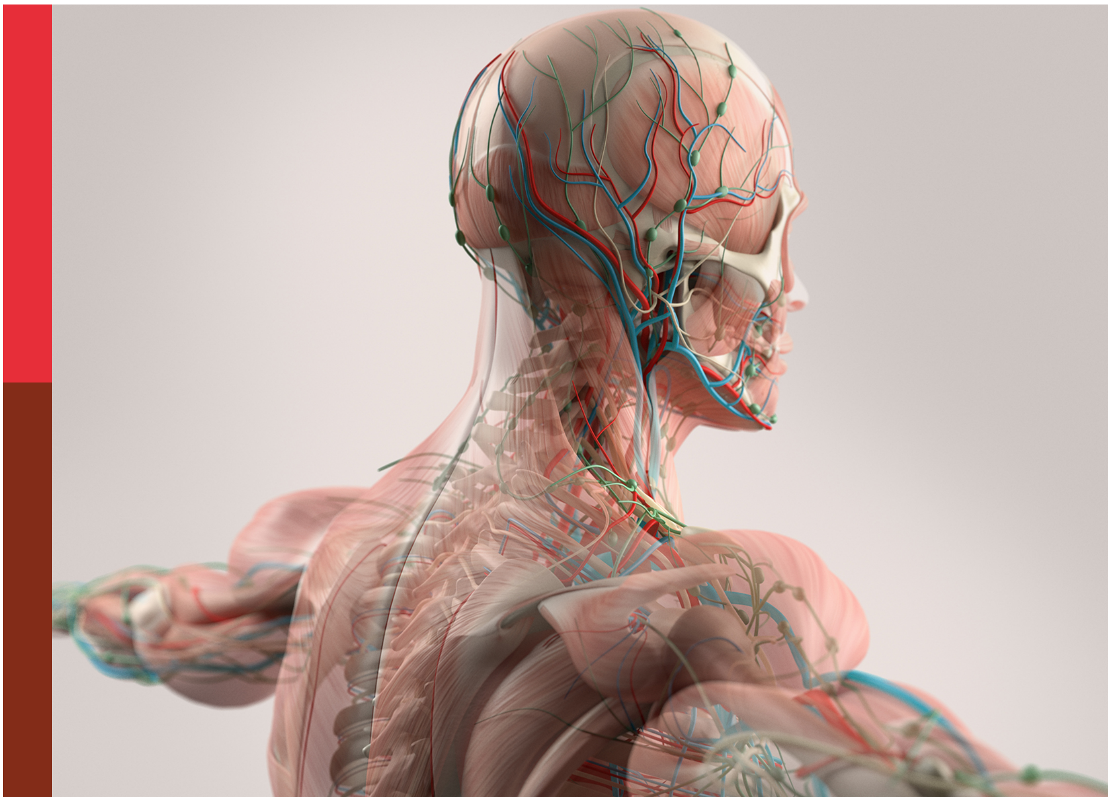
Research Topics

Microwave-induced thermoacoustic imaging/photoacoustic imaging

Edited by
Huan Qin, Jiao Li, Junhui Shi, Hao Yang and Lin Huang

Published in
Frontiers in Physiology
Frontiers in Physics



ISSN 1664-8714
ISBN 978-2-83251-447-4
DOI 10.3389/978-2-83251-447-4

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

Microwave-induced thermoacoustic imaging/photoacoustic imaging

Topic editors
Huan Qin — South China Normal University, China
Jiao Li — Tianjin University, China
Junhui Shi — Zhejiang Lab, China
Hao Yang — University of South Florida, United States
Lin Huang — University of Electronic Science and Technology of China, China

Citation
Qin, H., Li, J., Shi, J., Yang, H., Huang, L., eds. (2023). *Microwave-induced thermoacoustic imaging/photoacoustic imaging*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-83251-447-4

Table of contents

frontiers | Frontiers in Physics

ORIGINAL RESEARCH
published: 27 June 2022
doi: 10.3389/fphy.2022.939657

The Relationship Between Dielectric Properties, Thermoacoustic Signals and Temperature of Liver Tissues at 3.0 GHz

Jiawu Li[1], *Lin Huang*[2]*, *Yitong Peng*[2], *En Li*[2] *and Yan Luo*[1]*

[1]West China Hospital, Sichuan University, Chengdu, China, [2]School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China

Purpose: This study aimed to investigate the relationship between dielectric properties (permittivity and conductivity), thermoacoustic signals (TAS) and temperature of liver tissues at 3.0 GHz.

Materials and Methods: An open-ended coaxial probe was used to measure the dielectric properties of fresh porcine liver tissues *in vitro*, and a thermoacoustic imaging (TAI) system was used to collect the TAS. Porcine liver tissues were placed on a heating platform, and the dielectric properties measuring probe and temperature sensor were inserted into the liver tissues separated by 1.5 cm. First, the liver tissues were gradually heated by a heating platform from room temperature (30°C) to 60°C, and the dielectric properties and TAS were measured as the temperature increased approximately every 1°C. Second, after the temperature of the porcine liver tissue reached 60°C, the heating platform was turned off to naturally cool the porcine liver tissue to room temperature. During the process, the dielectric properties and TAS were also measured as the temperature decreased every 1°C. Finally, the changes in the dielectric properties and TAS of the liver tissues with temperature at 3.0 GHz frequency were analyzed.

Results: During the process of heating the tissues up to 60°C, the conductivity of the porcine liver tissues decreased while the permittivity and TAS of the porcine liver tissues increased, and the relationships were nonlinearly correlated. Meanwhile, during the cooling process, the conductivity of the porcine liver tissues increased, while the permittivity and TAS decreased, and the relationships were also nonlinearly correlated.

Conclusion: The dielectric properties and TAS of porcine liver tissue changed significantly with temperature, which makes it possible to differentiate the safety margin during liver thermal ablation with thermoacoustic imaging (TAI). While the relationship between temperature, dielectric properties and TAS needs to be further investigated, TAI has the potential to be utilized for safety margin screening during thermal ablation.

Keywords: conductivity, permittivity, thermoacoustic signal, temperature, liver

Edited by:
Yuancheng Fan,
Northwestern Polytechnical University, China

Reviewed by:
Guoqiang Liu,
Institute of Electrical Engineering (CAS), China
Ningfeng Bai,
Southeast University, China

Correspondence:
Lin Huang
lhuang@uestc.edu.cn
Yan Luo
yanluo@scu.edu.cn

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

Received: *09 May 2022*
Accepted: *13 June 2022*
Published: *27 June 2022*

Citation:
Li J, Huang L, Peng Y, Li E and Luo Y (2022) The Relationship Between Dielectric Properties, Thermoacoustic Signals and Temperature of Liver Tissues at 3.0 GHz. Front. Phys. 10:939657. doi: 10.3389/fphy.2022.939657

INTRODUCTION

Hyperthermia therapy is a medical treatment that exposes biological tissue to high temperatures to destroy and kill cancer cells, and has become an important tumor treatment modality in addition to surgical treatment, chemotherapy, radiotherapy and immunotherapy. The key to local thermal ablation is to produce a necrotic area that completely covers the tumor. Therefore, real-time evaluation of the ablation process and the safety margin is essential. Thermal ablation is mainly used to treat liver tumors smaller than 3 cm [1]; however, with the expansion of indications, single liver tumors smaller than, for example, 5 cm can also be treated with thermal ablation [2], which makes the evaluation of safety margins more important to ensure complete tumor ablation. The dielectric properties of biological tissues carry abundant physiological and pathological information, but they are sensitive to environmental factors, especially the influence of temperature [3, 4]. Therefore, changes in tissue dielectric properties must be considered during thermal ablation. Moreover, knowledge of the temperature-dependent dielectric properties of liver tissues is important to imaging modalities that generate images based on differences in dielectric properties.

Since the importance of the dielectric properties of biological tissues to the diagnosis and treatment of diseases has been recognized, many researchers have studied the dielectric properties of human or animal biological tissues [5–7]. Although several studies have explored the dielectric properties of liver tissues at different frequencies at room temperature [8, 9], the dielectric properties of liver tissues at different temperatures need further investigation. Recently, there have been studies on the correlation between the dielectric properties of liver tissues and temperature [10–15]. Although most of these studies were limited to a narrow frequency range of 460–480 kHz (RF frequencies), they indicated that the dielectric properties of liver tissue are temperature dependent. In recent years, the dielectric properties of biological tissues under microwave frequencies have also attracted more attention [14, 16]. However, the temperature-dependent dielectric properties of liver tissues within the temperature range of 30°C–60°C at 3.0 GHz have been only partially studied. In clinical thermal ablation, 2.45 GHz is mainly utilized. Thus, it is necessary to further explore the temperature-dependent dielectric properties of liver tissues during the heating and cooling process.

Thermoacoustic imaging (TAI) is a novel imaging modality that can reflect the biological characteristics (especially dielectric properties) of tissues *via* thermoacoustic effects, and has the advantages for high contrast of microwave imaging and high resolution for ultrasound imaging [17]. Some studies have investigated the relationship between thermoacoustic signals (TAS) and temperature in phantom and biological tissues [18, 19], while a few studies have explored the temperature-dependent TAS and dielectric properties simultaneously in liver tissues [19, 20]. Meanwhile, the combination of temperature, dielectric properties and TAS has certain guiding significance for the theoretical study of TAI for the detection of ablation safety boundaries. Furthermore, the frequency of TAI is 3.0 GHz, which is close to the frequency used for microwave ablation (2.45 GHz), and the dielectric properties of liver tissue are similar at these two frequencies [5]. Therefore, this study aimed to explore the dielectric property changes of *ex vivo* porcine liver tissue during the heating and cooling process, as well as the temperature dependence of TAS at 3.0 GHz.

MATERIALS AND METHODS

In this study, fresh porcine liver tissue was obtained from a local slaughterhouse. All porcine liver tissues used in the experiment were acquired from different individuals to minimize sampling error. Homogeneous blocks of porcine liver, at least 3*3 cm^2, were used for the experiments. Each *ex vivo* porcine liver tissue was wrapped in aluminum foil to keep it fresh before experiments. The initial temperature at the beginning of the experiments was approximately 30°C.

Calibration of the Dielectric Properties Measuring Device

The dielectric property measurement was conducted with an open-ended coaxial system [21], as shown in **Figure 1**. In this study, short standard calibration was attained by pressing the coaxial probe on a thin copper sheet. Then, an open standard was established by placing the probe in free space, and the load standard was completed by immersing the probe into deionized water. After calibration the performance of the coaxial probe was verified by measuring the conductivity of pure water during the heating and cooling process, which is shown in **Figure 2**.

Dielectric Property Measurement of Liver Tissues

Figures 1A,B shows the placement of the coaxial probe and thermometer. The *ex vivo* porcine liver tissues were placed on a heating platform (JF-956K, Tianjin, China), and the temperature of the heating platform was controlled by an external temperature regulating system. A digital thermometer (Omega Engineering, Inc., Stanford, CT) was used to measure the temperature of the liver tissues. The dielectric property measuring probe was inserted 0.5 cm deep into the center of the liver tissue, and two temperature probes were inserted at the same depth in the same plane approximately 1.5 cm away from the dielectric property testing probe. The measuring system was kept stationary throughout the experiment. In total, four liver tissues were used in the study. The probe was wiped with clean test paper before the measurement of the different liver tissues so that probe residues did not affect the test results.

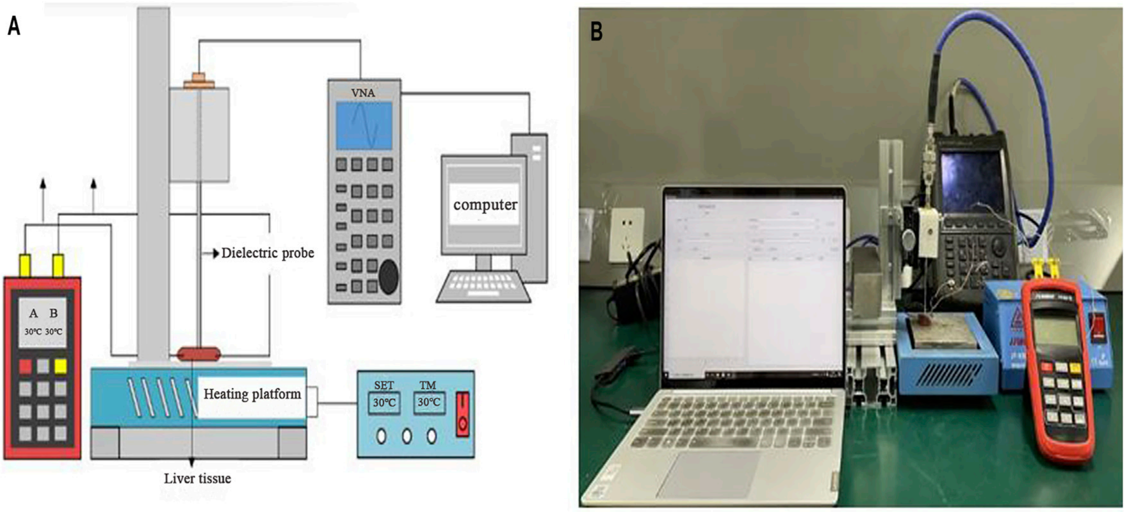


FIGURE 1 | Schematic diagram **(A)** and image **(B)** of the experimental setup for dielectric property measurement in *ex vivo* liver tissues. The porcine liver tissue was placed on the surface of a heating platform, and the dielectric properties testing probe and temperature probes were inserted into liver tissues at an interval of 1.5 cm.

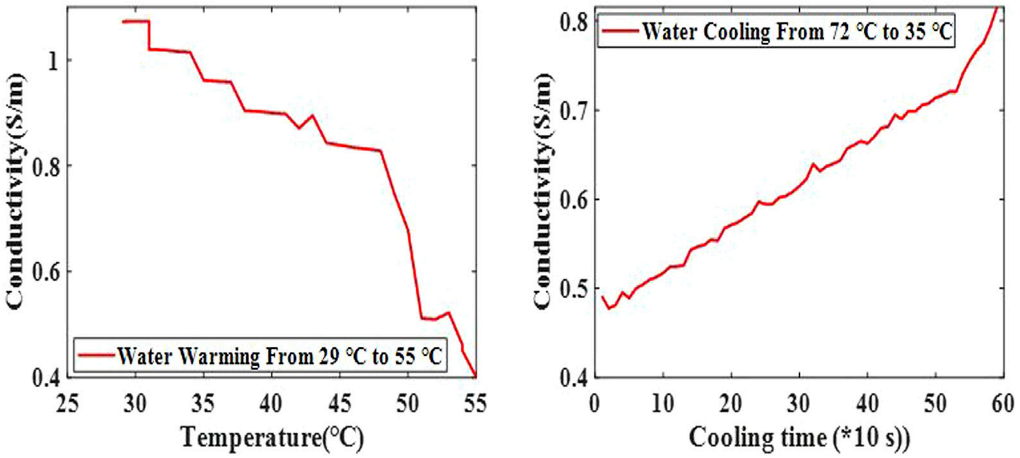


FIGURE 2 | Conductivity changes of pure water with temperature. The conductivity of pure water decreased gradually in the heating process and increased gradually in the cooling process.

Thermoacoustic Signal Measurement of Liver Tissues

The thermoacoustic signal (TAS) was collected with a previously reported array transducer-based system [22] [**Figure 3A** presents a schematic of the TA imaging system]. Briefly, pulsed microwaves emitted from a homemade microwave generator (frequency: 3.0 GHz, bandwidth: 50 MHz, peak power: 60 kW, pulse duration: 550 ns) were coupled to a handheld antenna [23] *via* a semirigid coaxial cable (1.5 m long with 2.2 dB insertion loss). The actual average microwave power density at the liver was only approximately 15.0 mW/cm^2 when a 50 Hz repetition frequency and 550 ns pulse duration were utilized, which is below the IEEE standard for safety levels (20 mW/cm^2 at 3.0 GHz) [24]. The excited TAS was captured by a 128-element linear array transducer (8.5 MHz center frequency, SH7L38, SASET. Inc., China) and further amplified by homemade 128 channel amplifiers (bandwidth: 0.2–2.5 MHz, gain: 56 dB). Finally, the TAS was averaged 50 times and recorded by two 32-channel acquisition cards (NI5752, NI Inc., United States) at a sampling rate of 50 MPs. The TA images were recovered by using a delay and sum algorithm created by MATLAB (Mathworks Inc., Natic, MA) [25]. To investigate the relationship between TAS and the temperature of liver tissue, the TA images were obtained (**Figures 3B,C**), and the sum of the area within the white square was regarded as the TAS for further analysis.

During the heating process, the temperature of the liver tissue was gradually increased to 60°C by adjusting the power of the heating plate. When the temperature of the liver tissue reached 60°C, the

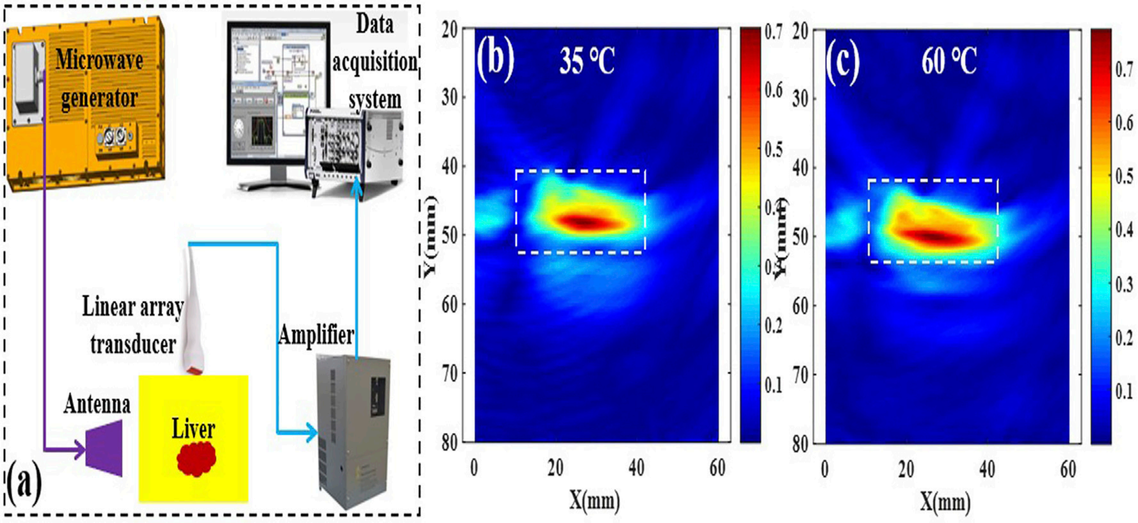


FIGURE 3 | Schematic of the TA imaging system **(A)**, and the corresponding TA images of porcine liver tissue at 35 and 60°C **(B,C)**.

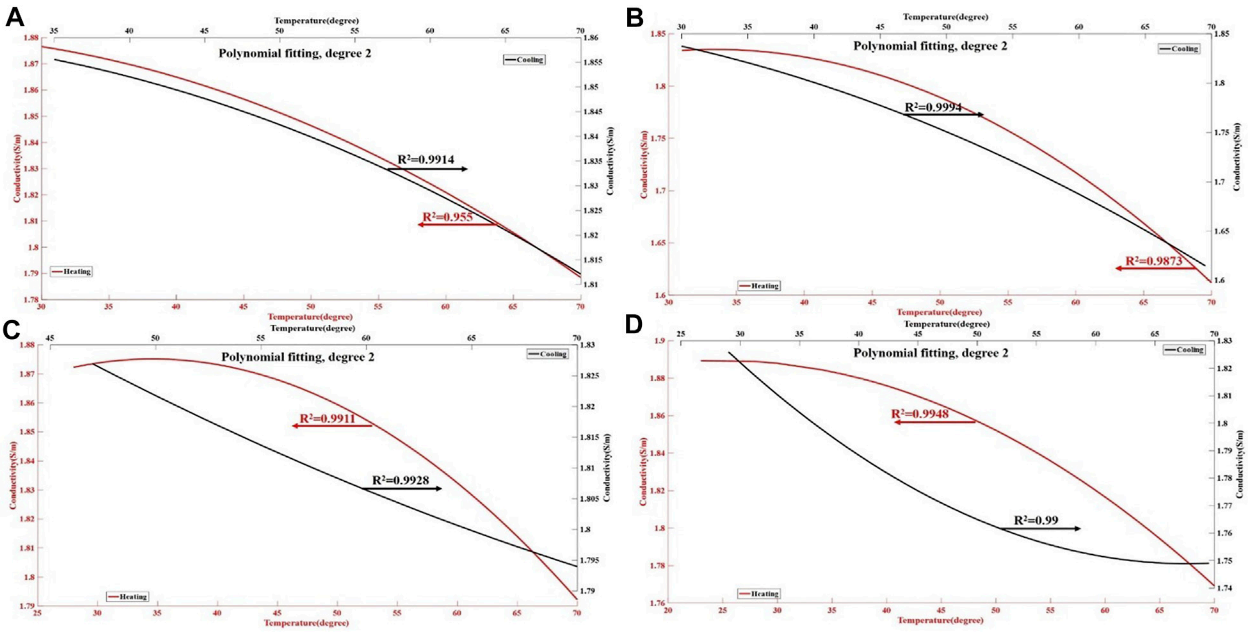


FIGURE 4 | Changes in the conductivity properties of *ex vivo* liver tissue during the heating and cooling process. The red lines in the figures **(A–D)** represent the polynomial curve fitting of the conductivity of liver tissue during the heating process, with an R^2 values of 0.955–0.9948. The black line represents the polynomial curve fitting of conductivity of liver tissue during the cooling process, with an R^2 of 0.99–0.9994. The results show an almost nonlinear relationship between the conductivity of liver tissues and temperature.

heating platform was turned off to naturally cool the liver tissue to room temperature. The dielectric properties and TAS of *ex vivo* porcine liver tissues at temperatures ranging from room temperature (30°C) to 60°C at a frequency of 3.0 GHz (the frequency used for thermoacoustic imaging) were measured. A temperature above 60°C was not attempted because significant physiological changes were expected in tissues above 56°C–60°C [26, 27], which is also an indicator of effective tumor destruction during thermal ablation [28]. The dielectric properties and TAS were measured every 1°C throughout the heating and cooling process. The relationship between dielectric properties (permittivity and conductivity), TAS and temperature of the liver tissue at 3.0 GHz frequencies was expressed by polynomial curve fitting. Higher correlations have correlation coefficient, R^2, closer to 1.

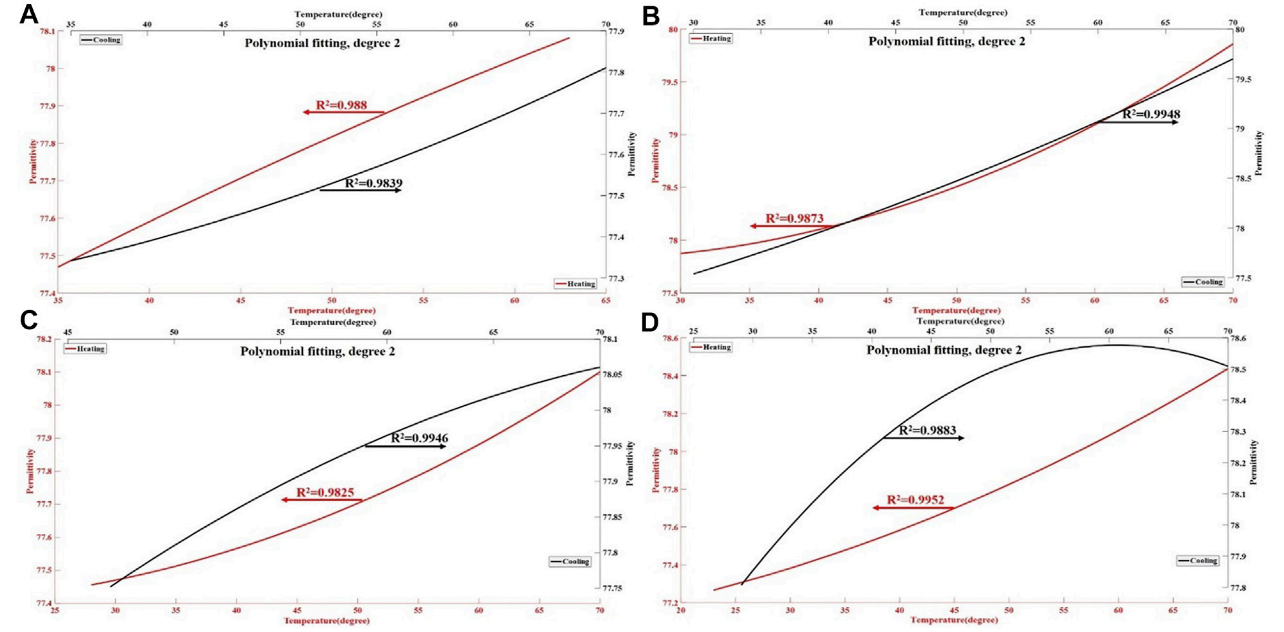


FIGURE 5 | The permittivity changes of *ex vivo* liver tissue during the heating and cooling process. The red lines in the figures **(A–D)** represent the polynomial curve fitting of the permittivity of liver tissue during the heating process, with an R^2 values of 0.9825–0.9952. The black line represents the polynomial curve fitting of permittivity of liver tissue during the cooling process, with an R^2 of 0.9839–0.9948. The results show an almost nonlinear relationship between the permittivity of liver tissue and temperature.

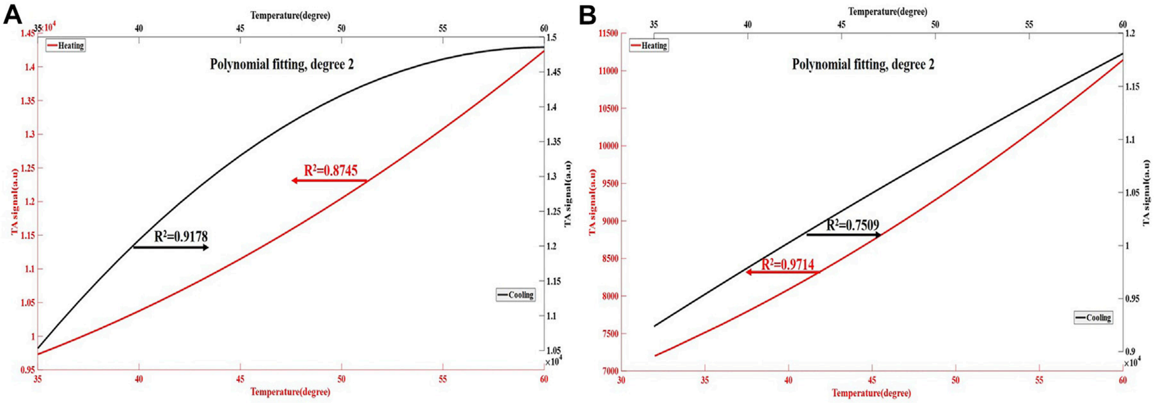


FIGURE 6 | The thermoacoustic signals changes of *ex vivo* liver tissue during the heating and cooling process. The red line in the figure **(A,B)** represents the polynomial curve fitting of thermoacoustic signal of liver tissue during the heating process, with an R^2 values of 0.8745–0.9714. The black line represents the polynomial curve fitting of thermoacoustic signal of liver tissue during the cooling process, with an R^2 of 0.7509–0.9178. The results show that almost a nonlinear relationship between thermoacoustic signal of liver tissue and temperature.

RESULTS

The Relationship Between the Conductivity of Pure Water and Temperature

The conductivity of pure water decreased gradually in the heating process and increased gradually in the cooling process, which is consistent with other studies [29] (**Figure 2**). This may be related to higher temperatures, resulting in more intense thermal motion of water molecules, making it more difficult to induce uniform polarization.

Temperature-Dependent Dielectric Properties and Thermoacoustic Signals of *Ex Vivo* Porcine Liver Tissue

In general, the dielectric properties and TAS changes of liver tissues with temperature were consistent in all the test groups. During the

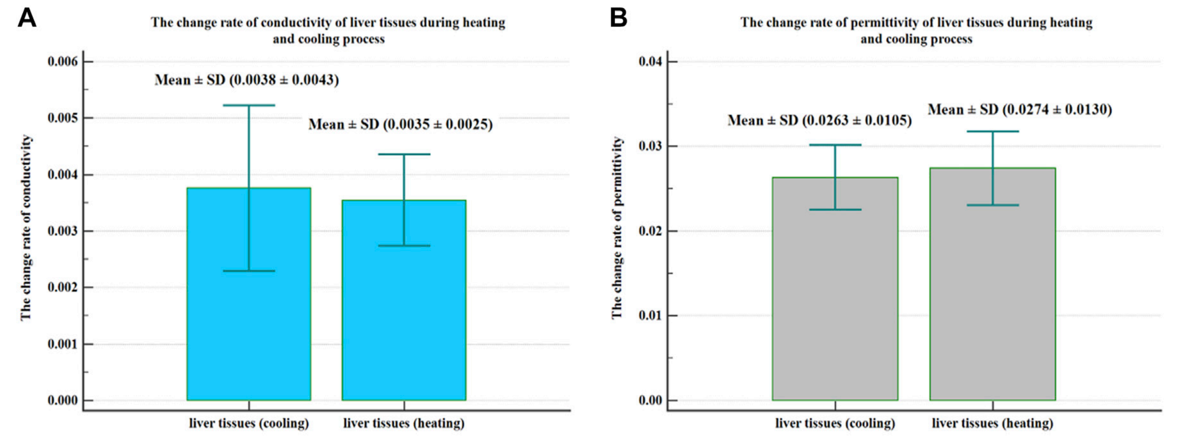


FIGURE 7 | The change rate of conductivity and permittivity of liver tissues during the heating and cooling process. During the heating and cooling process of liver tissues, the change rate of conductivity per degree was 0.0035 ± 0.0025 and 0.0038 ± 0.0043, respectively **(A)**, and the change rate of permittivity per degree was 0.0274 ± 0.0130 and 0.0263 ± 0.0105, respectively **(B)**.

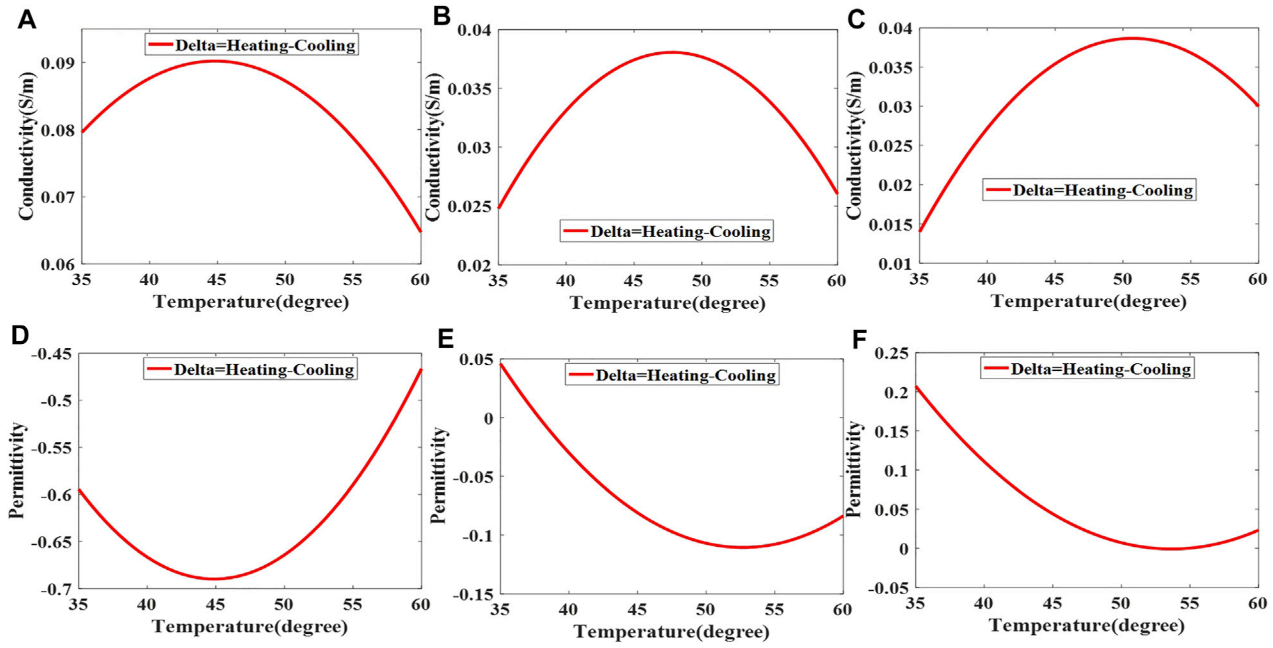


FIGURE 8 | The difference between conductivity **(A–C)** and permittivity **(D–F)** of liver tissues at the same temperature point during the heating and cooling process.

process of heating to 60°C, the conductivity of the porcine liver tissues decreased with temperature (**Figure 4**), the permittivity (**Figure 5**) and TAS (**Figure 6**) of the porcine liver tissues increased, and the relationships were nonlinearly correlated. Meanwhile, during the cooling process, the conductivity of the porcine liver tissues increased with temperature, while the permittivity and thermoacoustic signals decreased, and the relationships were also nonlinearly correlated.

In addition, during the processes of heating and cooling the liver tissues, the change rates of the conductivity per degree were 0.0035 ± 0.0025 and 0.0038 ± 0.0043, respectively. The change rates of the permittivity per degree were 0.0274 ± 0.0130 and 0.0263 ± 0.0105, respectively. The change rates of the conductivity and permittivity of the liver tissues during the heating and cooling process are presented in **Figure 7**.

During the process of heating and cooling, the conductivity and permittivity at the same temperature point are different, and **Figure 8** shows the change curve of the difference.

DISCUSSION

This study explored the change in dielectric properties and TAS of fresh *ex vivo* porcine liver tissue with temperature. The results showed that the permittivity, conductivity and TAS of porcine liver tissues changed regularly with temperature at 3.0 GHz. Recently, an increasing number of studies have been conducted on the dielectric properties of the liver in the radiofrequency (RF) and microwave (MW) frequency ranges with changes in temperature.

In the RF frequency range, [10] explored the temperature-dependent *in-vivo* and *ex-vivo* conductivity at 470 kHz, and the results showed that conductivity increased continuously and uniformly from body temperature to 77°C, and a continuous reduction in conductivity was observed during the cooling phase. In addition, [11] measured the dielectric properties of porcine liver *ex vivo* up to 100°C at 480 kHz and found increasing conductivity with temperature. Furthermore, [12] measured the conductivity of liver tissue at 500 kHz at temperatures ranging from 40 to 90°C. They found that the conductivity increased with temperature, and the change was irreversible above 60°C. The study of [30] indicated that at 915 MHz, the permittivity of *ex vivo* bovine liver decreased with temperature, while the conductivity gradually increased.

However, the frequency of RF ablation is lower than that of microwave ablation. To date, only a few studies have been found on the temperature-dependent conductivity properties of liver tissues at a higher frequency (1–3 GHz). [14] showed a significant decrease in conductivity (approximately 33%) in liver tissue at 2.45 GHz when the temperature increased to over 60°C, with a dramatic drop when the temperature was close to 100°C. [31] explored the changes in the dielectric properties in liver tissue at 915 and 2.45 GHz with temperatures up to 100°C. Their study indicated that permittivity and conductivity decrease substantially and irreversibly at high temperatures. Interestingly, as seen from the graphs presented in their results, when the temperature was below 60°C, the conductivity increased with increasing temperature at 915 MHz and decreased with increasing temperature at 2.45 GHz. The electrical conductivity results of liver tissues with temperature increases at higher frequencies are consistent with our study. It seems that at a lower frequency (below 1 GHz), the conductivity of liver tissue increases with temperature, and at a higher frequency (1–3 GHz), it decreases with temperature. This is related to the fact that the ionic loss increases with temperature, and the relaxation process is dominated by ionic losses at low frequencies and dielectric losses at higher frequencies.

Our study showed that the permittivity and conductivity changes of the liver tissues with temperature were nonlinear at 3.0 GHz. Furthermore, previous studies have reported that dielectric properties are not always linear over a wide range of temperatures and frequencies [15, 32]. However, studies have proposed that the temperature coefficients of permittivity and conductivity are linear, which can be used to further deduce the dielectric properties of biological tissues [30, 33]. Further studies are needed to determine whether there is a linear relationship between temperature and the permittivity and conductivity of biological tissues. In addition, we found that as the temperature increased, the dielectric properties of the liver tissues changed irreversibly; that is, at the same temperature point during the heating and cooling process, the conductivity and permittivity were different. This change is most obvious when the temperature reaches 45°C, which may be related to the aggravation of tissue cell damage after the temperature exceeds 45°C [34, 35], rather than tissue dehydration, as significant tissue dehydration occurs when the temperature reaches 100°C. In this study, we further explored the changes in the thermoacoustic signals of liver tissues with temperature. The results showed that the TAS of liver tissues increased with temperature and decreased during the cooling process. [19] explored the relationship between TAS and temperature in porcine livers *in vitro*, and the results showed that the thermoacoustic signal increased with temperature (30°C–50°C), and the relationship was nonlinear, which is in agreement with our results. The changes in TAS were consistent with the changes in permittivity with temperature, which could provide theoretical guidance for thermoacoustic imaging to distinguish ablation tissue from normal tissue.

However, the current study has some limitations. The microwave ablation uses 2.45 GHz for thermal ablation, which was slightly different from the frequency used for TAI (3.0 GHz). However, according to a previous study [36] and an open data site "Calculation of the Dielectric Properties of Body Tissues in the frequency range 10 Hz–100 GHz" [5], the dielectric properties of the liver at 2.45 GHz (s = 1.6864 S/m, εr = 43.035) and 3.0 GHz (s = 2.0755 S/m, εr = 42.165) are close. In addition, the results were obtained from *in vitro* porcine liver tissue, which may differ from *in vivo* experiments. Finally, the porcine liver tissues used in the experiment were acquired from different individuals, but the resulting dielectric properties and thermoacoustic signals showed a consistent trend with temperature.

CONCLUSION

In this study, we explored the temperature-dependent dielectric properties and TAS of porcine liver tissue at 3.0 GHz. The experimental results indicated that the dielectric properties and TAS of porcine liver tissue changed significantly with temperature, which makes it possible to differentiate the safety margin during liver ablation by thermoacoustic imaging. However, the relationship between temperature, dielectric properties and TAS needs to be further investigated, especially the quantitative correlation between TAS and temperature, which will be useful in the quantitative evaluation of clinical thermal ablation boundaries.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The animal study was reviewed and approved by the West China Hospital of Sichuan University.

AUTHOR CONTRIBUTIONS

Study design: All authors. Data Collection: JL, YP, and LH. Data analysis and interpretation: JL and LH. Manuscript writing: JL. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by National Natural Science Foundation of China (Nos 82001832, 81701702, and 82071940).

REFERENCES

1. Vogl T, Nour-Eldin N-E, Hammerstingl R, Panahi B, Naguib N. Microwave Ablation (MWA): Basics, Technique and Results in Primary and Metastatic Liver Neoplasms - Review Article. *Fortschr Röntgenstr* (2017) 189:1055–66. doi:10.1055/s-0043-117410
2. Deng M, Li S-H, Guo R-P. Recent Advances in Local Thermal Ablation Therapy for Hepatocellular Carcinoma. *The Am Surgeon* (2021) 000313482110545 Epub ahead of print. doi:10.1177/00031348211054532
3. Geddes LA, Baker LE. The Specific Resistance of Biological Material-A Compendium of Data for the Biomedical Engineer and Physiologist. *Med Biol Engng* (1967) 5:271–93. doi:10.1007/BF02474537
4. Zia G, Sebek J, Prakash P. Temperature-dependent Dielectric Properties of Human Uterine Fibroids over Microwave Frequencies. *Biomed Phys Eng Express* (2021) 7: 065038. doi:10.1088/2057-1976/ac27c2
5. Andreuccetti D, Fossi R, Petrucci C. *Calculation of the Dielectric Properties of Body Tissues in the Frequency Range 10 Hz-100 GHz Florence: IFAC-CNR 1997-2015*. (Florence (Italy): Institute for Applie Physics). (2021). http://niremf.ifac.cnr.it/tissprop/htmlclie/htmlclie.php.
6. Gabriel C, Peyman A, Grant EH. Electrical Conductivity of Tissue at Frequencies below 1 MHz. *Phys Med Biol* (2009) 54:4863–78. doi:10.1088/0031-9155/54/16/002
7. Yilmaz T, Ates Alkan F. *In Vivo* Dielectric Properties of Healthy and Benign Rat Mammary Tissues from 500 MHz to 18 GHz. *Sensors* (2020) 20:2214. doi:10.3390/s20082214
8. Wang L, Wang H, Xu C, Ji Z, Li J, Dong X, et al. Dielectric Properties of Human Active Liver, Kidney and Spleen Compared to Those of Respective Inactive Tissues, Porcine Tissues and the Data Provided by a Database in the Frequency Range of 10 Hz to 100 MHz. *IEEE Trans Biomed Eng* (2021) 68:3098–109. doi:10.1109/TBME.2021.3065016
9. O'Rourke AP, Lazebnik M, Bertram JM, Converse MC, Hagness SC, Webster JG, et al. Dielectric Properties of Human normal, Malignant and Cirrhotic Liver Tissue:in Vivoandex Vivomeasurements from 0.5 to 20 GHz Using a Precision Open-Ended Coaxial Probe. *Phys Med Biol* (2007) 52:4707–19. doi:10.1088/0031-9155/52/15/022
10. Zurbuchen U, Holmer C, Lehmann KS, Stein T, Roggan A, Seifarth C, et al. Determination of the Temperature-dependent Electric Conductivity of Liver Tissue *Ex Vivo* and *In Vivo*: Importance for Therapy Planning for the Radiofrequency Ablation of Liver Tumours. *Int J Hyperthermia* (2010) 26: 26–33. doi:10.3109/02656730903436442
11. Macchi EG, Gallati M, Braschi G, Persi E. Dielectric Properties of RF Heatedex Vivoporcine Liver Tissue at 480 kHz: Measurements and Simulations. *J Phys D: Appl Phys* (2014) 47:485401. doi:10.1088/0022-3727/47/48/485401
12. Ryan TP, Platt RC, Dadd JS, Humphries S. Tissue Electrical Properties as a Function of thermal Dose for Use in a Finite Element Model. *Adv Heat Mass Transfer Biotechnol* (1997) 355:167–71.
13. Fallahi H, Sebek J, Prakash P. Broadband Dielectric Properties of *Ex Vivo* Bovine Liver Tissue Characterized at Ablative Temperatures. *IEEE Trans Biomed Eng* (2021) 68:90–8. doi:10.1109/TBME.2020.2996825
14. Lopresto V, Pinto R, Lovisolo GA, Cavagnaro M. Changes in the Dielectric Properties Ofex Vivobovine Liver during Microwave thermal Ablation at 2.45 GHz. *Phys Med Biol* (2012) 57:2309–27. doi:10.1088/0031-9155/57/8/2309
15. Fu F, Xin SX, Chen W. Temperature- and Frequency-dependent Dielectric Properties of Biological Tissues within the Temperature and Frequency Ranges Typically Used for Magnetic Resonance Imaging-Guided Focused Ultrasound Surgery. *Int J Hyperthermia* (2014) 30:56–65. doi:10.3109/02656736.2013.868534
16. Silva NP, Bottiglieri A, Conceição RC, O'Halloran M, Farina L. Characterisation of *Ex Vivo* Liver Thermal Properties for Electromagnetic-Based Hyperthermic Therapies. *Sensors* (2020) 20: 3004. doi:10.3390/s20103004
17. Kruger RA, Reinecke DR, Kruger GA. Thermoacoustic Computed Tomography-Technical Considerations. *Med Phys* (1999) 26:1832–7. doi:10.1118/1.598688
18. Pramanik M, Wang LV. Thermoacoustic and Photoacoustic Sensing of Temperature. *J Biomed Opt* (2009) 14:054024. doi:10.1117/1.3247155
19. Lou C, Xing D. Temperature Monitoring Utilising Thermoacoustic Signals during Pulsed Microwave Thermotherapy: A Feasibility Study. *Int J Hyperthermia* (2010) 26:338–46. doi:10.3109/02656731003592035
20. Zhou T, Meaney PM, Hoopes PJ, Geimer SD, Paulsen KD. Microwave thermal Imaging of Scanned Focused Ultrasound Heating: Animal Experiments. *Proc SPIE* (2011) 7901:79010N. doi:10.1117/12.877584
21. La Gioia A, Porter E, Merunka I, Shahzad A, Salahuddin S, Jones M, et al. Open-Ended Coaxial Probe Technique for Dielectric Measurement of Biological Tissues: Challenges and Common Practices. *Diagnostics* (2018) 8:40. doi:10.3390/diagnostics8020040
22. Zheng Z, Huang L, Jiang H. Label-free Thermoacoustic Imaging of Human Blood Vessels *In Vivo*. *Appl Phys Lett* (2018) 113:253702. doi:10.1063/1.5054652
23. Huang L, Li T, Jiang H. Technical Note: Thermoacoustic Imaging of Hemorrhagic Stroke: a Feasibility Study with a Human Skull. *Med Phys* (2017) 44:1494–9. doi:10.1002/mp.12138
24. IEEE. IEEE Standard for Safety Levels with Respect to Human Exposure to Radio Frequency Electromagnetic fields 3 kHz to 300 GHz. *IEEE Stand* (1999) C95:1. doi:10.1109/IEEESTD.1999.89423
25. Hoelen CGA, de Mul FFM. Image Reconstruction for Photoacoustic Scanning of Tissue Structures. *Appl Opt* (2000) 39:5872–83. doi:10.1364/ao.39.005872
26. Bircan C, Barringer SA. Determination of Protein Denaturation of Muscle Foods Using the Dielectric Properties. *J Food Sci* (2002) 67:202–5. doi:10.1111/j.1365-2621.2002.tb11384.x
27. Nikfarjam M, Muralidharan V, Christophi C. Mechanisms of Focal Heat Destruction of Liver Tumors. *J Surg Res* (2005) 127:208–23. doi:10.1016/j.jss.2005.02.009
28. Brace CL, Laeseke PF, van der Weide DW, Lee FT. Microwave Ablation with a Triaxial Antenna: Results in *Ex Vivo* Bovine Liver. *IEEE Trans Microwave Theor Techn.* (2005) 53:215–20. doi:10.1109/TMTT.2004.839308
29. Artemov VG. A Unified Mechanism for Ice and Water Electrical Conductivity from Direct Current to Terahertz. *Phys Chem Chem Phys* (2019) 21:8067–72. doi:10.1039/c9cp00257j

30. Stauffer PR, Rossetto F, Prakash M, Neuman DG, Lee T. Phantom and Animal Tissues for Modelling the Electrical Properties of Human Liver. *Int J Hyperthermia* (2003) 19:89–101. doi:10.1080/0265673021000017064
31. Brace CL. Temperature-dependent Dielectric Properties of Liver Tissue Measured during thermal Ablation: toward an Improved Numerical Model. *Annu Int Conf IEEE Eng Med Biol Soc* (2008) 2008:230–3. doi:10.1109/IEMBS.2008.4649132
32. Lazebnik M, Converse MC, Booske JH, Hagness SC. Ultrawideband Temperature-dependent Dielectric Properties of Animal Liver Tissue in the Microwave Frequency Range. *Phys Med Biol* (2006) 51:1941–55. doi:10.1088/0031-9155/51/7/022
33. Chin L, Sherar M. Changes in Dielectric Properties Ofex Vivobovine Liver at 915 MHz during Heating. *Phys Med Biol* (2001) 46:197–211. doi:10.1088/0031-9155/46/1/314
34. Schildkopf P, J. Ott O, Frey B, Wadepohl M, Sauer R, Fietkau R, et al. Biological Rationales and Clinical Applications of Temperature Controlled Hyperthermia - Implications for Multimodal Cancer Treatments. *Cmc* (2010) 17:3045–57. doi:10.2174/092986710791959774
35. Oei AL, Vriend LEM, Krawczyk PM, Horsman MR, Franken NAP, Crezee J. Targeting Therapy-Resistant Cancer Stem Cells by Hyperthermia. *Int J Hyperthermia* (2017) 33:419–27. doi:10.1080/02656736.2017.1279757
36. Wolf M, Gulich R, Lunkenheimer P, Loidl A. Broadband Dielectric Spectroscopy on Human Blood. *Biochim Biophys Acta (Bba) - Gen Subjects* (2011) 1810:727–40. doi:10.1016/j.bbagen.2011.05.012

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.

