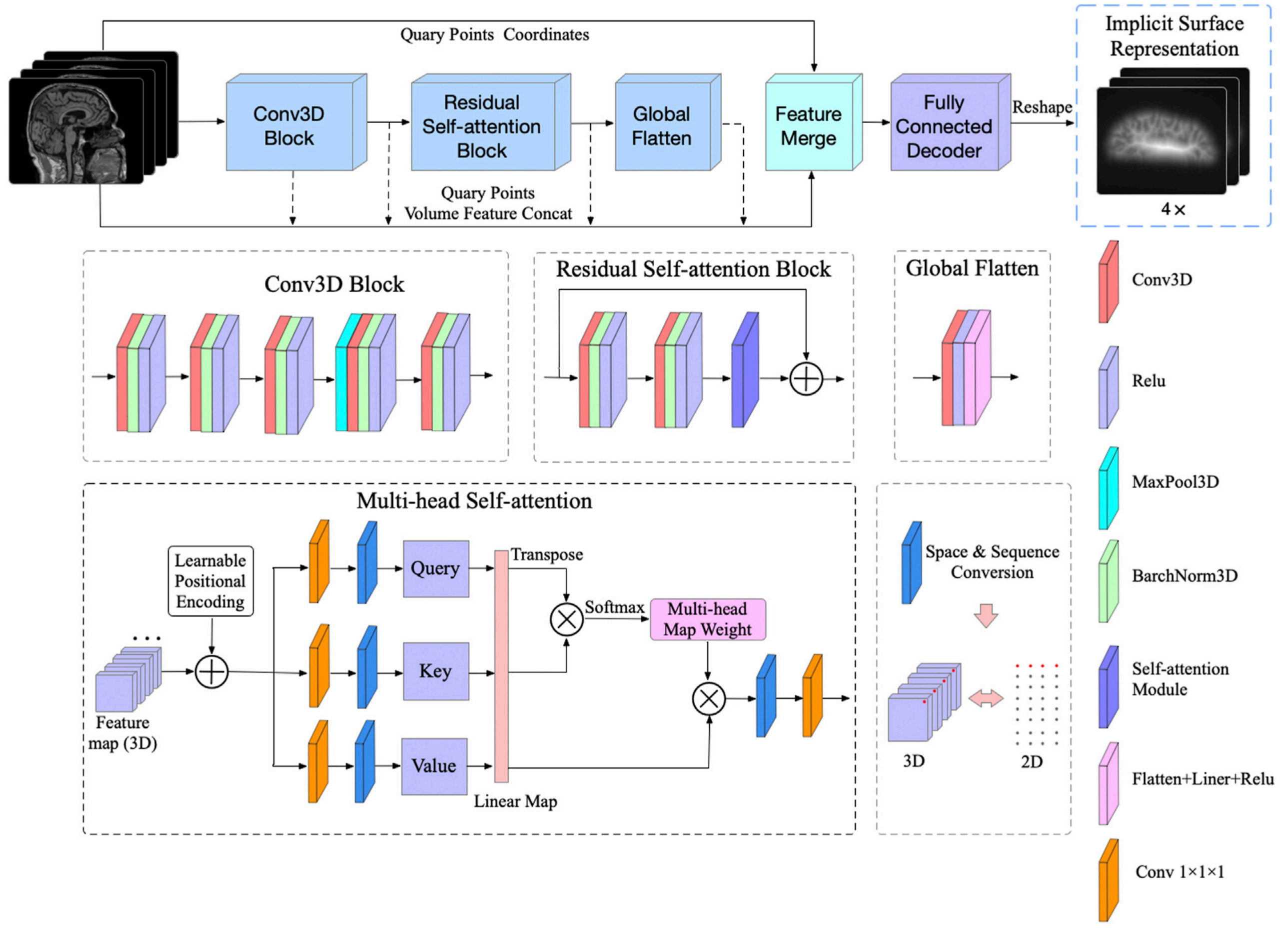


FIGURE 2
The proposed ResAttn-recon architecture for cortical surface reconstruction. The concept of the multi-head self-attention mechanism was introduced to our residual self-attention block.

512 × 512 × 512), which is capable of covering the registered brain MRI in MNI105 standard registration space. After that, the relative position coordinates of the predefined sampling points in MNI105 space as mentioned above are projected to the multiscale feature maps generated by the Conv3D block and the residual self-attention block. The interpolated values obtained from the projected locations and the global feature vector after the Global Flatten block are concatenated as the feature vector of the sampling point.

2.3.2 Residual self-attention mechanism

To better aggregate the feature representation of the sampled points, the widely used self attention mechanism is introduced to our proposed encoder-decoder network architecture, where the detail operation can be described as follows:

$$\hat{X} = \mathrm{Softmax}\left(\frac{\theta(X)\phi(X)}{\sqrt{d_{feat}}}\right)\varphi(X) \tag{3}$$

Where $X \in N \times L$ stands for the local and global feature representation extracted from encoder feature maps, N is the number of sampled points, and L is the dimension of the corresponding feature vector. After the 1×1 convolution and reshape operation, linear transformation of $\theta(\cdot)$, $\phi(\cdot)$ and $\varphi(\cdot)$ are implemented by three single-layer perceptrons (Linear map) in this work. The point-to-point affinity is calculated by the inner product of $\theta(X)$ and $\varphi(X)$.