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 22 August 2022
DOI 10.3389/fphys.2022.981463

OPEN ACCESS

EDITED BY
Lin Huang,
University of Electronic Science and Technology of China, China

REVIEWED BY
Sheng Ge,
Southeast University, China
Xiaomin Yang,
Sichuan University, China

*CORRESPONDENCE
Yu Pang,
pangyu@cqupt.edu.cn
Guo Zhang,
zhangguo@swmu.edu.cn

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

RECEIVED 29 June 2022
ACCEPTED 11 July 2022
PUBLISHED 22 August 2022

CITATION
Sun W, Chen J, Yan L, Lin J, Pang Y and Zhang G (2022), COVID-19 CT image segmentation method based on swin transformer.
Front. Physiol. 13:981463.
doi: 10.3389/fphys.2022.981463

COVID-19 CT image segmentation method based on swin transformer

Weiwei Sun[1], Jungang Chen[1], Li Yan[2], Jinzhao Lin[1], Yu Pang[1]* and Guo Zhang[1,2]*

[1]Chongqing University of Posts and Telecommunication, Chongqing, China, [2]School of Medical Information and Engineering, Southwest Medical University, Luzhou, China

Owing to its significant contagion and mutation, the new crown pneumonia epidemic has caused more than 520 million infections worldwide and has brought irreversible effects on the society. Computed tomography (CT) images can clearly demonstrate lung lesions of patients. This study used deep learning techniques to assist doctors in the screening and quantitative analysis of this disease. Consequently, this study will help to improve the diagnostic efficiency and reduce the risk of infection. In this study, we propose a new method to improve U-Net for lesion segmentation in the chest CT images of COVID-19 patients. 750 annotated chest CT images of 150 patients diagnosed with COVID-19 were selected to classify, identify, and segment the background area, lung area, ground glass opacity, and lung parenchyma. First, to address the problem of a loss of lesion detail during down sampling, we replaced part of the convolution operation with atrous convolution in the encoder structure of the segmentation network and employed convolutional block attention module (CBAM) to enhance the weighting of important feature information. Second, the Swin Transformer structure is introduced in the last layer of the encoder to reduce the number of parameters and improve network performance. We used the CC-CCII lesion segmentation dataset for training and validation of the model effectiveness. The results of ablation experiments demonstrate that this method achieved significant performance gain, in which the mean pixel accuracy is 87.62%, mean intersection over union is 80.6%, and dice similarity coefficient is 88.27%. Further, we verified that this model achieved superior performance in comparison to other models. Thus, the method proposed herein can better assist doctors in evaluating and analyzing the condition of COVID-19 patients.

KEYWORDS

COVID-19, CT image, deep learning, detection and recognition, lesion segmentation

1 Introduction

According to several studies (Zhu et al., 2019; Dong et al., 2020; Xu et al., 2020), computed tomography (CT) clearly displays the characteristic lung lesions of Covid-19 in patients. However, CT scans contain hundreds of slices, and CT images must be reconstructed and transmitted through an image archiving and communication system for doctors to interpret results and diagnose patients. Covid-19 and other types of pneumonia are generally identified by radiologists by simply processing images at communication system terminals, reading them, or projecting them through a lamp (Bai et al., 2020; Song et al., 2021; Bernheim et al., 2020; Rubin et al., 2020; Wong et al., 2020; Lee et al., 2001). Simultaneously, radiologists must be experienced to achieve sufficient detection results. Covid-19 has similar medical imaging characteristics to other types of pneumonia (Shi et al., 2020), and CT can be used to determine whether a patient is infected with viral pneumonia (Covid-19 is a viral pneumonia caused by the SARS-COV-2 virus) (Chen and Li, 2020). However, CT is unable to determine which virus causes viral pneumonia; the novel coronavirus or another virus, making it difficult to distinguish and diagnose the virus type. Considering these difficulties, quickly and accurately distinguishing between Covid-19 and other types of pneumonia is crucial to facilitating the screening process in clinical practice. Therefore, with the AI-assisted diagnosis of medical images, accurate and efficient recognition of Covid-19 lung CT images is of profound significance for controlling the epidemic (Ardila et al., 2019; Esteva et al., 2019; Esteva et al., 2017; Litjens et al., 2017; Mei et al., 2020; Qin et al., 2019; Topol 2019; Li et al., 2020; Jaiswal et al., 2020).

Locality is used by traditional convolutional networks to improve efficiency but at the cost of losing the connection in a global context. Convolutional architecture has an inherent induction bias, lacking an understanding of position dependence in images (Wang et al., 2020; Valanarasu et al., 2021). In a study by Dosovitskiy et al. (2020), the proposed vision transformer (ViT) was trained on large image datasets using location-embedded two-dimensional (2D) image patches as input sequences, thus achieving a performance comparable to that of convolutional networks. Based on the transformer architecture, a self-attention mechanism was utilized to encode the position dependence at a distance to learn efficient representations. However, most existing transformer-based network architectures require large datasets for training. Generalization may be inadequate if the training is performed using insufficient data. In a study by Hassani et al. (2021), a compact convolutional transformer (CCT) (Wang et al., 2021) was proposed to eliminate the misunderstanding of the requirement of a transformer for large amounts of data. It achieves comparable performance on small datasets; however, when the input dimension is large, the operational cost of the self-attention mechanism increases significantly. Global pooling does not use the location information in the process of extracting pneumonia symptoms, potentially causing loss of location information. For imaging tasks, it is important to obtain the spatial position structure of an image.

Therefore, we use a new method to solve the above problems in CT lesion segmentation of COVID-19. To solve the problem of detail loss, we add CBAM (Woo et al., 2018) and atrous convolution to the U-Net encoder part, and replace the partial convolution operation with the empty convolution operation. This can solve the problem of feature image detail loss caused by the decrease of resolution after repeated down-sampling operations. A Swin Transformer (Liu et al., 2021) is added to obtain local information in the CNN network, and the joint loss function is used for optimization during training. Thus, the segmentation of background regions, lungs, ground-glass opacities, and lung parenchyma in the chest CT images of patients is achieved. The results of ablation experiments demonstrate that this method achieved significant performance gain, in which the mean pixel accuracy is 87.62%, mean intersection over union is 80.6%, and dice similarity coefficient is 88.27%. The feasibility and effectiveness of this method are proved. Chest CT examination has a very important application prospect in clinical observation of treatment effect, monitoring of lesions and follow-up.

2 Materials and methods

Here, a new lesion segmentation method in chest CT images of COVID-19 patients is proposed, and the network structure is shown in Figure 1. The input is downsampled 4 times in total. The encoder performs a normal convolution and a dilated convolution operation before downsampling. The BN layer and the activation function layer are added to speed up the network convergence process. The CBAM mechanism is introduced in the downsampling process. After each downsampling iteration, the size of the feature vector is halved, and the number of channels is doubled. In the experiment, images with a height and width of 512 and three channels are used as input, that is,$512 \times 512 \times 3$ After being processed by the encoder part, a feature vector of size $32 \times 32 \times 512$ is output. Then, the downsampled feature images are flattened to fit the vector dimension of the Swin Transformer structure by linear embedding. The vector dimension does not change in the Transformer encoder structure, and a sequence vector of $1{,}024 \times 512$ dimensions is output. The sequence vector is restored to $32 \times 32 \times 512$ by the Reshape operation to fit the input dimension requirement of the segmentation network upsampling. Finally, the segmentation result whose height and width are consistent with the input is obtained after passing through the decoder for four upsampling iterations.

2.1 Convolution attention module

We use an attention mechanism in the network to perform weight adjustment on the feature vectors. This is similar to how

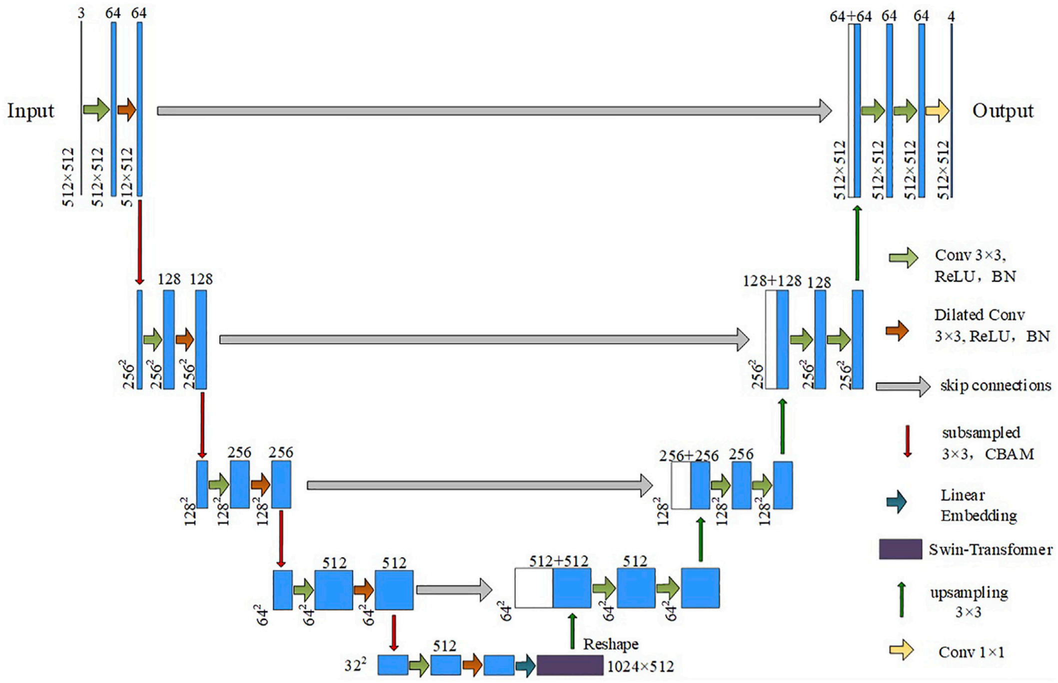


FIGURE 1
Segmentation network structure diagram.

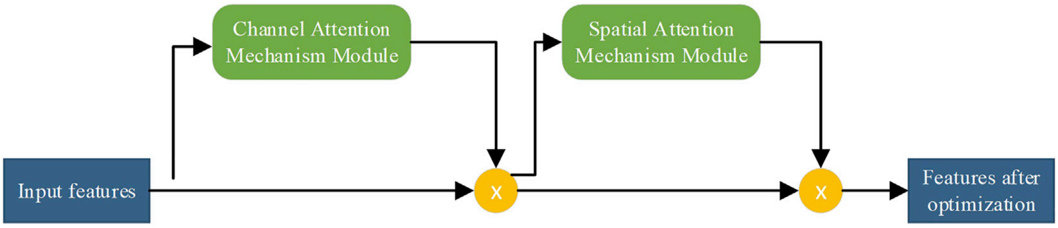


FIGURE 2
Flowchart of the attention module of the convolutional block.

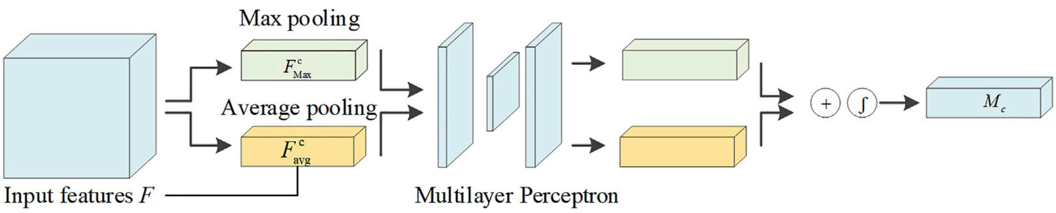


FIGURE 3
Flowchart of channel attention mechanism.

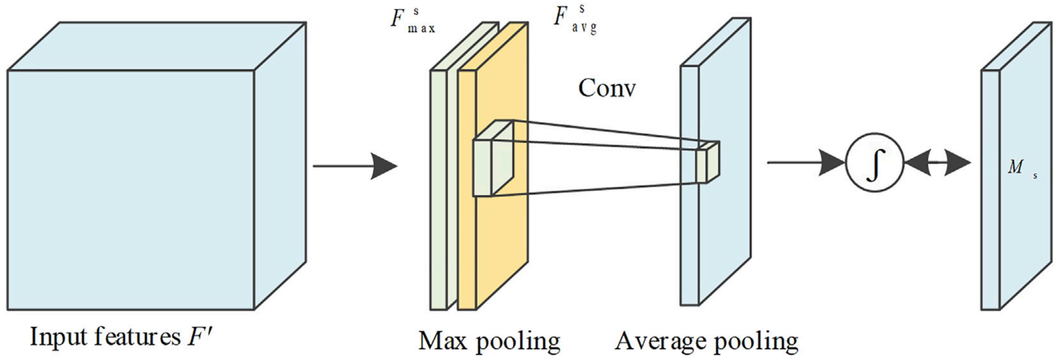


FIGURE 4
Flow chart of spatial attention mechanism.

the human brain focuses on important information. Important information is made more prominent, and other information is filtered. The convolutional attention module is composed of channel attention and spatial attention modules, which are used for the attention mechanism of the feature vector channel and space, respectively. The process is shown in Figure 2. Finally, the attention weights are multiplied by the input feature image to obtain the output feature image.

2.1.1 Channel attention mechanism

Channel attention assigns weights to each channel of the feature image. Valid channel weights are increased, and invalid channel weights are suppressed. The flow of the channel attention mechanism is shown in Figure 3. The input feature $F \in \mathbb{R}^{H\times W\times C}$ is average-pooled to generate the vector $F^{\mathrm{C}}_{\mathrm{avg}} \in \mathbb{R}^{1\times 1\times C}$, where C represents the channel. The vector $F^{\mathrm{C}}_{\mathrm{max}} \in \mathbb{R}^{1\times 1\times C}$ is generated by a max-pooling operation. Average pooling has the advantage of optimizing the spatial information of feature images. Max pooling can extract landmark information in feature images. The two output features are fed into a shared multilayer perceptron, and features with contextual descriptions are generated. Finally, the ReLU activation function is used to output the feature image channel weights. Feature images are summed and merged elementwise. The feature vector $M_{\mathrm{C}} \in \mathbb{R}^{\mathrm{C}\times 1\times 1}$ is output through the sigmoid activation function.

According to the above process, the calculation formula is as follows:

$$\begin{aligned} M_c(F) &= \sigma\left(\mathrm{MLP}\left(\mathrm{AvgPool}(F)\right) + \mathrm{MLP}\left(\mathrm{MaxPool}(F)\right)\right) \\ &= \sigma\left(W_1\left(W_0\left(F^c_{\mathrm{avg}}\right)\right) + W_1\left(W_0\left(F^c_{\mathrm{max}}\right)\right)\right) \end{aligned} \tag{1}$$

where σ represents the Sigmoid activation function, AvgPool represents the average pooling operation, MaxPool represents the maximum pooling operation, MLP represents the shared multi-layer perceptron, and W_0 and $W_1 \in \mathbb{R}^{C/r\times C}$ are the weights of the shared multi-layer perceptron.

2.1.2 Spatial attention mechanism

The spatial attention mechanism can measure some regions of the feature image to obtain higher responses, and the mechanism flow is shown in Figure 4. Suppose the feature vector optimized by the channel attention module is $F' \in \mathbb{R}^{H\times W\times C}$. F' generates the two-dimensional vector $F^{\mathrm{S}}_{\mathrm{avg}} \in \mathbb{R}^{H\times W\times 1}$ by the max pooling operation, and $F^{\mathrm{S}}_{\mathrm{max}} \in \mathbb{R}^{H\times W\times 1}$ is generated through average pooling, where S represents a channel. The two-dimensional vector information obtained by the pooling operation is concatenated. The feature information is fused through the convolution operation, and a two-dimensional spatial attention image is generated through the sigmoid activation function. Finally, the output of the spatial attention module is dot multiplied with the feature image at the pixel level to obtain the weighted feature image.

The equation of the above process is as follows:

$$\begin{aligned} M_s(F) &= \sigma\left(f^{7\times 7}\left(\left[\mathrm{MaxPool}(F'), \mathrm{AvgPool}(F')\right]\right)\right) \\ &= \sigma\left(f^{7\times 7}\left(\left[F^s_{\mathrm{max}}, F^s_{\mathrm{avg}}\right]\right)\right) \end{aligned} \tag{2}$$

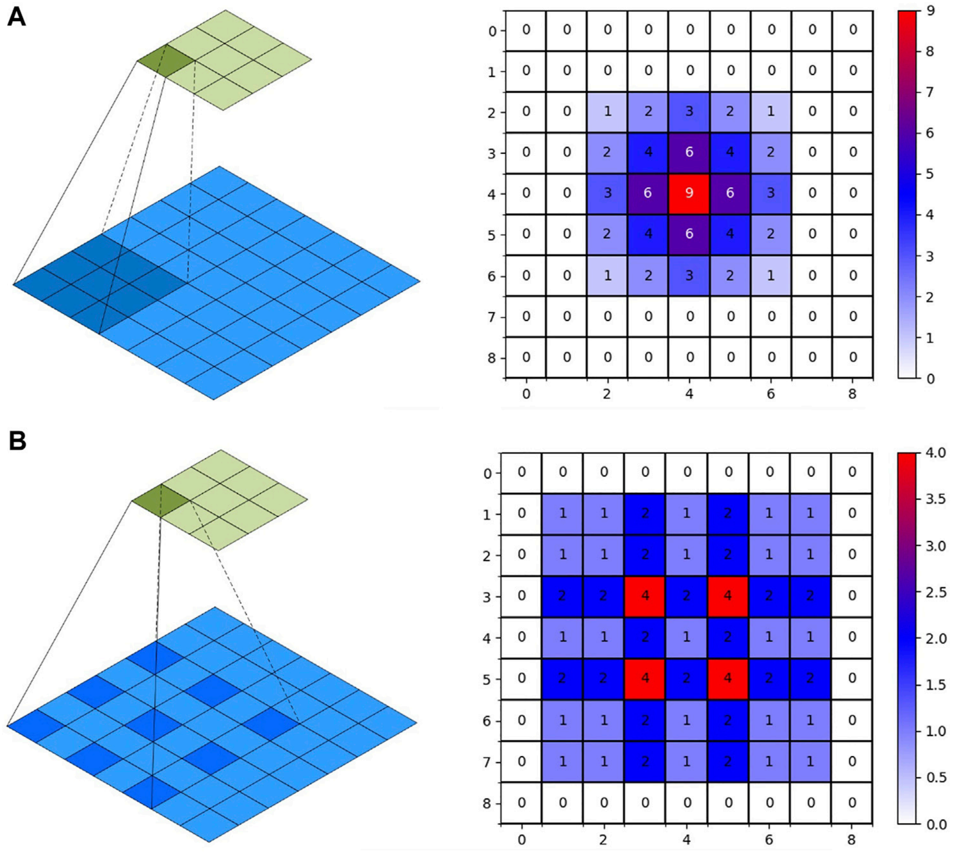
where σ is the sigmoid activation function, and $f^{7\times 7}$ indicates that the feature vector in parentheses is convolved with a convolution kernel of size 7 × 7.

F represents the feature image, the output F' is optimized by the channel attention module, and the output F'' is optimized by the spatial attention module. Therefore, feature F is optimized by the CBAM module:

$$F' = M_c(F) \otimes F \tag{3}$$

$$F'' = M_s(F') \otimes F' \tag{4}$$

where ⊗ represents that the elementwise multiplication.



A

0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
2	0	0	1	2	3	2	1	0	0
3	0	0	2	4	6	4	2	0	0
4	0	0	3	6	9	6	3	0	0
5	0	0	2	4	6	4	2	0	0
6	0	0	1	2	3	2	1	0	0
7	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	0

B

0	0	0	0	0	0	0	0	0	0
1	0	1	1	2	1	2	1	1	0
2	0	1	1	2	1	2	1	1	0
3	0	2	2	4	2	4	2	2	0
4	0	1	1	2	1	2	1	1	0
5	0	2	2	4	2	4	2	2	0
6	0	1	1	2	1	2	1	1	0
7	0	1	1	2	1	2	1	1	0
8	0	0	0	0	0	0	0	0	0

FIGURE 5
Convolution operation. **(A)**: 2D-Convolution. **(B)**: Dilated Convolution.

2.2 Atrous convolution

Feature information is extracted using U-Net model convolution operations. Due to device performance limitations, multiple pooling operations reduce the resolution of feature vectors. When using the convolution operation to extract higher-level features, the next convolution operation can obtain a larger receptive field. However, as the feature size decreases, feature information will be lost. The restoration detail information cannot be restored, while upsampling restores the size. Replacing ordinary convolution operations with atrous convolution can achieve a larger receptive field range within a limited convolution kernel. Therefore, the loss of detail information caused by the downsampling process can be solved. The ordinary convolution and atrous convolution methods and the obtained receptive fields are shown in Figure 5. In the right image of Figures 5A,B, the feature maps of 9 × 9 use a convolution kernel of size 3 × 3 and stride 1 for convolution operation. In the right picture of Figure 5A, the receptive field is obtained after two ordinary convolution iterations; the range is 5 × 5. In the right picture of Figure 5B, the receptive field is obtained after one ordinary convolution and one dilated convolution with a dilation factor of 2; the range is 7 × 7. It shows that a larger receptive field range is obtained after using atrous convolution. The numbers in the figure represent the number of times the pixels are convolved.

When using continuous atrous convolution, the dilation factor cannot be a common divisor greater than 1. And the expansion factor must satisfy the following formula:

$$M_i = \max[M_{i+1} - 2r_i, M_{i+1} - 2(M_{i+1} - r_i), r_i] \tag{5}$$

where M_i represents the maximum expansion factor of the i-th layer, and r_i is the expansion factor that represents the distance between adjacent elements in the hollow convolution kernel, which should be less than or equal to the size of the convolution kernel. In the atrous convolution operation, the convolution kernel size is fixed. When the dilation rate increases, the spacing of adjacent elements in the convolution kernel increases. It is also possible to keep the height and width of the original input feature map unchanged.

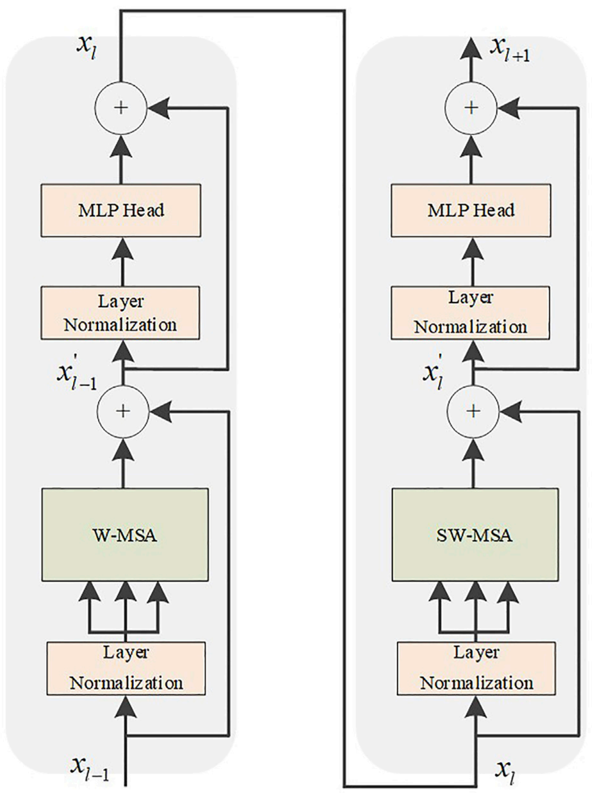


FIGURE 6
Swin Transformer module.

2.3 Swin transformer module

After improving the convolutional structure network of U-Net to extract feature information, we use the Swin Transformer to extract the global information from the feature information. We combine the CNN with the Transformer structural model. The insufficiency of context dependencies in the acquisition of low-level features by convolutional networks will be compensated. Compared with ViT, we improve the Transformer encoder by introducing windows multi-head self-attention (W-MSA) and shifted windows multi-head self-attention (SW-MSA) (Hatamizadeh et al., 2022). Assuming the input is x_{l-1} , the formula is as follows:

$$x'_{l-1} = x_{l-1} + \mathrm{W-MSA}\left(\mathrm{LN}\left(x_{l-1}\right)\right), \tag{6}$$

$$x_l = x'_{l-1} + \mathrm{MLP}\left(\mathrm{LN}\left(x'_{l-1}\right)\right), \tag{7}$$

$$x'_l = x_l + \mathrm{SW-MSA}\left(\mathrm{LN}\left(x_l\right)\right), \tag{8}$$

$$x_{l+1} = x'_l + \mathrm{MLP}\left(\mathrm{LN}\left(x'_l\right)\right). \tag{9}$$

where $l \in \{1, 2, \cdots, L\}$

According to the Swin Transformer formula, it can be concluded that the structure consists of two Transformer encoder modules. After the input is normalized by the layer, the attention value is calculated using W-MSA, and the residual structure is formed with the original input. After layer normalization and MLP operation, the encoder module with the SW-MSA calculation method is used to output the feature vector. The Swin Transformer structure is shown in Figure 6. Compared with the ViT, an encoder module is added, and the redesigned W-MSA and SW-MSA calculation methods greatly reduce the computational complexity.

In the W-MSA operation of the Swin Transformer, the feature map is divided into windows of the same size, which is equivalent to reducing the size of the patch. Thereby the computational complexity is reduced. We utilize the same self-attention mechanism as ViT inside each individual window. However, after dividing the feature map into separate feature windows, the attention mechanism values of the feature windows are calculated separately, and there is no information interaction between them. As a result, the self-attention mechanism cannot obtain global information. Therefore, the SW-MSA operation is increased, and the window operation is shifted. This solves the defect that information cannot be exchanged between W-MSA operation windows. The operation flow of common MSA, W-MSA, and SW-MSA is shown in Figure 7. The W-MSA window size is 4. In the SW-MSA operation, the feature window is divided into three different patch sizes, which are 2 × 2, 2 × 4, and 4 × 4 sizes. After combining four 2 × 2-sized windows and combining four 2 × 4-sized windows, two feature windows with patch size 2 × 2 are obtained. Then, the attention value is obtained by continuing the calculation of W-MSA. Finally, the original window dimensions are restored. As such, not only is the computational complexity reduced, but the interactive information between the windows can be obtained.

2.4 Optimization of loss function

In medical image segmentation, common loss functions include cross entropy loss (CE loss) and dice coefficient loss (Dice loss). The chest CT image segmentation method we proposed includes four categories: background region, lung region, ground glass opacity, and lung parenchyma. Figure 8 shows the chest CT images and the pixel distribution maps of different categories in the corresponding segmentation gold standard. The abscissa is the segmentation type, and the ordinate is the number of pixels. It can be seen that the proportion of ground glass and lung parenchyma is much smaller than the background and lung areas. This is common in mild and moderate patients, and there may even be no focal manifestations. Therefore, uneven data distribution will be caused in the experiment, which makes network training more difficult.

The cross-entropy loss function compares the pixel-predicted value output by the training model with the real

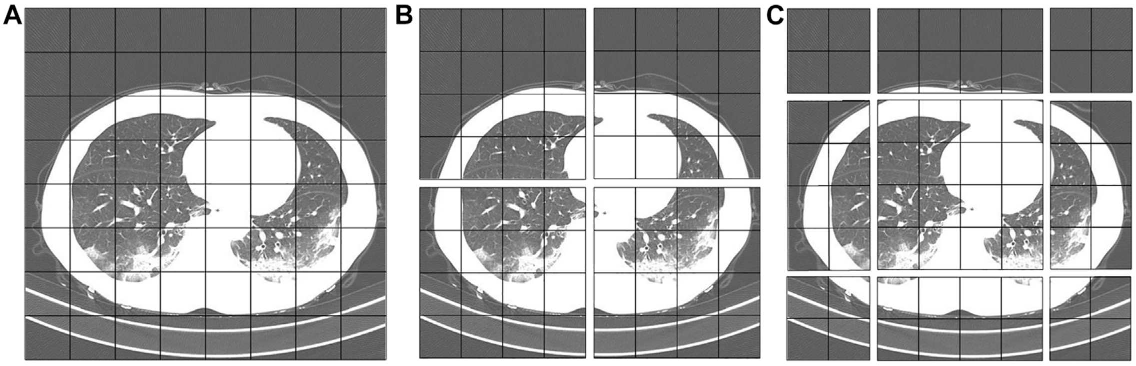


FIGURE 7
Operation diagram of self-attention mechanism. **(A)**: MSA. **(B)**: W-MSA. **(C)**: SW-MSA.

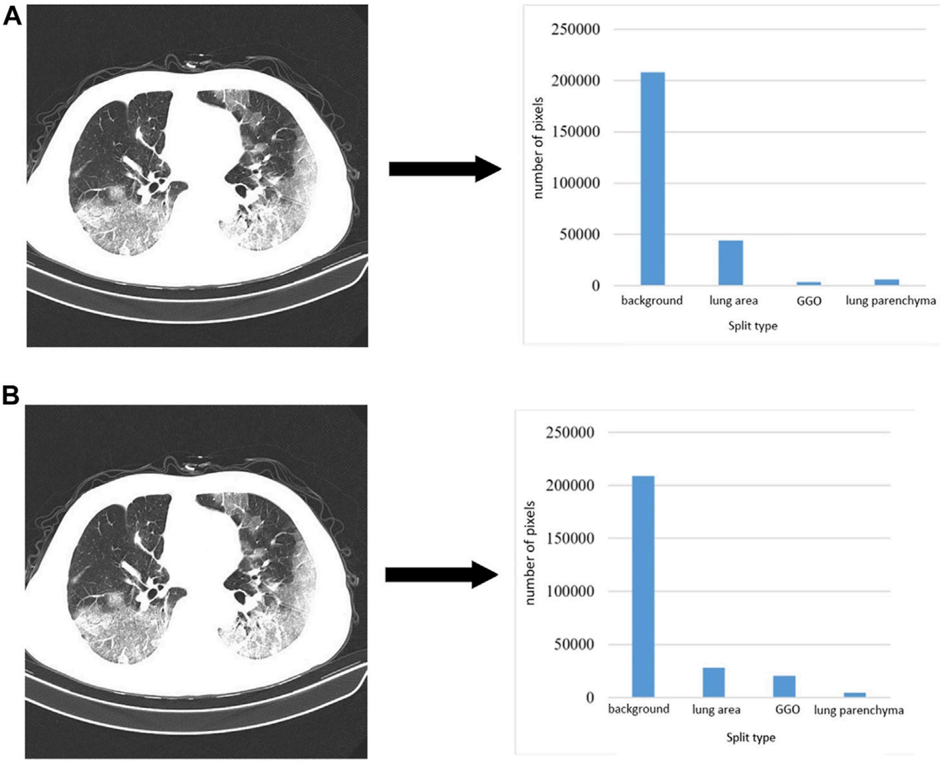


FIGURE 8
CT images and corresponding pixel category distributions.

value. In the case of training without overfitting, the smaller the loss value, the better the result. The formula is as follows:

$$CE\,loss = -y^{*}\log_2\left(y'\right) \tag{10}$$

where y is the real label paper, y' is the predicted value, and the loss function has the same prediction weight for each category. As shown in Figure 8, the background area accounts for a large proportion, and the factors leading to the final result will be biased towards the background area. After training, the performance value of the loss function is small, but it cannot reflect the classification effect of other categories through the loss value.

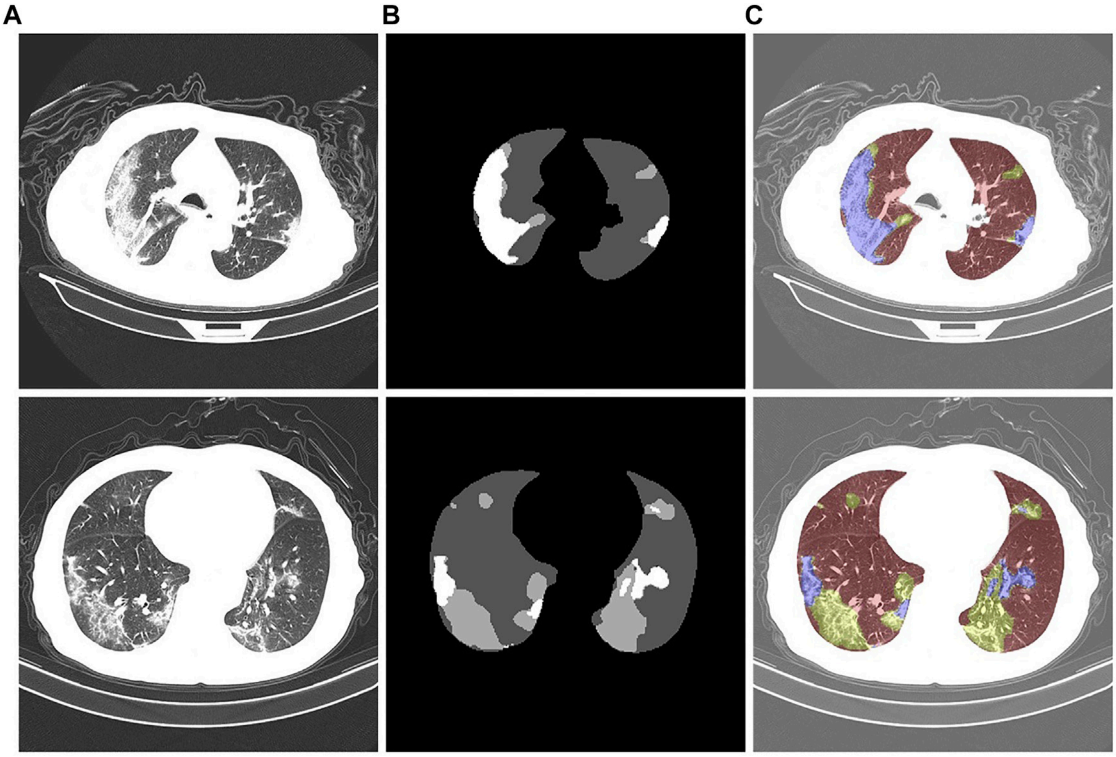


FIGURE 9
Chest CT images of COVID-19 patients. **(A)**: Initial CT image. **(B)**: Gold standard. **(C)**: Color-annotated segmentation results.

In dice loss, dice represents the dice similarly coefficient (DSC), which indicates the degree of similarity between two sample areas; the value is between 0 and 1; the larger the value, the higher the similarity. Assuming that A and B represent sets of regions, the DSC formula is as follows:

$$Dice = \frac{2|A \cap B|}{|A| + |B|} \tag{11}$$

where ∩ represents the intersection of sets, and the dice loss formula can be obtained according to the DSC. The formula is as follows:

$$Dice\,loss = 1 - \frac{2\left|Y \cap \bar{Y}\right| + 1}{|Y| + \left|\bar{Y}\right| + 1} \tag{12}$$

where Y is the real segmentation area, and $\bar{Y}$ is the model prediction area. We add 1 to the denominator and numerator to prevent the denominator from being zero and to reduce the possibility of overfitting during the training process. Compared with the CE loss function, dice loss is not affected by the background when the number of pixel categories is unevenly distributed. However, training instability occurs when the prediction is incorrect. Therefore, we combine the CE loss function and dice loss as a joint loss function and use the CE loss function to guide dice loss for training. The formula is as follows:

$$loss = 0.5 \times CE + Dice\,loss \tag{13}$$

$$loss = 0.5 \times Y \cdot \log_2\left(\bar{Y}\right) + 1 - \frac{2\left|Y \cap \bar{Y}\right| + 1}{|Y| + \left|\bar{Y}\right| + 1} \tag{14}$$

2.5 Datasets

We utilized a dataset from the China Consortium for Chest CT Imaging Research (CC-CCII) (Ai et al., 2020). The CC-CCII dataset contains 617,775 CT images from 6,752 CT scans of 4,154 patients. The study sample size was estimated by standard AI training and validation methods. Patients were randomly assigned to a training set (60%), an internal validation set (20%) or a test set (20%). We chose to use 750 annotated chest CT images selected from 150 COVID-19 patients by five radiologists with 15 years of experience. These images include background areas, lung areas, ground-glass opacities, and lung parenchyma. Mild patients mainly present with ground-glass opacity, which is distributed in the lower lobes of both lungs and adheres closely to the pleura. Ground-glass shadows are characterized by spreading

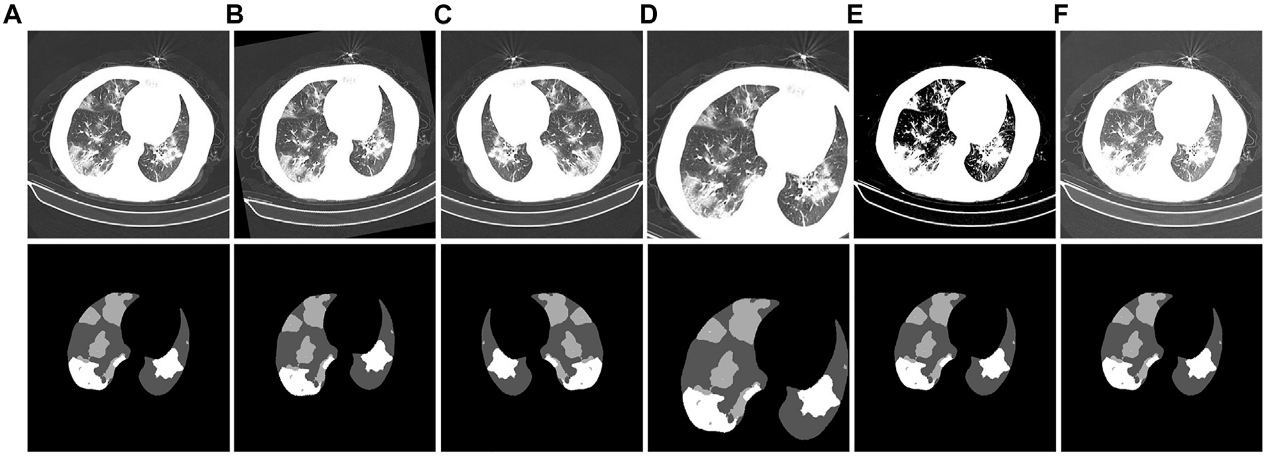


FIGURE 10
Example of CT image data-enhancement results. **(A)**: the original CT image and its corresponding gold standard. **(B)**: the CT image obtained after rotation and its corresponding gold standard. **(C)**: the CT image obtained after horizontal flipping and its corresponding gold standard. **(D)**: the CT image obtained after random cropping and its corresponding gold standard. **(E)**: the CT image obtained after adjusting the contrast and its corresponding gold standard. **(F)**: the CT image obtained after adjusting the brightness and its corresponding gold standard.

toward the center and blurring at the edges. In moderate patients, the number of lesions proliferated, and the lesions were markedly plaque-like. The patient is accompanied by a condition of cough and fatigue. In severe patients, the density of lung tissue increases and the lung parenchyma changes. The patient presented with fever and headache. An example of the segmentation of a chest CT image of a COVID-19 patient is shown in Figure 9. Figure 9A is the initial image, Figure 9B is the gold standard of the dataset, and Figure 9C is after the gold standard mask and the initial CT image are superimposed; the highlighted color is used to distinguish the segmentation results. The gray area is the background area of the patient, the red area is the lung area, the yellow area is the ground glass opacity, and the blue area is the lung parenchyma.

2.5.1 Data augmentation

CT images have different properties, such as brightness, saturation, and angle. Therefore, a data augmentation method is added in the preprocessing stage of experimental training to prevent overfitting of the training results. In this way, the model performance is increased, and the data augmentation is shown in Figure 10. Figure 10A is the initial image. Figures 10B–F are the corresponding labels of the original image after rotating, horizontally flipping, randomly cropping, adjusting saturation, and adjusting brightness, respectively.

2.5.2 Training parameters

The training set of the CC-CCII dataset is divided into ten groups, each time nine groups of images are used as the training set and one group is used as the validation set. They were used in ten-fold cross-validation experiments. After training and validation separately, we use the test set to test, repeat this process ten times, and finally take the average of the ten results as the evaluation of algorithm accuracy. All CT image pixels are resized to 512 × 512 pixels before being input into the training model. In the model training, the network adopts the mean square error loss function; The initial learning rate of the Adam optimizer is 0.0001; The batch size is set to 64; And the fully connected layer uses a dropout layer with probability 0.5. This deep learning method does not require much analysis of the threshold and gray value of CT images. Data augmentation is achieved by adjusting contrast, affine transformation, and color dithering to achieve better performance of the model. The details of the experimental training parameters are listed in Table 1.

TABLE 1 Training parameter settings.

Type	Setting
Batch size	64
Learning rate	0.0001
Optimizer	Adam
Iterations (Epoch)	100
Ubuntu 18.04	PyToch1.6.0

2.5.3 Evaluation indicators

To analyze the segmentation performance of the trained model, we used three common performance metrics: mean intersection over union (mIoU) (Rezatofighi et al., 2019), DSC (Huang et al., 2022), and mean pixel accuracy (mPA) (Paintdakhi et al., 2016). mIoU is the average of the ratios of the intersection and union of the results predicted by the

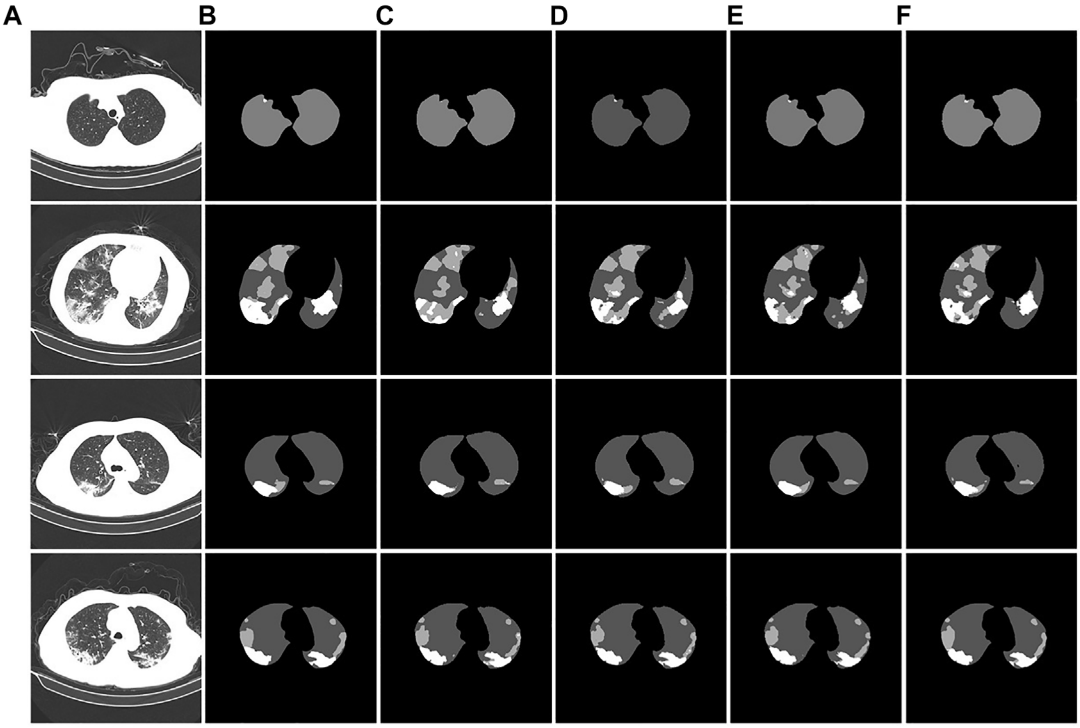


FIGURE 11
Example of ablation experiment comparison. **(A)**: the CT image of the COVID-19 patient. **(B)**: the gold standard of the CT image. **(C)**: the U-Net segmentation result. **(D)**: the U-Net segmentation result after introducing atrous convolution. **(E)**: the U-Net segmentation result after introducing atrous convolution and CBAM. **(F)**: the U-Net segmentation result after introducing the atrous convolution, CBAM, and Swin Transformer modules.

TABLE 2 Comparison of ablation experiments.

Modle	mPA/%	mIoU/%	DSC/%
U-Net	85.86	78.59	86.74
U-Net + Atrous convolution	86.22	79.11	87.15
U-Net + Atrous convolution + CBAM	87.41	80.41	87.49
U-Net + Atrous convolution + CBAM + Swin Transformer	87.62	80.64	88.27

model for each category and the true label, DSC is the similarity measure function, which can calculate the similarity between the true label and the predicted label, and mPA is represents the pixel accuracy of each category. The pixel accuracy is summed and averaged.

$$mIOU = \frac{1}{k+1}\sum_{i=0}^{k}\frac{TP}{TP+FN+FP} \tag{15}$$

$$DSC = \frac{1}{k+1}\sum_{i=0}^{k}\frac{2TP}{FP+2TP+FN} \tag{16}$$

$$mPA = \frac{1}{k+1}\sum_{i=0}^{k}\frac{TP}{TP+FN} \tag{17}$$

where k is the number of classes, TP is the number of pixels that are correctly predicted as positive examples, FN is the number of pixels that are incorrectly predicted as negative examples, and FP is the number of pixels that are incorrectly predicted as positive examples.

3 Results and discussion

3.1 Ablation experiment

To verify the segmentation effect of the improved U-Net model, we conducted ablation experiments. The segmentation

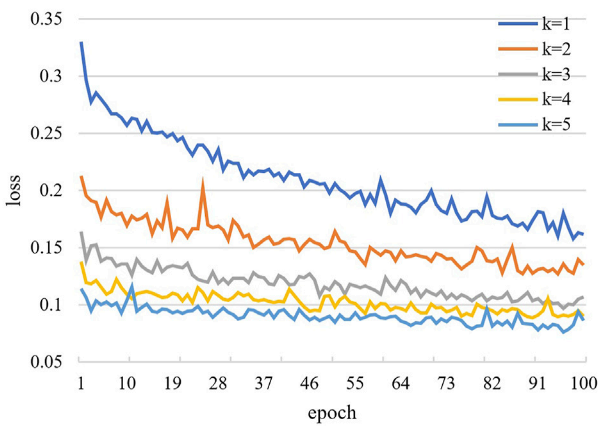


FIGURE 12
Training convergence loss curve.

TABLE 3 Performance comparison of different models.

Model	mPA/%	mIoU/%	DSC/%	FLOPs (G)
ResU-Net	87.05	80.1	87.54	1.46
Attention U-Net	86.26	78.31	86.47	1.95
TransU-Net	86.99	79.33	87.31	1.39
Ours	87.62	80.6	88.27	1.44

TABLE 4 Subjective evaluation scoring method.

Score	Features of the restored image
0	Severely distorted image
1	Image with severe distortion in some areas
2	Slightly distorted image
3	Difficult to spot distorted images
4	Images with better visual effects
5	Very sharp images

test results are shown in Figure 11. From the segmentation results of the CT image example, it can be observed that the original U-Net did not segment the tiny lesion details. The other improved models identified the lesions, but the segmentation effects were different. The U-Net segmentation result after adding atrous convolution is shown in Figure 11D. After adding CBAM, the effect is improved, as shown in Figure 11E. The model segmentation results after introducing the atrous convolution, CBAM, and Swin Transformer modules are significantly improved, as shown in Figure 11F. The segmentation performance of our proposed method achieved the best performance; especially in the case of a large number of lesion areas, the segmentation results of lesion and lung areas by this method are closer to the corresponding gold standard.

The experimental segmentation performance indicators are listed in Table 2. On the basis of U-Net, the atrous convolution mPA is added, and the mIoU and DSC indicators are increased by 0.36%, 0.52%, and 0.41%, respectively. After adding atrous convolution and CBAM, the corresponding indicators greatly improved. mPA, mIoU, and DSC metrics improved by 1.55%, 1.82%, and 0.75%, respectively. After adding atrous convolution, the corresponding indicators of CBAM and the Swin Transformer improved the most. The mPA, mIoU, and DSC metrics improved by 1.76%, 2.05%, and 1.53%, respectively. The corresponding metrics demonstrate the effectiveness and feasibility of our method.

The convergence effect of the training loss function of the new model is shown in Figure 12. The curves in the figure represent the training loss curves from the 1st to 5th fold, respectively. After the training method of cross-validation is used, we find that the training loss value of each epoch in fold 1 is the largest and the training loss value of each epoch in fold five is the smallest. The training loss value of each epoch in the next fold is smaller than that of the previous fold. The results show that the convergence effect of the new model is significantly improved.

3.2 Models comparison

We demonstrate the feasibility and effectiveness of the proposed improved method through ablation experiments. To further verify the segmentation ability of the model, we compared it with other models. The results of the segmentation experiment are shown in Figure 13. First, the ResU-Net model (Jha et al., 2019) adds a residual structure to the convolution operations of the encoder and decoder to improve model performance. In CT

TABLE 5 Subjective quality evaluation of different algorithms.

Method	Sharpness	Resolution	Invariance	Acceptability
ResU-Net	3.3 ± 0.21	3.5 ± 0.25	0.5 ± 0.39	3.8 ± 0.21
Attention U-Net	3.6 ± 0.24	3.9 ± 0.49	0.6 ± 0.16	3.9 ± 0.41
TransU-Net	3.7 ± 0.16	4.1 ± 0.21	0.6 ± 0.25	4.2 ± 0.24
Our method	3.9 ± 0.24	4.3 ± 0.07	0.7 ± 0.36	4.2 ± 0.81

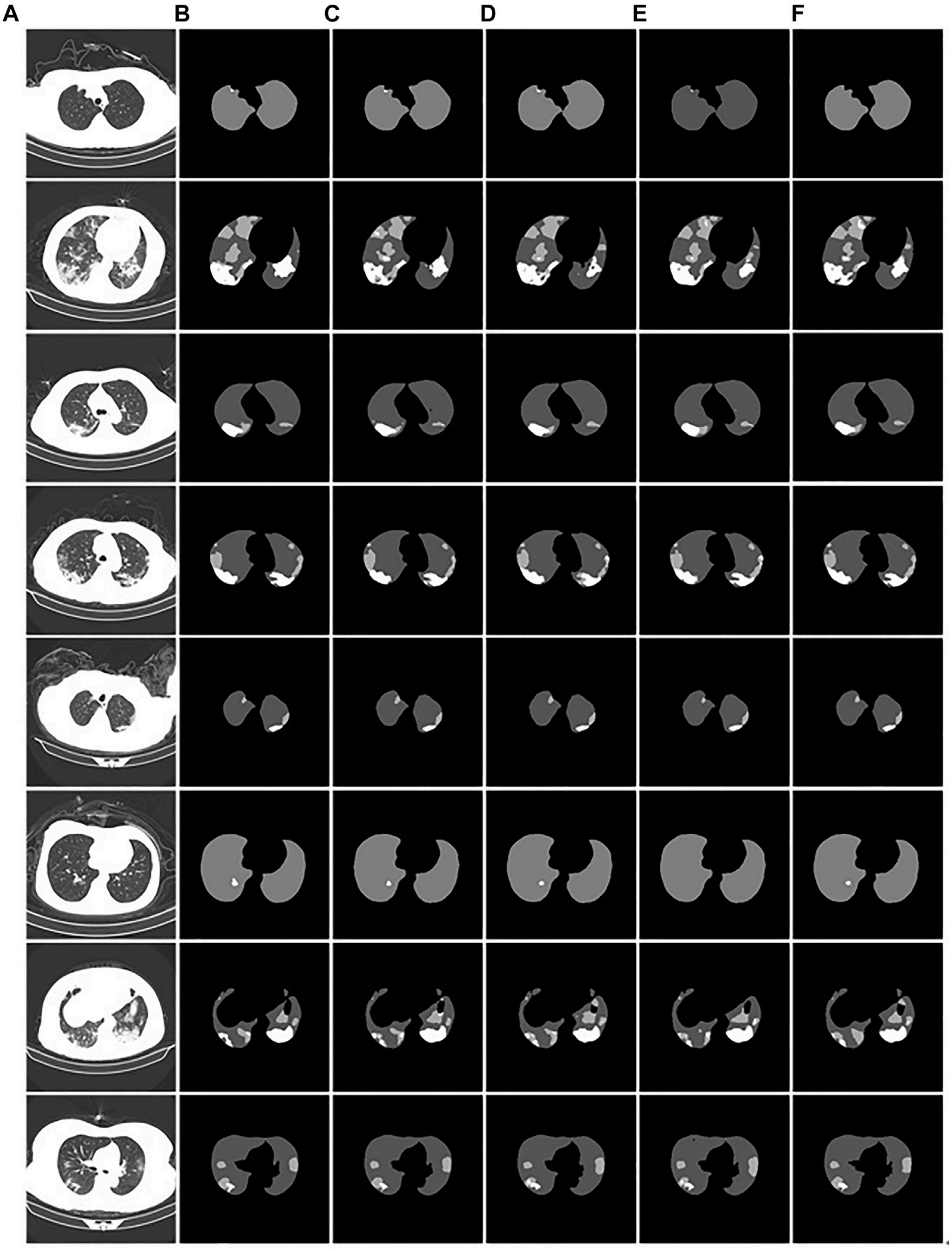


FIGURE 13
Example of comparison of different models. **(A)**: the CT images of COVID-19 patients. **(B)**: the gold standard of CT images. **(C)**: the segmentation results of our model. **(D)**: the ResU-Net segmentation results. **(E)**: the Attention U-net segmentation results. **(F)**: the TransU-net segmentation result.

images of mild patients, our method is compared with the ResU-net method, as shown in Figures 13C,D. In the figure, the performance of the two methods is comparable when segmenting smaller lesions. However, when the proportion of the lesion area is relatively large, the segmentation results show obvious voids, as shown in the second picture in Figure 13D. Second, Attention U-net (Oktay et al., 2018) introduces a soft attention mechanism, which is implemented by supervising the upper-level features through the next-level features. Our method is compared with the attention U-net method, as shown in Figures 13C,E. From the segmentation results, it can be seen that our method performs significantly better than the attention U-net method in terms of lesion segmentation accuracy in smaller regions. Further, leaky segmentation is present in the sixth picture of Figure 13E. Finally, TansU-Net (Chen et al., 2021) applies the Transformer encoder to image segmentation. Our method is compared with the TansU-Net method, as shown in Figures 13C,F. In the segmentation example, the TansU-Net method also appears similar to Attention U-net, failing to successfully identify smaller lesion areas. We used the Swin Transformer encoder structure before the segmentation network decoder. Although CBAM and hole convolution are added, the FLOPs are not much different, and the comprehensive segmentation ability is significantly better than TransU-Net. The effectiveness of our method is further demonstrated, and some complexity is reduced from the Transformer structure.

The comparison performance indicators of the above models are listed in Table 3. The performance metrics of the Attention U-Net method were the worst. The ResU-Net model outperformed TransU-Net in segmentation performance in the used test dataset. Compared with ResU-Net, our proposed segmentation method has improved performance indicators. mPA, mIOU, and DSC were improved by 0.57%, 0.5%, and 0.73%, respectively. Therefore, our proposed method performed the best among the compared models.

3.3 Subjective evaluation

For more specialized medical evaluation of segmentation models, clinical validation is required. We invited 10 chief physicians with more than 5 years of clinical experience in radiology to independently perform image analysis (sharpness, resolution, invariance, and acceptability). The scoring criteria for subjective evaluation are shown in Table 4. Ten groups of test samples were randomly constructed, and each group consisted of ten CT images of the lesion area. The subjective quality evaluation results of different algorithms utilized by radiologists are listed in Table 5.

As shown in Table 5, our proposed Atrous Convolution + CBAM + Swin Transformer model achieves the best subjective quality evaluations in terms of sharpness, resolution, invariance, and acceptability. The main reason is the benefit from introducing W-MSA and the exchange of information. Compared with other segmentation methods, our W-MSA fuses the mutual information and the multimodal features of CT images and has strong representation. The consistency of pathological information between segmented CT image and original CT image was guaranteed. This method achieves the best segmentation effect in terms of ground-glass opacity and visible plaque and lung parenchyma lesions.

4 Conclusion

Currently, a key approach to prevent the spread of the epidemic is to combine the chest CT images of patients for diagnosis. Therefore, this paper proposed an improved U-Net network for lesion segmentation in chest CT images of COVID-19. Atrous convolution was used as the convolution operation of each layer of the segmentation network encoder structure, and CBAM was introduced in the downsampling process to solve the problem of loss of lesion detail during the downsampling process. The Swin Transformer module was added to the encoder using the transformer structure to obtain global feature information. The primary improvement of the segmentation model framework is in the encoder part, which improved the model feature extraction performance. The results of the ablation experiments showed that the mPA, mIOU, and DSC reached 87.62, 80.6 and 88.27, respectively. In the subjective evaluation of radiologists, our method can effectively segment ground-glass opacity, visible plaque and lung parenchyma lesions, and maintain consistency with the original CT image pathological information. In future research, we will continue to refine the model. We aim to improve the screening process and the quantitative analysis of the disease and enhancing the efficiency of diagnosis and reducing infection.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: http://ncov-ai.big.ac.cn/download

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements. Written informed consent was not obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

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

Funding

This work was supported by the Doctoral Innovative Talents Project of Chongqing University of Posts and Telecommunications (BYJS202107). Chongqing Natural Science Foundation of China (grant number cstc2021jcyj-bsh0218); The National Natural Science Foundation of China (Grant No. U21A20447 and 61971079); The Basic Research and Frontier Exploration Project of Chongqing (Grant No. cstc2019jcyjmsxmX0666); Chongqing technological innovation and application development project (cstc2021jscx-gksbx0051); The Innovative Group Project of the National Natural Science Foundation of Chongqing (Grant No. cstc2020jcyj-cxttX0002), and the Regional Creative Cooperation Program of Sichuan (2020YFQ0025); The Science and Technology Research Program of Chongqing Municipal Education Commission (KJZD-k202000604).

Acknowledgments

We thank the School of Optoelectronic Engineering of Chongqing University of Posts and Telecommunications for their assistance in the research.

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

Ai, T., Yang, Z., Hou, H., Zhan, C., Chen, C., Lv, W., et al. (2020). Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (Covid-19) in China: A report of 1014 cases. *Radiology* 296, E32–E40. doi:10.1148/radiol.2020200642

Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., et al. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat. Med.* 25, 954–961. doi:10.1038/s41591-019-0447-x

Bai, H. X., Hsieh, B., Xiong, Z., Halsey, K., Choi, J. W., Tran, T. M. L., et al. (2020). Performance of radiologists in differentiating Covid-19 from non-Covid-19 viral pneumonia at chest CT. *Radiology* 296 (2), E46–E54. doi:10.1148/radiol.2020200823

Bernheim, A., Mei, X., Huang, M., Yang, Y., Fayad, Z. A., Zhang, N., et al. (2020). Chest CT findings in coronavirus disease-19 (Covid-19): Relationship to duration of infection. *Radiology* 295 (3), 200463. doi:10.1148/radiol.2020200463

Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., et al. (2021). Transunet: Transformers make strong encoders for medical image segmentation. *arXiv: 2102.04306*. doi:10.48550/arXiv.2102.04306

Chen, Y., and Li, L. (2020). SARS-CoV-2: Virus dynamics and host response. *Lancet. Infect. Dis.* 20 (5), 515–516. doi:10.1016/S1473-3099(20)30235-8

Dong, E., Du, H., and Gardner, L. M. (2020). An interactive web-based dashboard to track Covid-19 in real time. *Lancet. Infect. Dis.* 20 (5), 533–534. doi:10.1016/S1473-3099(20)30120-1

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. 2010:11929*. doi:10.48550/arXiv.2010.11929

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542 (7639), 115–118. doi:10.1038/nature21056

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., et al. (2019). A guide to deep learning in healthcare. *Nat. Med.* 25 (1), 24–29. doi:10.1038/s41591-018-0316-z

Hassani, A., Walton, S., Shah, N., Abuduweili, A., Li, J., and Shi, H. (2021). Escaping the big data paradigm with compact transformers. *arXiv. 2104:05704*. doi:10.48550/arXiv.2104.05704

Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H., and Xu, D. (2022). Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. *arXiv. 2201:01266*.

Huang, K., Xu, L., Zhu, Y., and Meng, P. (2022). AU-snake based deep learning network for right ventricle segmentation. *Med. Phys.* 49 (6), 3900–3913. doi:10.1002/mp.15613

Jaiswal, A., Gianchandani, N., Singh, D., Kumar, V., and Kaur, M. (2020). Classification of the Covid-19 infected patients using DenseNet201 based deep transfer learning. *J. Biomol. Struct. Dyn.* 39 (15), 5682–5689. doi:10.1080/07391102.2020.1788642

Jha, D., Smedsrud, P. H., Riegler, M. A., Johansen, D., Lange, T. D., Halvorsen, P., et al. (2019). "Resunet++: An advanced architecture for medical image segmentation." in Proceedings IEEE International Symposium on Multimedia (ISM). (San Diego, CA, USA: IEEE), 225–2255. doi:10.1109/ISM46123.2019.00049

Lee, Y., Hara, T., Fujita, H., Itoh, S., and Ishigaki, T. (2001). Automated detection of pulmonary nodules in helical CT images based on an improved template-matching technique. *IEEE Trans. Med. Imaging* 20 (7), 595–604. doi:10.1109/42.932744

Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., et al. (2020). Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy. *Radiology* 296, E65–E71. 200905. doi:10.1148/radiol.2020200905

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (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, 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:10.48550/arXiv.2103.14030

Mei, X., Lee, H. C., Diao, K., Huang, M., Lin, B., Liu, C., et al. (2020). Artificial intelligence–enabled rapid diagnosis of patients with Covid-19. *Nat. Med.* 26 (8), 1224–1228. doi:10.1038/s41591-020-0931-3

Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., et al. (2018). Attention u-net: Learning where to look for the pancreas. *arXiv:1804.03999*. doi:10.48550/arXiv.1804.03999

Paintdakhi, A., Parry, B., Campos, M., Irnov, I., Elf, J., Surovtsev, I., et al. (2016). Oufti: An integrated software package for high-accuracy high-throughput quantitative microscopy analysis. *Mol. Microbiol.* 99 (4), 767–777. doi:10.1111/mmi.13264

Qin, Z. Z., Sander, M. S., Rai, B., Titahong, C. N., Sudrungrot, S., Laah, S. N., et al. (2019). Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci. Rep.* 9 (1), 15000. doi:10.1038/s41598-019-51503-3

Rezatofighi, H., Tsoi, N., Gwak, J., Sadeghian, A., Reid, I., and Savarese, S. (2019). "Generalized intersection over union: A metric and a loss for bounding box regression." in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 15-20 June 2019. (Long Beach, CA, USA: IEEE), 658–666. doi:10.1109/CVPR.2019.00075

Rubin, G. D., Ryerson, C. J., Haramati, L. B., Sverzellati, N., Kanne, J. P., Raoof, S., et al. (2020). The role of chest imaging in patient management during the covid-19 pandemic: A multinational consensus statement from the fleischner society. *Chest* 296 (1), 106–116. doi:10.1016/j.chest.2020.04.003

Shi, H., Han, X., Jiang, N., Cao, Y., Alwalid, O., Gu, J., et al. (2020). Radiological findings from 81 patients with covid-19 pneumonia in wuhan, China: A descriptive study. *Lancet. Infect. Dis.* 20 (4), 425–434. doi:10.1016/S1473-3099(20)30086-4

Song, Y., Zheng, S., Li, L., Zhang, X., Zhang, X., Huang, Z., et al. (2021). Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 18 (6), 2775–2780. doi:10.1109/TCBB.2021.3065361

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nat. Med.* 25 (1), 44–56. doi:10.1038/s41591-018-0300-7

Valanarasu, J. M. J., Oza, P., Hacihaliloglu, I., and Patel, V. M. (2021). Medical transformer: Gated axial-attention for medical image segmentation. *arXiv. 2102: 10662*. doi:10.1007/978-3-030-87193-2_4

Wang, H., Cao, P., Wang, J., and Zaiane, O. R. (2021). Uctransnet: Rethinking the skip connections in u-net from a channel-wise perspective with transformer. *arXiv: 2109.04335*. doi:10.48550/arXiv.2109.04335

Wang, H., Zhu, Y., Green, B., Adam, H., Yuille, A., and Chen, L. C. (2020). "Axial-deeplab: Stand-alone axial-attention for panoptic segmentation." in Proceedings of the European conference on computer vision (ECCV), 29 October 2020. Berlin, Germany: Springer, 108–126. doi:10.1007/978-3-030-58548-8_7

Wong, H. Y. F., Lam, H. Y. S., Fong, A. H. T., Leung, S. T., Chin, T. W. Y., Lo, C. S. Y., et al. (2020). Frequency and distribution of chest radiographic findings in patients positive for Covid-19. *Radiology* 296 (2), E72–E78. doi:10.1148/radiol.2020201160

Woo, S., Park, J., Lee, J. Y., and Kweon, I. S. (2018). "Cbam: Convolutional block attention module." in Proceedings of the European conference on computer vision (ECCV), 06 October 2018. Cham: Springer, 3–19. doi:10.1007/978-3-030-01234-2_1

Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., et al. (2020). A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering* 6 (10), 1122–1129. doi:10.1016/j.eng.2020.04.010

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al. (2019). A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* 382, 727–733. doi:10.1056/NEJMoa2001017

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 01 September 2022
DOI 10.3389/fphys.2022.994343

OPEN ACCESS

EDITED BY
Huan Qin,
South China Normal University, China

REVIEWED BY
Teng Li,
Anhui University, China
Wu Jun,
Chinese Academy of Sciences (CAS),
China

*CORRESPONDENCE
Yu Pang,
pangyu@cqupt.edu.cn
Weiwei sun,
sunww@cqupt.edu.cn

†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 14 July 2022
ACCEPTED 01 August 2022
PUBLISHED 01 September 2022

CITATION
Zhang G, Huang Z, Lin J, Li Z, Cao E, Pang Y and sun W (2022), A 3D reconstruction based on an unsupervised domain adaptive for binocular endoscopy.
Front. Physiol. 13:994343.
doi: 10.3389/fphys.2022.994343

A 3D reconstruction based on an unsupervised domain adaptive for binocular endoscopy

Guo Zhang[1,2†], Zhiwei Huang[1,2†], Jinzhao Lin[3], Zhangyong Li[4], Enling Cao[5], Yu Pang[3*] and Weiwei sun[3*]

[1]School of Communication and Information Engineering, Chongqing University of Posts and Telecommunication, Chongqing, China, [2]School of Medical Information and Engineering, Southwest Medical University, Luzhou, China, [3]School of Optoelectronic Engineering, Chongqing University of Posts and Telecommunication, Chongqing, China, [4]School of Bioinformatics, Chongqing University of Posts and Telecommunication, Chongqing, China, [5]School of Software Engineering, Chongqing University of Posts and Telecommunication, Chongqing, China

In minimally invasive surgery, endoscopic image quality plays a crucial role in surgery. Aiming at the lack of a real parallax in binocular endoscopic images, this article proposes an unsupervised adaptive neural network. The network combines adaptive smoke removal, depth estimation of binocular endoscopic images, and the 3D display of high-quality endoscopic images. We simulated the smoke generated during surgery by artificially adding fog. The training images of U-Net fused by Laplacian pyramid are introduced to improve the network's ability to extract intermediate features. We introduce Convolutional Block Attention Module to obtain the optimal parameters of each layer of the network. We utilized the disparity transformation relationship between left- and right-eye images to combine the left-eye images with disparity in HS-Resnet to obtain virtual right-eye images as labels for self-supervised training. This method extracts and fuses the parallax images at different scale levels of the decoder, making the generated parallax images more complete and smoother. A large number of experimental research results show that the scheme can remove the smoke generated during the operation, effectively reconstruct the 3D image of the tissue structure of the binocular endoscope, and at the same time, preserve the contour, edge, detail, and texture of the blood vessels in the medical image. Compared with the existing similar schemes, various indicators have been greatly improved. It has good clinical application prospects.

KEYWORDS

adaptive, deep learning, binocular endoscopic, smoke, three-dimensional

1 Introduction

With the development of society, image processing (Li et al., 2016a; Li et al. 2016b; Li et al. 2018) is widely used in the medical field. During clinical surgery, the quality of medical images is degraded by noise. Noise is mainly composed of blood, light changes, specular reflection, smoke, etc. Among them, the smoke generated by laser and electrocautery-based human tissue ablations will significantly reduce the imaging quality of the lesion area. The results will affect the doctor's judgment, prolong the operation time, and increase the operation risk. Therefore, it is necessary to remove the smoke by physical means and purify it by image-processing algorithms (Kotwal et al., 2016; Yang and Sun, 2018; Chen et al., 2019a; Sidorov et al., 2020; Venkatesh et al., 2020). In addition, the particularity of the human tissue and imaging conditions are limited. Due to the influence of equipment light source and thermal noise acquisition, the quality of the collected endoscopic images is generally not high. Images obtained directly by endoscopy tend to have low imaging quality, resulting in the loss of some vascular tissue characteristics. Therefore, for the accuracy and convenience of later diagnosis, it is particularly important to improve the recognition ability of endoscopic images, filter out noise, and enhance the outline of the vascular tissue by reconstructing 3D details.

In the 3D display research based on the traditional stereo-matching method, the pixels of the left- and right-eye images have a parallax correspondence, and the 3D display can be performed after the parallax is obtained from the algorithm model (Hu et al., 2012; Besse et al., 2014; Yang and Liu, 2014; Penza et al., 2016). Compared with traditional algorithms, the method based on visual Simultaneous Localization and Mapping (SLAM) is slightly better in real-time performance. Most SLAM algorithms perform an inter-frame estimation and loop closure detection through feature point-matching techniques. Although the SLAM-based method only regards depth estimation as an intermediate product, its double-end depth estimation network provides a clear idea for subsequent research. Many subsequent articles have used its basic model (Mahmoud et al., 2016; Yi et al., 2016; Vijayanarasimhan et al., 2017; Wang et al., 2018a; Qiu and Ren, 2020). However, for the complex tissues and organs of the human body, traditional methods cannot meet the requirements of medical scenarios in terms of 3D reconstruction time and accuracy. In the research of 3D displays based on the neural network, researchers conducted supervised training on natural scene datasets containing depth labels. The final test can achieve the effect of real-time depth estimation (Antal, 2016; Kendall et al., 2017; Pang et al., 2017; Huang et al., 2018; Luo et al., 2019; Zhang et al., 2019). Since medical endoscopic images contain fewer datasets with depth labels (Penza et al., 2018), unsupervised learning is more suitable for 3D display of binocular laparoscopic images (Shurrab and Duwairi, 2022). A novel self-supervised learning strategy based on context restoration in order to better exploit unlabeled images (Chen et al., 2019b; Chen et al., 2022). The virtual viewpoint is obtained as a label through an implicit function, and the neural network is calculated and solved. Researchers can avoid a lot of dataset labeling work (Garg et al., 2016; Feng et al., 2017; Kendall et al., 2017; Zhou et al., 2017; Yin and Shi, 2018; Wang et al., 2019a; Tosi et al., 2019; Taleb et al., 2021).

In fact, it is often necessary to preprocess the image to remove various noises in the application of traditional methods and neural network schemes. Although the performance of neural networks on endoscopic images increases with the number of neurons, the complexity of convolution operations is very high. This leads to a blind increase in the size of the network and consumes a lot of training time. Therefore, combined with the real-time application requirements of clinical operations and the imaging characteristics of binocular endoscopes, we propose a 3D reconstruction method of binocular endoscope medical images based on adaptive neural network. The overall flow chart of the process is shown in Figure 1. The main contributions of this paper can be summarized as follows:

1) We proposed an improved U-NET adaptive network model for the smoke generated during laparoscopic surgery. We added training images fused by Laplacian pyramids at each layer of the encoder. A lightweight Convolutional Block Attention Module (CBAM) (Woo et al., 2018) attention mechanism module was added to the last five layers of the decoder to improve the network's ability to extract intermediate features. The processing time of a single image reaches 90.19 pfs, which can purify endoscopic surgical smoke in real time.
2) In view of the lack of true parallax in binocular endoscope images, we propose an improved HS-Resnet network. The left-eye image is combined with disparity to obtain a virtual right-eye image as a label for self-supervised training. In the process of feature extraction, multi-scale segmentation and synthesis are performed so that the network can effectively extract different scale features of various receptive fields. We reconstructed 3D structures with visibility and realism.
3) We proposed a color-difference 3D reconstruction scheme which separates the red component of the original image and combines the parallax, and fuses the combined red component with the blue–green component of the original image to obtain a 3D image. This can effectively reduce the details and color loss of the endoscopic image and retain the details of the medical images.

2 Methods

2.1 Smoke removal

2.1.1 Smoke synthesis

The improved U-Net (Zhou et al., 2021) model is used to realize the smoke purification function of endoscopic images. The steps of the smoke purification model are shown in Figure 2 below. Due to the lack of medical endoscopic image datasets containing real labels, we used the Render software to add smoke to real laparoscopy images as training images, and used the

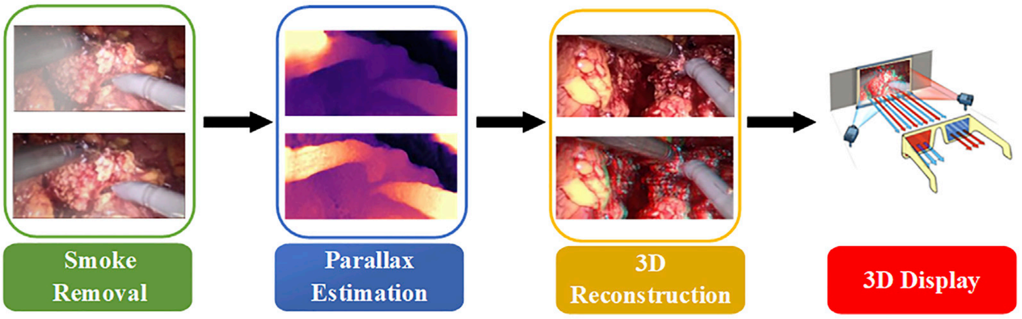


FIGURE 1
Overall process flow chart.

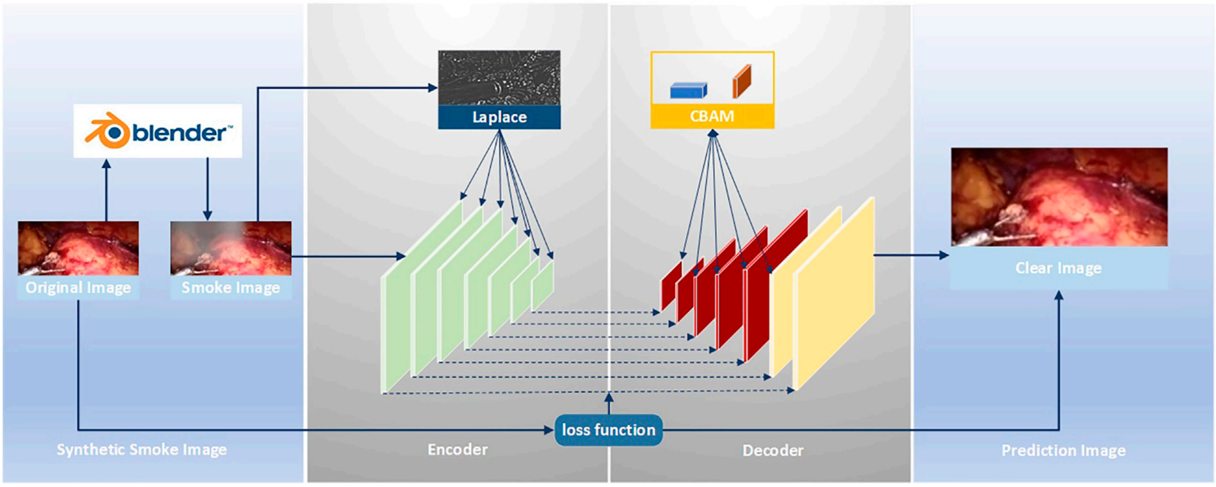


FIGURE 2
Smoke removal flow chart.

original images without smoke as labels. The loss function was obtained by comparing the purified image obtained by the model and the label, and back propagation reduced the loss to obtain the parameters of each layer of the network. In the network design, in order to increase the network's ability to retain image details and colors, we added the Rapp to the encoder. For the original fog image of Laplace fusion, the scale of Laplace transform is the same as that of the encoder. In order to improve the network's performance, we added the CBAM attention mechanism to the last five layers of the decoder to use the synthetic image containing smoke as the training set. The original image is sent to the improved U-Net model as a training set label for training. Through back propagation, each layer of the network obtains the corresponding parameters. Finally, the test set is sent to the model to predict the purified image.

The medical scene dataset in the field of smoke purification is relatively rare. There is currently no dataset containing real labels. Therefore, either unsupervised learning can be used or software can be used to add smoke to medical endoscopic images. Then, use the images without the added smoke as labels. In the two schemes, simple supervised learning can be used to solve the problem after smoke synthesis, and supervised learning is mature in the field of deep neural networks. So, we choose the scheme of artificially synthesizing smoke. The smoke generated in laparoscopic surgery is usually generated randomly and has nothing to do with the depth. The modern image-rendering engines have a complete built-in model. This can better simulate the shape of the smoke compared to physical solutions. Therefore, we used the 3D graphics-rendering engine. Render to the training

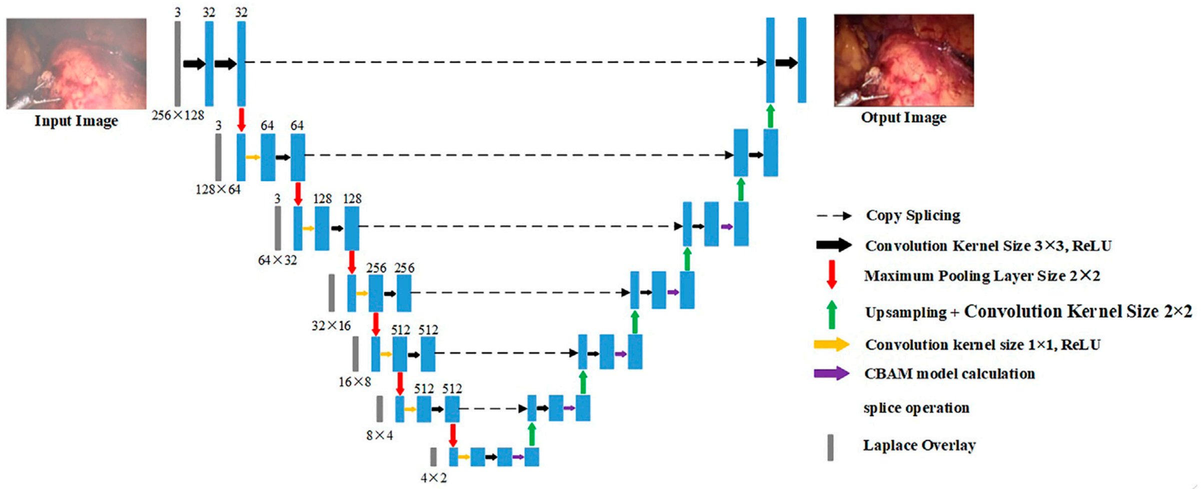


FIGURE 3
Encoder–decoder network.

images are obtained by rendering the smoke on laparoscopic images that do not contain smoke.

The smoke is rendered by the rendering engine and has local color and transparency. The smoke is controlled by the input parameters Trad, Drand, and Position, as shown in formula (1):

$$I_{\text{smoke}\,(x,y)} = \text{Blender}\left(T_{\text{rand}}, D_{\text{rand}}, P_{\text{rand}}\right) \tag{1}$$

Using Render to fog the laparoscopic image, the rendered smoke is similar to the real smoke. It has the characteristics of local pure white and transparency. The fogged image is superimposed by the original image and random smoke, as shown in formula (2):

$$I_{\text{s-image}}(x, y) = I_{\text{s-free}}(x, y) + I_{\text{smoke}} \tag{2}$$

The smoke added to the laparoscope is obtained by superimposing the luminance values of the rendered R, G, and B channels proportionally. The ratio is shown in formula (3):

$$I_{\text{mask}}(x, y) = \left(0.3^{*} I_{\text{smoke}}(x, y)^{R}\right) + \left(0.59^{*} I_{\text{smoke}}(x, y)^{G}\right) + \left(0.11^{*} I_{\text{smoke}}(x, y)^{B}\right) \tag{3}$$

To better simulate light smoke, fog, and thick smoke fog, we rendered two types of fog. Firstly, images without fog are selected as the original training set in the dataset. In rendering, the original dataset is randomly added fog using the data settings of the aforementioned formula. We added primary smoke as the light fog dataset. Then, the light fog dataset is sent into the rendering for secondary random adding fog to obtain the thick fog dataset. Finally, training is performed on the thick fog dataset and the light fog dataset, respectively.

2.1.2 Improved U-net network

For the original U-Net, it is found through experiments that it cannot effectively purify the smoke, or the image resolution decreases after purifying the smoke. This is due to the loss of image details in the process of up-sampling and down-sampling. But for medical scenes, the loss of detail information will seriously affect the doctor's judgment. Therefore, we added the training image fused by the Laplacian pyramid in the down-sampling part to compensate for the loss of details of the image during the down-sampling process. The image fusion of the Laplacian image pyramid is equivalent to a filter, which maps the image to different frequency bands. The features are learned, and fusion operations are performed on each frequency band, thereby effectively preserving image details on each frequency band. The U-Net model is improved according to the characteristics of medical endoscopy, as shown in Figure 3.

The down-sampling part on the left of Figure 3 is the encoder. The encoder can extract features of different scales of endoscopic images through convolution operations. Each layer of the encoder corresponds to splicing, two convolutions, and a max pooling operation. Laplace the superposition operation refers to the fusion of the Laplacian pyramid image for the input training image according to the size of each layer of the encoder. The transformed image and the features of each layer are spliced and sent to training. The seven convolution groups are named conv1 to conv7, respectively. The size of the convolution kernel of each layer is 7 × 7, 5 × 5, 3 × 3, 3 × 3, 3 × 3, 3 × 3, and 3 × 3. Each layer is convolved twice. The strides of the two convolutions are 1 and 2, respectively. The number of output layers per layer is 32, 64, 128, 256, 512, 512, and 512. Therefore, the encoder down-sampling factor is 64.

The decoder restores the down-sampled image to its original size. The CBAM attention module is inserted into the first five layers of the decoding part of the U-Net network, as shown in the up-sampling part on the right side of Figure 3. The decoder also adopts 7 sets of convolutions; each group contains two up-sampling layers with steps 1 and 2. The size of the convolution kernel is all 3 × 3, and the number of output layers is 512, 512, 256, 128, 64, 32, and 16, respectively. In addition, there are corresponding connections between the encoders and decoders where the features of the lower layers are connected with the features of higher layers. Information from the higher layers can be directly transmitted to the bottom layer of the network to prevent the loss of high-quality details.

The loss function of the improved U-Net network is the minimum absolute value deviation loss of the original image and the synthetic smoke image, as shown in formula (4):

$$L = \sum_{xy} \left| I_{\text{original}}(x, y) - I_{\text{desmoked}}(x, y) \right| \tag{4}$$

2.1.3 Laplacian image pyramid fusion

The maximum pooling operation is used in the down-sampling process. Due to the continuous down-sampling operation, the image details are lost in each frequency domain. In order to better preserve the image quality in the specified frequency domain, a Laplacian image is introduced in the encoder part of the pyramid fusion. This method uses the nearest point interpolation when up-sampling the image after Gaussian sampling. Especially in the place where the image gradient changes greatly, the problem of sudden change of the pixel value occurs easily. The image details are lost, and there may be mosaic or sawtooth noise (Wang et al., 2019b). This article uses bidirectional interpolation to replace the nearest neighbor interpolation to improve this problem. It processes the four direct neighbors near the sample point. The image quality is higher after processing.

The Laplacian-transformed smoke image is added before each convolutional layer in the encoder, and the main process of the Laplacian pyramid fusion is shown in formula 5:

$$L_i(I) = G_i(I) - \text{up}(\text{down}(G_i(I))) \tag{5}$$

where I represents the original image containing smoke; i represents the level pyramid. $\text{up}(\text{down}(G_i(I)))$ represents the up-sampled lower-layer Gaussian sampled image; and $G_i(I)$ represents the Gaussian sampled image.

As shown in Figure 4, to smoothly image the image to different frequency bands, we performed Gaussian down-sampling on the endoscopic image. As shown in the color endoscopic image, as the number of Gaussian sampling increases, the size of the endoscopic image becomes smaller. But it can retain the more important pieces of information in the image. For a Laplacian-transformed image of a specific size, Gaussian down-sampling is performed according to the specified scale, and then the Laplacian pyramid fusion image is obtained. As shown in the black and white image, the Laplacian pyramid fusion image (in order to make the image easy to observe, the brightness value of the Laplacian fusion image is increased) effectively retains the line and edge information of the image. The size is the same as the U-Net down-sampling size. Therefore, it can be directly superimposed and spliced with the input feature image in the network and then be sent to the network for training. Finally, this article splices it to the corresponding size of the convolutional layer to participate in training.

2.1.4 CBAM attention mechanism

The CBAM attention mechanism module is lightweight and effective. Therefore, we inserted the CBAM attention mechanism module in the last five layers of the decoder; following the network to participate in the training process to improve the feature selection performance of the model. For any input feature, CBAM obtains attention features along two independent dimensions of channel and space. The original input feature is optimized by multiplying the attention feature with the original input feature image. The specific process is as follows: for the input feature image $F \in R^{C*H*W}$ of any size, CBAM will calculate a one-dimensional channel feature image $M_c \in R^{C*1*1}$. A two-dimensional spatial feature image $M_c \in R^{1*H*W}$ is derived in the blue part of Supplementary Figure S1. The orange part is shown in formulas 6 and 7. The blue part of the channel information and the orange part of the spatial information are fused with the original input feature to obtain the optimized input feature. This feature is used as the next input to the convolutional layer.

$$F' = M_c(F) \otimes F \tag{6}$$

$$F'' = M_s(F') \otimes F' \tag{7}$$

We treat each channel of input features as a feature detector, as shown in the blue part of Supplementary Figure S1. Channel attention is used to pay attention to the content of the input image, so the features are compressed into a "pipe". Observe the image content that still exists after the reduction, and find its calculation method as follows: First, average the pooling and summing of the input features according to their feature-stacking direction. Convolution is performed after max pooling, followed by the activation of the result of the convolution, and finally the feature fusion. As shown in formulas 8 and 9:

$$M_c(F) = \sigma(MLP(AvgPool(F)) + MLP(MaxPool(F))) \tag{8}$$

$$M_c(F) = \sigma\left(W_1\left(W_0\left(F^c_{avg}\right)\right) + W_1(W_0(F^c_{\max}))\right) \tag{9}$$

where $W_0 \in R^{C/r*C}$ and $W_1 \in R^{C*C/r}$, using ReLU as the activation function after W_0.

Channel attention pays attention to the key positions of the image. Spatial attention compresses the feature dimension into

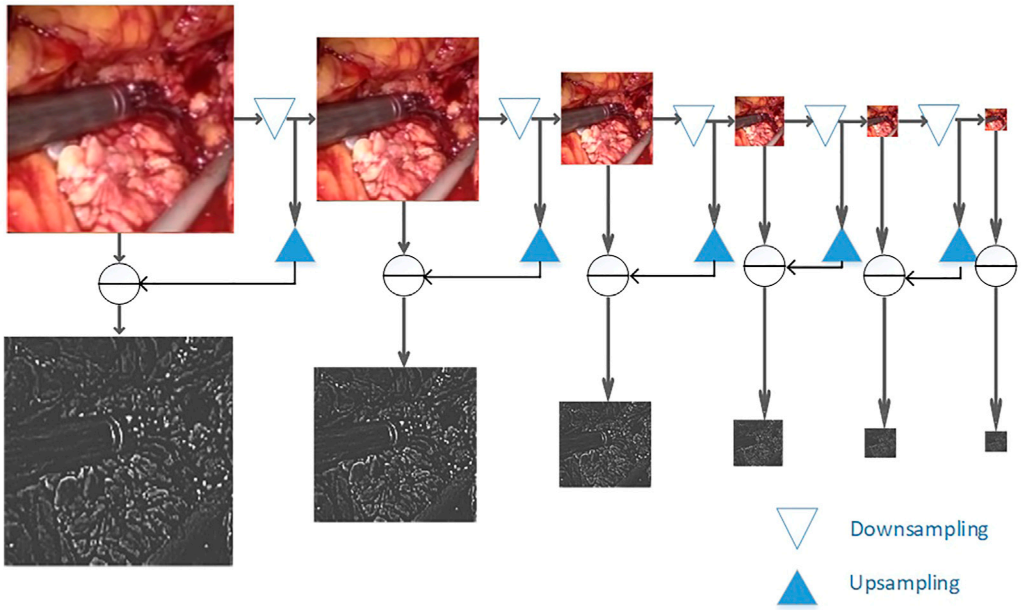


FIGURE 4
Laplacian image pyramid decomposition.

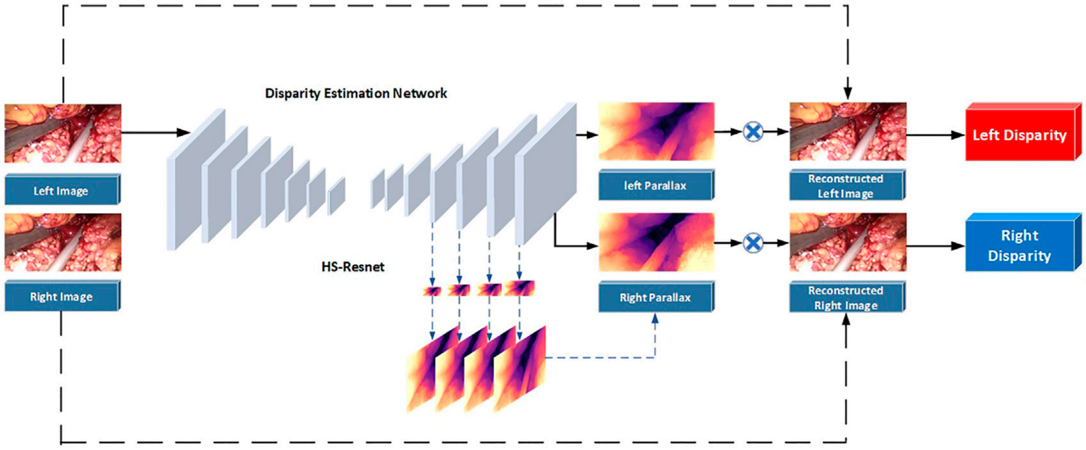


FIGURE 5
Flowchart of the parallax estimation algorithm.

an "image", which is convenient for the neural network to identify the position of the image object. As shown in the orange part of Supplementary Figure S1, two different feature descriptions $F^s_{max} \in R_{1*H*W}$ and $F^s_{avg} \in R_{1*H*W}$ are obtained by using max pooling and average pooling in the dimension of the channel, and then, the aggregation operation is used to generate the spatial feature image $M_s(F) \in R_{H*W}$. As shown in formulas 10 and 11:

$$M_s = \sigma\left(f^{7*7}\left(\left[AvgPool(F); MaxPool(F)\right]\right)\right) \qquad (10)$$

$$L = \sum_{xy}\left|I_{\text{original}}(x, y) - I_{\text{desmoked}}(x, y)\right| \qquad (11)$$

2.2 A method for estimating binocular disparity in endoscope is proposed

The parallax estimation method of a binocular endoscopic image based on self-supervised deep learning is shown in Figure 5. The corrected left and right images are used as inputs. The left image is used as a standard input into a convolutional neural network for training. The left and right original images are used as labels to provide supervision information for the network.

Step 1: The corrected left and right images are taken as the training images, and the left image is sent into the neural network for training. Then, the initial left and right parallax images are obtained by a CNN non-linear function fitting.

Step 2: The left and right parallax images obtained from the network can be regarded as the deviation between the left and right views. So, the virtual right image can be obtained by combining the original left image and the left parallax through bilinear interpolation. The virtual left image can be obtained by combining the original right image and the right parallax.

Step 3: Reverse propagation is carried out by comparing the difference between the virtual left view and the real left view and between the virtual right view and the real right view. Appropriate parameters can be obtained for each layer of the network.

2.2.1 CBAM attention mechanism

The encoder is used to construct the U-NET structure with ResNet as the convolutional layer of the network, extracting the features of endoscope images. The size of the images is restored to the original size through the decoder. Specifically, the encoder first preprocesses the convolution for the inputted RGB images, with a convolution kernel size of 7 × 7, step length of 2, and zero fill of 3. After preprocessing, the image is batch normalized, followed by 4 convolutions with a convolution kernel of 3 × 3. After 5 convolutions, the feature dimensions of the convolution kernel size are 16, 32, 64, 128, and 256.

Multi-scale features are particularly important in machine vision, which can image features to multiple frequency domains and be conducive to keeping detailed features of images. Focusing on medical endoscope images that require highly detailed features, an HS-Resnet containing multi-scale features is adopted (Godard et al., 2017). It contains a hierarchical separation module embedded in the convolutional module of the deep network, where HSB can effectively improve the performance of the network and HS-ResNet 50 can achieve 81.28% of the datasets on ImageNet, exceeding the current optimal effect of ResNet. As shown in the Figure 6, HS-Resnet is composed of multiple segmentation and splicing operations, of which the hierarchical segmentation and splicing operations together constitute the HSB multi-scale feature extractor.

HSB contains two main operations: split and concatenate. Among these two, split is used for feature grouping and to make the two groups after grouping have the same number of channels. When the number of features to be grouped is odd and the channel number of the two groups after the split operation is different, one part can be regarded directly as the output, equivalent to identity imaging, and the other part can be used as the input to the next convolution layer for a more in-depth and detailed feature extraction. The concatenate operation fuses features with the same size but different contents so that features with different convolution degrees can carry out information interaction. When the concatenate operation adopts a simple superposition operation, the characterization ability of the original features can be better ensured.

Figure 6 shows the structure of HSB, where HS-Resnet uses a deep residual module and HSB processes features in the 3 × 3 convolution layer. The input features are divided into S groups x_i with the same number of channels after 1 × 1 convolution. Then, after passing through a 3 × 3 convolution layer $F_i\,()$ times, x_i becomes y_i, which can be divided into $y_{i,1}$, $y_{i,2}$, $y_{i,3}$, $y_{i,4}$, and $y_{i,5}$. Among them, $y_{i,1}$ is added directly to the layer and to the output x_{i+1}, similar to the green feature on the top. $y_{i,2}$ is segmented into two groups with yellow features after the convolution operation, where one group is added to the layer and to output x_{i+1}. The other group is sent to the convolution layer after matching with $y_{i,3}$. Similarly, $y_{i,3}$ is divided into two groups with red features, where one group is added to the layer and the output x_{i+1}. The other group is sent to the convolution layer to obtain the green feature after matching with $y_{i,4}$. $y_{i,4}$ is processed same as $y_{i,3}$. Finally, the feature of $y_{i,5}$ after the convolution operation will be taken as the last part of the layer's output. After such continuous processing, the features are equivalent to more scale and deeper convolution. The small receptive field in the final output feature can pay attention to the detail part and enhance the processing ability of the network for small features.

Figure 6 shows the situation in which s is set to 5. In fact, a larger number of groups can achieve the extraction performance of more scales. A larger number of channels means richer features requiring more parameters. Therefore, it is necessary to choose between the number of parameters and the capability of feature extraction.

$$y_i = \begin{cases} x_i & i = 1 \\ f_i\left(x_i \oplus y_{i-1,2}\right) & 1 < i \leq s \end{cases} \tag{12}$$

HSB does not increase the number of parameters in the network. Compared with a standard convolution, it even has fewer parameters. The standard parameter complexity is shown in Formula 13:

$$P_{normal} = k \times k \times s \times w \times s \times w = k^2 \times s^2 \times w^2 \tag{13}$$

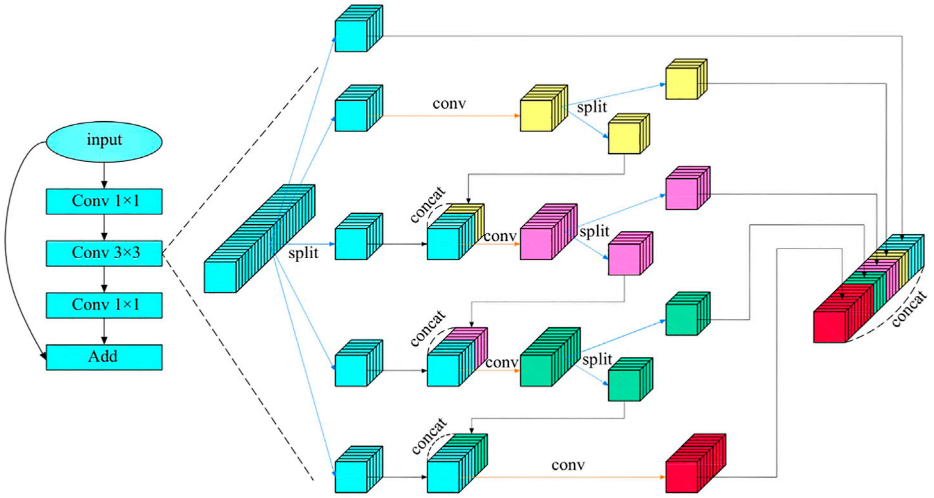


FIGURE 6
HSB module.