In this work, the multi-head self-attention mechanism [22] is applied to improve the expression ability of the attention module. For this 3D reconstruction work, as illustrated in Figure 2, the residual connection in the Residual Self-attention Block indicates that this block does not change the dimension of the input feature map, which equals $(batch\_size \times num\_channels \times d \times w \times h)$ for single image. From another point of view, the input feature map can be considered as a token array (with the shape of $(d \times w \times h)$), each token corresponds to a 128-dimensional feature embedding vector alone the channel direction (number of channels equals 128 in this case). Therefore, the input feature map of the Residual Self-attention Block is converted to a 2 days array squence input X with the shape of $((d \times w \times h), 128)$, by reshape and transpose operations. After that, mulit-head self-attention operation can be easily applied to X. In this case, the number of heads equals 4, therefore X is projected to subdimension space 4 times in parallel; after four self-attention operations, the outputs are concatenated and further projected. The sequence output is reshaped again by the reverse operation of "space

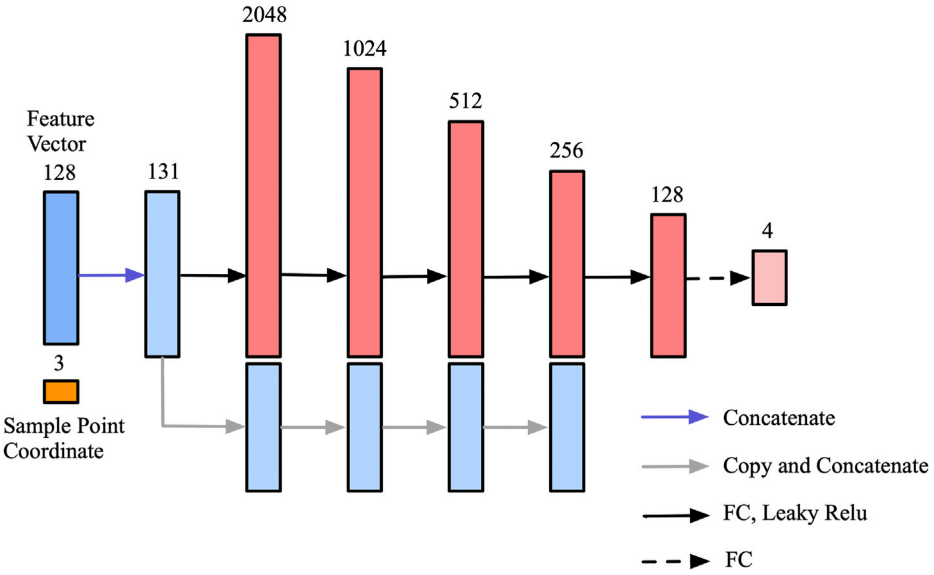


FIGURE 3
Decoder network for SDF values prediction. The skip-connection mechanism is introduced to make full use of the location information of the sampling points.

and sequence conversion" as shown in Figure 2, and output the attentioned 3D feature map, followed by a $1 \times 1 \times 1$ convolution. The "space and sequence conversion" operation means taking feature in the same spatial location across all channels.

$$\begin{gathered} \mathrm{MultiHead}(Q, K, V) = \mathrm{Concat}\,(\mathrm{head}_1, \ldots, \mathrm{head}_n) W^O \\ \text{where } head_i = \mathrm{Attention}\left(QW_i^Q, KW_i^K, VW_i^V\right) \end{gathered} \tag{4}$$

Here, Q, K and V represent the Query, Key and Value matrix as shown in Figure 2, and $n = 4$, $W_i^Q = W_i^K = W_i^V \in \mathcal{R}^{128\times 32}$, $W^O \in \mathcal{R}^{128\times 128}$ in this work.

It is worth noting that to fully take advantage of the spatial sequence information extracted by sampling, the absolute positional coding strategy is applied in this work, where the positional coding is directly added to the input of the self attention module. Specifically, one raw MRI input is firstly normalized and reshaped to $1 \times 182 \times 218 \times 182$, and the size of feature representation X becomes $128 \times 12 \times 14 \times 12$ after the Conv3D block, where the first dimension represents the number of channels. The shape of positional coding P can be expressed as: $(batch\_size \times embedded\_dim \times width \times height \times depth)$, where embedded_dim is the dimension of positional vector and equals 128 in this case. In this work, during network training, we initialized the positional coding P with standard normal distribution where $P \sim N(0, 1)$, and then taking P as trainable parameters to update with backpropagation. As seen in Figure 2, the learnable positional encoding matrix P was simply added to the input feature map, inspired by Sequence to Sequence Learning [34].

For the residual self-attention block, before the self-attention module, two 3D convolution operations are performed, each followed by 3D batch normalization and ReLU activation. Notably, we added the residual connection between the Conv3D block output and the self-attention module output to help improve the learning.

2.3.3 Decoder with skip connections

To further simplify the decoder network without losing performance, the feed-forward network is composed of six fully connected layers, and feature vectors extracted from feature maps and corresponding coordinates of sampling points are concatenated as the input of the decoder network. Since position coordinates are critical for 3D shape representation [21], we introduced skip-connection to maintain the proportion of location information. As shown in Figure 3, the input feature vector is concatenated with the intermediate output of the following four fully connected layers. The output vector of the decoder network representing the SDF values of four corresponding cortical surfaces for each voxel is a vector with length 4.