The complexity of HSB is shown in Formula 14:

$$P_{HSB} = \begin{cases} 0, & i = 1 \\ k^2 \times w^2 \times \left(\dfrac{2^{s-1}-1}{2^{s-1}} + 1\right) & 1 < i \le s \end{cases} \tag{14}$$

It can be seen from the comparison between Formulas 13 and 15 that the complexity analysis of HSB is actually smaller than that of an ordinary convolution.

$$k^2 \times w^2 \times \left(\frac{2^{s-1}-1}{2^{s-1}} + 1\right) \le k^2 \times w^2 \times \left(\frac{2^{s-1}-1}{2^{s-1}} + s - 1\right) < \\ k^2 \times w^2 \times (s-1+s-1) = k^2 \times w^2 \times (2s-2) < k^2 \times w^2 \times s^2 \tag{15}$$

2.2.2 Multi-scale decoder

The decoder is the deconvolution process of the encoder, aiming to restore the image to the original image size. The decoder up-samples the image, which includes a 3 × 3 deconvolution, to restore each layer of the image to the same size as the decoder. The output feature dimensions of each convolution are 256, 128, 64, 32, and 16. Bilinear sampling has gradient locality, and may not converge to the global minimum during the training process of the final disparity estimation. Therefore, the disparity is extracted from the last four layers of the filter during decoding. And then, the disparity calculation loss function of each layer is fused into the final loss function solution. Each layer calculates the loss function according to different image sizes. Due to the severe compression of low-resolution images, it is difficult to retain important details of the image. Parallax discontinuities are prone to occur in the weak repeating parts of the tissue structure, because the photometric errors at these locations are blurred and inaccurate. Inspired by binocular stereo vision, we improved the loss function and reconstructed the disparity image in the last four layers of the decoder with different image sizes. The loss functions of different scales were calculated.

2.2.3 Improved loss function

1) Photometric reconstruction loss

Self-supervised learning mainly uses the disparity relationship between the left and right images of the binocular endoscopic image to establish a loss function (Godard et al., 2019). The training loss is expressed as a photometric re-projection loss, which is used to describe the difference between the virtual viewpoint and the real view. The total loss is obtained by adding the losses of all pixel points. The calculation process of the loss function is shown in formula 16:

$$L(p) = \sum pt\left(I_t, I'_t\right) \tag{16}$$

I_t is the original image, I'_t is the virtual view, and pt is the difference between the two images. The total photometric loss is obtained by combining the differences of all images (Zhao et al., 2019). The structural similarity index SSIM is used to characterize the photometric reconstruction error. The specific calculation process is shown in formula 17:

$$pt\left(I_t, I'_t\right) = \frac{\alpha}{2}\left(1 - SSIM\left(I_t, I'_t\right)\right) + (1-\alpha)\left\|I_t, I'_t\right\|_1 \tag{17}$$

Among them, α is the weight coefficient between SSIM and L1 norms, which can be obtained from training experience. We set it as $\alpha = 0.85$. During training, the model extracts image features from the left image in the binocular laparoscopic image to obtain the initial disparity. Then, it use the left image and the original image to linearly translate to get the virtual right image, and then compare the real right image with the original right image to get the loss. Image sampling is performed using Spatial Transformer Networks (STN) (Jaderberg et al., 2015). The original image is sampled with the disparity image as the standard, and the STN takes the weighted sum of the surrounding four pixels for each sampling point. Its calculation process can be differentiated and can follow the neural network to participate in the process of back propagation.

As shown in formula 16, the existing literature generally averages the re-projection loss across all training images when calculating the photometric reconstruction error for self-supervised depth estimation. This has some problems in consecutive images. Certain matching feature points do not match in the occluded image. This leads to a large error in photometric reconstruction. However, the loss function is averaged, so that the two points cannot be correctly matched. Then, the obtained disparity image or depth image is blurred. Pixels that are easily occluded during continuous motion mainly come from the boundaries of moving objects. For example, in the process of laparoscopic surgery, the forceps move more frequently and there will be a long-term or short-term occlusion in the patient's body. The background in the human body cannot be matched. For the photometric reconstruction loss of the same pixel appearing in different images, this article adopts the minimum value instead of the average value to improve the photometric loss. As shown in formula 18:

$$L_p = \sum_{t'} pe\left(I_t, I_{t'\rightarrow t}\right) \Rightarrow L_p = \min_{t'} pe\left(I_t, I_{t'\rightarrow t}\right) \tag{18}$$

For all pixels in an image, it is not necessary to calculate the loss function in its entirety. We use an automatic masking scheme that preserves points that move relative to the camera and removes points that are stationary relative to the camera. For example, in laparoscopic surgery, when the abdominal lens is rotated, all pixels move with the lens. At this point, all pixel point losses are calculated. When the abdominal lens remains stationary, the background of the internal abdominal cavity that the endoscope can look into is fixed. As the forceps moves the abdominal tissue relative to the lens, only the moving portion is counted when calculating the loss. The rest of the points are removed, and the removed part is called a mask. The mask is computed by the network. Masked pixels can be characterized as a static camera, which is equivalent to being relatively stationary with the camera, or can represent low-texture areas.

This article uses the binary mask parameter $\mu \in \{0, 1\}$. Among all loss functions, μ is only related to the photometric reconstruction loss, as shown in formula 19:

$$\mu = \left[\min_{t'} pe\left(I_t, I_{t'\rightarrow t}\right) < \min_{t'} pe\left(I_t, I'_t\right)\right] \tag{19}$$

2) Left–right consistency loss

Our proposed photometric reconstruction error can examine the similarity between the original view and the virtual view. The left and right consistency loss is used to measure the similarity between the left and right disparity images generated by the network. The disparity acquisition module only has the left image as input, but needs to predict the left and right binocular disparity images. Therefore, the similarity between the left and right disparity images needs to be constrained. A virtual right disparity image can be obtained by linearly transforming the left image disparity on each pixel using right image disparity. The original right disparity image is compared with the virtual right disparity image, and the L1 norm is obtained as the left–right consistency loss. The left and right consistency losses can constrain the left and right parallaxes to ensure the accuracy and continuity of the parallax. In order to reconstruct the loss obtained from the right disparity, we also calculated the loss to reconstruct the left disparity during training, as shown in formula 20:

$$L^l_{lr} = \frac{1}{N}\sum_{i,j}\left|d^l_{ij} - d^r_{ij+d^l_{ij}}\right| \tag{20}$$

3) Edge-smoothing loss

There is a very strong connection between adjacent disparity images. Constraining the transformation magnitude of disparity through a loss function can effectively improve the problem of discontinuous disparity. Parallax can also be locally smoothed. We used the L1 norm to constrain the left and right disparities to ensure continuous and smooth binocular disparity, as shown in formula 21:

$$L^l_{ds} = \frac{1}{N}\sum_{i,j}\left|\partial_x d^l_{ij}\right| e^{-\left\|\partial_x I^l_{ij}\right\|} + \left|\partial_y d^l_{ij}\right| e^{-\left\|\partial_y I^l_{ij}\right\|} \tag{21}$$

To sum up, the improved loss function is composed of the aforementioned three types of loss functions, as shown in formula 22:

$$L = \mu\left(L^r_p + L^l_p\right) + \lambda\left(L^l_{lr} + L^r_{rl} + L^l_{ds} + L^r_{ds}\right) \tag{22}$$

2.3 Evaluation method

In clinical applications, the doctor's subjective evaluation is the most important factor in judging the image quality. There is no gold standard available for quantitative assessment especially in laparoscopic and endoscopic images (Zhang et al., 2022). Therefore, to verify the performance of tissue blood vessels, brightness, and color enhancement, we define two evaluation metrics: 1) Peak Signal-to-Noise Ratio (PSNR), and 2) Structural Similarity Index (SSIM).

PSNR and SSIM were used to evaluate image quality. PSNR is a measure of the quality of image reconstruction. The higher the PSNR value, the better the image quality will be. The formula is as follows:

$$MSE = \frac{1}{\mathrm{mn}} \sum_{i=0}^{\mathrm{m}-1} \sum_{j=0}^{n-1} \left[I(i,j) - K(i,j) \right]^2 \tag{23}$$

$$PSNR = 10 \times \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \tag{24}$$

where MSE represents the mean square error; MAX_i^2 represents the maximum possible pixel value of the image; $I(i,j)$ represents the original image; and $K(i,j)$ represents the noise image.

SSIM is used to measure the similarity of two images. The larger the SSIM value, the more similar the two images are. The formula is as follows:

$$\mathrm{SSIM}(x,y) = \left[l(x,y) \cdot c(x,y) \cdot s(x,y) \right] \tag{25}$$

$$l(x,y) = \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \tag{26}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \tag{27}$$

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x\sigma_y + c_3} \tag{28}$$

where μ represents the mean, σ represents the variance; and σ_{xy} represents the covariance of x and y; $c_1 = (k_1L)^2$ and $c_1 = (k_2L)^2$ represent two constants, with $k_1 = 0.01$ and $k_2 = 0.03$; and L represents the range of image pixels.

3 Results and discussion

3.1 Data set and training parameter settings

Our experimental conditions are 64-bit Windows 10 operating system, using Intel(R) Core(TM) i7-10750H CPU; 32 GB RAM; NVIDIA 12 GB 3080Ti GPU. Install CUDA9.0 and use cuDNN7.0 for acceleration. On this basis, the U-Net model is built on the Tensorflow1.10.0 framework, as shown in Supplementary Table S1.

The dataset adopts the updated laparoscopic binocular dataset from the Hamlin Center (Chen et al., 2017). The left eye image is used for smoke cleaning. The binocular data are used for disparity estimation. The experimental dataset has a total of 34,240 pairs of binocular laparoscopic training images and 7,000 pairs of test images. This article divides the training images, of which 30,000 pairs of laparoscopic images are used as training sets and 4,240 pairs of validation images. Since many images in the laparoscopic dataset originally contain images of smoke, we perform supervised learning after fogging the images. The fog in the original image will affect the performance of the model, so we selected images that do not contain fog from the dataset to add fog. To ensure the reliability of the experimental data, each round of experiments is tested on synthetic smoke images and real smoke images. They were used in ten-fold cross-validation experiments. After training and validation separately, we used the test set to test, repeat this process ten times, and finally take the average of the ten results as the evaluation of algorithm accuracy. The synthetic image test set contained 1,000 images and the real smoke dataset contained 129 images. After fog rendering was performed on each image as a training set, the rendered images were divided into two levels: light fog and dense fog. During the training process, all images were first resized to a fixed size of 256 × 128, and then input to the model, the mean square error loss function was used, Adam is used as the optimization, the batch is set to 16, and the initial learning rate is set to 0.0001. The experiment adopts the control variable method, and conducts four sets of experiments for two levels of fog: including U-Net network, U-Net network plus CBAM attention mechanism, U-Net network plus Laplace transform, U-Net network plus CBAM attention mechanism, and Laplace transform.

The average training time of each model group is 4.5 h. According to the different levels of smoke and different model combinations, when the average loss is reduced to 0.02–0.03 in the light fog image training set, it will no longer decrease, and overfitting will not occur. The average loss on the validation set drops to around 0.3 and no longer decreases. When the average loss is reduced to 0.03–0.04 in the training set of dense fog images, there is no drop and no overfitting. The average loss on the validation set drops to around 0.4 and no longer drops. After training, export the model. This article can perform smoke purification on the synthetic image dataset. In order to apply it to engineering practice, this article uses the real dataset containing smoke; this model can purify real smoke.

3.2 Experiment results of dehazing of laparoscopic images

In addition, the test results of the synthetic dataset are shown in Figure 7. Figure 7A shows a synthetic smoke image, which is characterized by thick smoke and blocking of the original tissue structure. Figure 7B shows the results of using the original U-NET. There is still some residual smoke and the effect is not good. We used the Laplace pyramid transform to completely purify the smoke in Figure 7C. But, the brightness and color saturation of the bright parts of the original image were reduced. The smoke can be effectively purified with good color retention after the Laplacian pyramid transform is added in Figure 7D and Figure 7E.

In order to verify the validity of the model, we used the light fog dataset to conduct comparison experiments with other parameter settings under the same control experiment. The results are shown in Table 1. The training loss of adding the

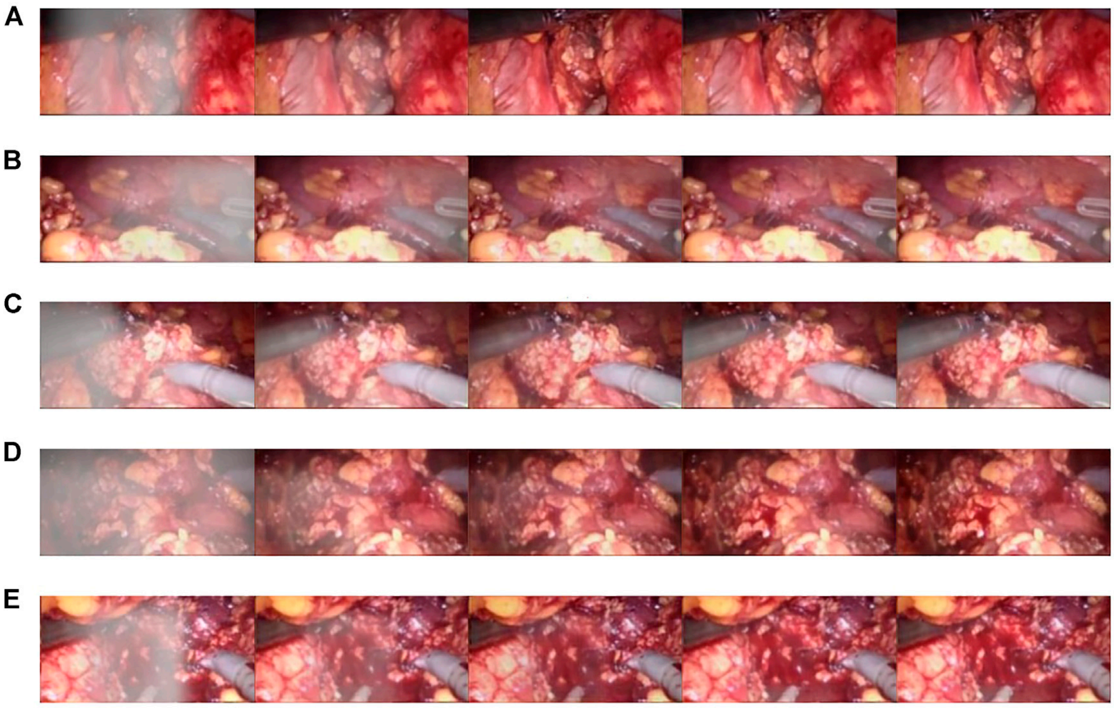


FIGURE 7
Synthetic smoke laparoscopic images and purified images. **(A)** Laparoscopic image with synthetic smoke and after purification; **(B)** Laparoscopic image with U-NET; **(C)** Laparoscopic image with U-NET + BAM; **(D)** Laparoscopic image with U-NET + Laplace; **(E)** Laparoscopic image with U-NET + CBAM + Laplace.

TABLE 1 Model performance verification.

Model	PSNR	SSIM	PFS	Loss
U-Net	30.522	0.936	72.256	0.045
U-Net + CBAM	31.435	0.966	106.40	0.023
U-Net + Laplace	31.126	0.977	74.074	0.038
U- Net + CBAM + Laplace	31.045	0.980	90.191	0.026

CBAM module alone is 0.023, and adding the Laplace transform alone is 0.038. In the case of Laplacian pyramid transform and the CBAM attention mechanism, the training loss of the model was 0.026. The CBAM module can better optimize the model. In terms of processing time, the CBAM module achieved a good result of 106.4pfs, and the best PSNR value was 31.435 dB. On the SSIM index, the experiment of adding the Laplacian pyramid obtained the best effect of 0.98.

3.3 Three-dimensional display experiment

Resnet50 was used for training; the training time was 7–8 h; the final loss obtained by training was 0.06. When HS-Resnet50 was used for training, the final loss was about 0.05. There was no overfitting in both schemes. The loss of HS-Resnet50 was lower, and the model training effect was better.

Qualitative test results are shown in Figure 8. Figure 8A and Figure 8C show the endoscope test images. Figure 8B and Figure 8D show the RGB parallax images obtained using the HS-Resnet model. It can be seen from the test images that the parallax images generated by the proposed model are complete and continuous, without any void phenomena. In the parallax images, a light-colored part is an object close to the camera and a dark-colored part is an object far from the camera. It can be confirmed from the original image that the distance relationship in the parallax images generated by this model is accurate.

As shown in Figure 9, the binocular endoscope depth-estimation algorithm based on the improved HS-Resnet model can effectively obtain the disparity image while retaining the image details. The blood vessels in the abdominal cavity in Figure 9A are well preserved in the parallax Figure 9B. The original tissue texture of the image can be observed through the parallax image. The blood vessel information is very important in medical images, highlighting the blood vessels in the image and more. More details can also prevent doctors from accidentally injuring patients.

Figure 10 shows the influence of smoke on disparity estimation. Figure 10A represents the synthetic smoke image, and it can be seen from the image that the smoke

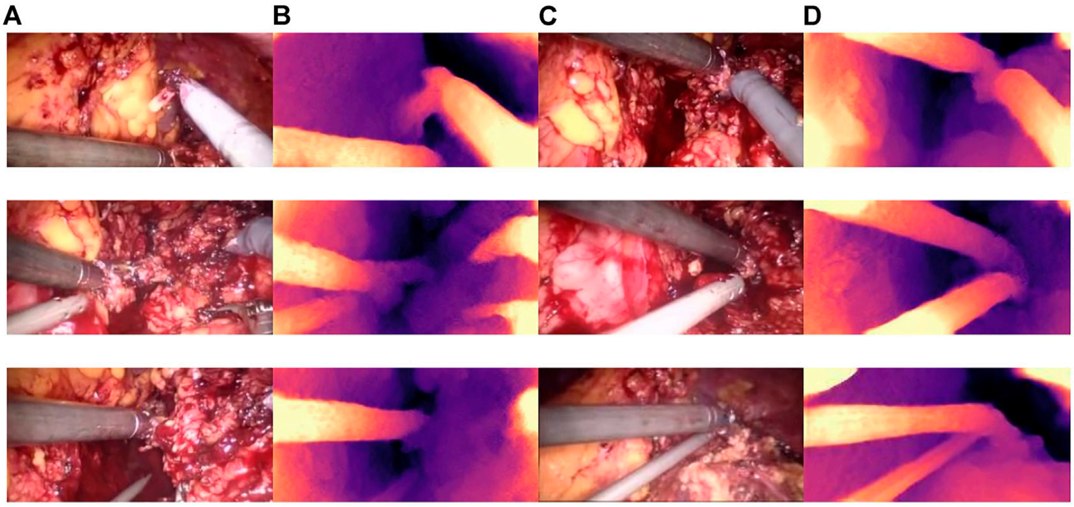


FIGURE 8
Parallax estimation results. **(A,C)**: endoscopic test images. **(B,D)**: the parallax image obtained using the HS-RESNET.

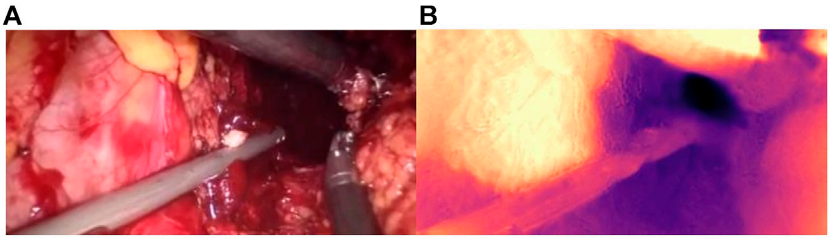


FIGURE 9
Details of the parallax estimation. **(A)**: Raw endoscope image; **(B)** proposed model parallax.

covers the front of the abdominal image. Figure 10B represents the image after chapter 3 smoke purification; it can be seen from the image that it no longer contains smoke. Figure 10C shows that the disparity value is obtained by performing a depth estimation on the image containing smoke. Due to the occlusion of the smoke, the disparity estimation is relatively blurred. The color is darker and it is difficult to distinguish the edge information. There are large black areas in the image that cannot be identified. Figure 11D shows the parallax estimation of the cleaned image compared with Figures 10C, D is lighter in color and easier to observe. The edge information image is clearer. The parallax estimation model in Figure 10C is occluded by smoke, which makes it difficult for the parallax estimation model to estimate the specific depth of human tissue. The parallax can be accurately estimated after the smoke is purified.

3.4 Smoke removal model performance verification

The CBAM attention module can effectively improve various indicators of the model. The Laplacian pyramid transform can better retain image details. The experimental results on real images are shown in Figure 11, and it can be seen from Figure 11A, Figure 11C, and Figure 11A that in surgery, real smoke generally blocks the doctor's sight and fuzzes up the real vision in the scene. After removal, Figure 11B, Figure 11D, and Figure 11F show that the image processed using this model can purify smoke in the figure so that the fuzzy images are clearer.

In each subimage, the left is a smoke image randomly captured from a real surgery video and the right is the smoke removal result of the CBAM + Laplace image pyramid fusion + U-NET model.

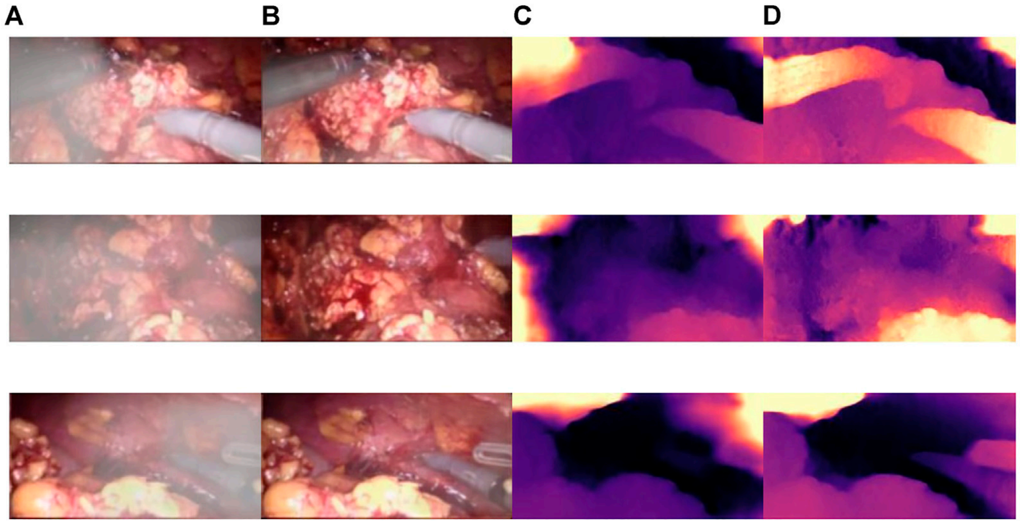


FIGURE 10
The effect of smoke on disparity estimation. **(A)** Synthetic smoke images; **(B)** smoke-removed images; **(C)** disparity images of the smoke containing images; **(D)** disparity images of the smoke-free images.

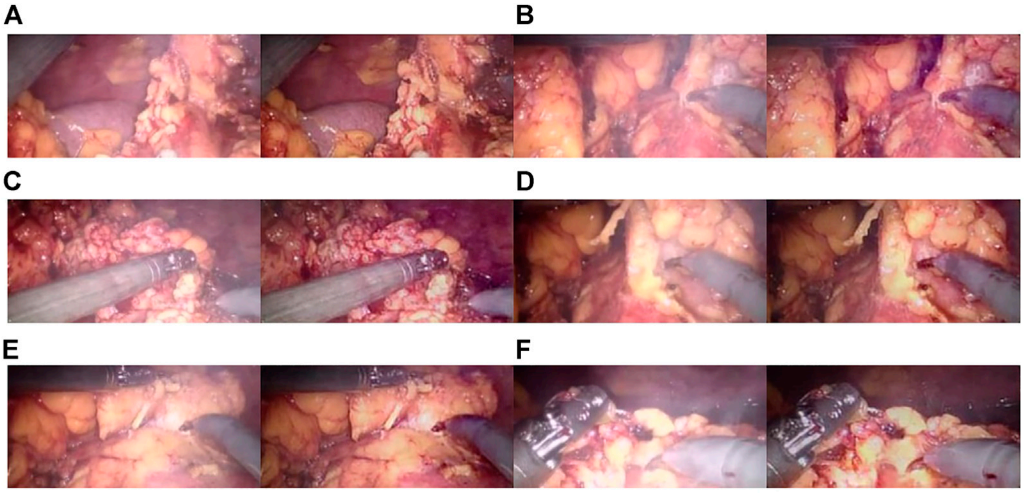


FIGURE 11
Laparoscopic images of real smoke and images after removal. **(A–F)** are randomly selected from the experimental results. Laparoscopic images of real smoke (left sub-panels) and images after smoke removal by U-Net + CBAM + Laplace pyramid fusion (right sub-panels).

The first image in each subimage is a smoke image randomly selected from a synthetic smoke dataset; the second image shows the smoke removal results of the basic U-NET model; the third image shows the smoke removal results of the CBAM + U-NET model; the fourth image shows the smoke removal results of the Laplacian image pyramid fusion + U-NET model; and the fifth image shows the smoke removal results of the CBAM + Laplacian image pyramid fusion + U-Net model.

To verify the effectiveness of this model, we compared the frames per second (fps) of this article with six other methods. As shown in Table 2, our 90.19 (Fps) is inferior to GAN. But it achieves the best results on two important metrics (PSNR and SSIM) in Figure 12. The requirements for clinical endoscopic surgery have been met.

Figure 12A shows the PSNR comparison between our method and the other six methods. Figure 12B shows the SSIM comparison between our method and the other six methods. Bolkar et al. (Bolkar

TABLE 2 Processing time comparison.

Methods	Model	Training images	Time (fps)	Platform
Bolkar et al.	CNN + DCP	Abdominal Cavity Images	32.40	Python(Caffe)
Chen et al.	CNN	Abdominal Cavity Images	89.14	Python(TensorFlow)
Shin et al.	physical method	Natural Images	1.28	Matlab
Wang et al.	U-Net	Abdominal Cavity Images	24.00	Python(Keras)
Isola et al.	GAN	Abdominal Cavity Images	120.0	Python(Pytorch)
Salazar et al.	GAN + DCP	Abdominal Cavity Images	92.19	Python(Pytorch)
Our Proposed method	U-Net + CBAM + Laplace	Abdominal Cavity Images	90.19	Python(TensorFlow)

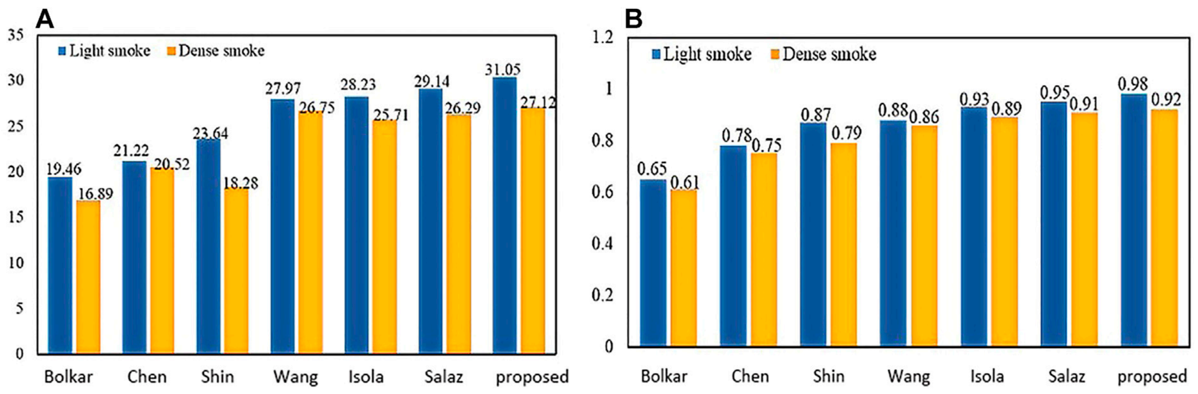


FIGURE 12
PSNR and SSIM comparison: **(A)** PSNR comparison; **(B)** SSIM comparison.

et al., 2018) derived the atmospheric diffusion model and implemented it with a neural network. It is an earlier classic method in the field of smoke purification, so the results of various indicators are lower compared with recent methods. Chen et al. used a synthetic dataset. The U-Net architecture is used to achieve smoke purification. Among the several methods compared, the time performance is better. But the purification effect on real smoke images is poor. Shin et al. (Shin et al., 2019) adopted the radiation reflectance optimization scheme. The processing speed of a single image is the slowest. Wang et al. adopted the U-Net architecture and improved the down-sampling part. Compared with the first three methods, the PSNR index is greatly improved. Isola et al. (2017) used the adversarial neural network method, and achieved the best results in time performance. Salazar et al. (Salazar-Colores et al., 2020) used an adversarial neural network and took the dark channel-detected image as input and achieved good performance in various indicators. We used PSNR and SSIM in PSNR and SSIM. The two indicators have achieved the best results among several methods. In terms of time performance, the time indicator can achieve a stable display playback without jitter, so it can be applied in real-time systems.

3.5 Three-dimensional model performance verification

There are few literature studies on disparity estimation of endoscopic images. The evaluation indicators are not unique. Basic (Ye et al., 2017) used DeConvNet as the basis of the model network and adopted a self-supervised scheme. The disparity image obtained by training endoscopic images and the original image are used as the comparison standard, taking the structural similarity SSIM as the indicator. ELAS (Geiger et al., 2010) triangulated the matching points of the binocular image, making the surrounding points easier to match. SPS (Yamaguchi et al., 2014) proposed a new target optimization algorithm to solve the occlusion problem. The algorithm preserved the connectivity of image segments and utilized shape regularization in the form of boundary lengths. The algorithm finally realized image segmentation and disparity estimation for natural scene images. Siamese (Xu et al., 2019) is a stereo-automatic encoding and decoding structure, which is similar to monocular. The input codec structure is Basic. The initial disparity image is obtained from the codec structure. Then,

TABLE 3 SSIM comparison.

Model	Basic	ELAS	SPS	Siamese	Our proposed
Mean SSIM	0.555	0.473	0.547	0.604	0.726
Std SSIM	0.106	0.079	0.092	0.106	0.085

TABLE 4 PSNR and SSIM comparison.

Model	Basic	Autoencoder	Our proposed
Mean SSIM	0.5414 ± 0.0709	0.8349 ± 0.0523	0.8826 ± 0.0678
Mean PSNR	7.7650 ± 1.3686	14.4957 ± 1.9676	17.2594 ± 1.6254

TABLE 5 MAE and RMSE comparison.

Model	Methods	MAE, mm	RMSE, mm
Heart 1	Godard et al	2.39 ± 0.62	2.99 ± 0.61
	Wang et al	2.16 ± 0.65	-
	Stoyanov et al	2.36 ± 0.92	3.88 ± 0.87
	Luo et al	1.84 ± 0.40	2.69 ± 0.58
	Our Proposed	1.65 ± 0.35	2.45 ± 0.52
Heart 2	Godard et al	1.79 ± 0.40	2.65 ± 0.28
	Wang et al	2.14 ± 0.83	-
	Stoyanov et al	3.20 ± 1.15	4.85 ± 1.82
	Luo et al	1.49 ± 0.41	1.90 ± 0.38
	Our Proposed	1.45 ± 0.40	1.62 ± 0.42

the virtual view is obtained by the STN network. The loss is obtained by comparing the difference between the real view and the virtual view. One layer gets suitable parameters. Compared with the method proposed, the Siamese results obtained by binocular images are better than the Basic results obtained by monocular images. The SSIM effect reaches 0.726 ± 0.085, which is better than the Siamese results as shown in Table 3:

The parallax image obtained by SLAM is the true value. Using SSIM and PSNR as standards, we compared the predicted parallax value with the true value. The results are shown in Table 4. Our proposed average SSIM and PSNR results were 0.8826 ± 0.0678 and 17.2594 ± 1.6254, respectively. The results showed that the proposed method is superior to other methods.

The experiments use the binocular heart data in the Hamlin endoscopy dataset. This dataset originally did not contain ground truth disparity values. Several algorithms are compared in Table 5. Godard et al. obtained the disparity image by extracting image features through CNN in the natural scene dataset. The parallax information from the left image to the right image is imaged to obtain the virtual view. The loss value is obtained by comparing the virtual view with the real view. The model results obtained from this training perform well on natural scene datasets. Wang et al. (Wang et al., 2018b) used variational disparity estimation technology to minimize the global energy function of the entire image. Based on the grayscale and gradient constants, they supposed that a data term and a local and non-local smoothing term were defined to construct the cost function. The real disparity image was obtained. Stoyanov et al. (Stoyanov et al., 2010) and Luo et al. (Luo et al., 2019) used two encoders and decoders to extract the disparity images for the left and right images, respectively. They used the traditional binocular algorithm AD-CENSUS to generate unsupervised training. The surrogate disparity labels, which guide the training process, achieved better results than the previous two literature studies on both MAE and RMSE metrics. This article compares the results with the aforementioned four methods. From the experimental results, we find that our result has a certain improvement in MAE. The RMSE index has a larger improvement than the aforementioned methods.

In endoscopic image evaluation, the doctor's subjective evaluation is still the important method to verify the image quality. The establishment of the quantitative assessment is a challenging task since there are no available gold standards. More specialized evaluations are needed to validate the effectiveness of 3D reconstruction methods for endoscopic images. Therefore, we invited 10 chief physicians from the Affiliated Hospital of Southwest Medical University with more than 5 years of laparoscopic surgery experience to score the 3D images. The subjective evaluation criteria referring to the Double Stimulus Continuous Scale (DSCQS), 3D effect, viewing comfort, and acceptability were rated on a scale of 1 (worst) to 5 (best). A score of 1 indicated a non-diagnostic image, and a score of 5 indicated an excellent diagnostic image quality. Pathological invariance was scored with 0 (change) or 1 (no change).

The evaluation of the 10 clinicians is shown in Table 6; our method received the best subjective quality evaluation. 3D effect, viewing comfort, and acceptability are better than other methods, and the improved loss function can better retain the details of medical images. It is proved that the proposed 3D reconstruction algorithm can be applied to clinical scenarios.

In this article, a total of 1,200 endoscopic images in the dataset were processed. The average processing time per image is 0.0275 s and 36fps was obtained. Therefore, it can meet the real-time requirements of a 3D display system.

The improved U-NET network applied to an original endoscopic image can obtain a better parallax image with higher accuracy to obtain a better three-dimensional display effect. Moreover, void-filling and reverse imaging can be performed on the parallax image to recover a better right view, color offset can be performed on the left and right views, and the 3D display effect can be seen by wearing red and blue lenses, as shown in Figure 13.

We validated the effectiveness of our method on the binocular laparoscopy dataset. For any image on the binocular laparoscopic dataset, an adaptive neural network endoscopic three-dimensional reconstruction method is proposed. If there is smoke, first use the

TABLE 6 Subjective evaluation and comparison of the 3D reconstruction effect. (Mean ± STANDARD, deviation).

Methods	3D Effect	Viewing comfort	Invariance	Acceptability
Godard[27]	3.8 ± 0.48	3.5 ± 0.11	0.3 ± 0.15	3.2 ± 0.21
Wang[40]	3.7 ± 0.72	3.6 ± 0.82	0.3 ± 0.55	3.5 ± 0.74
Stoyanov[41]	3.9 ± 0.91	3.8 ± 0.25	0.4 ± 0.51	3.8 ± 0.11
Luo[42]	4.1 ± 0.40	4.0 ± 0.35	0.6 ± 0.51	4.1 ± 0.63
Proposed	4.1 ± 0.69	4.1 ± 0.11	0.7 ± 0.12	4.2 ± 0.38

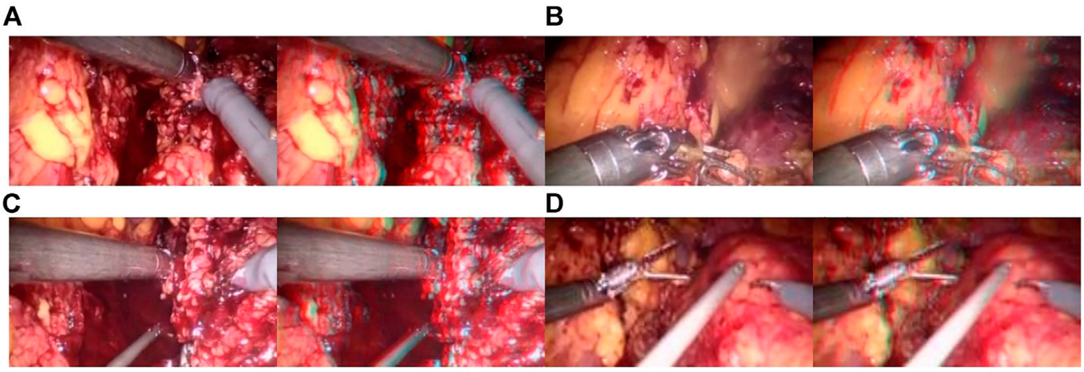


FIGURE 13
Red and blue 3D display images. The left image in each sub-image is the original view, and the right image is the chromatic 3D display image. **(A–D)** are randomly selected from the experimental results.

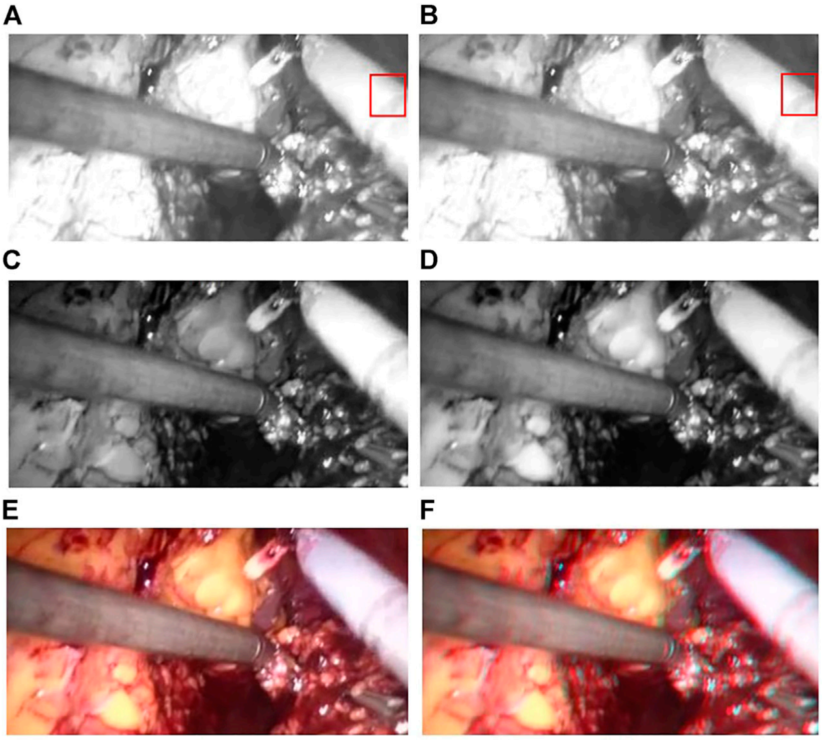


FIGURE 14
3D display of color difference; **(A)** Red Component; **(B)** Offset Red Component; **(C)** Blue Component; **(D)** Green Component; **(E)** Raw endoscope image; **(F)** 3D Display.

smoke purification algorithm to obtain the purified image, and secondly obtain the disparity image. The result of 3D display is shown in Figure 14. Figure 14A represents the red component of the original image. Figure 14B represents the red component after fusion of parallax. There is a slight difference between Figure 14A and Figure 14B. It is difficult to observe with the naked eye. We need to carefully observe the slight difference between the red boxes on the right side of the image. There are certain wrinkles in Figure 14B. It shows that the red component has moved after parallax stacking. Figure 14B is a virtual image from another viewpoint. Figure 14C and Figure 14D represent the blue and green components separated from the original image, respectively, and Figure 14E represents the original image. The RGB images are shown in Figure 14A, Figure 14C, and Figure 14D, respectively. Figure 14F represents the color-difference three-dimensional display image. From Figure 14, we can find that the red–blue parallax movement range becomes larger, which is more suitable for human eye observation.

4 Conclusion

To meet the practical application requirements of binocular endoscopic medical images, this article organically combines a global expansion with a local adaptive expansion of the network structure. Aiming at the lack of real parallax in unsupervised binocular endoscopic images, we proposed a 3D reconstruction scheme for adaptively processing the smoke images. Subjective evaluation and objective evaluation were used for verification. The 3D effects in the subjective evaluation obtained an optimal value of 4.2 ± 0.38. In the de-hazing tests on real datasets, our method achieved an SSIM of 0.980, a PSNR of 31.545 dB, an average running speed of 90.191 fps, and a much lower training time than similar methods. The proposed self-supervised disparity estimation method also outperformed the existing methods, with an SSIM of 0.726 ± 0.085 and a PSNR of 17.2594 ± 1.6254 dB; MAE 1.45 ± 0.40, RMSE 1.62 ± 0.42. It meets the needs of medical images in various indicators and solves the real-time problem of clinical operations. The present article can therefore guide the development of endoscopy devices.

Data availability statement

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

Author contributions

GZ: Investigation, Methodology, Software, Validation, Visualization, and Writing—original draft. ZH, EC, ZL, and JL: Investigation, Methodology, Software, and Supervision. YP and WS: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, and Writing—review and editing.

Funding

This work was supported by the Doctoral Innovative Talents Project of Chongqing University of Posts and Telecommunications ([BYJS202107). Chongqing Natural Science Foundation of China [grant number cstc2021jcyj-bsh0218]; The National Natural Science Foundation of China (Grant No. U21A20447 and 61971079); The Basic Research and Frontier Exploration Project of Chongqing (Grant No. cstc2019jcyjmsxmX0666); Chongqing technological innovation and application development project(cstc2021jscx-gksbx0051); The Innovative Group Project of the National Natural Science Foundation of Chongqing (Grant No. cstc2020jcyj-cxttX0002), and the Regional Creative Cooperation Program of Sichuan (2020YFQ0025); The Science and Technology Research Program of Chongqing Municipal Education Commission(KJZD-k202000604).

Acknowledgments

We would like to thank Smart Medical System and Core Technology Laboratory.