2.3.4 Loss function

In 3D object reconstruction, L1 loss is the most frequently used loss function, and the basic form of L1 loss for one cortical surface can be written as follows, where N represents the number of sampling points and B represents batch size during training:

$$\mathcal{L}(f_\theta(X), S) = \frac{1}{B}\sum_b^B \sum_{i=1}^N \left| f_\theta(\boldsymbol{x}_i) - s_i \right| \tag{5}$$

And for parallel training with four cortical surfaces, the formula becomes:

$$\mathcal{L}(f_\theta(X), S) = \frac{1}{B}\sum_b^B \sum_{i=1}^N \sum_{j=1}^4 \left| f_\theta\left(\boldsymbol{x}_i^j\right) - s_i^j \right| \tag{6}$$

For cortical surface representation with signed distance function values, sampling points close to the cortical surface are critical for reconstruction details, while sampling points away from the cortical surface contribute less to the reconstruction process. To help the training network capture more detailed information around the surface, the truncate interval $[-\delta, \delta]$ is applied to the ground-truth and the predicted signed distance values, where the truncated L1 loss can be expressed as follows:

$$T(\lambda, \delta) := min(max(\lambda, -\delta), \delta) \tag{7}$$

$$\mathcal{L}_{\text{truncate}}(f_\theta(X), S) = \frac{1}{B}\sum_{b}^{B}\sum_{i=1}^{N}\sum_{j=1}^{4}\left|T\left(f_\theta\left(\boldsymbol{x}_i^j\right), \delta\right) - T\left(s_i^j, \delta\right)\right| \tag{8}$$

$T(\cdot)$ is defined as the truncated function with parameter δ. Given proper small hyperparameter values of δ, the network can focus more on surface details, while for larger value of δ, more samples are used for model weight updates. In general, the basic form of L1 loss is the special case of its truncated form.

Furthermore, we also propose exploring the potential of the L1 loss function with a Gaussian decay coefficient. The Gaussian function in this case is of the following form:

$$G(s_i) = \frac{A}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\frac{(s_i - \mu)^2}{\sigma^2}\right) \tag{9}$$

where A is the amplitude coefficient. Due to the smooth decay characteristic of the Gaussian function, as the distance between the sampling point and the ground-truth surface increases, the loss weight of the corresponding sampling point decreases. We hope this design can perform dynamic loss constraint during network training, and help the network to better converge to the optimal solution. The Gaussian decay L1 loss can be described as follows:

$$\mathcal{L}_{gaussian}(f_\theta(X), S) = \frac{1}{B}\sum_{b}^{B}\sum_{i=1}^{N}\sum_{j=1}^{4}G\left(s_i^j\right)\cdot\left|f_\theta\left(\boldsymbol{x}_i^j\right) - s_i^j\right| \tag{10}$$

The amplitude coefficient A, μ and θ are hyperparameters based on experience. In this work, A equals 20, μ equals 0, and θ equals 1.5. For the netwok training, the Adam optimizer was employed with a fixed learning rate equals 0.0001.

2.4 Cortical surface extraction

For the cortical surface extraction pipeline, firstly, the input MRI volume is registered into the MNI105 space; secondly, given arbitrary reconstruction resolution, i.e., 512 × 512 × 512 by uniform sampling from MNI105 space, the attention-based encoder-decoder network outputs the SDF representation with the shape of 512 × 512 × 512, followed by a Gaussian filter smoothing with a standard deviation of 0.5.

To prevent grid self-intersection, and to ensure that the predicted signed distance function values is homeomorphic to a sphere, in this work, we apply a topology propagation algorithm using a fast marching technique proposed by bazin. et.al [35], that enforces the network prediction result to a desired topology.

```
M_{vertices}←[];
$M_v \leftarrow 0$;
While $n \leq$ N do
    if $l \geq V_n^i$ and $l \leq V_n^{\mathcal{N}(i)}$ $(i = 1..8)$ then
        $\Delta P \leftarrow \frac{(l - V_n^i)(|\boldsymbol{P}_n^{\mathcal{N}(i)} - \boldsymbol{P}_n^i|)}{V_n^{\mathcal{N}(i)} - V_n^i}$
        $\boldsymbol{M_v} \leftarrow \boldsymbol{P}_n^i + \Delta P$;/* Coordinate interpolation */
        $M_{vertices}.append(M_v)$;/* Accumulate mesh vertex coordinates */
    else
        continue
    end
end
```

Algorithm 1: Marching cubes pseudocode
Data: the SDF value of the i-th vertex of the n-th cube V_n^i, the 3D coordinate of the i-th vertex of the n-th cube P_n^i, N cubes to iterate, the scale surface level set l, the adjacency vertex index $N(i)$ of the index i in the query cube
Result: Extracted surface mesh vertices coordinates $M_{vertices}$

Then, the marching cubes algorithm proposed by [33] is further employed to cortical surface extraction, followed by Laplacian smoothing. The core idea of the marching cubes algorithm can be summarized as Algorithm 1 For vertices $v_i \in M$, where M represents the extracted surface mesh, the Laplacian smoothing operation can be described as follows:

$$\text{Smooth}(\boldsymbol{v}_i) = \sum_{k\in\mathcal{N}(i)}\frac{\boldsymbol{v}_k}{|\mathcal{N}(i)|} \tag{10}$$

where $N(i)$ is the adjacency vertices of the i-th vertex. Note that the post-processing operation of the four surfaces is performed in parallel, for efficiency.

4 Results

To verify the effectiveness of our method, we first designed a series of ablation studies to explore the importance of prior constraints and the skip connections mechanism, after which we evaluated the performance of three loss function. Finally, the precision analysis is performed compared with DeepCSR for challenging pial surface reconstruction.

4.1 Ablation experiment

As shown in Table 1, the ablation experiment is conducted to measure the importance of the proposed components.

The average absolute distance (AD) and the Hausdorff distance (HD) [36,37] are employed as the surface evaluation metrics, and the lower the values, the better the reconstruction results.