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.2022.994343/full#supplementary-material

References

Antal, B. (2016). Automatic 3d point set reconstruction from stereo laparoscopic images using deep neural networks. Available at: http://arXiv.org/abs/1608.00203. doi:10.48550/arXiv.1608.00203

Besse, F., Rother, C., Fitzgibbon, A., and Kautz, J. (2014). Pmbp: Patchmatch belief propagation for correspondence field estimation. *Int. J. Comput. Vis.* 110 (1), 2–13. doi:10.1007/s11263-013-0653-9

Bolkar, S., Wang, C., Cheikh, F. A., and Yildirim, S. (2018). "Deep smoke removal from minimally invasive surgery videos," in Proceeding of the 2018 25th IEEE International Conference on Image Processing, Athens Greece: ICIP, 3403–3407. doi:10.1109/ICIP.2018.8451815

Chen, L., Bentley, P., Mori, K., Misawa, K., Fujiwara, M., and Rueckert, D. (2019a). Self-supervised learning for medical image analysis using image context restoration. *Med. Image Anal.* 58, 101539. doi:10.1016/j.media.2019.101539

Chen, L., Tang, W., and John, N. W. (2017). Real-time geometry-aware augmented reality in minimally invasive surgery. *Healthc. Technol. Lett.* 4 (5), 163–167. doi:10.1049/htl.2017.0068

Chen, L., Tang, W., John, N. W., Wan, T. R., and Zhang, J. J. (2019b). De-smokeGCN: Generative cooperative networks for joint surgical smoke detection and removal. *IEEE Trans. Med. Imaging* 39 (5), 1615–1625. doi:10.1109/TMI.2019.2953717

Chen, X., Wang, X., Zhang, K., Fung, K. M., Thai, T. C., Moore, K., et al. (2022). Recent advances and clinical applications of deep learning in medical image analysis. *Med. Image Anal.* 79, 102444. doi:10.1016/j.media.2022.102444

Feng, Y., Liang, Z., and Liu, H. (2017). "Efficient deep learning for stereo matching with larger image patches," in Proceeding of the 2017 10th International Congress on Image and Signal Processing BioMedical Engineering and Informatics (CISP-BMEI), Shanghai, China: IEEE, 1–5. doi:10.1109/CISP-BMEI.2017.8301999

Garg, R., Bg, V. K., Carneiro, G., and Reid, I. (2016). Unsupervised cnn for single view depth estimation: Geometry to the rescue. *Eur. Conf. Comput. Vis.* 9912, 740–756. doi:10.1007/978-3-319-46484-8_45

Geiger, A., Roser, M., and Urtasun, R. (2010). "Efficient large-scale stereo matching," in Proceeding of the Asian Conference on Computer Vision, Berlin Heidelberg: Springer, 25–38. doi:10.1007/978-3-642-19315-6_3

Godard, C., Mac Aodha, O., and Brostow, G. J. (2017). "Unsupervised monocular depth estimation with left-right consistency," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu HI, USA: IEEE, 270–279. doi:10.1109/CVPR.2017.699

Godard, C., Mac Aodha, O., Firman, M., and Brostow, G. J. (2019). "Digging into self-supervised monocular depth estimation," in Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul Korea: IEEE, 3828–3838. doi:10.48550/arXiv.1806.01260

Hu, M., Penney, G., Figl, M., Edwards, P., Bello, F., Casula, R., et al. (2012). "Reconstruction of a 3D surface from video that is robust to missing data and outliers: Application to minimally invasive surgery using stereo and mono endoscopes. *Med. Image Anal.* 16, 597–611. doi:10.1016/j.media.2010.11.002

Huang, P. H., Matzen, K., Kopf, J., Ahuja, N., and Huang, J. B. (2018). "Deepmvs: Learning multi-view stereopsis," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Salt Lake City UT USA: IEEE, 2821–2830. doi:10.1109/CVPR.2018.00298

Isola, P., Zhu, J., Zhou, T., and Efros, A. A. (2017). "Image-to-Image Translation with Conditional Adversarial Networks," in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA: IEEE, 5967–5976. doi:10.1109/CVPR.2017.632

Jaderberg, M., Simonyan, K., and Zisserman, A. (2015). Spatial transformer networks." in Proceedings of the 28th International Conference on Neural Information Processing Systems 2, 2017–2025. doi:10.5555/2969442.2969465

Kendall, A., Martirosyan, H., Dasgupta, S., Henry, P., Kennedy, R., Bachrach, A., et al. (2017). End-to-end learning of geometry and context for deep stereo regression," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 66–75. doi:10.1109/ICCV.2017.17

Kotwal, A., Bhalodia, A., and Awate, S. P. (2016). "Joint desmoking and denoising of laparoscopy images," in Proceedings of the IEEE 13th International Symposium on Biomedical Imaging (ISBI), Prague Czech Republic: IEEE, 1050–1054. doi:10.1109/ISBI.2016.7493446

Li, T., Cheng, B., Ni, B., Liu, G., and Yan, S. (2016a). Multitask low-rank affinity graph for image segmentation and image annotation. *ACM Trans. Intell. Syst. Technol.* 7 (4), 1–18. doi:10.1145/2856058

Li, T., Meng, Z., Ni, B., Shen, J., and Wang, M. (2016b). Robust geometric ℓp-norm feature pooling for image classification and action recognition. *Image Vis. comput.* 55, 64–76. doi:10.1016/j.imavis.2016.04.002

Li, T., Wang, Y., Hong, R., Wang, M., and Wu, X. (2018). pDisVPL: probabilistic discriminative visual part learning for image classification. *IEEE Multimed.* 25 (4), 34–45. doi:10.1109/MMUL.2018.2873499

Luo, H., Hu, Q., and Jia, F. (2019). Details preserved unsupervised depth estimation by fusing traditional stereo knowledge from laparoscopic images. *Healthc. Technol. Lett.* 6 (6), 154–158. doi:10.1049/htl.2019.0063

Luo, W., Schwing, A. G., and Urtasun, R. (2016). Efficient deep learning for stereo matching. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 5695–5703. doi:10.1109/CVPR.2016.614

Mahmoud, N., Cirauqui, I., Hostettler, A., Doignon, C., Soler, L., Marescaux, J., et al. (2016). ORBSLAM-based endoscope tracking and 3D reconstruction. *Int. Workshop Computer-assisted Robotic Endosc.* 10170, 72–83. doi:10.1007/978-3-319-54057-3_7

Pang, J., Sun, W., Ren, J. S., Yang, C., and Yan, Q. (2017). "Cascade residual learning: A two-stage convolutional neural network for stereo matching," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Venice Italy: IEEE, 887–895. doi:10.1109/ICCVW.2017.108

Penza, V., Ciullo, A. S., Moccia, S., Mattos, L. S., and De Momi, E. (2018). Endoabs dataset: endoscopic abdominal stereo image dataset for benchmarking 3d stereo reconstruction algorithms. *Int. J. Med. Robot.* 14 (5), e1926. doi:10.1002/rcs.1926

Penza, V., Ortiz, J., Mattos, L. S., Forgione, A., and De Momi, E. (2016). Dense soft tissue 3D reconstruction refined with super-pixel segmentation for robotic abdominal surgery. *Int. J. Comput. Assist. Radiol. Surg.* 11 (2), 197–206. doi:10.1007/s11548-015-1276-0

Qiu, L., and Ren, H. (2020). Endoscope navigation with SLAM-based registration to computed tomography for transoral surgery. *Int. J. Intell. Robot. Appl.* 4 (2), 252–263. doi:10.1007/s41315-020-00127-2

Salazar-Colores, S., Jiménez, H. M., Ortiz-Echeverri, C. J., and Flores, G. (2020). Desmoking laparoscopy surgery images using an image-to-image translation guided by an embedded dark channel. *IEEE Access* 8, 208898–208909. doi:10.1109/ACCESS.2020.3038437

Shin, J., Kim, M., Paik, J., and Lee, S. (2019). Radiance–reflectance combined optimization and structure-guided $\ell _0$-Norm for single image dehazing. *IEEE Trans. Multimed.* 22 (1), 30–44. doi:10.1109/TMM.2019.2922127

Shurrab, S., and Duwairi, R. (2022). Self-supervised learning methods and applications in medical imaging analysis: a survey. *PeerJ Comput. Sci.* 8, e1045. doi:10.7717/peerj-cs.1045

Sidorov, O., Wang, C., and Cheikh, F. A. (2020). Generative smoke removal. machine learning for health workshop. *PMLR* 116, 81–92. doi:10.48550/arXiv.1902.00311

Stoyanov, D., Scarzanella, M. V., Pratt, P., and Yang, G. Z. (2010). "Real-time stereo reconstruction in robotically assisted minimally invasive surgery," in Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Berlin Heidelberg: Springer, 275–282. doi:10.1007/978-3-642-15705-9_34

Taleb, A., Lippert, C., Klein, T., and Nabi, M. (2021). Multimodal self-supervised learning for medical image analysis. *Int. Conf. Inf. Process. Med. Imaging* 12729, 661–673. doi:10.1007/978-3-030-78191-0_51

Tosi, F., Aleotti, F., Poggi, M., and Mattoccia, S. (2019). "Learning monocular depth estimation infusing traditional stereo knowledge," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach CA USA: IEEE, 9799–9809. doi:10.1109/CVPR.2019.01003

Venkatesh, V., Sharma, N., Srivastava, V., and Singh, M. (2020). Unsupervised smoke to desmoked laparoscopic surgery images using contrast driven Cyclic-DesmokeGAN. *Comput. Biol. Med.* 123, 103873. doi:10.1016/j.compbiomed.2020.103873

Vijayanarasimhan, S., Ricco, S., Schmid, C., Sukthankar, R., and Fragkiadaki, K. (2017). Sfm-net: learning of structure and motion from video. Available at; http://arXiv.org/abs/1704.07804. doi:10.48550/arXiv.1704.07804

Wang, C., Alaya Cheikh, F., Kaaniche, M., Beghdadi, A., and Elle, O. J. (2018a). Variational based smoke removal in laparoscopic images. *Biomed. Eng. Online* 17 (1), 139–218. doi:10.1186/s12938-018-0590-5

Wang, C., Cheikh, F. A., Kaaniche, M., and Elle, O. J. (2018b). Liver surface reconstruction for image guided surgery. *Med. Imaging 2018 Image-Guided Proced. Robotic Interventions, Model.* 10576, 576–583. doi:10.1117/12.2297398

Wang, C., Mohammed, A. K., Cheikh, F. A., Beghdadi, A., and Elle, O. J. (2019a) ,Multiscale deep desmoking for laparoscopic surgery, *Med. Imaging 2019 Image Process*. 10949, 505–513. doi:10.1117/12.2507822

Wang, Y., Lai, Z., Huang, G., Wang, B. H., Van Der Maaten, L., Campbell, M., et al. (2019b). "Anytime stereo image depth estimation on mobile devices," in Proceedings of the International Conference on Robotics and Automation (ICRA), Montreal QC Canada: IEEE, 5893–5900. doi:10.1109/ICRA.2019.8794003

Woo, S., Park, J., Lee, J. Y., and Kweon, I. S. (2018). Cbam: convolutional block attention module. *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 3–19. doi:10.1007/978-3-030-01234-2_1

Xu, K., Chen, Z., and Jia, F. (2019). Unsupervised binocular depth prediction network for laparoscopic surgery. *Comput. Assist. Surg.* 24 (1), 30–35. doi:10.1080/24699322.2018.1557889

Yamaguchi, K., McAllester, D., and Urtasun, R. (2014). Efficient joint segmentation, occlusion labeling, stereo and flow estimation. *Eur. Conf. Comput. Vis.* 8693, 756–771. doi:10.1007/978-3-319-10602-1_49

Yang, B., and Liu, C. (2014). Robust 3 D motion tracking for vision-based control in robotic heart surgery. *Asian J. Control* 16 (3), 632–645. doi:10.1002/asjc.785

Yang, D., and Sun, J. (2018). Proximal dehaze-net: a prior learning-based deep network for single image dehazing," in Proceedings of the European Conference on Computer Vision (ECCV), 729–746. doi:10.1007/978-3-030-01234-2_43

Ye, M., Johns, E., Handa, A., Zhang, L., Pratt, P., and Yang, G. Z. (2017). Self-supervised siamese learning on stereo image pairs for depth estimation in robotic surgery. Available at; http://arXiv.org/abs/1705.08260. doi:10.48550/arXiv.1705.08260

Yi, K. M., Trulls, E., Lepetit, V., and Fua, P. (2016). Lift: Learned invariant feature transform. *Eur. Conf. Comput. Vis. (ECCV)* 9910, 467–483. doi:10.1007/978-3-319-46466-4_28

Yin, Z., and Shi, J. (2018). "Geonet: Unsupervised learning of dense depth, optical flow and camera pose," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City UT USA: IEEE, 1983–1992. doi:10.1109/CVPR.2018.00212

Zhang, F., Prisacariu, V., Yang, R., and Torr, P. H. (2019). "Ga-net: Guided aggregation net for end-to-end stereo matching," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach CA USA: IEEE, 185–194. doi:10.1109/CVPR.2019.00027

Zhang, G., Lin, J., Cao, E., Pang, Y., and Sun, W. (2022). A medical endoscope image enhancement method based on improved weighted guided filtering. *Mathematics* 10 (9), 1423. doi:10.3390/math10091423

Zhao, S., Zhang, L., Shen, Y., Zhao, S., and Zhang, H. (2019). Super-resolution for monocular depth estimation with multi-scale sub-pixel convolutions and a smoothness constraint. *IEEE Access* 7, 16323–16335. doi:10.1109/ACCESS.2019.2894651

Zhou, J., Lu, Y., Tao, S., Cheng, X., and Huang, C. (2021). E-Res U-Net: an improved U-Net model for segmentation of muscle images. *Expert Syst. Appl.* 185, 115625. doi:10.1016/j.eswa.2021.115625

Zhou, T., Brown, M., Snavely, N., and Lowe, D. G. (2017). Unsupervised learning of depth and ego-motion from video." in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1851–1858. doi:10.1109/CVPR.2017.700

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 09 September 2022
DOI 10.3389/fphys.2022.950141

OPEN ACCESS

EDITED BY
Huan Qin,
South China Normal University, China

REVIEWED BY
Funeng Jiang,
South China University of Technology, China
Cui Li Gang,
Peking University Third Hospital, China
Li Junzheng,
Southern Medical University, China

*CORRESPONDENCE
Xinmin Guo,
Guoxm1509257@163.com

[†]These authors share first authorship

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

RECEIVED 23 May 2022
ACCEPTED 15 August 2022
PUBLISHED 09 September 2022

CITATION
Guo X, Huang Z, Chen J, He K, Lin J, Zhang H and Zeng Y (2022), Synergistic delivery of resveratrol and ultrasmall copper-based nanoparticles by aptamer-functionalized ultrasound nanobubbles for the treatment of nonalcoholic fatty liver disease.
Front. Physiol. 13:950141.
doi: 10.3389/fphys.2022.950141

Synergistic delivery of resveratrol and ultrasmall copper-based nanoparticles by aptamer-functionalized ultrasound nanobubbles for the treatment of nonalcoholic fatty liver disease

Xinmin Guo[1*†], Zhihui Huang[2†], Jialin Chen[1†], Kun He[1], Jianru Lin[1], Hui Zhang[1] and Yanying Zeng[1]

[1]Department of Ultrasound, Guangzhou Red Cross Hospital of Jinan University, Guangzhou, China, [2]Department of Nuclear Medicine, Guangzhou Red Cross Hospital of Jinan University, Guangzhou, China

Nonalcoholic fatty liver disease (NAFLD) is related to the production of reactive oxygen species (ROS) and oxidative stress, so antioxidant treatment can prevent its further development. Ultrasmall copper-based nanoparticles (CuNPs) have shown multiple enzyme-like activities for scavenging oxygen species, providing a new strategy for the treatment of inflammatory diseases. Resveratrol (Res), a natural polyphenol compound, has attracted much attention due to its ability to inhibit oxidative stress. We therefore aimed to first combine these two agents for the treatment of NAFLD. However, due to the poor water solubility and stability of Res, which is easily metabolized in the intestine, the development of a stable and effective carrier became the key to achieving a synergistic effect. Liver-targeted nanocarriers loaded with bioactive compounds may provide a more effective approach for the treatment of NAFLD. Therefore, we developed a novel ultrasonic nanobubble carrying nucleic acid aptamers with liver targeting properties, which has the advantages of a small molecular weight, no immunogenicity, a low cost of synthesis, and high stability through chemical modification. Res and the ultrasmall CuNPs were specifically delivered to liver tissue to maximize therapeutic efficiency. This study found that the combination of these two components can effectively treat inflammation in NAFLD and suggested that liver-targeted NAFLD-specific aptamer-mediated targeted ultrasound nanobubbles that can simultaneously deliver Res and CuNPs may be a safe and effective new platform for NAFLD and other liver diseases.

KEYWORDS

ultrasonic, nanobubble, resveratrol, ultrasmall, copper-based, nanoparticles, nonalcoholic fatty liver

1 Introduction

Nonalcoholic fatty liver disease (NAFLD) is now considered one of the most common chronic liver diseases, affecting approximately 25%–30% of people worldwide (Chachay et al., 2014; Berman et al., 2017). Its manifestations range from simple steatosis to liver injury, followed by liver fibrosis liver cirrhosis and liver cancer development, and a series of diseases, so early intervention has clinical significance (Kutlu et al., 2018). Recent studies have shown that NAFLD is associated with excessive production of reactive oxygen species (ROS) and oxidative stress, and hepatocytes are eventually damaged by oxidative stress and lipid peroxidation (Schwimmer, 2007; Lomonaco et al., 2013; Braud et al., 2017; Kasper et al., 2021; Ding et al., 2022).

At present, there is still a lack of effective therapeutic drugs for NAFLD in the clinic. From the perspective of pathogenic factors, the use of antioxidants can regulate adipogenesis, lipid oxidation and peroxidation, and inflammation. By improving the oxidative environment, liver damage and fibrosis may be improved, ultimately delaying, preventing, and reversing the progression of nonalcoholic steatohepatitis (NASH) and improving clinical outcomes. Many studies have shown that resveratrol (Res) (Abba et al., 2015; Rauf et al., 2017; Öztürk et al., 2017) has pharmacological effects on the nervous system, liver, and cardiovascular system and can reduce liver pathological damage by acting as an antioxidative stress factor. Recent research showed that ultrasmall copper-based nanoparticles have a variety of enzymatic activities and the ability to scavenge a variety of ROS and treat ROS-related diseases. At the same time, their ultrasmall size ensures that their strong ROS scavenging ability will reach any organ and might be rapidly subjected to renal clearance, avoiding damage to normal tissues. The synergy between these two drugs is expected to exert a strong antioxidative stress effect, but Res has poor water solubility and is not stable and easily metabolized by in intestinal tract, which limits its further clinical applications. Therefore, how to use an effective carrier to improve its stability and bioavailability is a current research hotspot (Francioso et al., 2014; Zupančič et al., 2015). In recent years, nanobased drug delivery systems have become good drug and gene carrier systems because of their stability and novel physical properties; they can quickly reach the target tissue through the vascular endothelial system. Some scholars reported that the use of poly(lactic-co-glycolic acid) (PLGA) nanoparticles as a carrier system for Res successfully achieved curative effects for nonalcoholic fatty liver (Wan et al., 2018).

Ultrasonic nanobubble-based drug delivery systems have attracted the attention of researchers because of their small size, excellent stability, and novel physical and surface properties, which allow them to quickly pass through the vascular endothelium and enter the target tissue for aggregation (Suzuki et al., 2016; Endo-Takahashi and Negishi, 2020; Sun et al., 2020). Although this type of nanobubble has the ability to achieve ultrasonic targeted release, the targeting specificity needs to be improved. At present, targeting is mainly achieved through the modification of specific targeting groups, such as antibodies, folic acid, and aptamers (Zhu et al., 2018; Shen et al., 2019; Yu et al., 2020). Nucleic acid aptamers have attracted much attention because of their unique performance advantages, such specific affinity equal to or even higher than that of antibodies, high stability, and high cost-effectiveness for mass production. Some scholars have constructed aptamer-modified nanobubbles for ultrasound imaging of various tumor cells (Zhu et al., 2018). To date, there are no reports on the synthesis of aptamer-modified nanobubbles loaded with drugs and nanoparticles for the treatment of NAFLD.

In this context, this study aimed to synthesize a nanobubble using thin-film hydration and mechanical oscillation, and the nanobubbles were modified with NAFLD-specific aptamers and loaded with Res and ultrasmall copper-based nanoparticles for ultrasonic targeted therapy for NAFLD. This study is expected to provide new ideas and a basis for the synergistic treatment of NAFLD with aptamer-functionalized nanobubbles for the delivery of drugs and nanoparticles.

2 Materials and methods

2.1 Main materials and instruments

The HepG2 human hepatoma cell line was purchased from the American Type Culture Collection (ATCC), and anhydrous copper chloride, L-ascorbic acid and sodium hydroxide were purchased from Sinopharm Group. Perfluoropropane was purchased from Sichuan 273 Degrees Environmental Protection Technology Co., Ltd.; DSPE-mPEG2K, DSPE-mPEGK-Mal and DOTAP were purchased from Qiyue Biological Co., Ltd.; Res was purchased from McLean; and lecithin was purchased from Aladdin. PBS was obtained from Gibco. The sulfhydryl-modified NAFLD01 aptamer sequence was purchased from Shanghai Shenggong Biological Co., Ltd. Other reagents not specifically mentioned were analytically pure. The magnetic stirrer (Gongyi Yuhua Instrument Co., Ltd.), centrifuge (Pingfan Technology Co., Ltd.), rotary evaporator (Changchun Technology Co., Ltd.), cell breaker (Ningbo Xinzhi Bio), multifunctional microplate reader (Tecan), laser nanoparticle sizer (Brookhaven), and ultraviolet absorber (Shanghai Jinghua Technology Co., Ltd.), X-ray diffraction (XRD) diffractometer, Rigaku D/max 2550) used in the study were purchased from Rigaku Company of Japan. A transmission electron microscope (FEI Company of America), X-ray photoelectron spectrometer (Thermo Fisher, ESCALAB 250Xi), fluorescence spectrometer (Hitachi, Japan), and inductively coupled plasma atomic emission spectrometer (ICP–AES, 7000DV, PerkinElmer) were also used.

2.2 Experimental methods

2.2.1 Establishment of a cell model of nonalcoholic fatty liver induced by high-oil fatty acids

The human hepatocellular carcinoma cell line HepG2 was obtained from the ATCC. Cells were cultured in Dulbecco's modified Eagle's medium containing 10% fetal bovine serum without penicillin–streptomycin. HepG2 cells were cultured with free fatty acid (FFA) DMEM for approximately 48 h. FFA components were composed of palmitic acid and oleic acid (Sigma-Aldrich) at a ratio of 1:2 and diluted with FBS-free DMEM to a final concentration of 4 mM. The cells were stained with oil red O to confirm fat deposition (Pu et al., 2021).

2.2.2 Evaluation of the targeted binding ability of aptamers in the context of nonalcoholic fatty liver disease

2.2.2.1 Laser confocal microscopy experiment

The groups were as follows. There were five groups in total: ① NAFLD cells + NAFLD01 aptamer sequence; ② NAFLD cells + nonsense aptamer sequence; ③ HepG2 cells + NAFLD01 aptamer sequence; ④ human normal hepatocytes + NAFLD01 aptamer sequence; and ⑤ human embryonic fibroblasts + NAFLD01 aptamer sequence. The methods were as follows. HepG2 cells were cultured in accordance with the above methods. First, oil red O staining was performed to confirm fat deposition. Then, NAFLD01 aptamer and nonsense aptamer sequences labeled with 250 nM FITC were incubated with NAFLD cells or HepG2 cells (number of cells: $1{\times}10^6$ cells) in 1 ml of binding buffer (BB) at 4°C for 1 h, stained with DAPI, and finally washed with washing buffer three times, and immunofluorescence images were captured.

2.2.2.2 Flow cytometry

The experimental groupings were the same as that described in Section 2.2.2.1. The methods were as follows. All cells ($5{\times}10^5$) were incubated with NAFLD01 labeled with FITC at 250 nM and BB at 200 µl at 4°C for 1 h. FITC-labeled nonsense sequences were used as controls. After incubation, the unbound adapter was washed away, and the cells were suspended in 500 µl WB. An Attune NxT flow cytometer was used to measure the fluorescence intensity.

2.2.3 Synthesis and characterization of aptamer-modified ultrasonic nanobubbles loaded with resveratrol and ultrasmall copper-based nanoparticles

2.2.3.1 Synthesis of $Cu_{5.4}O$ NPS

$CuCl_2$ (10 mM) was dissolved in 50 ml of deionized water and stirred magnetically for 10 min in an 80°C oil bath. Then, l-ascorbic acid (100 mM, 50 ml) was slowly added. After that, the pH was adjusted to 8.0–9.0 with sodium hydroxide, and the mixture was stirred at 80°C for 12 h. At the end of the reaction, the large aggregated particles were removed by centrifugation (6,577 × g, 15 min), and the supernatant was dialyzed for 2 days (molecular weight: 10,000 Da) to remove small molecules. Purification of $Cu_{5.4}O$ NPs was concentrated by centrifugation.

2.2.3.2 Synthesis of ultrasonic nanobubbles

DPPC and DSPE-PEG (2000) were dissolved in 2 ml chloroform at a mass ratio of 10:4. The mixture was transferred to a 25 ml glass flask with a rotating evaporator. A dry lipid film was produced in a commercially available rotary evaporator at 55°C and 130 RPM for 10 min by rotary evaporation. The phospholipid mixture was dried in 1.5 ml of hydrated solution (10% glycerol and 90% 1 × PBS, V/V). Then, the flask was placed in a culture shaker at 37°C and 120 RPM for 60 min to prepare the liposome membrane suspension. Subsequently, the liposome suspension was placed in a vial sealed with a rubber cap. A 50-ml syringe with a long needle was used to extract the gas in the vial, and then C_3F_8 gas (the bubble core was formed later) was added to the vial until the pressure in the vial was balanced. Finally, the vials were placed in a mechanical vibrator and subjected to mechanical vibration for 90 s. The stock suspension of ultrasonic nanobubbles was immediately placed on ice. The whole process was performed in the absence of light.

2.2.3.3 Aptamer-functionalized nanobubbles

1) DSPE-PEG (2000) was replaced with DSPE-PEG (2000)-Mal during the synthesis of the nanobubbles. 2) Next, 10 mM EDTA, 10 mM Tris (2-carboxyethyl) phosphine hydrochloride and 2 mM NAFLD01 aptamer were reacted at 37°C for 1 h, and then nanobubbles were added and incubated for 2 h. Then, the aptamer-modified nanobubbles floating in the upper layer of the suspension were collected by washing and centrifugation with cold PBS solution three times (300 RPM, 3 min). 3) The NBs@Res/Apt-NBs@Res preparation process was the same as that described for the nanobubbles/Apt-NBs, and Res was added at to a mass ratio of 10:4:0.3. 4) The NBs@$Cu_{5.4}O$ NPs/Apt-NBs@$Cu_{5.4}O$ preparation process was the same as that described for the nanobubble/Apt-NBs, with $Cu_{5.4}O$ addition at a 10:4:0.3 mass ratio. 5) The NBs@Res@$Cu_{5.4}O$ NPs/Apt-NBs@Res@$Cu_{5.4}O$ preparation process was the same as that described for nanobubbles/Apt-NBs; the mass ratio was set to 10:4:0.3:0.3, and Res (1 mg/ml) and $Cu_{5.4}O$ NPs were added at the same time (1 mg/ml, PBS). The preparation process of Apt-NBs@Res@$Cu_{5.4}O$ NPs is shown in Scheme 1. 6) Transmission electron microscopy (TEM), XRD, X-ray photoelectron spectroscopy (XPS), UV–VIS spectroscopy, energy-dispersive X-ray spectroscopy (EDS) and other approaches were used for characterization.

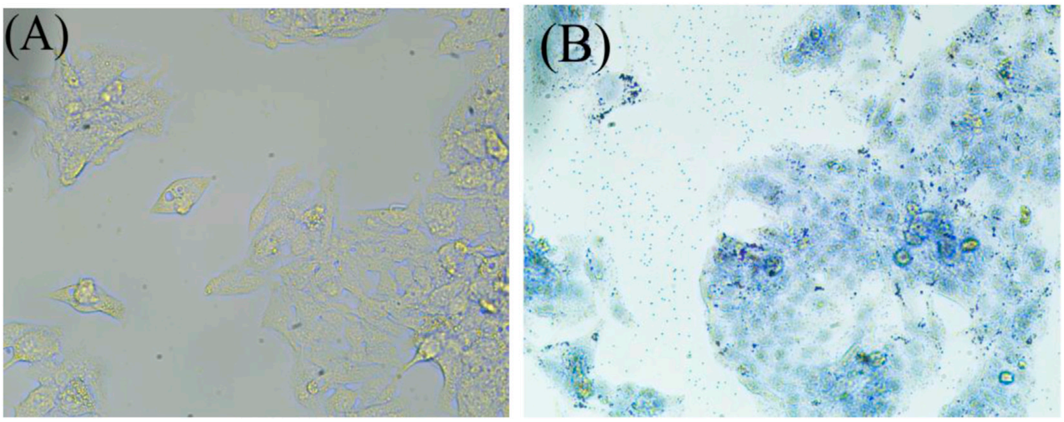


FIGURE 1
Oil red before staining **(A)**, and Oil red after staining **(B)**. The NAFLD cell model was established by oil red O staining. After being stained with oil red O and observed under a microscope with magnification of ×40, obvious lipid droplets were formed.

2.2.4 Aptamer-modified ultrasonic nanobubbles loaded with resveratrol and ultrasmall copper-based nanoparticles can be used for effective treatment of nonalcoholic fatty liver disease

2.2.4.1 CCK8 analysis of the of Apt-NBs@Res@ $Cu_{5.4}O$ NPs

HepG2 cells (1×10^3) were inoculated into a 96-well plate and cultured in DMEM complete medium containing 620, 310, 155 or 77.5 nM N-bromosuccinimide (NBS) (the initial concentration of NBS was 620 nM) for 48 h. Then, the cells were washed with PBS 3 times and incubated with 10 µl CCK-8 reagent (Bimake) for 2 h. Finally, the cell viability was measured with a microplate reader (BioTek) at an absorption wavelength of 450 nm. The experiment was repeated 3 times, and care was taken to avoid light during the experiment.

2.2.4.2 Aptamer-modified ultrasonic nanoparticles loaded with resveratrol and ultrasmall copper-based nanoparticles for the treatment of nonalcoholic fatty liver

The experimental groups were as follows: 1) HepG2 cells + ultrasound; 2) HepG2 cells + high oil fatty acid + ultrasound; 3) HepG2 cells + NBs + high-oil fatty acid + ultrasound; 4) HepG2 cells + NBs@Res@$Cu_{5.4}$ONPs + high-oil fatty acid + ultrasound; 5) HepG2 cells + Apt-NBs@Res@$Cu_{5.4}$ONPs + ultrasound. The HepG2 cells were placed in a 6-well plate, and high-oil fatty acidscontaining different nanobubbles were added (the maximum concentration of the nanobubbles without notable cytotoxicity was selected as the concentration for administration), and after incubation for 24 h, ultrasound irradiation (WED-100 ultrasound) was performed. After that, the cells were washed with cold PBS and collected. The expression of TNF-α and IL-10 was detected by Western blotting, and the level of total triglycerides (TGs) was determined by ELISA. The concentration of Res in each group was 90.6 mM (initial concentration), and the experiment was performed in triplicate.

2.3 Statistical analysis

SPSS 22.0 was used for statistical analysis. All data are expressed as mean ± SD, and one-way ANOVA was used for comparisons among groups. The LSD method was used for homogeneity of variance, and the Welch method or Brown-Forsythe method was used for correction of homogeneity of variance. Dunnett's T3 method was used for comparisons between groups. $p < 0.05$ indicated a statistically significant difference.

3 Results and discussion

NAFLD is associated with the production of ROS and oxidative stress (Schwimmer, 2007; Lomonaco et al., 2013; Braud et al., 2017; Kasper et al., 2021; Ding et al., 2022). The clinical use of broad-spectrum antioxidants such as N-acetylcysteine and acetyl-L-carnitine, which are currently used to clear ROS, is limited due to their poor bioavailability, low stability and low efficacy. One promising strategy is to develop nanozymes to maintain the natural redox balance in biological systems. CuNPs have a wide range of enzyme-like activities, the ability to clear ROS and ameliorate ROS-related diseases, excellent biocompatibility and high renal clearance. They have been shown to be effective in the treatment of a wide range of ROS-related diseases and to have no significant toxic effects (Liu et al., 2020). In addition, CuNPs are currently a research hotspot because they are easy to combine with other

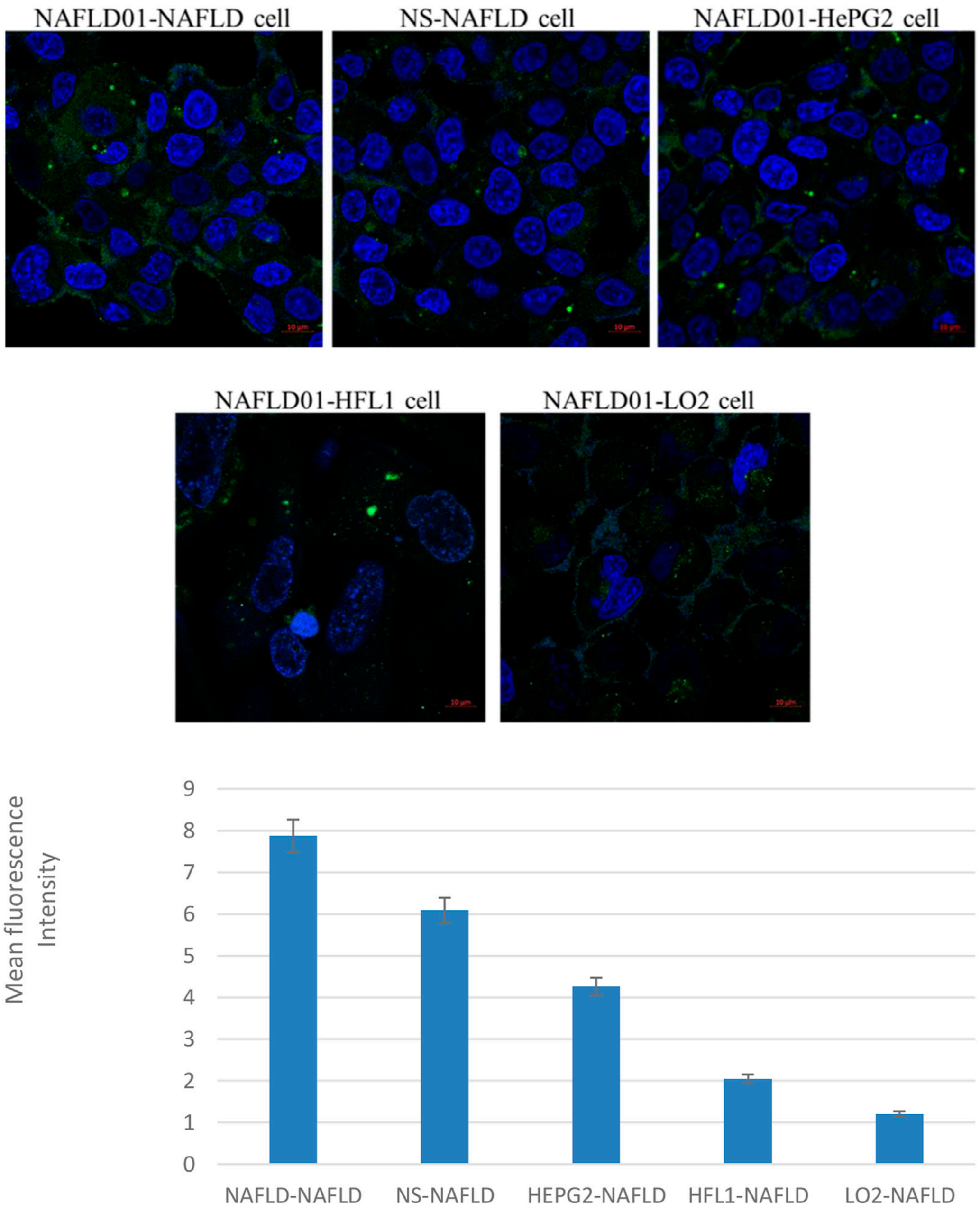


FIGURE 2
Confocal fluorescence verifies that aptamers can target NAFLD cells. Compared to HePG2 cells, LO2 cells, and HFL1 cells, NAFLD cells showed stronger green fluorescence signals. In addition, stronger green fluorescent signals were also shown around NAFLD01-NAFLD cells compared to NS-NAFLD cells. Quantification analysis of fluorescence intensity of the five groups (*$p < 0.05$).

polymers and show better stability than other nanoparticle polymers, with a large surface area and a unique morphology. Res has been widely studied by researchers because of its inhibitory effect on oxidative stress. To solve the problem of Res's poor water solubility, instability and ease of metabolism in the intestinal tract, Wan et al. (2018) developed PLGA nanoparticle-based RSV-PLGA-NPs that can be loaded with Res for successful treatment of NAFLD. The authors concluded that RSV-PLGA NPs can release RSV continuously and have high stability, water solubility and nontargeting properties. Teng et al. (2019) used liver-targeted oxidized starch lysozyme (OSL) as a nanocarrier to target the liver with covalently coupled galactose (Gal), which can be recognized by specifically expressed salivary glycoprotein

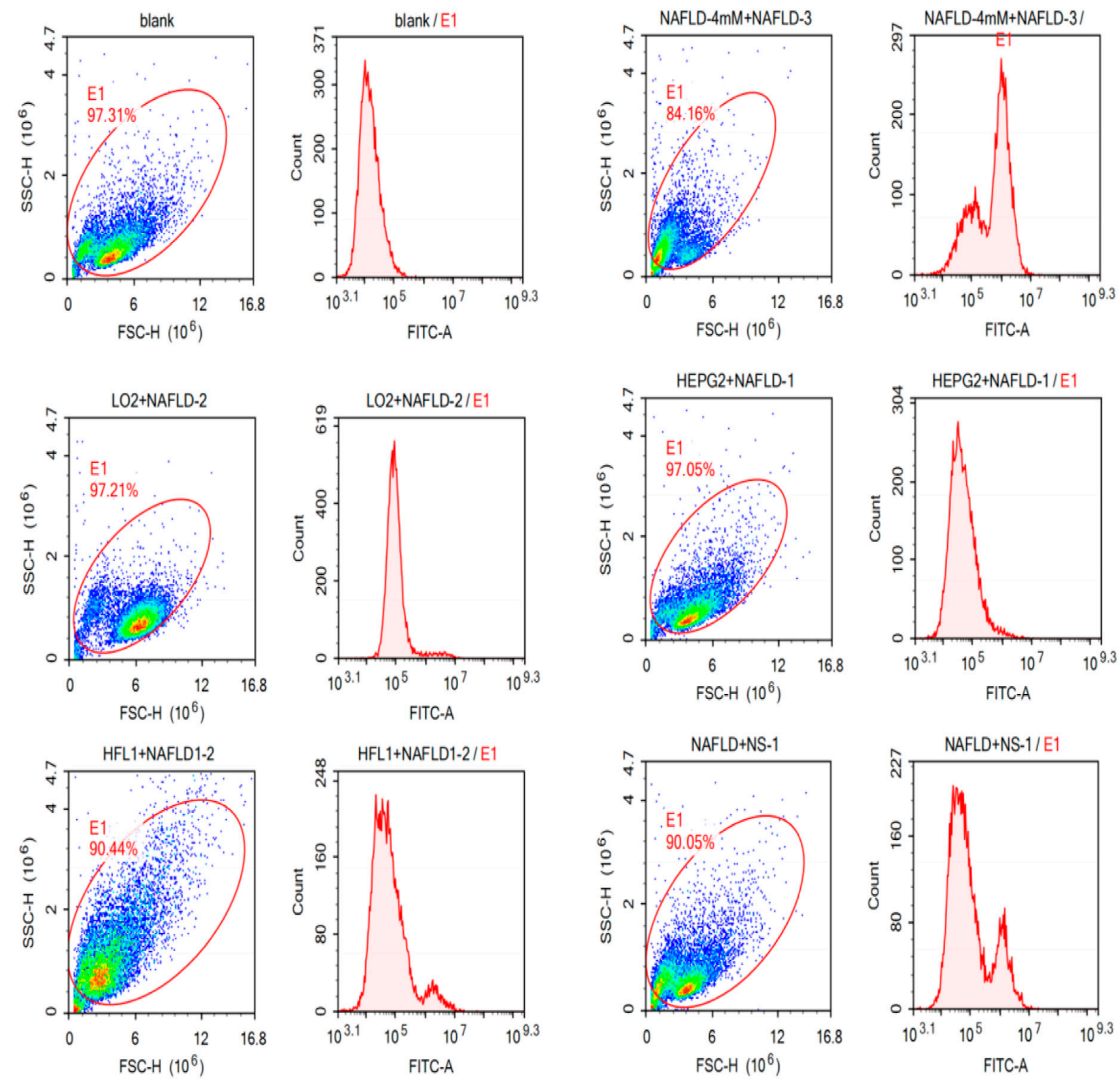


FIGURE 3
NAFLD01 aptamer's specific binding to NAFLD was verified by flow cytometry. The NAFLD01 aptamer sequence exhibited stronger fluorescence on NAFLD cells than the antisense aptamer sequence or other LO2, HFL1, and HepG2 cells.

receptors in liver cells. The nanocarrier specifically delivers Res to the liver to increase its therapeutic effectiveness. Although ultrasound nanobubbles are more suitable for delivering drugs than the abovementioned nanocarriers because of their ability to enter extravascular tissue, support specific ultrasound imaging, and achieve targeted drug release, it is necessary to modify them with specific targeting groups to improve their therapeutic efficacy. In view of the high affinity and selectivity of aptamers, as well as their advantages of good stability, easy chemical modification and coupling, and low cost, we used aptamer-functionalized ultrasound nanobubbles as nanoparticle carriers to design and synthesize particles loaded with Res and CuNPs for synergistic treatment of NAFLD.

3.1 Establishment of a nonalcoholic fatty liver disease cell model induced by high-oil fatty acids

As shown in Figure 1, HepG2 cells were treated with high-oil fatty acids (palmitic acid/oleic acid ratio of 1/2, final concentration of 4 mM), and many lipid droplets formed after

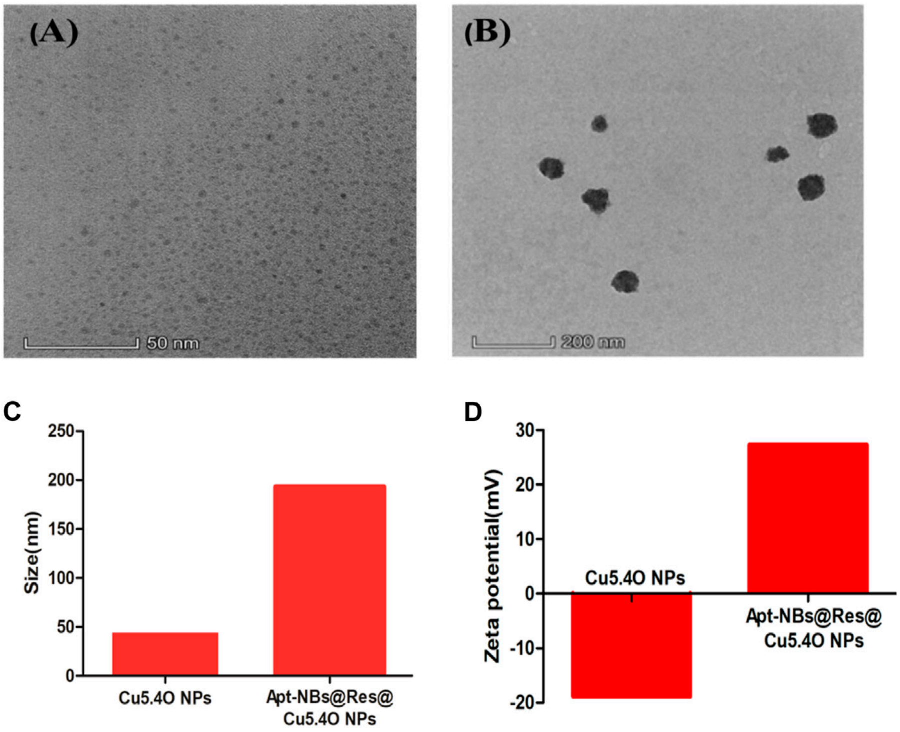


FIGURE 4
The TEM spectra of Cu$_{5.4}$ONPs **(A)**, and Apt-NBs@Res@Cu$_{5.4}$ONPs **(B)**; the size **(C)** and Zeta **(D)** of Cu$_{5.4}$ONPs and Apt-NBs@Res@Cu$_{5.4}$ONPs.

48 h, which confirmed the successful establishment of the NAFLD cell model.

3.2 Ability of the aptamer to target nonalcoholic fatty liver cells

There is a need to develop new effective methods and therapies for NAFLD, such as molecular probes that specifically recognize NAFLD cells. Nucleic acid aptamers (Zhou and Rossi, 2017), single-chain oligonucleotide molecules, have the advantages of a low molecular weight, no immunogenicity, a low synthesis cost and high stability after chemical modification. Aptamers can selectively bind to target sites through their unique three-dimensional structures and have a high affinity and a specificity similar to that of antigen-antibody interactions. Therefore, aptamers have been widely used in diagnosis, screening and drug development. Some aptamers not only recognize disease target molecules but also can be modified as drugs. Pu et al. (2021) developed a NAFLD cell-specific aptamer called NAFLD01 that specifically recognizes cultured steatotic hepatocytes and fatty liver tissue slices. It also improved the fatty acid degradation of NAFLD cells and increased the expression of PPAR-α (peroxisome activating receptor-α). Therefore, in this study, liver targeting was achieved by modifying the NAFLD01 aptamer. In our study, the NAFLD01 aptamer was used to modify ultrasound nanobubbles, and the specific binding ability of the NAFLD01 aptamer to NAFLD cells was verified through multiple experiments.

3.2.1 Confocal fluorescence microscopy imaging demonstrated specific binding of the NAFLD01 aptamer to nonalcoholic fatty liver disease cells

As shown in Figure 2, the fluorescence intensity of the fluorescein-labeled aptamer showed that NAFLD01 exhibited the strongest binding ability to NAFLD cells compared to nonsense sequences or other cells, which demonstrates that the NAFLD01 aptamer binds specifically to NAFLD cells and has potential for targeted treatment of NAFLD.

3.2.2 Flow cytometry verified the specific binding of the NAFLD01 aptamer to nonalcoholic fatty liver disease cells

As shown in Figure 3, the NAFLD01 aptamer sequence exhibited stronger fluorescence on NAFLD cells than the antisense aptamer sequence or other LO2, HFL1, and HepG2 cells. The fluorescence signal intensity decreased in the order NAFLD cells > NS-1 cells > HFL1 cells >

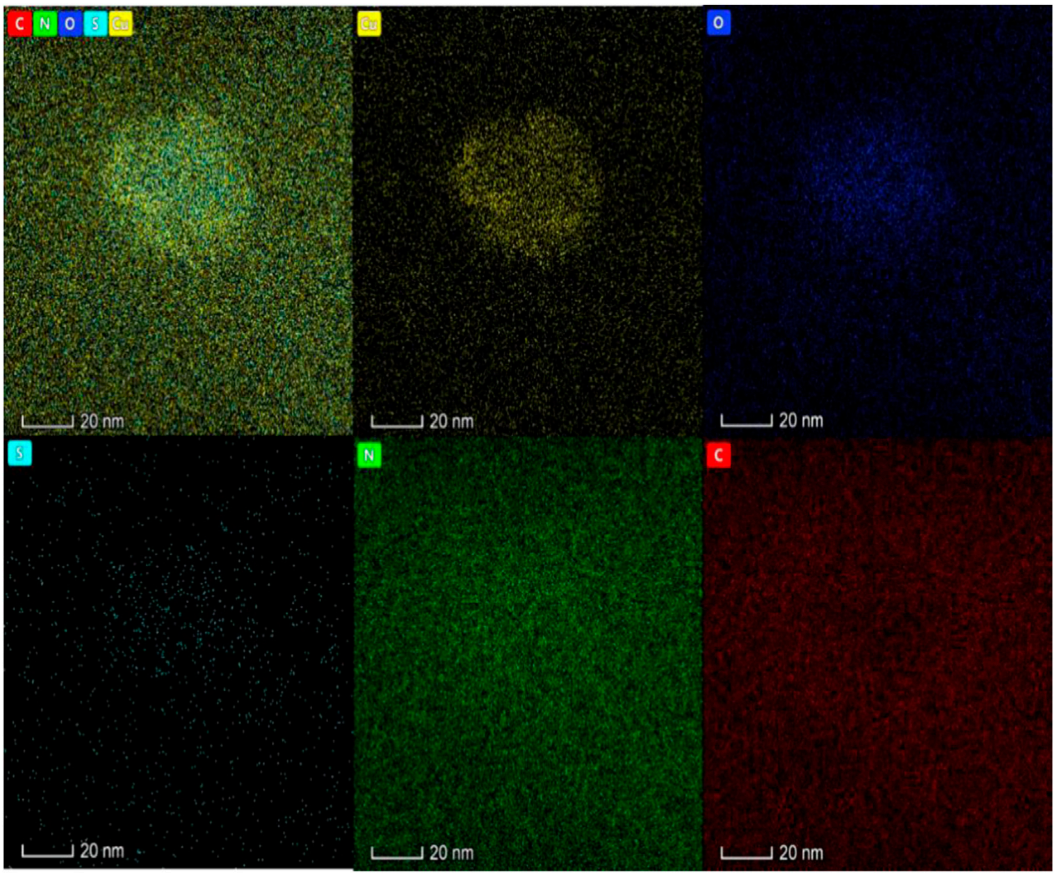


FIGURE 5
The Energy dispersive spectra of Apt-NBs@Res@Cu$_{5.4}$ONPs shows that the material also has five elements: Cu, N, O, S and C, in which Cu is mainly located in the core of the material and N is distributed in the shell.

HepG2 cells > LO2 cells > blank group, which further proved that the NAFLD01 aptamer binds specifically to NAFLD cells and has potential for targeted treatment of NAFLD.

3.2.3 Synthesis and characterization of aptamer-functionalized ultrasound nanobubbles loaded with Res and CuNPs

Zhu et al. (2018) constructed aptamer-modified nanobubbles for ultrasound imaging of various tumor cells. To date, there has been no research on the construction of NAFLD aptamer-modified nanobubbles or the synthesis of nanobubbles loaded with Res and CuNPs. We synthesized a nanobubble by thin-film hydration and mechanical oscillation, which is a relatively simple, green and economical method. As shown in Figures 4, 5, the morphology and size potential distribution of the materials were characterized by TEM and a Malvern particle size potentiometer. CuNPs have a small and uniform particle size distribution, and their particle size and zeta potential were 46.13 nm and −18.9 mV, respectively. When Res and CuNPs were loaded into the nanobubbles, the particle size and zeta potential of the nanobubbles changed to 127.6 nm and 35.73 mV, respectively. Then, the aptamer with NAFLD was covalently modified with maleamide and sulfhydryl groups, and the size and zeta potential of the nanobubbles changed to 193.73 nm and 27.39 mV, respectively. EDS, XRD, and XPS spectra of ultrasonic nanobubbles were also obtained. The results are shown in Figures 6–9 and indicate that the final synthesized aptamer-functionalized nanobubbles contained Cu, N, C, O, and S, among which Cu was mainly distributed in the inner position of the nanobubbles due to its hydrophilic properties. The C, N, and O from Res are mainly distributed in the middle of nanobubbles due to their hydrophobic properties. Encapsulation of Res in PLGA as a carrier can enhance its stability, solubility and pharmacological potential. The loading capacity of Res was measured using XPS and XRD. The results showed that the loading capacity of Res was 206.8 mg, and the loading rate was 94%, while the loading capacity of CuNPs was 6 mg (10 ml). We prepared nanoparticles with an encapsulation

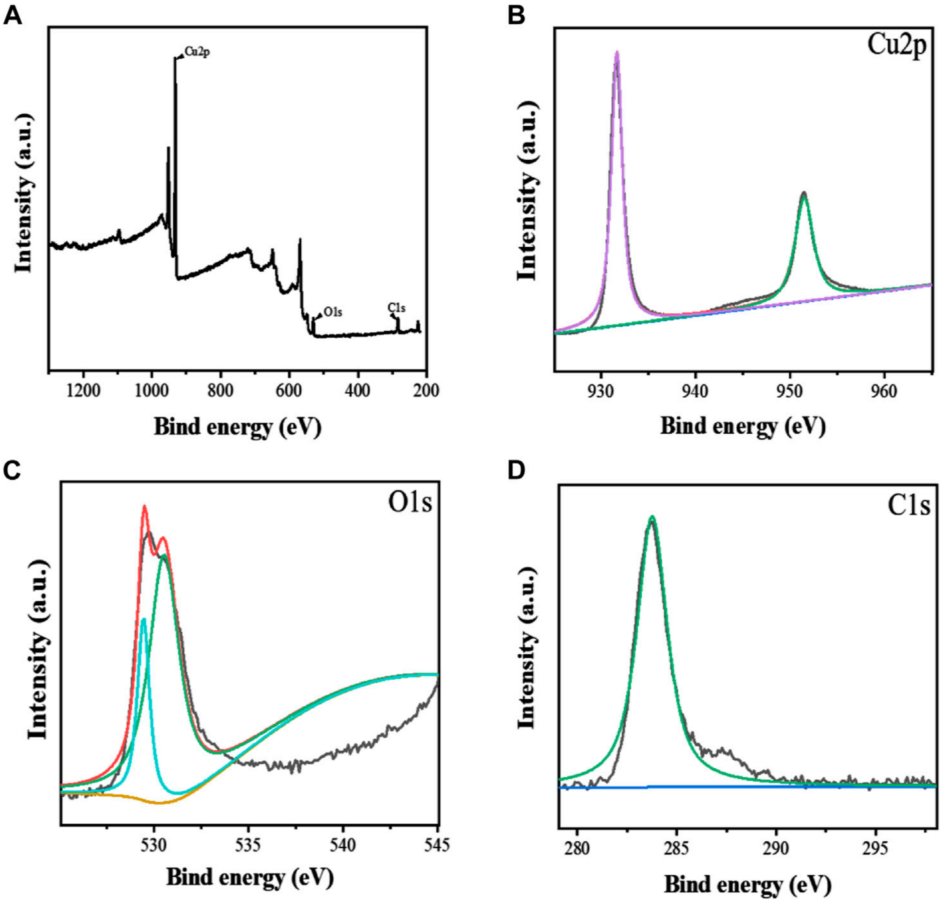


FIGURE 6
The XPS spectra of Apt-NBs@Res@$Cu_{5.4}O$ NPs **(A)**, The Cu2p **(B)**, O1s **(C)**, C1s **(D)** spectra of Apt-NBs@Res@$Cu_{5.4}$ONPs.