4.1.1 Decoder network with skip connections

In this experiment we explored the expressive capability of fully connected layers in the decoder network architecture with the skip connections mechanism. As shown in Table 1, row 4 records the baseline encoder-decoder network where the decoder network is the fully connected layers without skip connections mechanism, and the feature vectors and location coordinates are simply concatenated from the decoder input. It was found that the decoder with skip-connection (row 4) shows lower AD and HD indicators than that without skip connections (row 5). The results indicate that the location information of the sampling points make a considerable contribution to reconstruction work. Nevertheless, there is still much room for improvement in the reconstruction indicators, compared with the DeepCSR framework (row 3).

In order to further validate the performance of the proposed lightweight decoder network, different indicators are used to compare with the decoder network in DeepCSR framework. As

TABLE 1 Results of comparison analysis with DeepCSR and the ablation study on the proposed ResAttn-Recon framework for cortical surface reconstruction. Including the white matter surface and the pial matter surface, where AD = Average symmetric surface distance, HD = Hausdorff distance.

Method	Left white matter surface		Right white matter surface		Left pial matter surface		Right pial matter surface	
	AD(mm)	HD(mm)	AD(mm)	HD(mm)	AD(mm)	HD(mm)	AD (mm)	HD (mm)
DeepCSR(OCC)	0.669 (±0.543)	2.718 (±0.607)	0.601 (±0.482)	2.648 (±1.060)	0.298 (±0.149)	0.998 (±1.082)	0.291 (±1.082)	0.880 (±0.231)
DeepCSR(SDF)	0.280 (±0.054)	0.586 (±0.131)	0.273 (±0.047)	0.565 (±0.124)	0.292 (±0.073)	0.898 (±0.351)	0.290 (±0.063)	0.937 (±0.375)
Voxel2Mesh	0.389 (±0.251)	0.996 (±0.427)	0.403 (±0.187)	1.005 (±0.602)	-	-	-	-
GAN	0.429 (±0.107)	1.094 (±0.133)	0.448 (±0.207)	1.146 (±0.192)	0.641 (±0.251)	2.518 (±0.426)	0.675 (±0.170)	2.704 (±0.332)
3D CNN encoder +Fully connected decoder only	0.389 (±0.045)	1.056 (±0.241)	0.391 (±0.137)	1.102 (±0.255)	0.413 (±0.120)	1.103 (±0.271)	0.398 (±0.036)	1.122 (±0.292)
3D CNN encoder +Skip-connection decoder	0.318 (±0.103)	0.829 (±0.252)	0.359 (±0.128)	1.011 (±0.380)	0.352 (±0.090)	0.972 (±0.347)	0.343 (±0.131)	0.981 (±0.334)
3D CNN encoder + Skip-connection decoder +Truncated L1 loss	0.318 (±0.023)	0.705 (±0.023)	0.356 (±0.070)	0.750 (±0.101)	0.304 (±0.066)	0.968 (±0.467)	0.308 (±0.045)	0.834 (±0.164)
ResAttn-Recon (Proposed framework with residual self-attention block)	0.278 (±0.042)	0.591 (±0.075)	0.272 (±0.030)	0.557 (±0.064)	0.223 (±0.063)	0.614 (±0.277)	0.242 (±0.033)	0.720 (±0.214)

TABLE 2 Lightweight decoder analysis.

Methods	Params	Params	Inference	Computational
	Num (decoder)	Size M)	Speed (ms)	Cost (basic L1 loss)
DeepCSR	3549188	13.53	6.34 ± 0.02	1,452.85 ± 28.98
Proposed	2636527	10.05	0.47 ± 0.02	1,291.89 ± 10.50

shown in Table 2, the proposed decoder network outperforms DeepCSR's decoder network in the number of parameters (2636527 vs. 3549188), parameters size (10.05 MB vs. 13.53 MB), and the average inference speed (0.47 ± 0.02 vs. 6.34 ± 0.02) for testing dataset. The average L1 Loss (computational cost shown in formula 5) of the proposed framework is 1,291.89 ± 10.50, lower than that of the DeepCSR framework (1,452.85 ± 28.98). By the way, the computational cost looks larger, because the distance space is not normalized to the [-0.5,0.5] interval.

4.1.2 Loss function with prior constraints

The performance of the proposed framework between employing prior constraints (trained with a truncated L1 loss or Gaussian decay L1 loss) and without prior constraints (trained with a basic L1 loss) are compared, as shown in row five and row six of Table 1, Additionally, the reconstruction results improved with the help of prior constraints imposed on the loss function from the perspective of AD and HD indicators. For fairness, the training process of the DeepCSR-SDF network also employs the L1 loss and its variants, and take the optimal result for performance comparison and analysis.

4.2 Methods comparision

For further comparison and analysis, besides the DeepCSR framework, Voxel2Mesh and GAN [38] were added to the experiments. Since the Voxel2Mesh framework cannot reconstruct the inner and outer surfaces in parallel, analysis of white matter surface reconstruction performance is considered for convenience.

For GAN model, the 3D-UNet is employed as the backbone of the generator network. The generator predicts the SDF representation (output channel equals 4) relevant to four cortical surfaces. The discriminator takes the MRI volume and the ground truth/predicted SDF representation as input, and discriminate true or false label, where cross entropy loss is used as the discriminator loss function. As shown in Table 3, the AD and HD values derived from Voxel2Mesh and GAN model are significantly higher than that from the proposed framework. The GAN model shows the worst performance for AD and HD indicators.