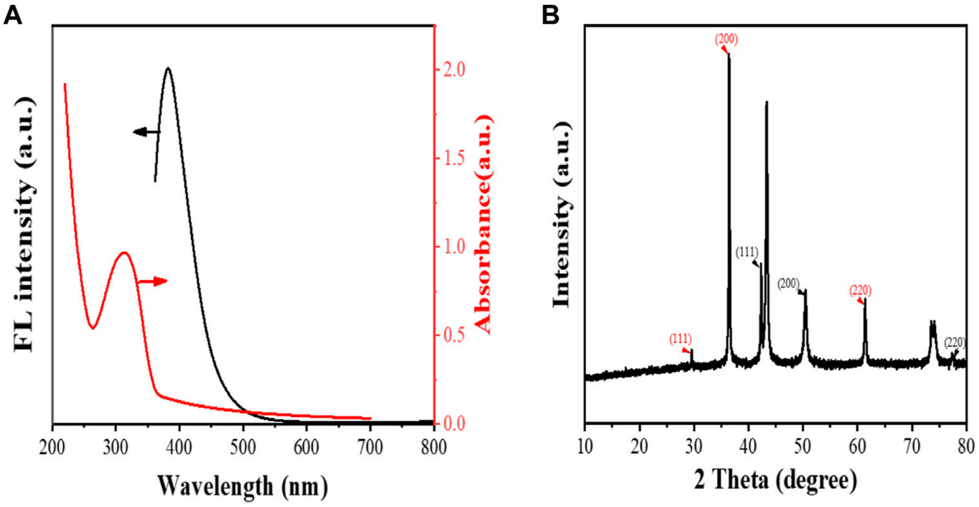


FIGURE 7
The fluorescence and UV-Vis spectra **(A)**, and XRD spectra of Apt-NBs@Res@$Cu_{5.4}O$ NPs **(B)**.

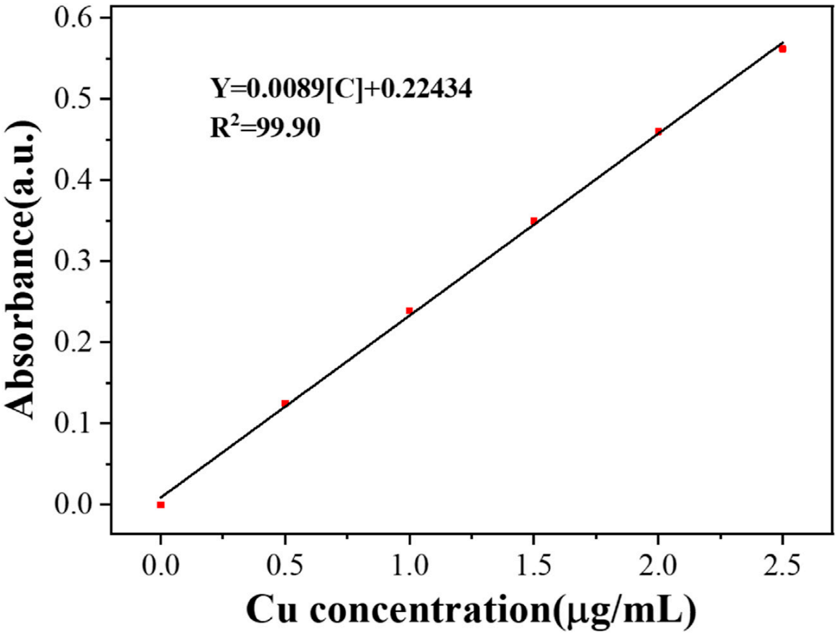


FIGURE 8
The linear calibration plot for Cu concentration by using Atomic Absorption Spectroscopy.

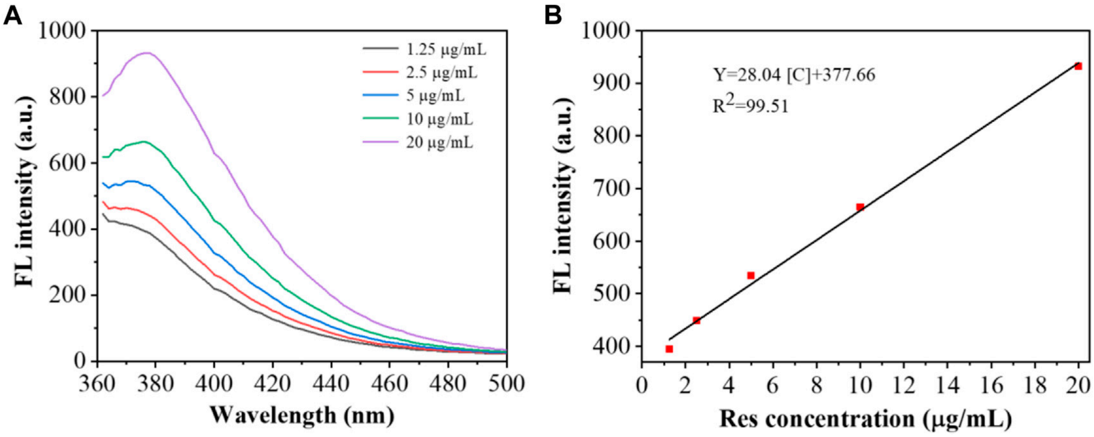


FIGURE 9
Fluorescence spectra **(A)** and corresponding linear calibration plot for Res **(B)**.

efficiency of up to 90% or higher, which showed excellent drug loading ability. Zeta potential is an important indicator of electrode stability. Colloidal dispersions can resist aggregation. The higher the zeta potential is (positive or negative), the more stable the system. As the absolute zeta potential is greater than the electrostatic repulsion, the nanobubble potential was 27.39 mV, which indicated that apt-NBs@Res@$Cu_{5.4}O$ NPs have high colloidal stability (Chen et al., 2016; Carbone et al., 2018). All of the above data effectively proved the successful synthesis of Res and CuNPs supported by adaptor-functionalized ultrasound nanobubbles.

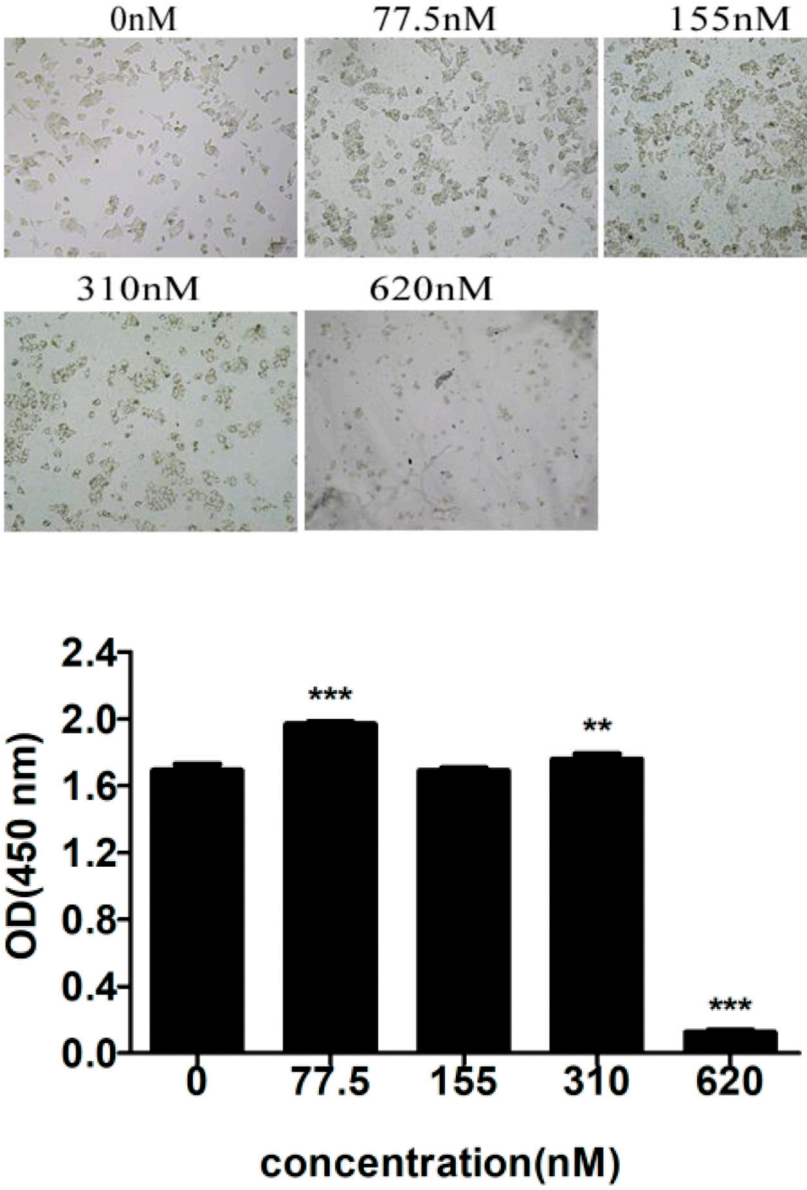


FIGURE 10
Cytotoxicity analysis of Apt-NBs@Res@$Cu_{5.4}$ONPs by CCK8.

3.3 Aptamer-modified ultrasound nanobubbles loaded with Res and CuNPs could be used to treat nonalcoholic fatty liver disease effectively

3.3.1 *In vitro* cytotoxicity test

Cytotoxicity analysis of Apt-NBs@Res@$Cu_{5.4}$O NPs with CCK8. As shown in Figure 10, nanobubbles no larger than 310 nM (Apt as the measurement unit) were basically not cytotoxic, and this concentration was used for subsequent treatment of NAFLD.

3.3.2 The expression levels of the inflammatory factors TNF-α and Il-10 were detected by Western blotting, and the level of total TGs was detected by ELISA

The results are shown in Figures 11–13. Compared with that in the other groups, the expression of the inflammatory factors TNF-α and IL-10 in the HepG2 + Apt-NBs@$Cu_{5.4}$O NPs@Res + FFA + ultrasound group was the lowest, while the level of total TGs in the HepG2 + Apt-NBs@$Cu_{5.4}$O NPs@Res + FFA + ultrasound group was lower than that in the FFA + ultrasound group, and there was no significant difference between them.

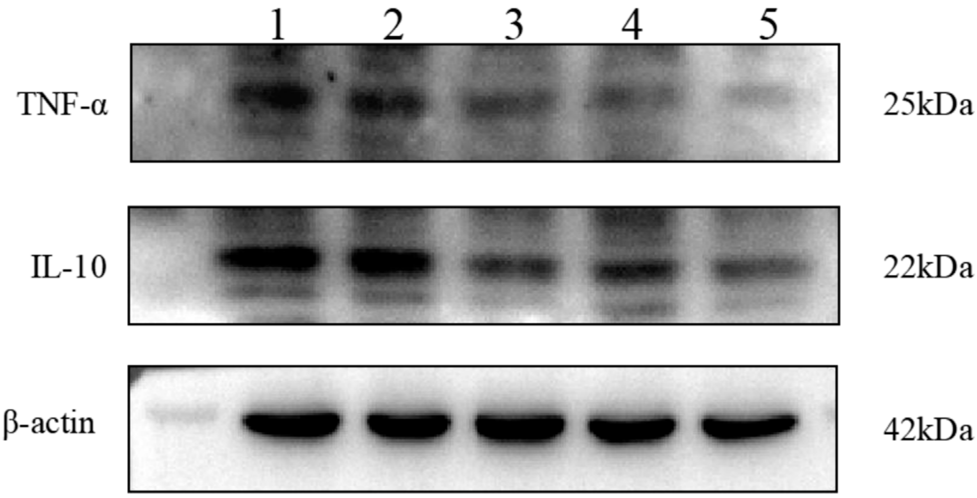


FIGURE 11
The effect of different treatment on IL-10 and TNF-α expression of HepG2 cell. 1) ultrasound; 2) FFA + ultrasound; 3) NBs + FFA + ultrasound; 4) NBs@$Cu_{5.4}O$ NPs@Res + FFA + ultrasound; 5)Apt-NBs@$Cu_{5.4}$ONPs@Res + FFA + ultrasound.

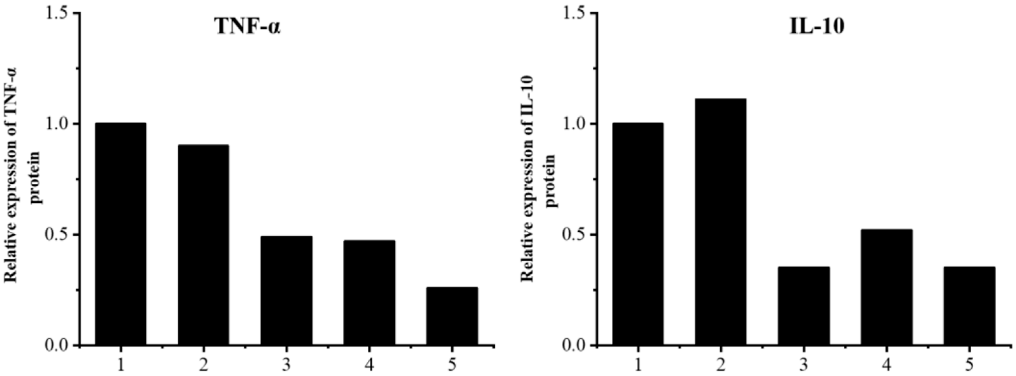


FIGURE 12
Relative levels of TNF-α and IL-10 expression after indicated treatments. 1) ultrasound; 2) FFA + ultrasound; 3) NBs + FFA + ultrasound; 4) NBs@$Cu_{5.4}$ONPs@Res + FFA + ultrasound; 5) Apt-NBs@$Cu_{5.4}$ONPs@Res + FFA + ultrasound.

NASH, an inflammatory phase of NAFLD, has increased in prevalence in recent years and has become the third major risk factor for the occurrence of hepatocellular carcinoma (HCC). Many studies have found that oxidative stress in NASH hepatocytes induces the production of a large number of inflammatory factors that induce gene expression and activate a self-targeted aggressive killing effect of $CD8^+$ T-cells, which may be an important mechanism of NASH hepatocyte injury and progression to HCC. Whether Res and CuNPs can effectively inhibit oxidative stress, reduce inflammatory factor levels and reverse the inflammation associated with NAFLD is an important therapeutic question. The results suggest that aptamer-modified nanobubbles loaded with drugs and CuNPs have a significant anti-inflammatory effect.

The findings of this study are as follows: 1. NAFLD01 was verified to be a strong adapter that specifically binds to nonalcoholic fatty liver (NAFL) cells. 2. Aptamer-modified targeted nanobubbles were constructed to specifically bind to NAFL cells. 3. Synthesis of NAFLD-specific aptamer-modified targeted ultrasound nanobubbles loaded with Res and CuNPs was achieved. 4. The cytotoxicity and efficacy of aptamer-modified nanobubbles loaded with drugs and CuNPs in the treatment of NAFLD were demonstrated *in vitro*.

Until now, there has been no study on the combination of aptamer-modified nanobubble-loaded drugs and CuNPs for the treatment of NAFLD.

Currently, liver-targeted drug delivery systems (HTDDS) have been widely used against liver injury and liver diseases,

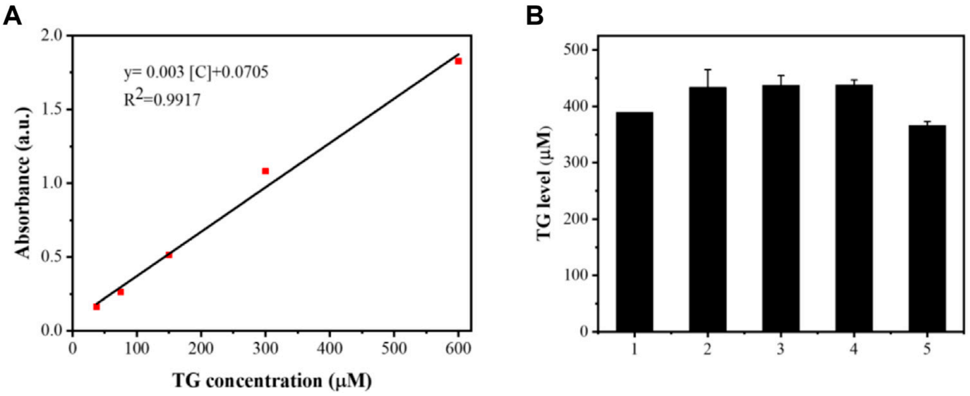
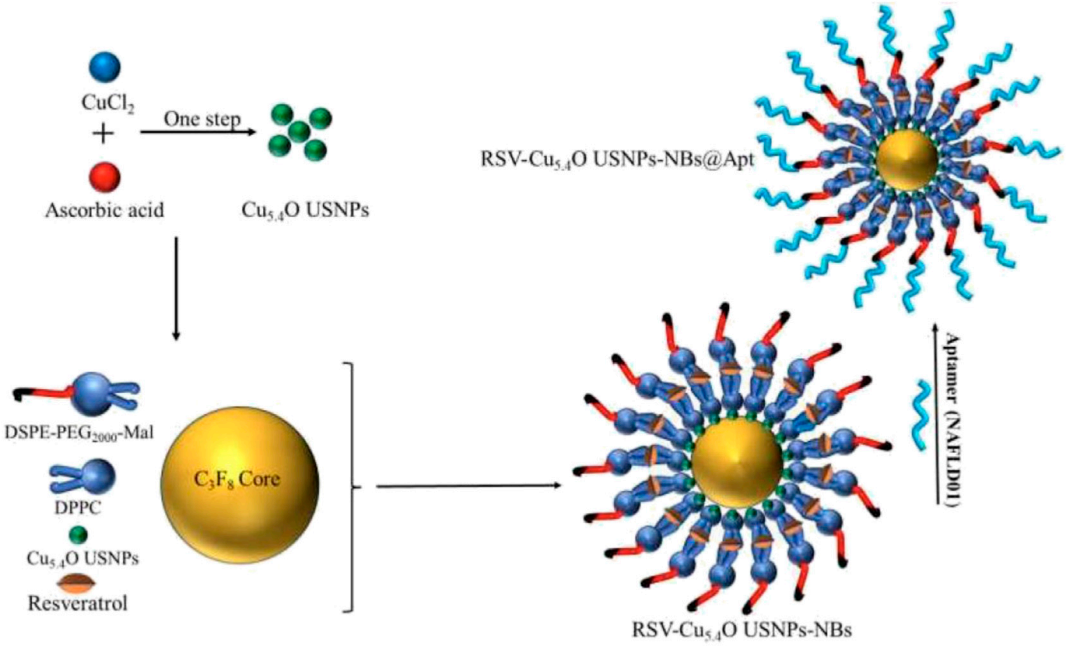


FIGURE 13
The linear calibration plot for TG concentration by ELISA method **(A)** and TG level after different treatment **(B)**; 1) ultrasound; 2) FFA + ultrasound; 3) The NBs + FFA + ultrasound; 4) NBs@$Cu_{5.4}$ONPs@Res + FFA + ultrasound; 5) Apt-NBs@$Cu_{5.4}$ONPs@Res + FFA + ultrasound.



SCHEME 1
Schematic diagram of the preparation of Apt-NBs@Res@$Cu_{5.4}$O NPs.

such as liver fibrosis, and HCC. HTDDS not only prolong the blood circulation time of the drug, but also deliver the drug exclusively to the liver, reducing side effects. These systems use distinct ligands that can specifically recognize and interact with receptors that are particularly overexpressed on the surface of hepatocytes. Studies on more effective liver-targeting molecules is presently a research hotspot. However, because nucleic acid aptamers are more economical, stable, and have higher affinity and specificity than other ligands without immunogenicity, some researchers have found that NAFLD01 is an aptamer that specifically binds to NAFL cells. However, its targeting ability has not been generally recognized. Our research results affirm that aptamer-mediated nanobubbles can effectively target and bind to NAFL cells.

Our synthesized novel nanobubbles can address the inefficiency and side effects of Res treatment of NAFLD while simultaneously delivering CuNPs for a synergistic effect. Moreover, our findings suggest that Res and CuNPs effectively and synergistically inhibit NAFLD inflammation, and Apt-NBs@Res@$Cu_{5.4}O$ NPs are feasible as an effective therapeutic to reverse NAFLD at an early stage. Liver-targeted NAFLD-specific aptamer-mediated targeted ultrasound nanobubbles may become a safe and promising platform for the treatment of NAFLD and other liver diseases.

4 Conclusion

In this study, we successfully prepared NAFLD01 aptamer-functionalized ultrasound nanobubbles loaded with Res and CuNPs, named Apt-NBs@Res@$Cu_{5.4}O$ NPs, which can be used for the targeted treatment of NAFLD. A nanobubble concentration less than 310 nM can effectively treat early NAFLD, which provides a new approach for the synergistic treatment of NAFLD with aptamer-functionalized nanobubbles and nanoparticles.

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 study was reviewed by the Medical Ethics Committee of Guangzhou Red Cross Hospital and did not require approval for the use of purchased human cell lines.

Author contributions

XG, ZH, and JC conceived and designed the study. JC, KH, JL, and HZ performed experiments. ZH, JL, KH, and YZ analyzed the data and drafted the manuscript. All authors have read and approved the version to be submitted.

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

Abba, Y., Hassim, H., Hamzah, H., and Noordin, M. M. (2015). Antiviral activity of resveratrol against human and animal viruses. *Adv. Virol.* 2015, 184241. doi:10.1155/2015/184241

Berman, A. Y., Motechin, R. A., Wiesenfeld, M. Y., and Holz, M. K. (2017). The therapeutic potential of resveratrol: A review of clinical trials. *NPJ Precis. Oncol.* 1, 35. doi:10.1038/s41698-017-0038-6

Braud, L., Battault, S., Meyer, G., Nascimento, A., Gaillard, S., de Sousa, G., et al. (2017). Antioxidant properties of tea blunt ROS-dependent lipogenesis: Beneficial effect on hepatic steatosis in a high fat-high sucrose diet NAFLD obese rat model. *J. Nutr. Biochem.* 40, 95–104. doi:10.1016/j.jnutbio.2016.10.012

Carbone, C., Martins-Gomes, C., Caddeo, C., Silva, A. M., Musumeci, T., Pignatello, R., et al. (2018). Mediterranean essential oils as precious matrix components and active ingredients of lipid nanoparticles. *Int. J. Pharm.* 548, 217–226. doi:10.1016/j.ijpharm.2018.06.064

Chachay, V. S., Macdonald, G. A., Martin, J. H., Whitehead, J. P., O'Moore-Sullivan, T. M., Lee, P., et al. (2014). Resveratrol does not benefit patients with nonalcoholic fatty liver disease. *Clin. Gastroenterol. Hepatol.* 12, 2092–2103. e1-6. doi:10.1016/j.cgh.2014.02.024

Chen, Z., Tai, Z., Gu, F., Hu, C., Zhu, Q., and Gao, S. (2016). Aptamer-mediated delivery of docetaxel to prostate cancer through polymeric nanoparticles for enhancement of antitumor efficacy. *Eur. J. Pharm. Biopharm.* 107, 130–141. doi:10.1016/j.ejpb.2016.07.007

Ding, L., Oligschlaeger, Y., Shiri-Sverdlov, R., and Houben, T. (2022). Nonalcoholic fatty liver disease. *Handb. Exp. Pharmacol.* 270, 233–269. doi:10.1007/164_2020_352

Endo-Takahashi, Y., and Negishi, Y. (2020). Microbubbles and nanobubbles with ultrasound for systemic gene delivery. *Pharmaceutics* 12, 964–977. doi:10.3390/pharmaceutics12100964

Francioso, A., Mastromarino, P., Masci, A., d'Erme, M., and Mosca, L. (2014). Chemistry, stability and bioavailability of resveratrol. *Med. Chem.* 10, 237–245. doi:10.2174/15734064113096660053

Kasper, P., Martin, A., Lang, S., Kütting, F., Goeser, T., Demir, M., et al. (2021). NAFLD and cardiovascular diseases: A clinical review. *Clin. Res. Cardiol.* 110, 921–937. doi:10.1007/s00392-020-01709-7

Kutlu, O., Kaleli, H. N., and Ozer, E. (2018). Molecular pathogenesis of nonalcoholic steatohepatitis- (NASH-) related hepatocellular carcinoma. *Can. J. Gastroenterol. Hepatol.* 2018, 8543763. doi:10.1155/2018/8543763

Liu, T., Xiao, B., Xiang, F., Tan, J., Chen, Z., Zhang, X., et al. (2020). Ultrasmall copper-based nanoparticles for reactive oxygen species scavenging and alleviation of inflammation related diseases. *Nat. Commun.* 11, 2788. doi:10.1038/s41467-020-16544-7

Lomonaco, R., Sunny, N. E., Bril, F., and Cusi, K. (2013). Nonalcoholic fatty liver disease: Current issues and novel treatment approaches. *Drugs* 73, 1–14. doi:10.1007/s40265-012-0004-0

Öztürk, E., Arslan, A., Yerer, M. B., and Bishayee, A. (2017). Resveratrol and diabetes: A critical review of clinical studies. *Biomed. Pharmacother.* 95, 230–234. doi:10.1016/j.biopha.2017.08.070

Pu, Y., Xiang, J., Zhang, X., Deng, Y., Liu, H., and Tan, W. (2021). CD36 as a molecular target of functional DNA aptamer NAFLD01 selected against NAFLD cells. *Anal. Chem.* 93 (8), 3951–3958. doi:10.1021/acs.analchem.0c04866

Rauf, A., Imran, M., Suleria, H., Ahmad, B., Peters, D. G., and Mubarak, M. S. (2017). A comprehensive review of the health perspectives of resveratrol. *Food Funct.* 8, 4284–4305. doi:10.1039/c7fo01300k

Schwimmer, J. B. (2007). Definitive diagnosis and assessment of risk for nonalcoholic fatty liver disease in children and adolescents. *Semin. Liver Dis.* 27, 312–318. doi:10.1055/s-2007-985075

Shen, Y., Lv, W., Yang, H., Cai, W., Zhao, P., Zhang, L., et al. (2019). FA-NBs-IR780: Novel multifunctional nanobubbles as molecule-targeted ultrasound contrast agents for accurate diagnosis and photothermal therapy of cancer. *Cancer Lett.* 455, 14–25. doi:10.1016/j.canlet.2019.04.023

Sun, X., Guo, L., Shang, M. M., Shi, D., Liang, P., Jing, X., et al. (2020). Ultrasound mediated destruction of LMW-HA-loaded and folate-conjugated nanobubble for TAM targeting and reeducation. *Int. J. Nanomedicine* 15, 1967–1981. doi:10.2147/IJN.S238587

Suzuki, R., Oda, Y., Omata, D., Nishiie, N., Koshima, R., Shiono, Y., et al. (2016). Tumor growth suppression by the combination of nanobubbles and ultrasound. *Cancer Sci.* 107, 217–223. doi:10.1111/cas.12867

Teng, W., Zhao, L., Yang, S., Zhang, C., Liu, M., Luo, J., et al. (2019). The hepatic-targeted, resveratrol loaded nanoparticles for relief of high fat diet-induced nonalcoholic fatty liver disease. *J. Control. Release* 307, 139–149. doi:10.1016/j.jconrel.2019.06.023

Wan, S., Zhang, L., Quan, Y., and Wei, K. (2018). Resveratrol-loaded PLGA nanoparticles: Enhanced stability, solubility and bioactivity of resveratrol for non-alcoholic fatty liver disease therapy. *R. Soc. open Sci.* 5, 181457. doi:10.1098/rsos.181457

Yu, Z., Wang, Y., Xu, D., Zhu, L., Hu, M., Liu, Q., et al. (2020). G250 antigen-targeting drug-loaded nanobubbles combined with ultrasound targeted nanobubble destruction: A potential novel treatment for renal cell carcinoma. *Int. J. Nanomedicine* 15, 81–95. doi:10.2147/IJN.S230879

Zhou, J., and Rossi, J. (2017). Aptamers as targeted therapeutics: Current potential and challenges. *Nat. Rev. Drug Discov.* 16, 181–202. doi:10.1038/nrd.2016.199

Zhu, L., Wang, L., Liu, Y., Xu, D., Fang, K., and Guo, Y. (2018). CAIX aptamer-functionalized targeted nanobubbles for ultrasound molecular imaging of various tumors. *Int. J. Nanomedicine* 13, 6481–6495. doi:10.2147/IJN.S176287

Zupančič, Š., Lavrič, Z., and Kristl, J. (2015). Stability and solubility of trans-resveratrol are strongly influenced by pH and temperature. *Eur. J. Pharm. Biopharm.* 93, 196–204. doi:10.1016/j.ejpb.2015.04.002

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 17 October 2022
DOI 10.3389/fphys.2022.1044308

OPEN ACCESS

EDITED BY
Huan Qin,
South China Normal University, China

REVIEWED BY
Xiong Wang,
ShanghaiTech University, China
Kedi Xiong,
South China Normal University, China

*CORRESPONDENCE
Dan Wu,
wudan@cqupt.edu.cn
Jun Yang,
imdyang@qq.com
Huabei Jiang,
hjiang1@usf.edu

†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 14 September 2022
ACCEPTED 04 October 2022
PUBLISHED 17 October 2022

CITATION
Wu D, Xu N, Xie Y, Shen Y, Fu Y, Liu L, Chi Z, Lu R, Xiang R, Wen Y, Yang J and Jiang H (2022), Noninvasive optoacoustic imaging of breast tumor microvasculature in response to radiotherapy.
Front. Physiol. 13:1044308.
doi: 10.3389/fphys.2022.1044308

Noninvasive optoacoustic imaging of breast tumor microvasculature in response to radiotherapy

Dan Wu[1]*†, Nan Xu[2]†, Yonghua Xie[1], Yang Shen[1], Yunlu Fu[1], Liang Liu[1], Zihui Chi[1], Runyu Lu[1], Renjie Xiang[1], Yanting Wen[1,3], Jun Yang[2]* and Huabei Jiang[4]*

[1]School of Optoelectric Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China, [2]Department of Radiology, The Third Affiliated Hospital of Kunming Medical University, Yunnan Cancer Hospital, Kunming, China, [3]Ultrasonic Department, The Fifth People's Hospital of Chengdu, Chengdu, China, [4]Department of Medical Engineering, University of South Florida, Tampa, FL, United States

Detailed insight into the radiation-induced changes in tumor microvasculature is crucial to maximize the efficacy of radiotherapy against breast cancer. Recent advances in imaging have enabled precise targeting of solid lesions. However, intratumoral heterogeneity makes treatment planning and monitoring more challenging. Conventional imaging cannot provide high-resolution observation and longitudinal monitoring of large-scale microvascular in response to radiotherapy directly in deep tissues. Herein, we report on an emerging non-invasive imaging assessment method of morphological and functional tumor microvasculature responses with high spatio-temporal resolution by means of optoacoustic imaging (OAI). *In vivo* imaging of 4T1 breast tumor response to a conventional fractionated radiotherapy at varying dose (14 × 2 Gy and 3 × 8 Gy) has been performed after 2 weeks following treatment. Remarkably, optoacoustic images can generate richful contrast for the tumor microvascular architecture. Besides, the functional status of tumor microvasculature and tumor oxygenation levels were further estimated using OAI. The results revealed the differential (size-dependent) nature of vascular responses to radiation treatments at varying doses. The vessels exhibited an decrease in their density accompanied by a decline in the number of vascular segments following irradiation, compared to the control group. The measurements further revealed an increase of tumor oxygenation levels for 14 × 2 Gy and 3 × 8 Gy irradiations. Our results suggest that OAI could be used to assess the response to radiotherapy based on changes in the functional and morphological status of tumor microvasculature, which are closely linked to the intratumor microenvironment. OAI assessment of the tumor microenvironment such as oxygenation status has the potential to be applied to precise radiotherapy strategy.

KEYWORDS

optoacoustic imaging, radiotherapy, breast tumor, tumor microvasculature, oxygenation status

1 Introduction

Tumor microenvironment, such as oxygenation state, has a great influence on the radiosensitivity of tumor cells. Tumor microenvironment is closely related to the function of tumor microvessels. Therefore, a detailed understanding of radiation-induced changes in tumor microvessels is essential to maximize the efficacy of radiation therapy for cancer (Park et al., 2012). The main manifestations of tumor microvessels response to irradiation are decreased perfusion, increased vascular wall permeability, decreased number of functional vessels, decreased vascular network density, vascular contraction, and decreased blood flow velocity (Pedro et al., 2019; Anamitra et al., 2022). Research has found that the therapeutic ratio in treating cancer with radiation could be increased by delivering the radiation in multiple fractions, fractionated radiotherapy has been an almost universally accepted clinical practice (Brand Douglas et al., 2019). The effect of radiation on tumor microvessels varies with the changes of total dose, dose rate, dose fraction and fraction number, as well as biological factors such as tumor type, tumor site and tumor growth stage (Yoon Stephanie et al., 2021).

Studying the mechanisms of tumor tissue destruction by different doses of radiation, especially its microvascular response, is crucial for optimizing treatment plan, prognosis and follow-up. Noninvasive imaging methods offer the potential for longitudinal monitoring of dynamic temporal changes occurring in the tumor microenvironment and allow us to map the spatial heterogeneity of tumor microvessels and tumor oxygenation. Conventional imaging methods such as magnetic resonance imaging (MRI) (Morgan Tiffany et al., 2019; Turkkan et al., 2022), positron emission tomography (PET) (Abravan et al., 2017; Jiang et al., 2021) and computed tomography (CT) (Wang et al., 2011; Crane Christopher and Koay Eugene, 2016) have been used to study tumor blood vessels and oxygenation in animal models and patients. However, these technologies are either more expensive or require the use of ionizing radiation or radioactive isotopes. The penetration depth of Doppler ultrasound imaging (UI) is appropriate, but the contrast and spatial resolution are insufficient to clearly identify slow blood flow microvessels (McNabb et al., 2020). Therefore, it is of great clinical value to study a cheap, reliable and simple method of tumor microvessels and oxygenation imaging.

Optoacoustic imaging (OAI) is a hybrid imaging technology that combines the sensitivity of optical imaging with the resolution of ultrasonic imaging (Jiang, 2014; Beard Paul and Cox Ben, 2016; Vasilis and Buehler, 2017; Liu et al., 2019). In OAI, tissues are irradiated with near-infrared light, which is absorbed by endogenous chromophores such as oxygenated and deoxygenated hemoglobin, causing thermoelastic expansion and producing broad-band pressure waves that are detected as acoustic signals (Wang, 2016; Duan et al., 2021; Wang et al., 2021; Gong et al., 2022). Interestingly, the absorption properties of hemoglobin are affected by whether or not it binds to oxygen. By comparing the PA signals of deoxyhemoglobin (Hb), oxyhemoglobin (HbO) and oxygen saturation (sO_2) could be calculated in intratumoral and peritumoral areas (Huang et al., 2016; Rich Laurie and Mukund, 2016; Dai et al., 2020; Jiao et al., 2020; Qian et al., 2020). In this way, OAI uses endogenous contrast mechanisms to visualize tumor microvascular structure and hemodynamics. As a result, there is widespread interest in developing OAI for medical applications, particularly in cancer research.

In this study, we used the OAI system based on a semicircular array detector to monitor the response of tumor vascular networks to multiple exposures to different doses of ionizing radiation. It meets our requirements for real-time monitoring and quantitative analysis of tumor vessels, especially to improve the imaging sensitivity of microvessels. Using the xenograft mouse model of 4T1 breast cancer, we demonstrated the differential (size dependent) nature of *in vivo* microvascular response to radiotherapy. Multispectral measurements were further used to monitor tumor hemodynamic parameters.

2 Methods and materials

2.1 *In vivo* models and cell lines

The experiments were carried out on nine 5-6-week-old female Balb/c mice (18–20 g). The animals were obtained from the department of experimental animals, Kunming Medical University (Kunming, China). The study was performed on the model of murine breast cancer cells 4T1 (Chinese Academy of Sciences Stem Cell Bank). The cells were cultured in RPMI-1640 containing 10% fetal bovine serum and penicillin-streptomycin. 1×10^6 cells in 200 μL of PBS were injected subcutaneously into the outer side of the right lower limb. The animal studies were approved by the Ethics Committee of Chongqing University of Posts and Telecommunications and Kunming Medical University.

2.2 Irradiation treatment and experimental protocols

Irradiation was performed with RS2000 Biological X-ray Biological Irradiator (VGeorgia, United States) in the 6X SRS mode with an accelerating voltage of 225 kV and a dose rate of 1.8 Gy/min. The mice were immobilized so that an irradiation field of 8×8 m^2 was formed at the tumor and the rest of the body was covered with a lead screen to reduce radiation to the normal tissue. The tumor-bearing mice were randomly divided into three groups: low dose radiation group, high dose radiation group and control group. Animals (three mice in a group) received 14 × 2 Gy (low dose) or 3 × 8 Gy (high dose) radiation for two

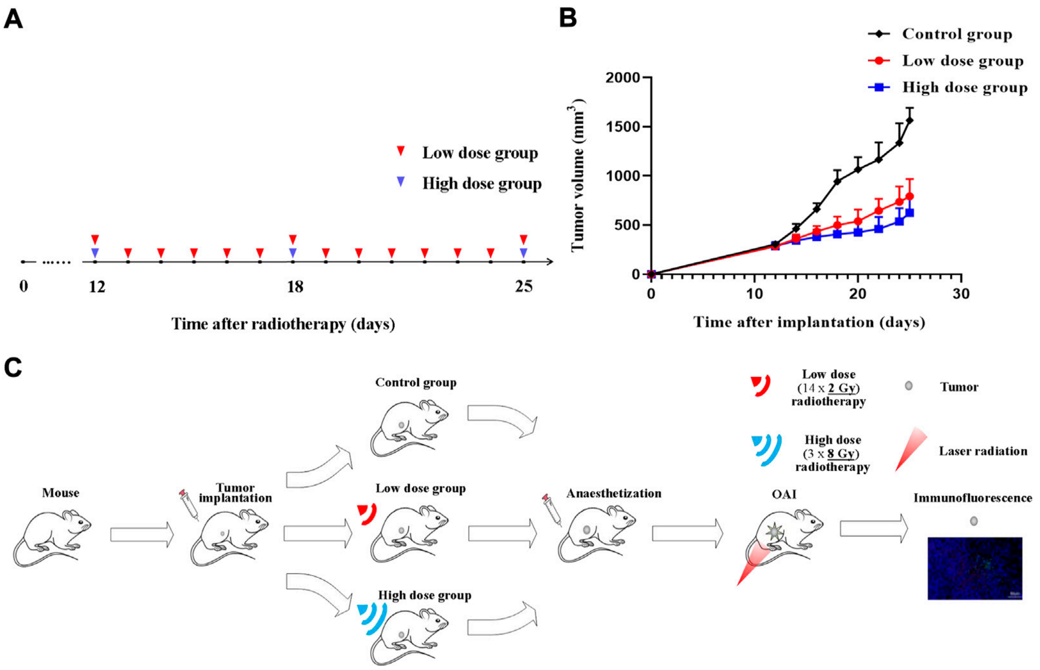


FIGURE 1
Radiotherapy protocol during tumor growth process. Experiment schedule **(A)** and dynamics of 4T1 volume after single-dose irradiation at different doses **(B,C)** Illustrative overview of experimental protocols. Time point of tumors radiotherapy is indicated by blue triangle (high dose) and red triangle (low dose).

consecutive weeks (Figure 1A). Three mice were used as an untreated control. Tumor volume and body weight were measured every 2 days. Tumors were measured along two perpendicular directions with a caliper, and their volumes were calculated as follows: $V = 0.5 \times a \times b^2$, where V is the volume and a and b are the two corresponding diameters (Figure 1B).

OA imaging was performed after tumor fractionated radiotherapy. Mice were anesthetized with chloral hydrate by 10% chloral hydrate solution which was injected intraperitoneally to the animal with a dose of 4 ml/kg body weight in all the experiments. After OA investigation, all mice were sacrificed by spinal dislocation. The tumor tissues were collected for immunofluorescence staining. Figure 1C shows an illustrative overview of experimental protocols.

2.3 Optoacoustic imaging system and data processing

An *in vivo* laboratory-built OAI system was used in this study. The schematic diagram of OAI system for monitoring the effects of radiotherapy on tumor microvessels in this study is shown in Figure 2. In this system, a pulsed Ti:Sapphire laser (Surelite OPO, Continuum United States) with 6–7 ns pulsed duration, 20 Hz pulse repetition rate and wavelengths of 690–960 nm was used as the excitation source. The maximum fluence at 690–960 nm was approximately 18 mJ/cm^2 during the experimentation. A custom-made fiber bundle with line-shaped illumination pattern (40 × 1 mm^2) was applied to deliver the light from a pulsed laser. The energy of pulsed laser was controlled below 8 mJ/cm^2 at the surface of the animal's skin.

A 128 element ultrasound transducer array (center frequency 5 MHz, bandwidth 90%, Japan Probe Co., Ltd.) to receive the photoacoustic signal. The received photoacoustic signal is amplified by a home-made preamplifier (gain 54 dB, frequency range: 200 kHz-15 MHz), and the amplified OA signals were then collected by 128-channel data acquisition cards at a sampling rate of 50 MS/s and 12-bit digital resolution (PXIe5105, National Instrument, United States). Ten-time averaging of the signal can minimize the unstability of laser energy, and thus the image quality is improved, making the calculation of tumor area more accurate. One complete frame of data from single wavelength was acquired in 0.5 s. During the acquisition process, real-time imaging can be realized on the Labview panel. The OAI tumor images were reconstructed from the photoacoustic signals using a multispectral quantitative reconstruction algorithm. The in-plane spatial resolution of the system is about approximately 150 μm according to our previous research (Yang et al., 2020a).

PAT tumor images were reconstructed from photoacoustic signals using a delay and sum reconstruction algorithm. According to the algorithm in our previous article (Yang

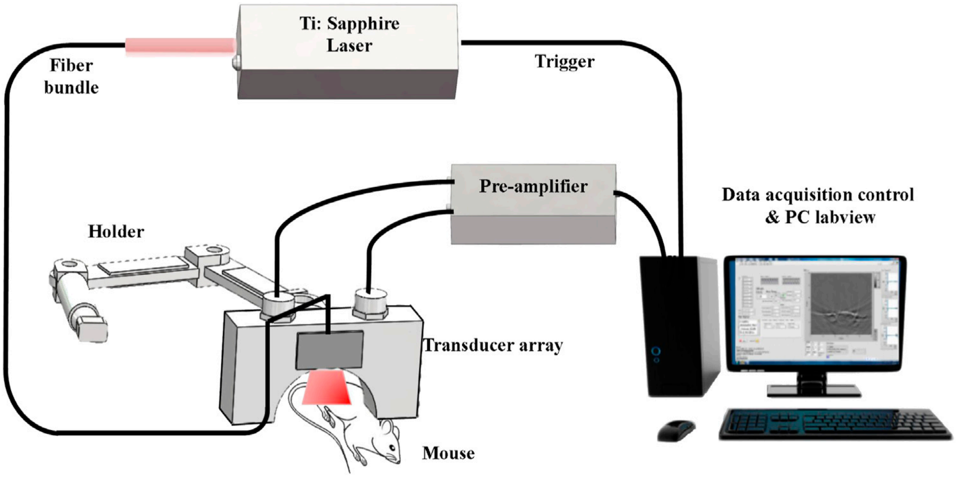


FIGURE 2
Schematic of our OAI system for *in vivo* imaging tumor-bearing mouse.

et al., 2020a; Yang et al., 2020b), we can obtain the tumor deoxyhemoglobin (Hb), oxyhemoglobin (HbO), oxygen saturation (sO_2) and water (H_2O) changes of tumor regions at three different wavelengths (760 nm, 840 nm, and 930 nm).

In each OA parameter image, an area of 20 × 20 pixels on the normal tissue surrounding the tumor, the boundary of the tumor and internal tumor was selected as three regions of interest (ROI I, ROI II, and ROI III). The mean pixel value of the ROIs was calculated. The ROI I was used as a baseline to calculate the relative changes on ROI II and ROI III. The resulting OA images were processed by background elimination and normalization. In the method of background elimination, the OA signals below the baseline were partly removed, and only the rest of the signals were retained (incomplete background elimination). This method can improve the signal-to-noise ratio. Furthermore, to conveniently calculate the pixel value within the ROIs, regions outside the ROIs were assigned a pixel value of zero. We can calculate the pixel value according to the following formula: $SUM = J \times P$, where P and J are the pixel matrix of the reconstructed image and the binary image matrix, respectively. First, the template image of the regions of interest (ROIs) is transformed into the binary image matrix J, where regions outside the ROIs were assigned a pixel value of zero. We then obtained the pixel value, that is, the normalized optical contrast value, by multiplying J and P. Finally, the pixel values for each ROI were calculated.

2.4 Tumor immunofluorescence

Tumor tissues were fixed with 4% paraformaldehyde and cut into 4-μm-thick sections after dehydration and embedment. In brief, specimens were incubated with anti-CD31 (1:500, Abcam, ab182981), anti-α-SMA (1:200, Servicebio, GB13044) antibodies for staining. The slices were imaged with a NIKON ECLIPSE C1 fluorescence microscope and scanned and analyzed with a PANNORAMIC panoramic slice scanner and Image-Pro Plus 6.0 analysis software. Tumor slices were divided into three equal areas, and the microvessel density (MVD) of each area was detected by CD31 staining of perivascular cells, and the area was quantified on the basis of the total number of microvessels per unit area; similarly, the coverage of the perivascular cells stained for α-SMA in each area was calculated. The vascular maturity index (VMI) refers to the percentage of blood vessels stained with the anti-α-SMA antibody compared to the total number of blood vessels stained with CD31.

2.5 Statistical analysis

All measurement data obtained for tumor growth are shown as mean ± SD; deoxyhemoglobin (Hb), oxyhemoglobin (HbO) and oxygen saturation (sO_2) are shown as means ± SEM. Parameters of microvessel density (MVD), microvessel segments number, and size are shown as the line chat including means, minimum and maximum values of the data set. SPSS 25.0 software (SPSS Company, Chicago, Illinois, United States) was used for statistical analysis. Statistically significant value was taken as $p \leq 0.05$.

3 Results

3.1 Tumor growth after irradiation

The volume in control group increased from 306.38 ± 21.33 mm^3 (on day 12) to 1,566.37 ± 125.91 mm^3 (on day 25).