From another persperctive, according to the data format being processed, the GAN model reconstruct the cortical surface based on

TABLE 3 The maximum IOU values with different loss functions: the L1 loss, Gaussian decay L1 loss and Truncate L1 loss functions.

Loss function	L1 loss	Gaussian decay L1 loss	Truncated L1 loss
IOU peak value	0.920	0.935	0.948

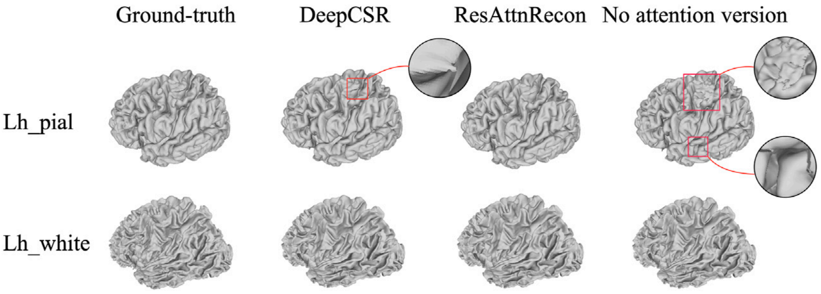


FIGURE 4
Visual assessment of cerebral cortical reconstruction results. The first row corresponds to the pial matter surface of the left hemisphere, and the second row corresponds to the white surface of the left hemisphere.

MRI voxels, 1) Due to the partial volume effect, the reconstruction accuracy is limited by the resolution of MRI volume, 2) The idea of generative adversarial network is to infer whether each voxel is inside or outside the surface of the cerebral cortex by training the generator, and then give a true or false judgment through the decision network. Each voxel has only binary information (inside or outside), while ignoring the distance information to the surface. Mesh based Voxel2Mesh model deform predefined sphere template meshes to cortical surfaces. 1) Despite fast processing, theoretical guarantees are lack to prevent self-intersections of the surface mesh. 2) Besides, a Voxel2Mesh model only able to reconstruct a single surface once (outer or inner surface). The proposed framework reconstruct the surface based on sampling points, and therefore has the theoretical support to reconstruct cortical surfaces of arbitrary resolution.

4.3 Visualization analysis

As illustrated in Figure 4, the reconstruction of the pial surface is more challenging than that of the white matter surface. The reconstructed pial surface of the left hemisphere is shown in the first row. The proposed ResAttn-Recon framework achieves the best reconstruction result, and the reconstruction result with DeepCSR has obvious surface bumps in the red rectangle area. The encoder-decoder architecture without the attention module (no attention version) has several reconstruction defects. The upper rectangular frame area shows severe reconstruction noise near the surface, and the lower rectangular frame area has multiple grid self-intersections. For the white matter reconstruction surface of the left hemisphere shown in the second row, there is no major visual difference between the three network structures in this case.

It is also worth noting that the ground-truth surfaces handled by FreeSurfer look rougher than the actual physiological surface. To address this, both the proposed framework and the DeepCSR framework have moderately smoothed the predicted surface through a post-processing algorithm. To further visualize the robustness of the proposed model, another eight examples of cortical surfaces (prediction and corresponding ground-truth) were illustrated in Figure 5.

4.4 Loss function evaluation

As seen in Figure 6, we compared the convergence curve of the intersection over union (IOU) based on the three loss functions mentioned above. By the way, for IOU, the SDF representation of the cortical surface could be further converted to the binarized SDF mask, 0 for inner surface points and one for outer surface points, then the binarized SDF mask based IOU could be further calculated between the ground-truth SDF and the predicted SDF. It is clear that the IOU curve based on the truncated L1 loss is distributed over the other two curves throughout the training process. For the other two curves, before 10 k training steps, the slope of the IOU curve based on the basic L1 loss is larger than that based on the Gaussian decay L1 loss, after which the IOU curve of the latter surpassed that of the former.

As shown in Table 3, the maximum IOU value of the trained with truncated L1 loss is 0.948 after convergence, followed by the loss function based on the Gaussian decay coefficient with a maximum IOU equals 0.935. The maximum IOU value based on the basic L1 loss is 0.92, which is considerably lower than the former. Also, thecorresponding convergence trend of the training curve can be confirmedfrom Figure 6.

The experimental results show that training with the truncated L1 loss and Gaussian decay L1 loss outperform the basic L1 loss, and among them, the truncated L1 is more effective. For the Gaussian decay L1 loss, the loss weighting coefficients tend to zero for sampling points away from the ground-truth surface, which leads to these points making less of a contribution for network training.

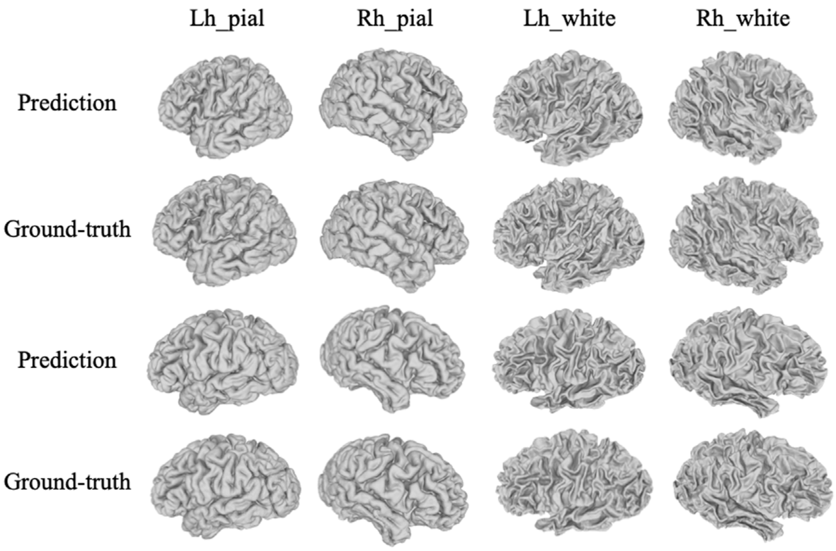


FIGURE 5
More visualization examples of the prediction results by the proposed method with corresponding ground-truth, including pial and white matter surfaces of left and right hemispheres.