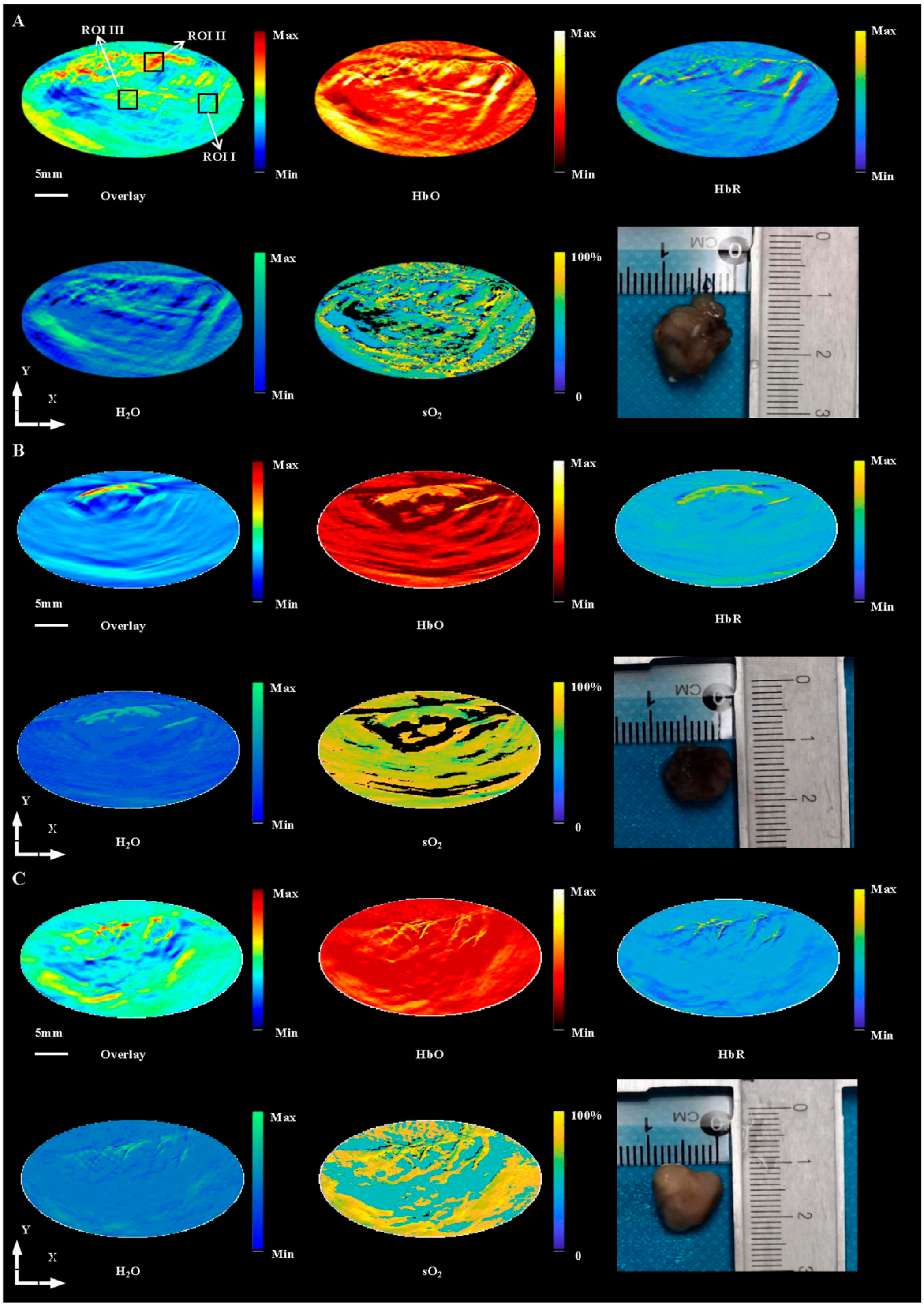


FIGURE 3
In vivo imaging of 4T1 breast tumors at HbT (overlay), HbO, HbR, H_2O and sO_2 distribution in control group **(A)**, 14 × 2 Gy group **(B)** and 3 × 8 Gy group **(C)** at 25 days after radiotherapy, and histological photograph of three collected tumor tissues in different groups. ROI I, ROI II, and ROI III in (A) indicates the baseline, peritumor and intratumor respectively. Scale bar = 5 mm for the images.

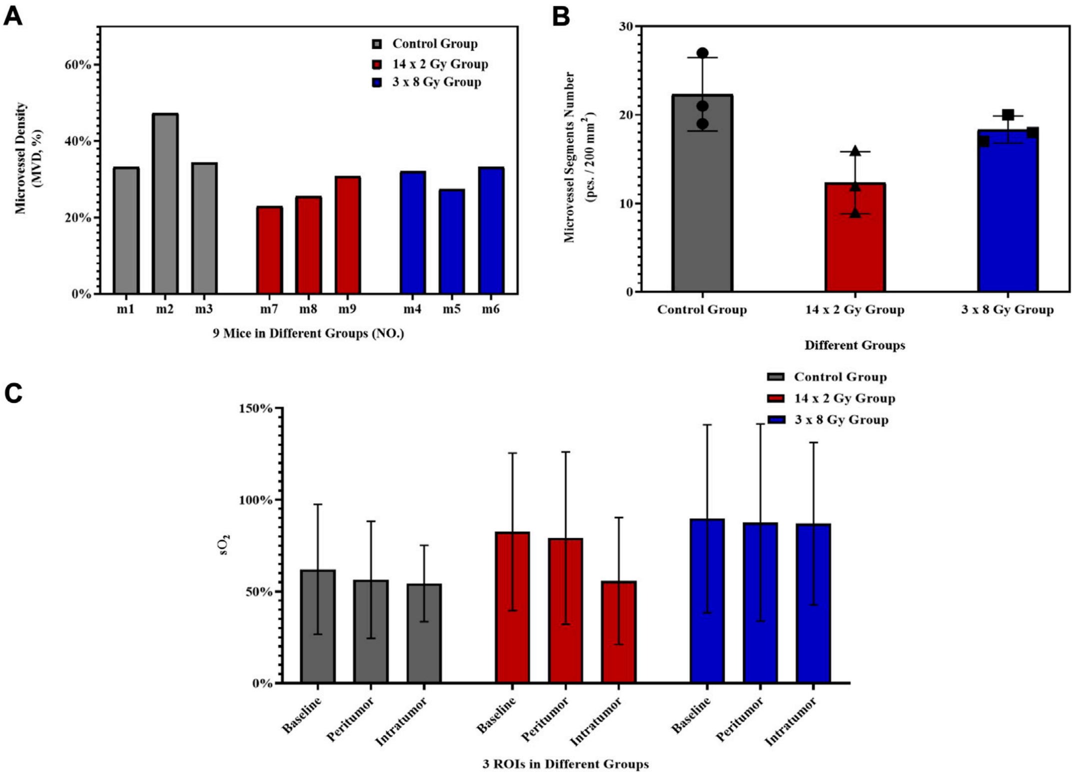


FIGURE 4
Histogram of OA parameters. Histogram of the percentage variation (SEM) of microvessel density (MVD, 100%) **(A)** and microvessel segments number **(B)** of 4T1 breast tumors after irradiation at different doses. **(C)** Histogram of the percentage variation (SEM) of sO_2 at baseline, peritumor and intratumor at different doses. The error bar indicated the standard deviation.

Irradiated tumors appeared to be of smaller size as compared to untreated ones at the time of monitoring completion. Differences between treated and untreated tumors on 20th day of the experiment were assessed for all groups. Tumor growth inhibition index (TGI) were calculated as follows: TGI = 1-Relative Tumor Volume (irradiation group)/Relative Tumor Volume (control group). TGI was 49.28% for 14 × 2 Gy-irradiated tumors and 60.01% for tumors irradiated with 3 × 8 Gy. Compared with the control group, radiotherapy significantly inhibited the tumor progression. The tumor growth in the high dose group was lower than that in the low dose group (Figure 1B).

3.2 Optoacoustic imaging

Using OAI system, nine mice in control group (Figure 3A), 14 × 2 Gy group (Figure 3B) and 3 × 8 Gy group (Figure 3C) were scanned at 25 days after radiotherapy. One representative mouse was selected in each group. HbT (an overlay of 760 nm image, 840 nm image and 930 nm image), HbO, HbR, H_2O, and sO_2 distribution images were shown in Figure 3. As evidenced by the OA images, the tumors presented as regions containing small, broken, irregular, randomly distributed blood vessels that were markedly different from the normal microvessels in the surrounding healthy tissue. Untreated 4T1 breast tumor was characterized by a more extensive distribution of hemoglobin in the tumor peripheral area and inside the tumor. Elongated small vascular structures formed from the inside to the surface of the tumor, followed by superficial scabs at later stages of tumor development. The sO_2 distribution on the surface and inside the tumor was relatively uniform. Histological photographs of three collected tumor tissues in different groups indicated that the tumor size in the control group was larger than that in the irradiation group. After irradiation, only short fragments of blood vessels retained in the tumor area, and the tumor microvessel density and the water content was decreased overall. Compared with the control group, the hemoglobin concentration at the tumor boundary decreased slightly, while the hemoglobin concentration inside the tumor decreased significantly in the irradiation group. After irradiation, the internal sO_2 distribution is very uneven. The response was

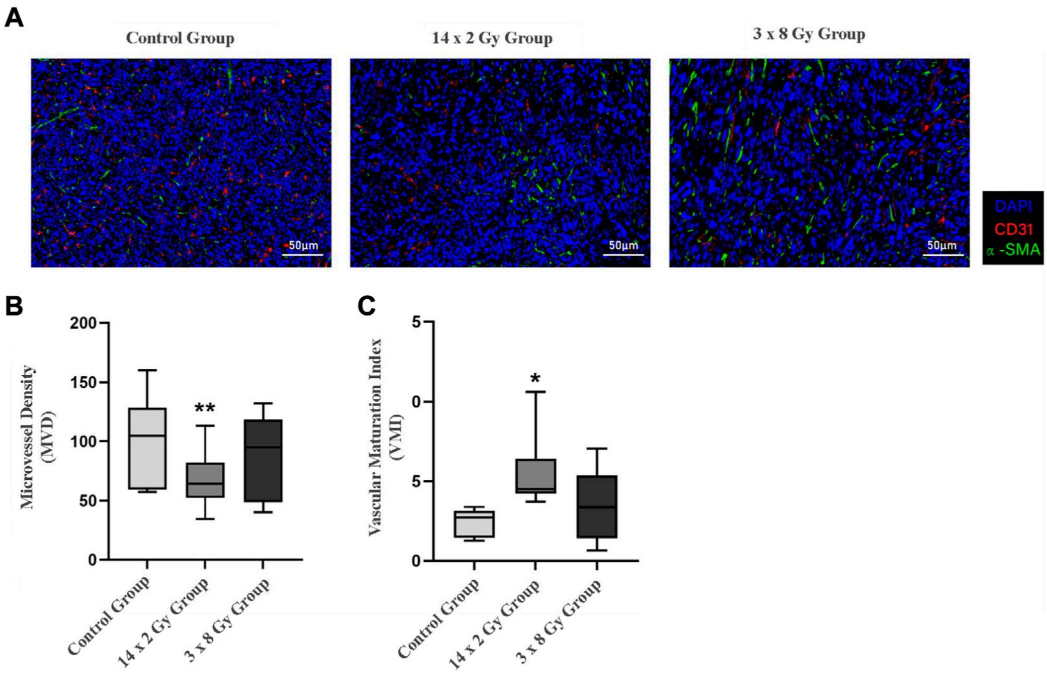


FIGURE 5
Immunofluorescence analysis of 4T1 breast tumors after irradiation at different doses. **(A)** Shows double-marker immunofluorescence of CD31 (red) and α-SMA (green), and the blue area represents the nucleus. **(B)** Indicates the quantitative expression of microvessel density (MVD) in tumor after radiotherapy. **(C)** Indicates the change of vascular maturity index (VMI) after radiotherapy. Compared with control the control group, *$p < 0.05$; Compared with control the control group, **$p < 0.01$.

more pronounced at 14 × 2 Gy irradiation dose than at 3 × 8 Gy irradiation dose. OA data of all tumors in different groups were statistically analyzed. Microvessel density (MVD, 100%) was determined as a percentage of tumor area occupied by microvessels. Number of microvessel segments was calculated as a ratio between the number of microvessel fragments and the corresponding vessel area. MVD, microvessel segments number and sO_2 were used to evaluate the effect of fractionated radiotherapy at varying dose on tumor angiogenesis. Compared with the control group, the number of blood vessels decreased, the 14 × 2 Gy group decreased to 30.89%, the 3 × 8 Gy group decreased to 19.18%. There was a significant difference in MVD between the irradiation group and the control group (Figure 4A). In addition, compared with the control group, radiotherapy decreased the number of microvessel segments, the 14 × 2 Gy group decreased by 44.78%, the 3 × 8 Gy group decreased by 17.91%, there was a more significant difference in microvessel segments number in the 14 × 2 Gy group and the increase in 3 × 8 Gy group was not statistically significant (Figure 4B). Finally, the percentage variation (SEM) of sO_2 at ROI I, ROI II and ROI III at different doses were statistically analyzed. ROI I, ROI II and ROI III in Figure 3A represents the baseline, peritumor and intratumor respectively. Compared with the control group, sO_2 in irradiation group increased remarkably. There was a significant difference in sO_2 at intratumor between 3 × 8 Gy group and the control group. Besides, there was a significant difference in sO_2 at baseline and peritumor between the irradiation group and the control group (Figure 4C).

Tumor sections were stained with CD31-α-SMA to detect tumor microvessels (Figure 5A). The red area represents CD31, the green area representsα-SMA (green), and the blue area represents the nucleus. Microvessel density (MVD) and vascular maturity index (VMI) were used to evaluate the effect of radiotherapy segmentation on tumor angiogenesis. As the tumor grew, compared with the control group, the number of blood vessels decreased, the 14 × 2 Gy group decreased to 37.52%, the 3 × 8 Gy group decreased to 11.54%. However there was a significant difference in MVD between the 14 × 2 Gy group and the control group ($p < 0.01$) (Figure 5B). In addition, compared with the control group, radiotherapy could increase VMI, the 14 × 2 Gy group increased to 40.48%, the 3 × 8 Gy group increased to 19.35%, there was a more significant difference in VMI in the 14 × 2 Gy group ($p < 0.05$) and the increase in 3 × 8 Gy group was not statistically significant ($p > 0.05$) (Figure 5C).

4 Discussion and conclusion

A detailed understanding of radiation-induced microvascular changes in tumors is important to maximize the efficacy of radiotherapy. In particular, studying the mechanism of tumor response to a single irradiation is a pressing challenge in radiobiology. A key aspect is the dynamic assessment of the subtle interplay between tumor vascular responses and changes in oxygen status and the contribution of oxygen status changes to the process of driving tumor cell death. In this study, we used high spatiotemporal resolution OA imaging to achieve noninvasive assessment of tumor microvascular morphologic and functional responses. Abundant tumor microvessels can be observed clearly in OA images. The nature of the different (size-dependent) responses of 4T1 breast tumor microvasculature to different doses of radiation as seen from the obtained OA images. Compared with the control group, the vascular density decreased and the number of vascular segments decreased after irradiation. MVD in OA images is the relative value, while that in immunofluorescence images is absolute value. Our results of the overall trend of change of MVD in OA images were roughly consistent with those in immunofluorescence images. The measurement results further showed that the tumor oxygenation level and vascular maturation index increased under 14 × 2 Gy and 3 × 8 Gy irradiation.

Many aspects of vascular biology are affected by radiotherapy and vary according to the radiotherapy protocol. Radiotherapy can cause varying degrees of endothelial cell death and redistribute neoplastic neovascularization by pruning structurally disordered/dysfunctional vessels, but such vascular changes may not translate systematically into long-term effects on the tumor. In a word, Our results indicate that different doses of radiotherapy remodeling tumor blood vessels make tumor vascular structure and function tend to be dynamic balance, improve tumor angiogenesis, and improve the internal reoxidation of tumors. Our findings show the potential of OAI in monitoring the early radiotherapy response of tumor and evaluating the changes of tumor mircovessels and microenvironment. Successful development and application of OAI for tumor imaging could potentially guide the study design of radiotherapy and assess the dose rate, dose fraction and fraction number. Scientific assessment of functional microenvironment changes in response to radiotherapy could also determine the treatment plan in cancer patients.

Data availability statement

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

Ethics statement

The animal study was reviewed and approved by the Ethics Committee of Chongqing University of Posts and Telecommunications and Kunming Medical University.

Author contributions

DW and NX designed experiments and contributed equally to this manuscript. YX, YS, YF, and LL carried out experiments. ZC, RL, and RX helped in the experiments. JY and HJ gave advice. DW supported the findings. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by grants from the Youth Fund Project of Chongqing Municiple Education Commission (Nos KJQN202000607; KJQN202000610), Chongqing post doctoral research project (special funding project, Nos 2021XM3040; 2021XM 2026) and National Natural Science Foundation of China (No. 82060313).

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

Abravan, A., Knudtsen, I. S., Eide, H. A., Løndalen, A. M., Helland, A., van Luijk, P., et al. (2017). A new method to assess pulmonary changes using 18F-fluoro-2-deoxyglucose positron emission tomography for lung cancer patients following radiotherapy. *Acta Oncol.* 56, 1597–1603. doi:10.1080/0284186X.2017.1349336

Anamitra, M., Allam, N., Jeffrey, Z. W., Demidov, V., Flueraru, C., and Vitkin, I. A. (2022). Binary dose level classification of tumour microvascular response to radiotherapy using artificial intelligence analysis of optical coherence tomography images. *Sci. Rep.* 12, 13995. doi:10.1038/s41598-022-18393-4

Beard Paul, C., and Cox Ben, T. (2016). *Biomedical photoacoustic imaging: Theory, practice and applications[M]*. Hoboken: Wiley-IEEE Press.

Brand Douglas, H., Tree Alison, C., Peter, O., van der Voet, H., Loblaw, A., Chu, W., et al. (2019). Intensity-modulated fractionated radiotherapy versus stereotactic body radiotherapy for prostate cancer (PACE-B): Acute toxicity findings from an international, randomised, open-label, phase 3, non-inferiority trial. *Lancet. Oncol.* 20, 1531–1543. doi:10.1016/S1470-2045(19)30569-8

Crane Christopher, H., and Koay Eugene, J. (2016). Solutions that enable ablative radiotherapy for large liver tumors: Fractionated dose painting, simultaneous integrated protection, motion management, and computed tomography image guidance. *Cancer* 122, 1974–1986. doi:10.1002/cncr.29878

Dai, Y., Xiang, Y., Jianshuang, W., Zeng, F., Li, Y., Yang, X., et al. (2020). Metastatic status of sentinel lymph nodes in breast cancer determined with photoacoustic microscopy via dual-targeting nanoparticles. *Light Sci. Appl.* 9, 164. doi:10.1038/s41377-020-00399-0

Duan, Y., Cheng, Z., Qiu, T., Wen, L., and Xiong, K. (2021). Spherical-matching hyperbolic-array photoacoustic computed tomography. *J. Biophot.* 14, e202100023. doi:10.1002/jbio.202100023

Gong, X., Tian, J., Wang, Y., Zhang, R., Qi, W., and Xi, L. (2022). Photoacoustic microscopy visualizes glioma-induced disruptions of cortical microvascular structure and function. *J. Neural Eng.* 19, 026027. doi:10.1088/1741-2552/ac5fcc

Huang, L., Cai, W., Zhao, Y., Wu, D., Wang, L., Wang, Y., et al. (2016). *In vivo* tumor detection with combined MR–Photoacoustic-Thermoacoustic imaging. *J. Innov. Opt. Health Sci.* 09 (5), 1650015. doi:10.1142/s1793545816500152

Jiang, H. (2014). *Photoacoustic tomography[M]*. Boca Raton: CRC Press.

Jiang, Y.-Q., Gao, Q., Han, C., Shi, X. X., Wu, J. B., Chen, Y., et al. (2021). Positron emission tomography-based short-term efficacy evaluation and Prediction in patients with non-small cell lung cancer treated with Hypo-fractionated radiotherapy. *Front. Oncol.* 11, 590836. doi:10.3389/fonc.2021.590836

Jiao, L., Andrei, C., Prakash, J., Glasl, S., Vetschera, P., Koberstein-Schwarz, B., et al. (2020). Spatial heterogeneity of oxygenation and haemodynamics in breast cancer resolved *in vivo* by conical multispectral optoacoustic mesoscopy. *Light Sci. Appl.* 9, 57. doi:10.1038/s41377-020-0295-y

Liu, L., Chen, Q., Wen, L., Li, C., Qin, H., and Xing, D. (2019). Photoacoustic therapy for precise eradication of glioblastoma with a tumor site blood-brain barrier permeability upregulating nanoparticle[J]. *Adv. Funct. Mater.* 29 (11), 1808601. doi:10.1002/adfm.201808601

McNabb, E., Al-Mahrouki, A., Law, N., McKay, S., Tarapacki, C., Hussein, F., et al. (2020). Ultrasound-stimulated microbubble radiation enhancement of tumors: Single-dose and fractionated treatment evaluation. *PLoS One* 15, e0239456. doi:10.1371/journal.pone.0239456

Morgan Tiffany, M., David, Z., Switchenko Jeffrey, M., Eaton, B. R., Crocker, I. R., Ali, A. N., et al. (2019). Fractionated radiotherapy is Associated with lower rates of treatment-related Edema than stereotactic Radiosurgery in magnetic resonance imaging-Defined Meningiomas. *World Neurosurg.* 121, e640–e646. doi:10.1016/j.wneu.2018.09.179

Park, H. J., Griffin Robert, J., Hui, S., Levitt, S. H., and Song, C. W. (2012). Radiation-induced vascular damage in tumors: Implications of vascular damage in ablative hypofractionated radiotherapy (SBRT and SRS). *Radiat. Res.* 177, 311–327. doi:10.1667/rr2773.1

Pedro, R.-B., Pablo, D.-B., AraceliFeijoo, G.-A., Neira, S., Guiu-Souto, J., and López-Pouso, Ó. (2019). A model of indirect cell death caused by tumor vascular damage after high-dose radiotherapy[J]. *Cancer Res.* 79, 6044–6053. doi:10.1158/0008-5472.CAN-19-0181

Qian, Y., Huang, S., Wu, Z., Zheng, J., Chen, X., and Nie, L. (2020). Label-free Visualization of early cancer Hepatic Micrometastasis and Intraoperative image-guided Surgery by photoacoustic imaging. *J. Nucl. Med.* 61, 1079–1085. doi:10.2967/jnumed.119.233155

Rich Laurie, J., and Mukund, S. (2016). Photoacoustic monitoring of tumor and normal tissue response to radiation. *Sci. Rep.* 6, 21237. doi:10.1038/srep21237

Turkkan, G., Bilici, N., Sertel, H., Keskus, Y., Alkaya, S., Tavli, B., et al. (2022). Clinical utility of a 1.5 T magnetic resonance imaging-guided linear accelerator during conventionally fractionated and hypofractionated prostate cancer radiotherapy. *Front. Oncol.* 12, 909402. doi:10.3389/fonc.2022.909402

Vasilis, N., and Buehler, A. (2017). Multispectral optoacoustic tomography. *Scholarpedia* 12 (3), 42449. doi:10.4249/scholarpedia.42449

Wang, L. V. (2016). WE-H-206-01: Photoacoustic tomography: Multiscale imaging from Organelles to patients by Ultrasonically Beating the optical Diffusion Limit. *Med. Phys.* 43 (6), 3847–3848. doi:10.1118/1.4958016

Wang, Z., Lu, X., Zhou, G., Yan, L., Zhang, L., Zhu, Y., et al. (2011). Multiphase-computed tomography-based target volume definition in conventional fractionated radiotherapy of lung tumors: Dosimetric and reliable comparison with the technique using addition of generic margins. *Tumori* 97, 603–608. doi:10.1700/989.10719

Wang, Z., Yang, F., Cheng, Z., Zhang, W., Xiong, K., and Yang, S. (2021). Photoacoustic-guided photothermal therapy by mapping of tumor microvasculature and nanoparticle. *Nanophotonics* 10 (12), 3359–3368. doi:10.1515/nanoph-2021-0198

Yang, J., Zhang, G., Chang, W., Chi, Z., Shang, Q., Wu, M., et al. (2020). Photoacoustic imaging of hemodynamic changes in forearm skeletal muscle during cuff occlusion. *Biomed. Opt. Express* 11 (8), 4560–4570. doi:10.1364/BOE.392221

Yang, J., Zhang, G., Shang, Q., Wu, M., Huang, L., and Jiang, H. (2020). Detecting hemodynamic changes in the foot vessels of diabetic patients by photoacoustic tomography. *J. Biophot.* 13, e202000011. doi:10.1002/jbio.202000011

Yoon Stephanie, M., Chu, F.-I., Ruan, D., Steinberg, M. L., Raldow, A., and Lee, P. (2021). Assessment of Toxic effects Associated with dose-fractionated radiotherapy Among patients with cancer and Comorbid Collagen vascular Disease. *JAMA Netw. Open* 4, e2034074. doi:10.1001/jamanetworkopen.2020.34074

frontiers | Frontiers in Physiology

TYPE Review
PUBLISHED 19 October 2022
DOI 10.3389/fphys.2022.1036621

EDITED BY
Lin Huang,
University of Electronic Science and Technology of China, China

REVIEWED BY
Zhen Yuan,
University of Macau, China
Zhu Zheng,
Shenzhen Mindray Bio-Medical Electronics Co., Ltd, China
Shuangli Liu,
Southwest University of Science and Technology, China

*CORRESPONDENCE
Yanting Wen,
1375825374@qq.com
Huabei Jiang,
hjiang1@usf.edu
Dan Wu,
wudan@cqupt.edu.cn

†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 04 September 2022
ACCEPTED 05 October 2022
PUBLISHED 19 October 2022

CITATION
Wen Y, Guo D, Zhang J, Liu X, Liu T, Li L, Jiang S, Wu D and Jiang H (2022), Clinical photoacoustic/ultrasound dual-modal imaging: Current status and future trends.
Front. Physiol. 13:1036621.
doi: 10.3389/fphys.2022.1036621

Clinical photoacoustic/ultrasound dual-modal imaging: Current status and future trends

Yanting Wen[1,2*†], Dan Guo[1†], Jing Zhang[1,2], Xiaotian Liu[1], Ting Liu[1], Lu Li[1], Shixie Jiang[3], Dan Wu[2*] and Huabei Jiang[4*]

[1]Department of Ultrasound Imaging, The Fifth People's Hospital of Chengdu, Chengdu, China, [2]School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China, [3]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, United States, [4]Department of Medical Engineering, University of South Florida, Tampa, FL, United States

Photoacoustic tomography (PAT) is an emerging biomedical imaging modality that combines optical and ultrasonic imaging, providing overlapping fields of view. This hybrid approach allows for a natural integration of PAT and ultrasound (US) imaging in a single platform. Due to the similarities in signal acquisition and processing, the combination of PAT and US imaging creates a new hybrid imaging for novel clinical applications. Over the recent years, particular attention is paid to the development of PAT/US dual-modal systems highlighting mutual benefits in clinical cases, with an aim of substantially improving the specificity and sensitivity for diagnosis of diseases. The demonstrated feasibility and accuracy in these efforts open an avenue of translating PAT/US imaging to practical clinical applications. In this review, the current PAT/US dual-modal imaging systems are discussed in detail, and their promising clinical applications are presented and compared systematically. Finally, this review describes the potential impacts of these combined systems in the coming future.

KEYWORDS

photoacoustics, ultrasound, multimodal, review, imaging, clinical

Introduction

Photoacoustic tomography (PAT) is an emerging method that provides a sub-millimeter spatial resolution image with a penetration depth of several centimeters, which is achieved by the combination of optical excitation and acoustic detection (Beard, 2011; Dai et al., 2017a; Yang and Ghim, 2021). Photoacoustic effect is induced by a nanosecond pulsed laser source. With the light illuminating the targeted tissue, it is absorbed by molecules inside the targeted tissues, leading to a temperature rise and thermoelastic expansion. In response, with the generated broadband acoustic waves, the signals are subsequently detected by conventional ultrasound (US) transducers. Multispectral optoacoustic tomography (MSOT) is based on the principle of PAT, which distinguishes absorbers based on their spectral signatures, due to various optical absorbers such as endogenous material (hemoglobin, melanin, lipids, water, and other chromophores in human) or exogenous contrast agents (Ntziachristos and Razansky, 2010; Bayer et al., 2013; Li M et al., 2018).

Over the recent years, the number of studies on PAT systems in the literature has been increasing rapidly (Su et al., 2010; Beard, 2011; Menke, 2015; Wang et al., 2016; Upputuri and Pramanik, 2017; Dong et al., 2017; Choi et al., 2018; Jo et al., 2018; Li X et al., 2018; Lin et al., 2018; Manohar and Dantuma, 2019; Nyayapathi and Xia, 2019; Steinberg et al., 2019; Zhao et al., 2019; Duan et al., 2020; Manohar and Gambhiret, 2020; Deán-Ben and Razanskyet, 2021; Gröhl et al., 2021; Li et al., 2021; Wen Y. T. et al., 2022; Yang and Ghim, 2021; Wen YT. et al., 2022). In these studies, the system designs have in common that an optimized imaging system achieved a higher spatial and temporal resolution, better penetration in tissue with reduced artifacts. Consequently, the advancements in PAT have enabled a wide applications ranging from small animal studies to clinical imaging, including imaging of breast (Becker et al., 2018; Xu et al., 2019; Yang et al., 2020), thyroid (Dima and Ntziachristos, 2016; Sinha et al., 2017; Roll et al., 2019), skin (Petri et al., 2016; Dahlstrand et al., 2020), tumors (Li et al., 2015; Yamada et al., 2020; Karmacharya et al., 2021; Knorring and Mogensen, 2021; Wang C et al., 2021), cardiovascular (Taruttis et al., 2013; Karlas et al., 2021a), functional neuroimaging (Wang et al., 2003; Wu et al., 2019a), eyes (Liu and Zhang, 2016) and others (Nagae et al., 2018; Yang et al., 2018; Liang et al., 2021; Yan et al., 2021). Therefore, PAT imaging has broader clinical translational potential than other forms of pure optical imaging, indicating its ability to provide potent structural, functional, and molecular information *in vivo* (Yao and Wang, 2018; Wu M et al., 2021). Notably, although the PAT image shows the heterogenous localization in tumors, combining US with these images can provide the exact anatomical co-localization and establish suspect region of interest (ROI), which allows for a more detailed PAT analysis of these ROIs (Manohar and Dantuma, 2019; Park et al., 2017a). Hence, it suggests the necessity of combining PAT with another structural imaging modality with deep tissue penetration to achieve feasible clinical applications.

Ultrasound, as a common imaging technique with deep tissue penetration, high spatial resolution, and properties of real-time imaging, has been widely used in clinics (Bene et al., 2022; Paola et al., 2022). Nonetheless, its non-whole body imaging, poor osseous and gas-containing organ penetration capabilities limit further development. In response to these issues, various multi-modal imaging systems have been integrated to provide complementary information, which consequently boost the sensitivity and specificity for disease diagnostics.

Since the hybrid nature of PAT makes it easy to integrate with the existing US imaging systems (Jennings and Long, 2009), several studies have been conducted on PAT/US dual-modal systems. As an obvious reason for this combination, it allows for the use of the same piezoelectric transducer, as well as the same data acquisition (DAQ) process for PAT/US signal detection (Das et al., 2009; Bouchard et al., 2014). Consequently, the data for both images is acquired simultaneously and gathered. Additionally, morphologic information produced by US imaging, including tissue or lesion boundaries (due to different sound speed and acoustic impedance) helps PAT image reconstruction and facilitates its semi-quantitative or quantitative assessment by tissue functional/molecular parameters (Beard, 2011; Dong et al., 2017).

In the last two decades, several studies on PAT/US imaging in phantom and animals were reported (Aguirre et al., 2009; Nam et al., 2012; Kruizinga et al., 2013; Karpiouk et al., 2012; Gerling et al., 2014; Luke et al., 2014; Luke and Emelianov, 2015; Mallidi et al., 2015a; Mallidi et al., 2015b; Raes et al., 2016; Salehi et al., 2016; Kang et al., 2017; Yamaleyeva et al., 2018; Dumani et al., 2019; Hartman et al., 2019; Dadkhah and Jiao, 2021; Kang et al., 2022). Moreover, by combining PAT with US imaging, it has the potential to obtain accurate and clinically relevant imaging data for diagnostic and therapy-monitoring purposes (Mallidi et al., 2015b). Several comprehensive researches for preclinical/clinical applications on PAT/US systems, such as PAT/US for healthy tissue (Lou et al., 2017; Kim and Chang, 2018; Mennes et al., 2018), tumors and metastasis (Garcia-Uribe et al., 2015; Neuschler et al., 2017; Yang et al., 2017; Becker et al., 2018; Li M et al., 2018; Shiina et al., 2018; Xavierselvan et al., 2021), bones and joints (van den Berg et al., 2017; Feng et al., 2020) and cardiovascular (Kruizinga et al., 2013; Karlas et al., 2021a; Karlas et al., 2021b; Wu M et al., 2021; Wu Y et al., 2021) were reported.

In addition, multifunctional contrast agents for PAT and US imaging were synthesized decades ago (Kim et al., 2010a). To date, as there are still barriers for clinical use of dual-modal contrast agents, most studies have been focused on the preclinical stage so far. Potentially, simultaneous PAT/US imaging enhanced by contrast agents such as microbubbles or nanobubbles can be a valuable tool for intraoperative assessment of tumor boundaries and resection margins. Thus, it will open a novel avenue of translating PAT/US imaging to clinical applications. Furthermore, given the nonionizing radiation, wide availability, portability, and low-cost nature of PAT/US imaging, this technology is thusly unique when combined with photodynamic therapy (PDT), providing multiparametric anatomical and functional information in therapeutic process. In addition, due to its label-free nature, PAT/US is capable of long-term longitudinal monitoring for image-guided treatment.

This review will focus on the progress of PAT/US systems with a potential of being translated into clinical applications. The review is organized as following: in Section 2 we discuss the overall comparison of technological advances in translational clinical PAT/US systems; in Section 3 is dedicated to PAT/US systems towards various clinical applications; in Section 4; we will briefly conclude and discuss the future directions of PAT/US for clinical translations.

PAT/US system for clinical translation

Development of PAT/US imaging systems has been proceeded in all aspects, including hardware, instrument design, and image reconstruction algorithms (Poudel et al., 2019; Tian et al., 2020). In this section, we present feasible dual-modal platforms from a perspective of radiologists and clinicians.

Graphical user interface

In order to achieve clinical utility of an imaging system, a friendly GUI of the platform is important. In recent decades, the integration technologies of PAT/US platform with friendly GUI have been extensively studied (Aguirre et al., 2009; Gerling et al., 2014; Dumani et al., 2019; Kang et al., 2022). So far, the combination of PAT and US imaging system in clinical translation has been demonstrated by different research groups (Niederhauser et al., 2005; Kolkman et al., 2006; Wang et al., 2011; Neuschler et al., 2017; Kim et al., 2020; Han et al., 2022). User-friendly operation is a critical requirement for successful clinical translation of a PAT/US system, such that radiologists or clinicians can use the system by themselves.

The first generation PAT/US system utilizes a duplex PAT imaging device with a single handheld duplex probe (Jennings and Long, 2009; Neuschler et al., 2017). The main feature of these devices are a programmable software, which allows one to modify the operation sequence of the system. There are two modes of GUI on the study devices: 1) US mode, generating only grayscale US images, and 2) PAT/US mode, generating grayscale US images fused with functional PAT data. In this way, the received PAT signals are coregistered with grayscale US signals, generating real-time and pseudo-colored maps of relative oxygenated hemoglobin (HbO_2) and deoxygenated (HbR). Accordingly, total hemoglobin (HbT) concentration can be calculated through the system. However, they suffered from the limitation of lack of real-time modification of the imaging parameters by users, hence the user-friendly GUI needs to be explored in further study to help the user to optimize the images.

A system having capability to support real-time parameter modification during imaging was designed, which can optimize the images by using the parameter control software, including time gain compensation coefficients adjustment, dynamic range for image optimization, the frequencies (center frequency, cutoff frequency, and filter size) demodulation and the beamforming options (beamforming method, apodization window type, and speed of sound) modification. Accordingly, the users could optimize the images in real-time by a MATLAB-based software without pausing the imaging operation during the image acquisition process (Kim et al., 2020). This system advances PAT/US device one more step closer towards the translation clinical use.

Design of PAT/US probes

Various research groups have implemented integrated PAT and US probes for clinical applications (Levi et al., 2014; Dai et al., 2015; Kim et al., 2016; Miranda et al., 2018; Kothapalli et al., 2019; Zhang and Wang, 2020; Pang et al., 2022). A programmable US system can utilize various US transducers (Figure 1) to generate PAT/US images (Kim et al., 2016). In order to acquire images of different depth of human *in vivo*, the PAT/US platform is fully compatible with several different probes. A novel clinical PAT/US system with various US probes were designed as follows: 1) A linear array (L3-12); 2) A convex array (SC1-6); 3) A phased array (SP1-4); and 4) An intracavity transducer (EC3-10H). Accordingly, the PAT/US system can be applied to a wide range of clinical applications by selecting suitable sensors, from superficial tissues to deep organs at different depths in human.

The design of the integrated PAT/US imaging probe is also updated for efficient light delivery and optimized the geometry of the imaging probe. An arc-array was established based on commercial US systems (iThera Medical, Germany) (Levi et al., 2014), which was suitable for small organs detection *in vivo*. Due to the relatively small field of view (FOV) of the arc-array transducer, it might not be suitable for the large region. Meanwhile, the linear-array transducers were developed for preclinical and clinical trials, including iU22 (Philips Healthcare) (Garcia-Uribe et al., 2015) and EC-12R (Alpinion Medical Systems, Republic of Korea) (Kim et al., 2016). However, these high-frequency transducers suffered from a shallow imaging depth, which were not suitable for deep organ detection, thus were limited for general clinical applications.

The developed system can be used to provide new biological information in diagnostic fields such as uterus, bowels, vascular and organs adjacent to cavity viscera (prostate, pancreas, etc.). The endoscopic image modalities have been proposed (Dai et al., 2017b; Li et al., 2019). Emerging endoscopy techniques refers to PAT is photoacoustic endoscopy (PAE), which incorporates PAT in a small probe to visualize internal organs through intra-cavity introduction (Wang, 2008). In addition to PAE, researchers are exploring a PAT/US "mini-probe" integrated into a conventional endoscope's instrument channel, which is a promising strategy to achieve clinical benefits (Yang et al., 2012). In this study, a simultaneous PAE/EUS dual-modal system with a streamlined shape probe (3.8-mm-diameter) was designed to image internal organs *in vivo* (Yang et al., 2012). However, it could not provide a

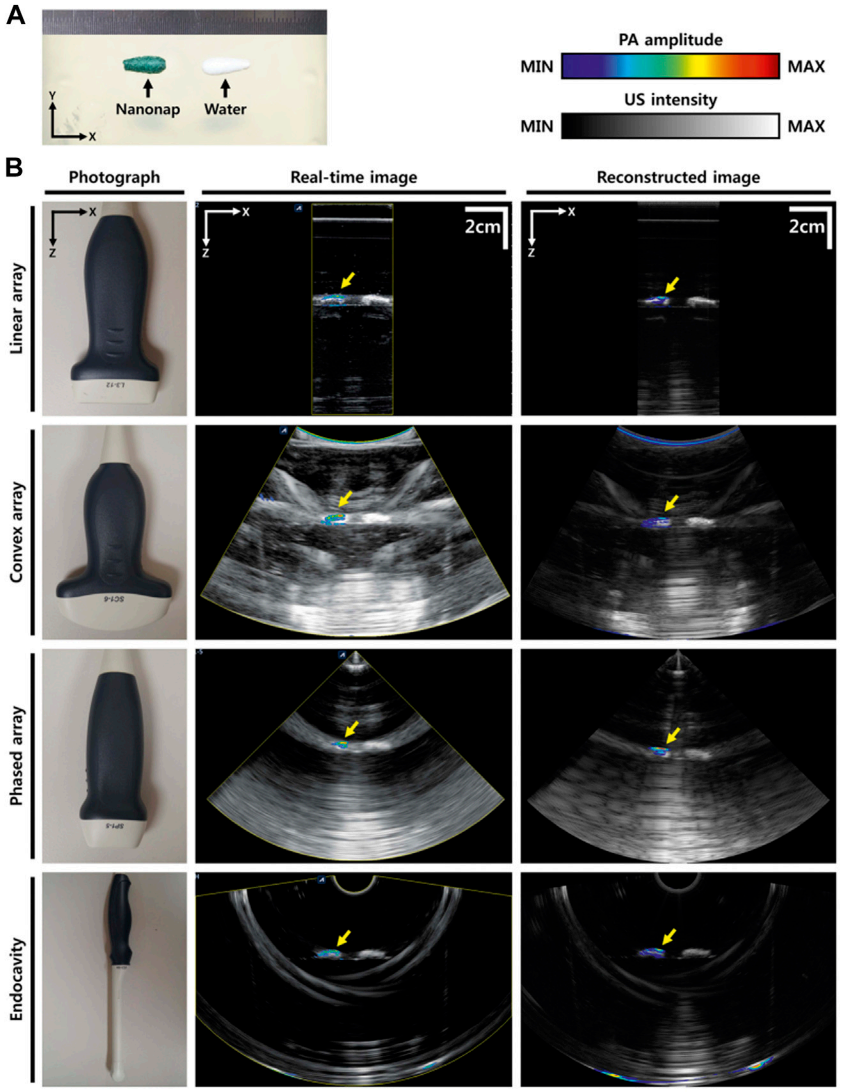


FIGURE 1
PAT/US images using different transducers. **(A)** Photograph of the phantom. **(B)** Photographs of different transducers (linear array, convex array, phased array, and intracavity transducers), real-time PAT/US images, and reconstructed PAT/US images acquired by these transducers. PAT, photoacoustic tomography; US, ultrasound. Reprinted with permission from [Kim J., Park S., Jung Y., Chang S., Park J., Zhang Y., et al. (2016). Programmable Real-time Clinical Photoacoustic and Ultrasound Imaging System. Sci Rep.6,35137. doi: 10.1038/srep35].

360° field of vision which is often needed in clinical settings. Later, a miniaturized PAE/EUS system with a 2.5-mm-diameter probe was conducted (Li X et al., 2018). This catheter provided a full (360°) field-of-view cross-section images, which was comparable with the 2.8-mm instrument channel of conventional clinical endoscopes. A novel intrauterine PAT/US imaging probe (linear array, 15-MHz) was designed to detect endometrial diseases *in vivo* (Miranda et al., 2018). For application within the human uterus, the intra-cavity probe (2-mm-diameter) is comparable with an endometrial suction curette, which is a catheter-like device with a diameter of <3 mm. As a custom designed probe for deep tissue, the imaging depths up to several centimeters have been achieved. A miniaturized capacitive micromachined ultrasonic transducer (CMUT) array for simultaneous imaging of transrectal PAT/US system were combined for human prostate detection *in vivo* (Kothapalli et al., 2019).

Moreover, with the system equipped with a three-dimensional (3D) detection aperture, the spiral 3D images were reconstructed for PAT/US images. Compared to traditional sensors, this transducer had a higher bandwidth, and the signal-to-noise ratio (SNR) was improved as well. It has the potential to be introduced in clinics in the future. As well, current acoustic-resolution PAE/USE generally employs a point-

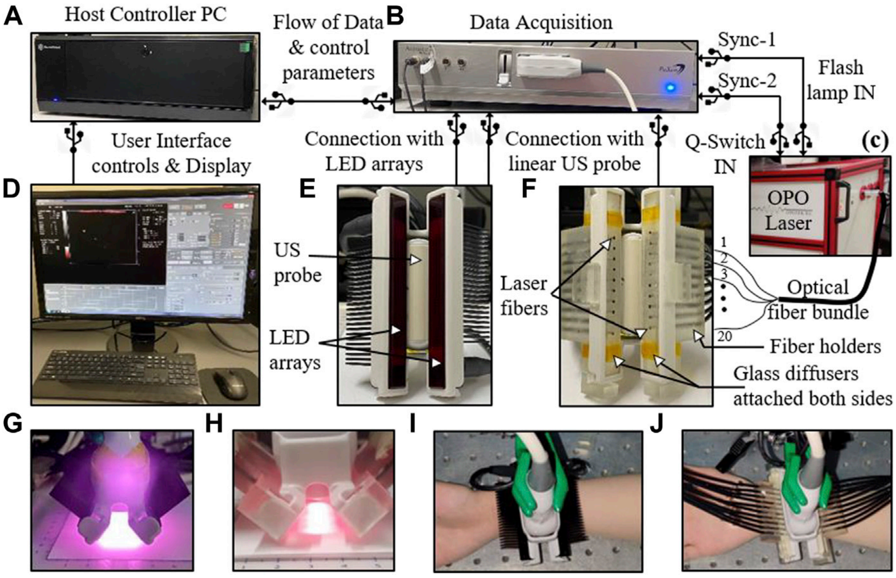


FIGURE 2
The experimental setup designed for comparing light-emitting diode (LED)-based and high-power laser-based PAT and US imaging. **(A)** A host controller PC; **(B)** DAQ hardware; **(C)** A portable high-power laser with its output coupled to the input end of an optical fiber bundle; **(D)** Computer display: displays PAT (red scale), B- mode US (grayscale), and coregistered PAT/US (overlaid red PAT on gray US); **(E)** Arrangement of two 850 nm LED arrays around the US probe; **(F)** Arrangement of twenty laser fibers inserted into the two fiber holders around the US probe; **(G,H)** Optical illumination profile achieved with two LED array sources and laser sources, respectively; **(I,J)** Pictures of a human wrist under imaging with the LED and laser arrangements, respectively. LED: light-emitting diode; PAT, photoacoustic tomography; US, ultrasound; DAQ: data acquisition; PC: personal computer. Reprinted with permission from [Agrawal S., Singh M. K. A., Johnstonbaugh K., Han D. C., Pameijer C. A., and Kothapalli S-R. (2021). Photoacoustic Imaging of Human Vasculature Using LED versus Laser Illumination: A Comparison Study on Tissue Phantoms and *In vivo* Humans. Sensors. 21,424. doi: 10.3390/s21020424].

focused transducer which is only sensitive in its focal region. As a result, the sensitivity and lateral resolution dramatically reduce when the targets move out of its focus. A designed line-focused transducer emits a more uniform sound field, as compared to a point-focused transducer, resulting in the original signal intensity and SNR of the adjacent targets to be closer in the radial direction, which improves the uniformity of target signals in hybrid imaging (Pang et al., 2022). Further study is need to evaluate the diagnostic ability and accuracy of abovementioned PAE/USE transducers in larger clinical trials.

Miniaturization of PAT/US device

In order to achieve certain clinical applications, miniaturized PAT/US imaging system were designed by several groups (Kothapalli et al., 2019; Kim et al., 2020). For instance, a devised portable PAT/US system which can visualize vascular distribution without injecting any contrast agent was designed (Kim et al., 2020). While US and certain optical technologies are available in the size of a mobile phone (Ahn et al., 2015), PAT/US systems still need to be improved to reach that level of portability, mainly focusing on advanced laser source and high performance DAQ system.

Conventionally, a high-power laser source is employed for most dual-modal PAT/US systems (Kothapalli et al., 2019; Regensburger et al., 2019; Agrawal et al., 2021a). In order to translate PAT/US technology to clinical applications, especially for point of care (POC) detection, a significant reduction in both cost and size of laser source is required. To address this issue, several cost-effective alternatives for the optical excitation/detection (Liu and Zhang, 2016; Upputuri and Pramanik, 2017 have been explored. Recently, the light-emitting diodes (LEDs) have emerged as portable optical sources for PAT *in vivo* (Anas et al., 2018; Jo et al., 2018; Xia et al., 2019; Farnia et al., 2020; Maneas et al., 2020). However, these state-of-art LED arrays carry significantly lower optical energy (<0.5 mJ/pulse) and high pulse repetition frequencies (PRF) (4 KHz) compared to the high-power laser sources (100 mJ/pulse) with low PRFs of 10 Hz. To enhance the performance of LEDs, an arrayed arrangement of LED elements was developed (Zhu et al., 2018; Zhu et al., 2020), thereby increasing the pulse energies from a few μJ to hundreds of μJ. In addition, higher PRF of the LED allowed a sufficient and fast PAT signal averaging which led to significant SNR improvements for deep tissue imaging. A commercial LED based PAT/US system (Figure 2) was designed (Agrawal et al., 2021b) with a lower mean noise compared to the laser based PAT/US system. However, the SNR value for the laser-based PAT image was about 6 dB lower than the SNR with the LED array acquisition. Therefore, due to the low power of LED arrays, a higher frame averaging is required to image deep tissue targets. Nevertheless, LED-PAT/US systems have strong potential to be a mobile health care technology for clinical applications.

Toward clinical translation PAT/US device

In order to achieve highly translatable to the clinical field, several PAT/US systems based on commercial US system were conducted without the need of additional hardware and algorithms to obtain data and reconstruct images, making it compatible with most commercial US platforms (Park et al., 2021). However, direct use of a US commercial system to acquire quality PAT image is difficult. The most challenging issue is how to recover weak PAT signals from expected subsurface tissues, since the signal level in US are generally much higher than those in PAT, and neither the SNR nor the digitizer threshold of most commercial US systems are suitable for directly PAT imaging.

At the initial stage towards a clinical PAT system, several commercial PAT/US platforms have been used in preclinical applications, including the iU22, Phillips Healthcare, Netherlands (Kim et al., 2010b), the VevoLAZR series (FujiFilm VisualSonics, Canada) (Needles et al., 2013), the MSOT Acuity series (iThera Medical, Germany) (Levi et al., 2014) and the Vintage series (Verasonics, United States) (Kothapalli et al., 2019). However, for general clinical research, these systems suffered from several limitations, such as immobile laser, unprogrammable US machine and without approvement of the United States Food and Drug Administration (FDA). Later, in order to overcome the limitations, a clinically applicable PAT/US system was developed (Kim et al., 2015). A portable pulsed laser (Phocus, OPOTEK, United States) and a FDA-approved clinical US machine (EC-12R, Alpinion Medical Systems, Republic of Korea) were integrated in the system. However, the abovementioned systems only combine PAT with US for imaging, without utilizing other optical imaging modalities such as photoacoustic microendoscopy (PAM) imaging, which providing microvascular networks map in superficial tissue. Consequently, a linear transducer combined PAT, PAM, and B-mode US imaging into one commercial US platform was design (Wang S. L et al., 2021). As compared to existing multi-modality systems based on PAT and US, this system provides more complementary morphological and functional information of tissue *in vivo*. It has potential to achieve the best benefits of integrated PAT/US and promised for multi-scale and multi-functional imaging for clinical applications in the future.