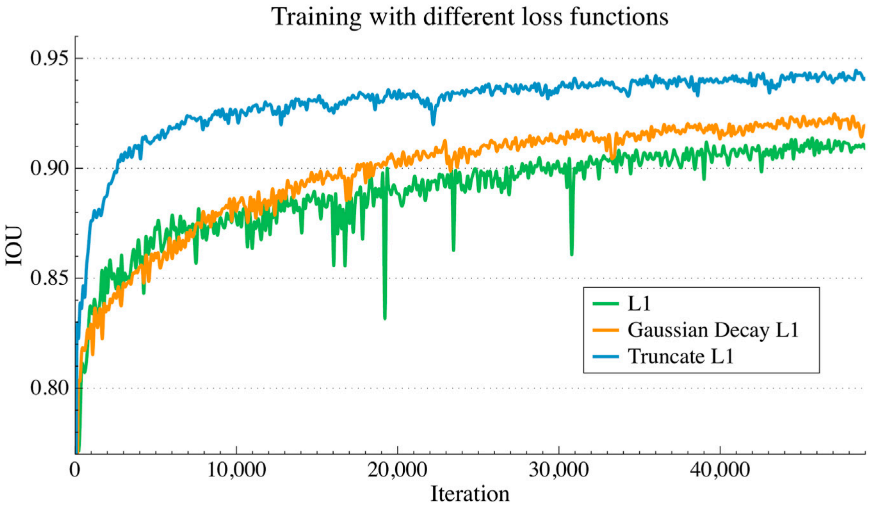


FIGURE 6
IOU convergence curves during training process with different loss functions. L1 loss, the truncated L1 loss and the Gaussian decay L1 loss were compared for analysis.

4.5 Precision analysis

To measure the precision of the reconstructed cortical surface, the average absolute distance (AD) is employed in millimeters (mm). Mesh vertices with an AD greater than 1 mm, 2mm and 5 mm as a percentage of the total number of mesh vertices are calculated.

Due to the highly folded and curved geometry, pial surface reconstruction is much more challenging than white surface reconstruction. Moreover, 60 ADNI test datasets with pial surface ground-truth are enrolled for precision analysis. As shown in Table 4, compared with DeepCSR, the AD percentage greater than 1 mm (3.616 vs 5.582 for left pial), 2 mm (1.467 vs 2.432 for left pial) and 5 mm (0.131 vs 0.249 for the left pial) are lower than those of the DeepCSR framework, indicating that the proposed method has a good robustness of the overall reconstruction.

5 Discussion

We successfully developed a residual self-attention-based architecture to reconstruct the inner and outer surface of the left

TABLE 4 Precision analysis for left pial matter surface and right pial matter surface: with range greater than 1mm, 2mm and 5 mm.

Method	Left pial matter surface			Right pial matter surface		
	AD (%>1 mm)	AD (%>2 mm)	AD (%>5 mm)	AD (%>1 mm)	AD (%>2 mm)	AD (%>5 mm)
DeepCSR	5.582	2.432	0.249	5.615	2.455	0.269
Proposed	3.616	1.467	0.131	3.588	1.468	0.153

and right hemispheres in parallel. In the current work of cortical surface reconstruction, including the voxel-based, mesh-based, and the implicit surfaces representation-based approaches, none of them considered the importance of capturing long range feature dependencies, during voxelwise or mesh vertex wise or sampling pointwise feature extraction, which is essential for 3D geometric surfaces representation is not considered. Given enough sampling points, the implicit surface representation-based approach has the theoretical support to reconstruct cortical surfaces of arbitrary resolution, our study is developed based on this direction and compared with the typical DeepCSR architecture. To adapt the residual connected multi-head self-attention to the 3D shape representation task, the 1 × 1 convolution is embedded in this module, with a learnable positional encoding matrix.

Ablation studies are undertaken to evaluate the impact of the residual self-attention module, the skip connections trick in decoder network, and the loss function with prior constraints. Experiments show that these components are effective in improving the reconstruction performance. We proposed the truncated L1 loss and the Gaussian decay weighted L1 loss to explore the effect of loss function on model expression potential. The truncated L1 loss achieves optimal results compared to simply applying the basic L1 loss function, and the training process achieved a higher IOU value (0.948 vs. 0.920) with the proposed truncated L1 loss. The average symmetric surface distance (AD) for the inner and outer surfaces is 0.253 ± 0.051, the average Hausdorff distance (HD) is 0.629 ± 0.186, which is lower than that of DeepCSR, whose AD equals 0.283 ± 0.059, and HD equals 0.746 ± 0.245. In addition, to measure the robustness of the overall reconstruction process, the AD greater than 1 mm, 2 mm and 5 mm as a percentage of the total number of mesh vertices are calculated, and we evaluated the challenging pial cortical surface result compared with DeepCSR, in 1 mm (3.616 vs. 5.582) and 2 mm (1.467 vs. 2.432) and 5 mm (0.131 vs. 0.249). From the perspective of visual analysis, the proposed ResAttn-Recon outperforms DeepCSR and the simple encoder-decoder architecture without the attention module. From Figure 5, it was found that the SDFs of pial matter surfaces are harder to approximate than the white matter surfaces during network training in parallel. Our proposed framework can better capture the surface details in a limited data size. Thus, the proposed residual self-attention-based framework can be a promising approach for improving the cortical surface reconstruction performance.

Our study has one main limitation: due to the lengthy processing time by the FreeSurfer pipeline, only a total of 560 T1 weighted images are enrolled in our dataset. In the future, more MRI volumes will be retrieved to expand the training dataset. Furthermore, it is desirable to enroll a larger pool of multicenter data to demonstrate the clinical value of this framework. Since topology correction algorithm during post processing usually takes a few minutes to enforce the prediction result to a desired topology, we will also pay more attention to the optimization of the post processing algorithm to further shorten the time of cortical surface reconstruction pipeline.

6 Conclusion

In this paper, we proposed ResAttn-Recon for challenging cerebral cortical surface reconstruction tasks. Firstly, we explored the concept of residual self-attention to the encoder-decoder architecture. Secondly, a lightweight decoder network with skip connection is developed to simplify the network without losing performance. In addition, experiments show that the proposed truncated L1 loss and Gaussian decay weighted L1 loss function contribute to the network training and performance improvement. The superior performance is achieved by the proposed framework compared with DeepCSR and a simple encoder-decoder framework without an attention block. The proposed framework can be a promising approach for improving the cortical surface reconstruction performance. We hope our work can inspire insights and show new directions toward cortical surface reconstruction study.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Boards of all of the participating institutions. Informed written consent was obtained from all participants at each site. The patients/participants provided their written informed consent to participate in this study.

Author contributions

MA performed all the experiments and wrote the manuscript, ZL analyzed and optimized the proposed algorithm. JC collected and cleaned the raw data. YC and KT conceived and supervised the project. JW, CW, and KZ reviewed the content of the manuscript and revised the grammar, they also provided valuable suggestions for modification of figures and tables. All authors contributed to the critical reading and writing of the manuscript.

Acknowledgments

The authors express special thanks to the Alzheimer's Disease Neuroimaging Initiative for available data providing, and all the medical students for assisting with the collection of the data.

Conflict of interest

Author KZ was employed by the company Inspur Electronic Information Industry Co., Ltd.

The remaining 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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2023.1003874/full#supplementary-material

References

1. Du A-T, Schuff N, Kramer JH, Rosen HJ, Gorno-Tempini ML, Rankin K, et al. 'Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia'. *Brain* (2006) 130(4):1159–66. doi:10.1093/brain/awm016

2. Rimol LM, Nesvag R, Hagler DJ, Bergmann O, Fennema-Notestine C, Hartberg CB, et al. 'Cortical volume, surface area, and thickness in schizophrenia and bipolar disorder'. *Biol Psychiatry* (2012) 71(6):552–60. doi:10.1016/j.biopsych.2011.11.026

3. Cruz RS, Lebrat L, Bourgeat P, Fookes C, Fripp J, Salvado O. DeepCSR: A 3D deep learning approach for cortical surface reconstruction, In: 2021 IEEE Winter Conference on Applications of Computer Vision (WACV); January 5-9, 2021; Waikoloa, HI, USA (2021). 806–15. doi:10.1109/WACV48630.2021.00085

4. Duan D, Xia S, Rekik I, Meng Y, Wu Z, Wang L, et al. 'Exploring folding patterns of infant cerebral cortex based on multi-view curvature features: Methods and applications'. *NeuroImage* (2019) 185:575–92. doi:10.1016/j.neuroimage.2018.08.041

5. MacDonald D, Kabani N, Avis D, Evans AC. 'Automated 3-D extraction of inner and outer surfaces of cerebral cortex from MRI'. *NeuroImage* (2000) 12(3):340–56. doi:10.1006/nimg.1999.0534

6. Kriegeskorte N, Goebel R. 'An efficient algorithm for topologically correct segmentation of the cortical sheet in anatomical MR volumes'. *NeuroImage* (2001) 14(2):329–46. doi:10.1006/nimg.2001.0831

7. Shattuck DW, Leahy RM. 'BrainSuite: An automated cortical surface identification tool'. *Med Image Anal* (2002) 6:129. doi:10.1016/s1361-8415(02)00054-3

8. Kim JS, Singh V, Lee JK, Lerch J, Ad-Dab'bagh Y, MacDonald D, et al. 'Automated 3-D extraction and evaluation of the inner and outer cortical surfaces using a Laplacian map and partial volume effect classification'. *NeuroImage* (2005) 27(1):210–21. doi:10.1016/j.neuroimage.2005.03.036

9. Fischl B. 'FreeSurfer'. *NeuroImage* (2012) 62(2):774–81. doi:10.1016/j.neuroimage.2012.01.021

10. Dahnke R, Yotter RA, Gaser C. 'Cortical thickness and central surface estimation'. *NeuroImage* (2013) 65:336–48. doi:10.1016/j.neuroimage.2012.09.050

11. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. *Neuroimage* (2012) 9:197. doi:10.1006/nimg.1998.0395

12. Rao C, Liu Y. 'Three-dimensional convolutional neural network (3D-CNN) for heterogeneous material homogenization'. *Comput Mater Sci* (2020) 184:109850. doi:10.1016/j.commatsci.2020.109850

13. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. (2016). *3d u-net: learning dense volumetric segmentation from sparse annotation. In International conference on medical image computing and computer-assisted intervention (Springer)*, 424–432.

14. Segonne F, Pacheco J, Fischl B. 'Geometrically accurate topology-correction of cortical surfaces using nonseparating loops'. *IEEE Trans Med Imaging* (2007) 26(4):518–29. doi:10.1109/tmi.2006.887364

15. Henschel L, Conjeti S, Estrada S, Diers K, Fischl B, Reuter M. 'FastSurfer - a fast and accurate deep learning based neuroimaging pipeline'. *NeuroImage* (2020) 219:117012. doi:10.1016/j.neuroimage.2020.117012

16. Gopinath K, Desrosiers C, Lombaert H. 'SegRecon: Learning joint brain surface reconstruction and segmentation from images'. In: *Medical image computing and computer assisted intervention – miccai 2021*. Cham: Springer International Publishing (2021). p. 650–9.

17. Gonzalezballester M. 'Estimation of the partial volume effect in MRI'. *Med Image Anal* (2002) 6(4):389–405. doi:10.1016/s1361-8415(02)00061-0

18. Ma Q, Robison EC, Kainz B, Rueckert D, Alansary A. (2021). *Pialnn: A fast deep learning framework for cortical pial surfacereconstruction. In International Workshop on Machine Learning in ClinicalNeuroimaging (Springer)*, 73–81.

19. Wickramasinghe U, Remmelli E, Knott G, Fua P. (2020). Voxel2mesh: 3d mesh model generation from volumetric data. InInternational Conference on Medical Image Computing and Computer-Assisted Intervention (Springer), 299–308.

20. Bongratz F, Rickman AM, Polserl S, Wachinger C. Vox2Cortex: Fast explicit reconstruction of cortical surfaces from 3D MRI scans with geometric deep neural networks. In: 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); June 18 2022 to June 24 2022; New Orleans, LA, USA (2022). 20741–51. doi:10.1109/CVPR52688.2022.02011

21. Park JJ, Florence P, Straub J, Newcombe R, Lovergrove S. (2019). Deepsdf: Learning continuous signed distance functions for shape representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 165–174.

22. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. (2017) *Attention is all you need. Advances in neural information processing systems*, 30.

23. Shi H, Gao J, Ren X, Xu H, Liang X, Li Z, et al. (2021) *Sparsebert: Rethinking the importance analysis in self-attention. In International Conference on Machine Learning (PMLR)*, 9547–9557.

24. Shen T, Zhou T, Long G, Jiang J, Pan S, Zhang C. (2018) DiSAN: Directional self-attention network for RNN/CNN-free language understanding, arXiv.

25. Ramachandran P, Parmar N, Vaswani A, Bello I, Levskaya A, Shlens J, et al. (2019) *Stand-alone self-attention in vision models. Advances in Neural Information Processing Systems*, 32.

26. Zhao H, Jia J, Koltun V. Exploring self-attention for image recognition. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); June 13 2020 to June 19 2020; Seattle, WA, USA (2020). 10073–82. doi:10.1109/CVPR42600.2020.01009

27. Vaswani A, Ramachandran P, Srinivas A, Parmar N, Hetchman B, Shlens J. Scaling local self-attention for parameter efficient visual backbones. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); June 20 2021 to June 25 2021; Nashville, TN, USA (2021). p. 12889–99. doi:10.1109/CVPR46437.2021.01270

28. 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*.

29. Yuan L, Chen Y, Wang Y, Yu W, Shi Y, Jiang Z, et al. Tokens-to-Token ViT: Training vision transformers from scratch on ImageNet. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV); Oct. 11 2021 to Oct. 17 2021; Montreal, QC, Canada (2021). p. 538–47. doi:10.1109/ICCV48922.2021.00060

30. 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.

31. Petersen RC, Aisen PS, Beckett LA, Donohue MC, Gamst AC, Harvey DJ, et al. ‘Alzheimer’s disease neuroimaging initiative (ADNI): Clinical characterization’. *Neurology* (2010) 74(3):201–9. doi:10.1212/wnl.0b013e3181cb3e25

32. Mazziotta JC, Toga AW, Evans A, Fox P, Lancaster J. ‘A probabilistic atlas of the human brain: Theory and rationale for its development’. *Neuroimage* (1995) 2(2): 89–101. doi:10.1006/nimg.1995.1012

33. Lorensen WE, Cline HE. Computer graphics. *J Comp* (1987) 21, 7.

34. Gehring J, Auli M, Grangier D, Yarats D, Dauphin YN. (2017) ‘Convolutional sequence to sequence learning’. arXiv. doi:10.48550/arXiv.1705.03122

35. Bazin P-L, Pham DL. ‘Topology correction of segmented medical images using a fast marching algorithm’. *Comp Methods Programs Biomed* (2007) 88(2):182–90. doi:10.1016/j.cmpb.2007.08.006

36. Tosun D, Rettmann ME, Naiman DQ, Resnick SM, Kraut MA, Prince JL. ‘Cortical reconstruction using implicit surface evolution: Accuracy and precision analysis’. *NeuroImage* (2006) 29(3):838–52. doi:10.1016/j.neuroimage.2005.08.061

37. Taha AA, Hanbury A. ‘Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool’. *BMC Med Imaging* (2015) 15(1):29. doi:10.1186/s12880-015-0068-x

38. Ning Z, Zhang Y, Pan Y, Zhong T, Liu M, Shen D. (2020). “Ldgan: Longitudinal-diagnostic generative adversarial network for disease progression prediction with missing structural MRI,” in M Liu. (eds) *Machine learning in medical imaging*. Cham: Springer International Publishing(Lecture Notes in Computer Science), 170–9. doi:10.1007/978-3-030-59861-7_18