Clinical applications of PAT/US system

For clinicians or radiologists, “seeing is believing”. In the last several years, several clinical studies on PAT/US dual-modality were reported, including human thyroid (Dima and Ntziachristos, 2016; Yang et al., 2017; Kim et al., 2021), breast (Garcia-Uribe et al., 2015; Becker et al., 2018; Neuschler et al., 2017; Nyayapathi and Xia, 2019; Kelly et al., 2020; Goh et al., 2019; Yang et al., 2020), skin (Oraevsky et al., 2018; Park et al., 2021), extremities (Xia et al., 2015; Mercep et al., 2015; Liu and Zhang, 2016; Kim et al., 2016; Jo et al., 2017; Oeri et al., 2017; van den Berg et al., 2017; Feng et al., 2020; Daoudi et al., 2021), prostate (Agrawal et al., 2020; Kothapalli et al., 2019), bowels (Knieling et al., 2017; Leng et al., 2018; Yang et al., 2019), vascular (Karpiouk et al., 2012; Wu et al., 2015; Andrei et al., 2021), placenta (Xia et al., 2015; Maneas et al., 2020) and others (Jose et al., 2009; Gonzalez et al., 2021; Mozaffarzadeh et al., 2021). With the combination of the two modalities in one imaging system, it is acceptable for radiologists or clinicians to adapt and associate morphological features with functional information. On the other hand, US detection is naturally registered with PAT detection; therefore, researches using US-guided PAT detection were included in this review. We compared the properties of each dual-modality system from device performance to clinical trials, and fully discuss their advantages and disadvantages, as well as possible clinical applications in future.

Thyroid imaging

Thyroid tumors are common tumors in the head and neck. Newly-diagnosed thyroid cancer cases have increased due to advancements in diagnostic imaging techniques such as US, X-rays, and magnetic resonance imaging (MRI) (Vaccarella et al., 2016). Although the incidence of malignant cases of all the discovered thyroid nodules were about 10%, not all nodules need to be treated immediately. Hence, it is estimated that over 560,000 patients were overdiagnosed over the last two decades (Ahn et al., 2014; Vaccarella et al., 2016). The American Thyroid Association recommended US as a routine thyroid examination for all patients with thyroid lesions and for healthy people (Gharib et al., 2016). However, conventional color Doppler US has limited capacity in discriminating untypical benign and malignant nodules. An accurate diagnosis of thyroid disease can be aided by reliable vascular information. In recent years, contrast-enhanced ultrasound (CEUS) has been applied for the clinical evaluation of the thyroid nodule. However, CEUS is rather invasive for the intravenous injection of contrast agent. Therefore, a noninvasive functional imaging modality, with the ability of evaluating the morphological and functional information simultaneously, will be beneficial to the early diagnosis and clinical management of thyroid tumors. PAT is a novel hybrid imaging modality, which relies on sensitive tissue optical properties. Therefore, PAT can provide important functional information, such as the oxyhemoglobin saturation (SO_2).

PAT combined with US can provide important clue in the diagnosis of thyroid disease with reliable vascular information in

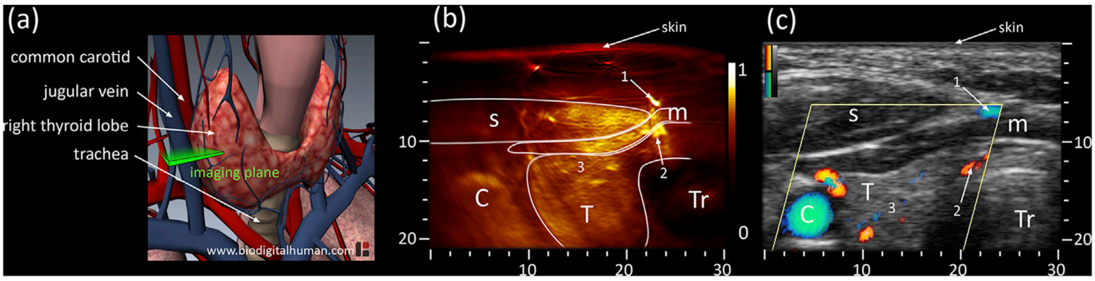


FIGURE 3
(A) Anatomy of thyroid including cardio-vascular and respiratory system; the cross-sectional imaging plane is highlighted in green. **(B)** PAT and **(C)** US cross-sections of the left thyroid lobe of the first volunteer. PAT: Photoacoustic tomography, US: Ultrasound; C: Carotid T: Thyroid, Tr: Trachea, s: sternocleidomastoid muscle, m: infrahyoid muscle; axes in mm. Reprinted with permission from [Dima A., Ntziachristos V. (2016). *In-vivo* handheld optoacoustic tomography of the human thyroid[J]. Photoacoustics, 4(2):65–69. doi: 10.1016/j.pacs.2016.05.003].

an initial clinical study (Banaka et al., 2011). Then some thyroid nodules studies have been conducted using costumed PAT/US system (Dima and Ntziachristos, 2016; Yang et al., 2017; Kim et al., 2021). A curved US array (Dima and Ntziachristos, 2016) and linear array (Yang et al., 2017) were employed to delivery high-fidelity performance in human thyroid *in vivo*, respectively. Both studies demonstrated that it was possible to detect the thyroid's outline and identify vascular features. In addition, it was found that PAT was more efficient at detecting blood vessels compared to colored Doppler US (Figure 3). Later, multi-spectral acquisition was used to further improve accuracy by tissue oxygenation parameter in thyroid imaging (Kim et al., 2021). All the single parameter analyses showed encouraging results with statistically differentiable distributions. Furthermore, they successfully visualized real-time PAT/US images of thyroid nodules.

Taken together, previous clinical trials have demonstrated that PAT with a high-quality clinical US system can provide high-quality morphological and functional images in human thyroid *in vivo*. Subsequent developments are needed to further update the dual-modal system in several regards: 1) A large FOV is required for efficient navigation during monitoring and needle guidance; 2) improving the duplex probe for easier hand-held operation; 3) To improve PAT SNR, optimized laser delivery is needed to achieve higher spatial resolution and less artifacts; 4) Motion artifact causing by the target and operator movement between multiple laser pulses; 5) Analyses of MSOT parameters and comparisons between different types of thyroid cancer needs further study; 6) Multispectral parameters processed and displayed in GUI in real time.

Breast imaging

In 2022, over 110,000 new cases of breast cancer will be diagnosed in the United States, making it the most common cancer for women worldwide (Porter, 2009; Ghoncheh et al., 2016; Siegel et al., 2021). Conventional breast imaging techniques, like X-ray mammography and US, primarily focus on morphological changes of breast tissue to discriminate benign from malignant tissue. X-ray mammography is not reliable for women with dense breasts with ionizing radiation and US is strongly operator dependent (Heijblom et al., 2012). The hybrid nature of PAT breast imaging provides both structural information and hemoglobin-related functional information within the breast, which can aid clinical diagnosis. In addition, since breasts have protruding geometry in the superficial region, they are optically transparent compared to other organs, making them ideal to image with PAT. Several clinical studies indicate that angiogenesis begins at an early stage of breast carcinoma *in situ*, with this understanding, the dual-modal PAT/US imaging based on two different contrast mechanisms (functional optical and anatomical US) can achieve greater clinical performance with a merit of radiation-free, breast-compression-free, and relatively inexpensive.

In recent years, combined PAT/US technology has demonstrated its clinical feasibility in human breast cancer diagnosis *in vivo* (Becker et al., 2018; Oraevsky et al., 2018; Han et al., 2022) and (Goh et al., 2019) *ex vivo* in humans. Advanced systems explored for real-time PAT/US breast imaging with high temporal feature have been designed (Becker et al., 2018; Oraevsky et al., 2018). The individual images are reconstructed at possible rate of 25 Hz for single wavelength imaging and the MSOT image is possible at a 5 Hz refresh rate per multispectral image (Becker et al., 2018). MSOT/US image was conducted to acquire functional information of breast cancer, revealing increased signals for HbR and HbO_2 in invasive breast carcinoma (Figure 4). Another real-time PAT/US system named Imagio™ was developed for testing in a multicenter clinical trial termed PIONEER (Oraevsky et al., 2018). A spatial-temporal coregistration of functional and anatomical images is explored in clinical trials (Figure 4). In

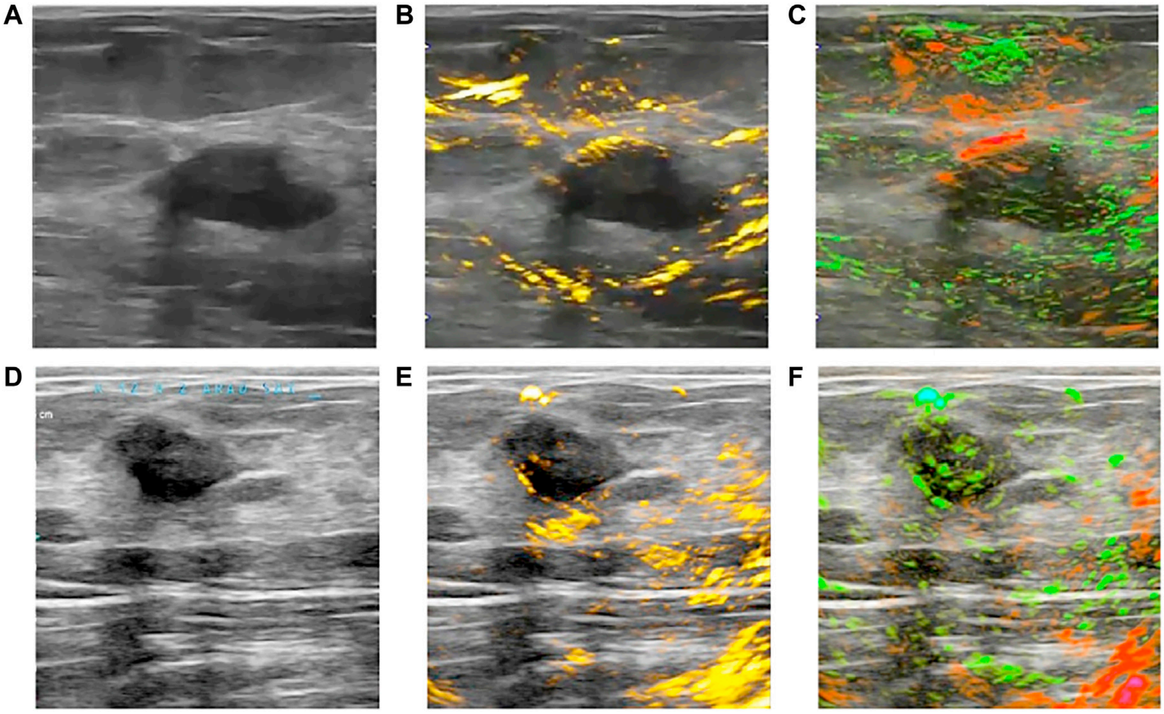


FIGURE 4
Clinical OA/US imaging of two benign breast tumors. **(A,D)** show US of two breast tumors; **(B)** reveals a lack of angiogenesis microvasculature within the breast tumor in OA/US image; **(C)** shows the majority of the tumor is normally oxygenated in OA/US image; **(E, F)** display another similar breast tumor in OA/US image. OA: optoacoustic; US: ultrasound. Reprinted with permission from Oraevsky et al. (2018).

this system, radiologists can evaluate the vascular pattern around tumors, the microvascular density of lesions, and the relative values of HbT, SO_2 to adjacent tissues. While the Imagio™ lacks the advantages of US, such as speed of sound and acoustic attenuation measurement, further study would focus on improvement of both PAT and US performance on this system (Zalev et al., 2012; Oraevsky et al., 2018; Stephens et al., 2021).

Large sample study and data set for PAT/US system is as a bridge from trials to clinical applications. Analyses in two large samples of PAT/US breast imaging were performed (Neuschler et al., 2017; Kim et al., 2020), enrolling 2105 women and 2000 patients, respectively. A comparison on Breast Imaging Reporting and Data System (BI-RADS) categories using PAT/US versus US alone was completed (Neuschler et al., 2017). With a similar sensitivity (US: 96% vs. PAT/US: 98.6%), the specificity of PAT/US exceeded that of US by 14.9%. While PAT will not replace any US functions, but provides important complementary information for US imaging. The result demonstrated it may be possible to improve the specificity of breast mass assessment by using PAT/US, thereby reducing the number of false-positive examinations and biopsies. However, it also had limitation that led to some false-negative interpretations. Later, another multicenter clinical trials demonstrated the clinical feasibility with a hand-held duplex technology (Kim et al., 2020). Complex signal processing and image reconstruction algorithms in the software enable real-time coregistration of PAT/US imaging.

In comparison to two-dimensional (2D) functional PAT/US imaging, 3D functional PAT/US imaging has several advantages. Using 3D PAT/US imaging can provide quantified results based on 2D PAT imaging, which can better represent the overall functional imaging features of breast tumor. To verify this, Yang et al. explored a quantitative method to analyze characteristics of breast tumors using 3D volumetric data obtained from a 3D PAT/US functional imaging system (Yang et al., 2020). The analysis of the 3D distribution of vessels could provide a more comprehensive description of the tumor vasculature than 2D analysis. Furthermore, 3D quantification of PAT functional information may be able to minimize intra-observer differences compared to previous 2D PAT/US imaging studies. Limitations of this research include the “limited view” problem, causing most of the reconstructed vessels to have an orientation that tended to be parallel with the scanning direction.

Taken together, PAT/US imaging has provided meaningful information for a radiologist to accurately diagnose malignancy.

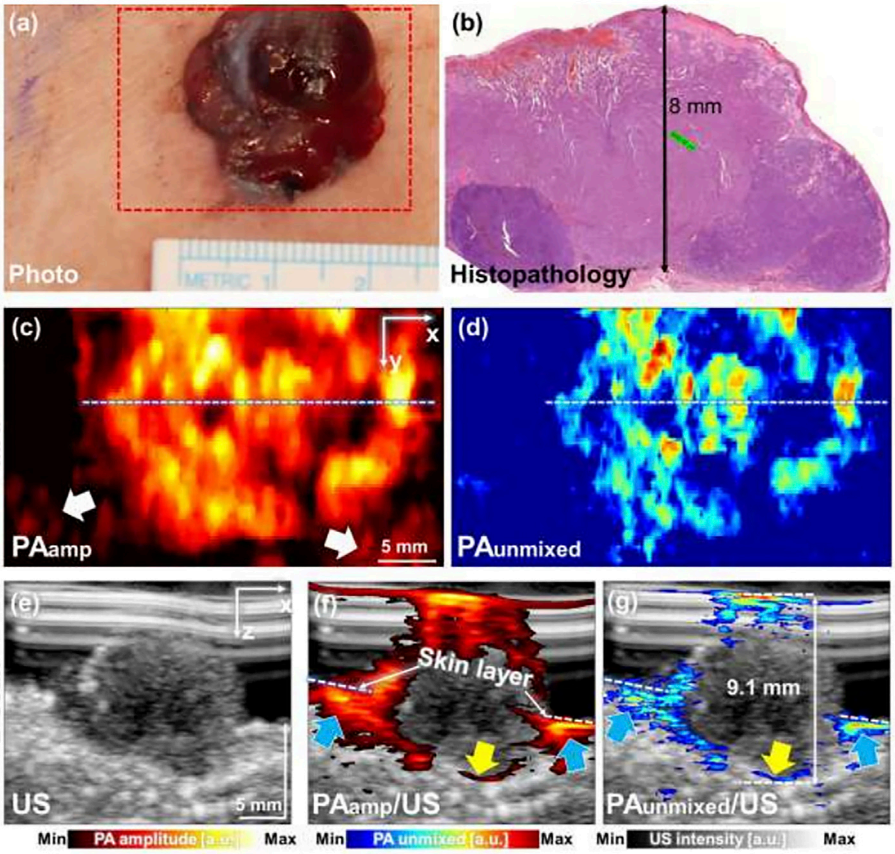


FIGURE 5
Combined PAT/US imaging of skin cancers. PAT image of a skin melanoma acquired simultaneously with a US image. **(A)** Clinical photo and **(B)** histopathology of melanoma on human chest, **(C)** PAamp MAP and **(B)** PAunmixed MAP images along the area of red dashed box in **(A)**; **(E)** US image of melanoma; **(F)** PAamp/US image and **(G)** PAunmixed/US image along the white dashed lines in **(C,D)**, respectively. PAT: photoacoustic tomography; US: ultrasound; PAamp, photoacoustic amplitude; PAunmixed, photoacoustic unmixed melanoma; MAP, maximum amplitude projection. Reprinted with permission from Park et al. (2021).

Notably, for patients with breast cancer, PAT/US is accelerating its clinical translation in macroscopic and microscopic imaging. To define the role of PAT/US in clinical practice, further research should be conducted into feature analysis and interpretation strategies. Moreover, further technical advances of the technology will be envisioned in the direction of quantitatively accurate PAT/US image and the 3D PAT/US systems with large FOV for human breast.

Skin imaging

Skin cancer is one of the most common types of cancers affecting around one out of five people in most developed countries (Deán-Ben and Razanskyet, 2021). The most aggressive type of skin tumors is malignant melanomas. The melanoma patients have a very poor prognosis if it is not identified and treated in early stage accurately. Recently, noninvasive imaging techniques, such as high-frequency ultrasonography (HFUS), reflectance confocal microscopy (RCM), and optical coherence tomography (OCT) have been developed to improve the diagnostic sensitivity and accuracy for skin melanoma (Heibel et al., 2020). However, those methods are neither sufficient to measure the accurate depth of the melanoma nor accurately estimate the real invasive depth of the tumor (Crisan et al., 2013). PAT image is gaining great attention as a noninvasive and nonionizing diagnostic method to visualize skin melanomas, due to the presence of strong melanin contrast in tumor. The spectral PAT images were also integrated to a pulse-echo US image serving as anatomical reference. In the last couple of years, several groups have conducted experiments of skin melanoma in animals to confirm its feasibility using PAT imaging system (Neuschmelting et al., 2016; Hindelang et al., 2019).

Detection and quantification of melanoma depth have been reported with several types of PAT/US imaging systems *in vivo* (Li M et al., 2018; Wang S. L et al., 2021; Park et al., 2021). This unique PAT/US imaging here opens unprecedented capabilities for high-resolution skin imaging at scalable depths *in vivo*. A pilot study showed six melanoma patients examined *in vivo* using the 3D MSOT imaging system (Figure 5). By using a MSOT/US system, melanoma of various sizes, locations (chest, thigh, heel, feet, and palm) and forms (1.3–30 mm lateral diameter, 0.6–9.1 mm depth) were detected by US technology. Feeding vessels were visualized in the melanoma using 3D PAT image, suggesting the neovascularization in the tumor. An analysis of those MSOT data confirmed a high correlation between the depth of the melanoma and its histology (Park et al., 2021). Thus, this PAT(MSOT)/US system, in particular with 3D reconstruction, will possibly serve as an important noninvasive imaging tool in determining the stage of skin cancer, in deciding the excision region of the cancer in surgery. Furthermore, it would improve the prognosis of the skin melanoma patients in near future. However, these works were still preliminary attempt for the diagnosis. In order to further explore the potential of PAT/US in early diagnosis of melanoma, more experiments *in vivo* are necessary with different subtypes and stages of melanoma. Further, deep learning (DL) and artificial intelligence (AI) algorithms can be combined to detect the invasion depth and boundary of melanoma more precisely.

Extremities imaging

Monitoring the microcirculation in target tissue is crucial in assessing bone diseases, inflammation of the synovium and peripheral vascular diseases, such as diabetic foot, synovitis, rheumatoid arthritis, and arterial embolization in lower extremity. However, conventional imaging modalities are focused on diagnosis in major arteries, and are limited to provide microvascular information in early stages of the disease. PAT imaging reliably quantify vascular parameters noninvasively in human extremities (Kruizinga et al., 2013; Xu et al., 2013; Jo et al., 2017; van den Berg et al., 2017; Choi et al., 2022). However, the localization of the PAT signal may require another modality. A unique advantage of US is its good localization, non-ionizing nature, and ability to penetrate soft tissues. Given the abovementioned advantages of PAT and US technique, PAT/US image system has potential for clinical extremities health assessment.

To facilitate clinical use, a costumed 3D PAT/US system has been conducted for finger imaging in a healthy human *in vivo* (Oeri et al., 2017). Tomography consists of four separate and fully automated removable curved sensors that can image all three finger joints. This study has provided new opportunities in finger diagnostics. However, noises (similar frequency as signals) were observed in raw data of some arcs, yielding streaking artifacts in reconstructed images in certain regions.

Recently, researchers challenged the use of US-guided PAT to visualize human bones (Feng et al., 2020). The results suggested that it can distinguish PAT signal of human cortical and trabecular bones *in vivo*, as well as the surrounding soft tissue. However, this work has not provided a quantitative assessment based on trabecular bone PAT signal. And then, MSOT/US was applied for assessment of 17 systemic sclerosis (SSc) patients (5 out of 17 was in early phase) with nailfold damage, 5 primary Raynaud's phenomenon (PRP) and 9 health controls (Daoudi et al., 2021). Since MOST can quantitatively evaluate capillary density and hemoglobin (Hb) contents of the third fingers and US is capable of measure skin thickness of the lesions, this hybrid method could help to distinguish early SSc from PRP individuals and health controls in both Hb contents and skin thickness.

Furthermore, with the increase need of fast diagnosis devices, POC technique has gained popularity and also been studied for possible application in clinical settings. The feasibility of a portable PAT/US system was evaluated for clinically evident synovitis (van den Berg et al., 2017). The proximal interphalangeal joints of the inflamed and non-inflamed joints of ten patients were examined and compared with the joints of 7 healthy volunteers. PAT scan, power Doppler US (US-PD) were performed (Figure 6). The PAT probe in this study is sensitive to vessels and vascular networks of 0.2 mm in size. PAT signals in inflamed joints increased significantly, compared with contralateral non-inflamed joints and with joints from healthy volunteers, which was highly correlated with US-PD (Figure 7). Therefore, combined PAT with US using a compact handheld probe is able to detect clinically manifest synovitis. However, one of the technical limitations of this system was the lack of shared access to high-quality PAT and US images. The short delay between the two may have contributed to inaccuracy caused by the unexpected movement of the fingers. The restriction on this setting will be addressed in future versions, resulting in almost simultaneous access to PAT and US images.

Future applications of PAT/US system in extremities can take advantage of its multi-spectral imaging capabilities, allowing the estimation of the SO_2 of lesion tissues. In next, quantitative parameters acquired from MOST on different bone mineral densities of long bones will be available. The bone and joint evaluation platform is needed to provide both microstructural and metabolic information, which is highly valuable for diagnosis and grading as well as monitoring of osteoporosis therapy and other osteoarticular diseases. Furthermore, targeted PAT/US contrast agents based on molecular markers need to be investigated, providing information about inflammation similar to positron emission tomography.

Large blood vessels imaging

Lipids in plaques are an important marker of atherosclerosis. Among endogenous contrast compositions,

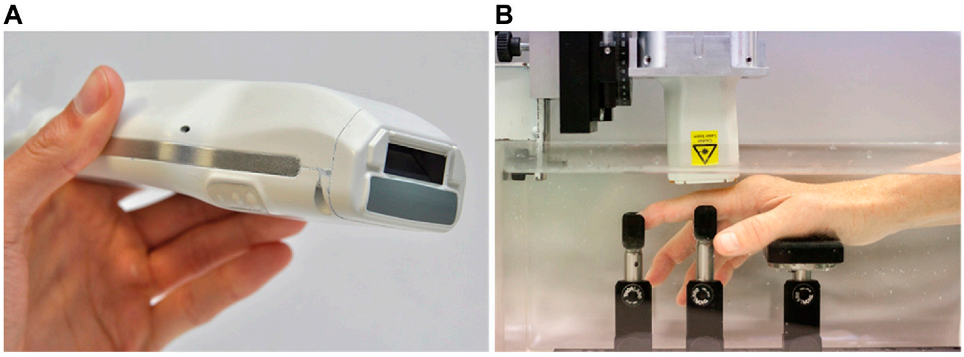


FIGURE 6
(A) A PAT/US probe with a front-end view showing the light delivery window (dark aperture) and gray acoustic lens in medium; **(B)** The patient's hand is submerged in water and is supported by a series of braces. The sensor is mounted on a motorized 2-axis stage and positioned above the joint. PAT: photoacoustic tomography; US: ultrasound. Reprinted with permission from [van den Berg P. J, Daoudi K., Bernelot Moens H. J., and Steenbergen W. (2017). Feasibility of photoacoustic/ultrasound imaging of synovitis in finger joints using a point-of-care system. Photoacoustics. 8,8–14. doi: 10.1016/j.pacs.2017.08.002].

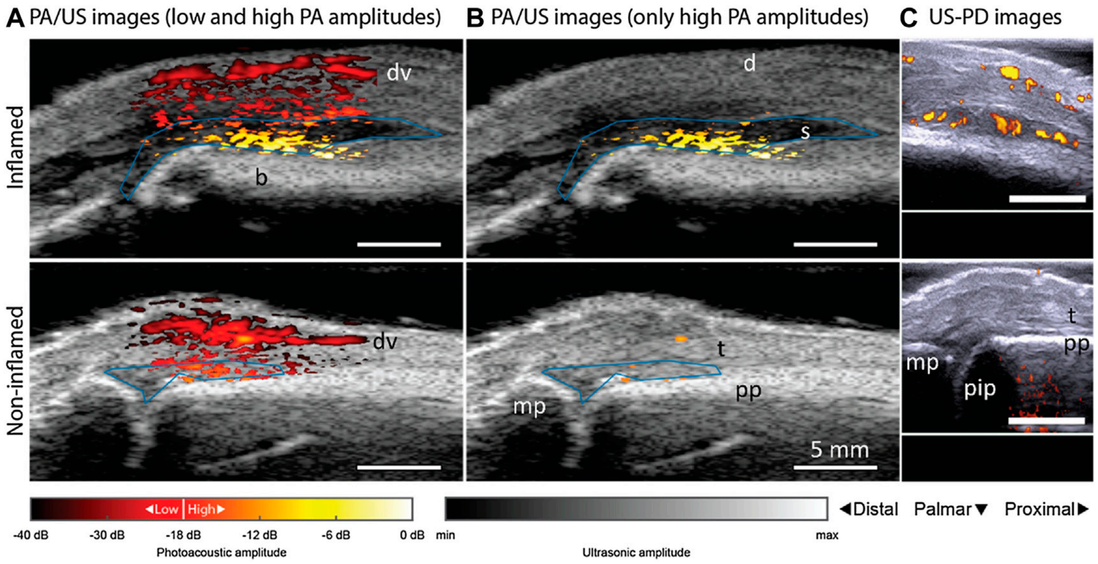


FIGURE 7
PAT/US and US/PD images of an inflamed contra-lateral (bottom row) and non-inflamed (upper row) joint from a patient with rheumatoid arthritis. PAT/US images in **(A)** show a difference in color between inflamed and non-inflamed corresponding to an increase in amplitude levels. If low PAT amplitudes are discarded in **(B)**, only features in the inflamed joint are seen. Figure **(C)** shows the corresponding US-PD images. Blue lines in PAT/US images indicate ROIs used for PAT feature quantification in synovial space. The 0 dB level is the maximum PAT amplitude produced by an inflamed joint. d, dermis; dv, dorsal vein; pp, proximal phalanx; pip, proximally located interphalangeal joint; mp, midline phalanxes; s, synovium; t, extensor tendon; PAT, photoacoustic tomography; US, ultrasound; PD, powered Doppler. Reprinted with permission from [van den Berg P. J., Daoudi K., Bernelot Moens H. J., and Steenbergen W. (2017). Feasibility of photoacoustic/ultrasound imaging of synovitis in finger joints using a point-of-care system. Photoacoustics. 8,8–14. doi: 10.1016/j.pacs.2017.08.002.

lipid is one of the most commonly used PAT biomarkers and has been intensively studied. An intravascular PAT/US probe was applied to visualize blood lipid in arteries, which was able to successfully detect and distinguish plaque lipids in human coronary arteries from adventitial fat *ex vivo* (Wu et al., 2015). This imaging technique has demonstrated its ability to identify features of plaque instability but come with limitations, such as the use of contrast agents, long examination times and poor portability. Recently, five patients with carotid atherosclerosis, five healthy

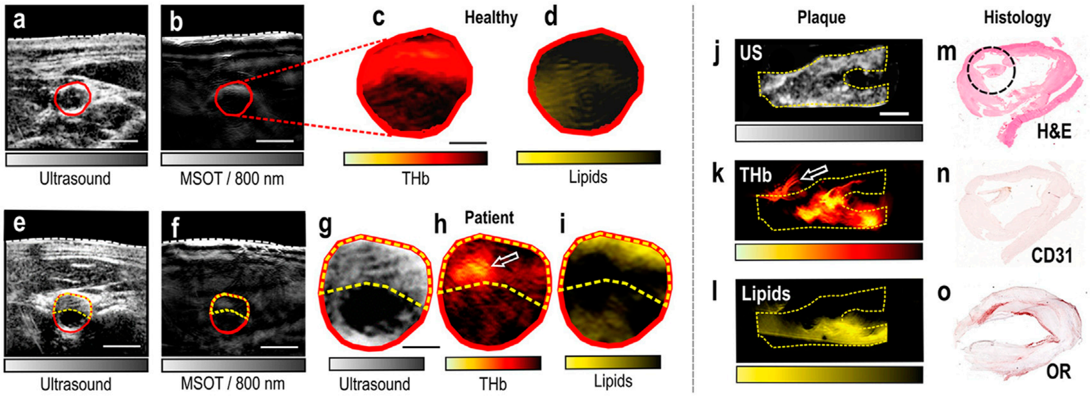


FIGURE 8
MSOT/US imaging of the carotid artery of a healthy volunteer and a patient, MSOT/US imaging and histology of an excised plaque. **(A)** US image of the transverse section of the carotid artery of a healthy volunteer. **(B)** Corresponding MSOT image at 800 nm. White dashed line: skin surface. The carotid lumen area is demarcated in red. Scale bars: 5 mm. **(C)** Magnification of the lumen area of the spectrally unmixed MSOT image in **(B)** showing the HbT signal. **(D)** Same magnification showing the Lipids signal. Scale bar: 2.5 mm. **(E)** US image of a patient with carotid atherosclerosis. **(F)** Corresponding MSOT image at 800 nm. White dashed line: skin surface. The lumen is demarcated in red and the plaque area with a yellow dashed line. Scale bars: 5 mm. **(G)** Magnification of the arterial cross-section in US. **(H)** Magnification of the arterial cross-section of the spectrally unmixed MSOT image in **(F)**, showing the HbT signal. White arrow: region of increased HbT content. **(I)** Same magnification showing the Lipids-signal. Scale bar: 2.5 mm. **(J)** Sagittal view of the excised plaque in US. **(K)** Same view of the plaque in corresponding unmixed HbT image. White arrow: postoperatively attached suture. **(L)** Same view of the plaque in unmixed MSOT Lipids image. Scale bar: 4 mm. **(M)** Histological view with H&E staining. The black circle shows a region with thrombotic and erythrocytes components. **(N)** Histological view with CD31-staining targeting the neovascularization. **(O)** Histological view with OR-staining showing the lipid content of the plaque. MSOT: multispectral optoacoustic tomography; US: ultrasound. Reprinted with permission from [Karlas A., Kallmayer M., Bariotakis M., Fasoula N. A., Liapis E., Hyafil F., et al. (2021). Multispectral optoacoustic tomography of lipid and Hb contrast in human carotid atherosclerosis. Photoacoustics.23,100283. doi: 10.1016/j.pacs.2021.100283].

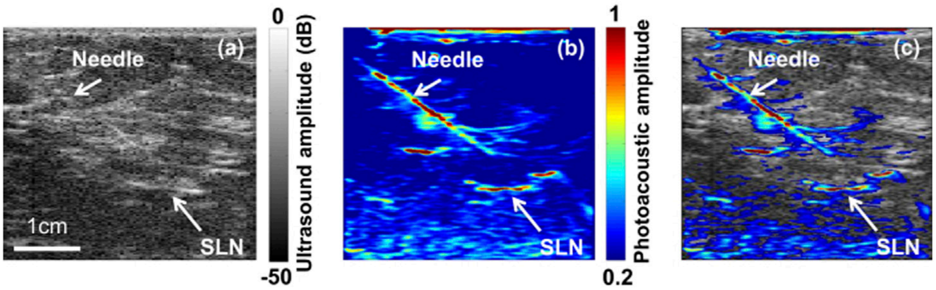


FIGURE 9
Images of the SLN and the needle acquired using PAT/US *in vivo*. **(A)** US image showing the lymph node and needle *in vivo*. **(B)** PAT image of the SLN and needle *in vivo*. **(C)** Coregistered PAT/US image of the SLN (long white arrow) and needle (short white arrow). SLN: sentinel lymph node; PAT: photoacoustic tomography; US: ultrasound. Reprinted with permission from [Garcia-Uribe A., et al. (2015). Dual-Modality Photoacoustic and Ultrasound Imaging System for Noninvasive Sentinel Lymph Node Detection in Patients with Breast Cancer. Sci Rep.5, 15748. doi: 10.1038/srep15748].

volunteers and two excised plaques, were scanned with handheld MSOT noninvasively (Karlas et al., 2021a). Spectral unmixing allowed visualization of lipid and Hb content within three ROIs: whole arterial cross-section, plaque and arterial lumen. This finding introduces MSOT as a new tool for molecular imaging of human carotid atherosclerosis and opens new opportunities for research and clinical evaluation of carotenoid plaques. Following this, MSOT, US and colored Doppler imaging of the carotid arteries in healthy individuals was performed, along with blood flow and oxygenation measurements. (Wu Y et al., 2021). This work demonstrates that multimodality has the potential to provide comprehensive information with increasing accuracy (Figure 8). However, lipids generated just moderate PAT signals at wavelength of 1734 nm, making PAT images at this wavelength vulnerable to noise. To tackle this problem, noise reduction and probe sensitivity will improve accuracy and reliability of lipid identification.

Lymph system

Detecting regional lymph node metastasis is important in cancer staging, as it guides patient prognosis and treatment strategy. Sentinel lymph node biopsy (SLNB) heroine is an accurate and less invasive alternative to axillary lymph node dissection (Aragon et al., 2022). Since high sensitivity to dyes, high spatial, contrast and temporal resolution, enough imaging depth are key requirements for SLNB (Song et al., 2008; Kim and Chang, 2018), blue dye is an ideal contrast agent for PAT due to its strong optical absorption. Accurate identification of sentinel lymph node (SLN) by PAT/US can enable SLN sampling using fine needle aspiration biopsy (FNAB) for a minimally invasive approach to axillary staging (Garcia-Uribe et al., 2015). As a non-ionizing hybrid imaging method, coregistered PAT/US imaging can detect SLNs and lymphatic vessels using methylene blue dye (Figure 9). However, 5 fps of coregistered images is not sufficient for real-time biopsy guided by PAT/US in clinical applications. To achieve a higher frame rate of reconstructed PAT/US images, DAQ computer is necessary to be improved, aiming for real-time visualization in preoperative evaluation in patients with newly diagnosed invasive breast cancer in the future.

Molecular imaging

The biomedical PAT is based on the absorption spectra of intrinsic absorbers present in human tissue, such as HbO_2 and HbR (Laufer et al., 2005; Pan et al., 2010), lipids (Guggenheim et al., 2015; Kole et al., 2019), melanin (Kratkiewicz et al., 2019), water (Xu et al., 2011), RNA and DNA (Yao et al., 2010). When these intrinsic chromophores are not sufficient to reveal the disease, extrinsic contrast agents can be utilized to target different diseased biomarkers to increase molecular sensitivity and specificity. Applications of PAT imaging in the clinical research have shown promising results with endogenous contrast (Wu et al., 2014; Wu et al., 2019b), including in inflammatory bowl (Knieling et al., 2017), dermatology (Masthoff et al., 2018) and breast cancer (Nyayapathi and Xia, 2019). For US imaging, Optison (human serum albumin stabilized perflutren microspheres), Definity (perflutren lipid microspheres), SonoVue (phospholipid-stabilized microbubble), and Sonazoid (F-butane encapsulated in a lipid shell) have been approved for clinical use by the FDA (Cohen et al., 1998; Morel et al., 2000; Halpern et al., 2002; Datta et al., 2008).

Dual-modal contrast agents for PAT/US imaging have been studied to improve the diagnostic sensitivity and specificity. In order to enable molecular PAT/US detection, contrast agents can be specifically tailored to molecular targets of relevance to tumor metastasis, including those biomarkers expressed during lymphangiogenesis, as well as those expressed by tumors. Studies on exogenous contrast agents for PAT/US imaging biomedical application have been reported (Kim et al., 2010a; Xu et al., 2010; Jeon and Kim, 2014; Paproski et al., 2016; Zhang et al., 2016; Bayer et al., 2017; Park et al., 2017b; Chitgupi and Lovell, 2018; Shi et al., 2021). Angiogenesis, the formation of new blood vessels, is a hallmark of many diseases, including tumor and inflammation. In response, those lesions can be detected by PAT based on Hb, which is one of the major chromophores. The treatment of the bubble surface to target specific molecules can broaden the applications of these contrast agents using combined PAT and US imaging. Current contrast agents for PAT/US dual-modality include the conjugating NBs with cancer-targeting ligands (Xu et al., 2010), Texas Red dye in poly (lactic-co-glycolic acid) (PLGA) NBs (Kim et al., 2010c) and gold nanotracers (Au NTs) labeled with mesenchymal stem cells (MSCs) for monitoring disease processing (Nam et al., 2012), encapsulated gold nanorod human serum albumin (HSA)-shelled microbubbles (AuNR-HSA) for thermotherapy (Wang et al., 2012), liquid perfluorocarbon (PFC) nanodroplets with encapsulated plasmonic nanoparticles and encapsulated-ink microbubbles for biomedical application (Jeon M., and Kim, C., 2014). As mentioned above, most of contrast agents are applicable in preclinical trails, only few of them have potential opportunities towards clinical translation, such as particular microbubbles loaded with ICG. In addition, by targeting specific molecules on the bubble surface, these contrast agents can be used for simultaneous PAT and US imaging, including tumor borders, intracranial imaging, and molecular imaging of primary and metastatic tumors in the future.

Other clinical applications

In the last few decades, through intensive study and elaboration, PAT/US dual-modal technology can therefore visualize the human structures or tissues extending to whole-body from the organelles to organ range, including prostate *in vivo* (Kothapalli et al., 2019; Agrawal et al., 2020), placenta *ex vivo* (Xia et al., 2019; Maneas et al., 2020), bowels *in vivo* and *ex vivo* (Knieling et al., 2017; Leng et al., 2018 Yang et al., 2019), periodontal health *in vivo* (Mozaffarzadeh et al., 2021) and spine in human cadaver (Gonzalez et al., 2021).

In prostate and some of bowels researches, intracavity PAT/US probes were conducted *in vivo*. In order to image the entire prostate and circumferential bowel wall, the FOV probes should be designed as large as possible, which in the range of 130°–150° at least (Agrawal et al., 2020). However, compared to the

TABLE 1 Comparison of various configurations for clinical PAT/US dual-modal systems.

Transducers	f_0 (MHz)	No. of elements	FOV	Organs/Tissues	References
Intravascular, full-ring array	20–50	Single	360°	coronary atherosclerosis, *ex vivo*	Karpiouk et al. (2012)
Intravascular, full-ring array	45	—	360°	Coronary artery, *ex vivo*	Wu et al. (2014)
Handheld duplex probe	—	—	—	Breast, *in vivo*	Neuschler et al. (2017)
Handheld, linear array	3	256	125°	Breast, *in vivo*	Becker et al. (2018)
Handheld, linear array	0.1–12	128	—	Breast, *in vivo*	Oraevsky et al. (2018)
Rotate, concave array	10	384	360°	Breast, *in vivo*	Kelly et al. (2020)
Handheld, linear array	5.8	192	3D	Breast, *in vivo*	Yang et al. (2020)
Handheld, linear array	5	256	125°	Breast, *ex vivo*	Goh et al. (2019)
Handheld, linear array	5–12	256	—	Sentinel lymph node, *in vivo*	Garcia-Uribe et al. (2015)
Handheld, curved array	~7.5	64	172°	Thyroid, *in vivo*	Dima and Ntziachristos, (2016)
Handheld, linear array	5.8	192	—	Thyroid, *in vivo*	Yang et al. (2017)
Handheld, linear array	8.5	128	—	Thyroid, *in vivo*	Kim et al. (2021)
Transrectal, linear array	5	64	± 20°	Prostate, *in vivo*	Kothapalli et al. (2019)
Transrectal, linear array	4–8	128	135°	Prostate, *in vivo*	Agrawal et al. (2020)
Handheld, linear array, EC-12R, Alpinion	—	—	—	Melanomas, *in vivo*	Park et al. (2021)
Handheld, linear array,CL15-7, Philips	8.9	—	180°	Joint, *in vivo*	Jo et al. (2017)
Fixed, full-ring	5	512	360°	Fingers, *in vivo*	Xia et al. (2015)
Handheld, spherical array	4	256	90°	Fingers, *in vivo*	Xu et al. (2013)
Handheld, linear probe	7.5	128	~30°	Fingers, *in vivo*	Daoudi et al. (2021)
Fixed, curvilinear array	4–5	64–512	135°–270°	Fingers, *in vivo*	Mercep et al. (2018)
Rotated	3.5	2	~10°	Fingers, *in vivo*	Liu and Zhang, (2016)
Fixed, linear array	14	128	—	Fingers, *in vivo*	van den Berg et al. (2017)
Fixed, arc-like array	10	768	360°	Fingers, *in vivo*	Oeri et al. (2017)
Handheld, linear array	3–12	128	—	Forearm, *in vivo*	Kim et al. (2016)
Fixed	0.5	—	—	Bone, *in vivo*	Feng et al. (2020)
Handheld, linear array	—	—	—	Bowel, *in vivo*	Knieling et al. (2017)
Endorectal probe	20	—	360°	Bowel, *in vivo*	Leng et al. (2018)
Endorectal probe	6–10	—	—	Bowel, *in vivo*	Yang et al. (2019)
Handheld, linear array	10, 20.5, and 40	128	4.5 × 3.5 cm	Periodontal, *in vivo*	Mozaffarzadeh et al. (2021)
Handheld, linear array, EC-12R, Alpinion	4	128	—	Spinal pedicle, *ex vivo*	Gonzalez et al. (2021)
Handheld, linear array, SonixMDP	10	128	62°	Placenta, *in vivo*	Xia et al. (2019)
Handheld, linear array, AcousticX	9	128	3D	Placenta, *ex vivo*	Maneas et al. (2020)

PAT, photoacoustic tomography; US, ultrasound; FOV, field of view; 3D, three-dimensional; --, not mentioned.

conventional clinical PAT/US devices, these PAT/US probes were lack of sufficient elements within the transducers and high performance DAQ systems. On the other hand, the probe should be small enough to relieve the patient's pain during dectetion. Accordingly, these two restrictions limit the spatial and contrast resolutions of intracavity images. The advanced PAT/US system with higher performance intracavity probe could solve this issue, due to its clear visualization of microvasculature distribution in prostate (Kothapalli et al., 2019; Agrawal et al., 2020) and bowel diseases (Knieling et al., 2017; Leng et al., 2018; Yang et al., 2019). Furthermore, in order to achieve better clinical results in the future, a wider FOV of multi-modal probe with a 360° visualization, a better resolution of the system and 3D image reconstruction are needed in the second generation PAT/US device. Meanwhile, further study on the large cohort, multi-center datasets of this approach is needed in monitoring clinical efficacy.

To investigate the ability of PAT/US to image deeper tissue, research groups have performed PAT/US imaging on *ex vivo* human placentas (Xia et al., 2019; Maneas et al., 2020). It is suggested that PAT imaging combined with US tracking could provide a useful method for detecting the placental vasculature during minimally invasive fetal surgery (Xia et al., 2019). Moreover, the 3D dual-modal PAT/US imaging appears to be promising for visualizing human placental vasculature in healthy and twin-to-twin transfusion syndrome (TTTS) treated placentas (Maneas et al., 2020). However, limitations of proposed method include insufficient sensitivity in detecting vessels at all depths in the placenta, and poor spatial resolution in detecting the smallest vessels. It is one of the key challenges for the application in future studies. Further MSOT imaging could be used to discriminate between coagulated and non-coagulated blood based on their different absorption spectral during the treatment of TTTS.

Another challenge application in deep tissue is utilized for a human cadaver vertebra imaging by PAT/US guidance system. This combined system was promising to assist surgeons with identifying and avoiding impending bone breaches during pedicle cannulation in spinal fusion surgeries (Gonzalez et al., 2021). However, the research is still conducted on human cadavers. Extensive preclinical trials are needed before it can be used in humans *in vivo*.

A pilot study of simultaneous visualization of the teeth and periodontium is of significant clinical interest for image-based monitoring of periodontal health (Mozaffarzadeh et al., 2021). It was found that a successful visualization periodontal anatomy and periodontal pocket depths in humans using a dual-modal PAT/US imaging system for the first time. This work demonstrated that 3D PAT/US images allow for accurate measurement and visualization of periodontal features, including the periodontal anatomy, enamel pigmentation, and pocket depth. Efforts are made to remove shaking artifacts from 3D PAT/US images by a specific algorithm. In spite of this, the calculation is complicated and needs to be improved in the future work.

Summary and future perspectives

In this review, we focus on the application and advances in dual-modal PAT/US imaging technology in clinical translations and discussed in details in a systematic way. We explained the principle of PAT/US dual-modality and also discussed a variety of existing PAT/US systems and summarized their key characteristics in Table 1. In the next, we detail clinical applications of current PAT/US system. Compared to traditional PAT imaging, PAT/US detection owns several advantages in clinical applications, such as optical transparency, material flexibility and anti-electromagnetic interference. Several advanced PAT/US imaging technologies have taken another step forward for clinical translation. However, further improvements are needed to make it more clinically compatible.

Future work of this hybrid system may include: 1) Real-time 3D reconstruction technology would be a good future for advanced PAT/US system, which can produce a stereo vision of lesions for radiologists and clinicians. Thus, PAT/US imaging systems with 3D imaging require new methods and materials for systematic testing which can help in decision-making during clinical translation; Furthermore, 3D printed biomodel for simulation of anatomic angioarchitecture of lesions will facilitate surgeon to develop and evaluate a surgical protocol; 2) An advanced PAT/US imaging reconstruction algorithm need to be further proposed based on the previous studies (Jiang, 2015; Wang et al., 2022) to improve lateral resolution of hybrid image, as well as temporal resolution aiming to simultaneously displaying structural, functional, and molecular information; 3) The miniaturization of a PAT/US dual-modal system with significant improvements in the performance of portable laser pump sources, high performance DAQ computers and handhold transducer with smaller size and lighter weight; In next, POC would refer to PAT/US examination outside the lab, such as bedside care, in emergency departments, surgery monitoring or ambulant first aid. It will be a widely used tool for imaging and therefore reducing the time in clinical decision making, emergency and medical education in the coming future. 4) To explore other human organs researches and translations, such as the lungs, pediatric heart, fetus, uterine, neck organs, and others. Due to the dependent optical attenuation depth and wavelength and unknown optical and acoustic heterogeneities limit PAT/US imaging performance in deep tissue regions; therefore, efficient deep tissue's energy transfer system should be developed in next work; 5) PAT/US dual-modal exogenous contrast agents, such as particular microspheres loaded with ICG, have potential opportunities in future to monitor tumor process, metastasis in different part of human body; 6) The multifunctional nanocomposites with dual-modal PAT/US imaging and synergistic therapy will have great application value in different clinical fields involving tumor, vascular plaque, antimicrobial therapy and others in the coming future; 7) PAT/US dual-modal system integrated with AI applications have a broad research prospect in the diagnosis and treatment of human diseases. The previous simulation platform has the potential to generate large-scale application-specific training and test datasets for AI, enhancing AI assisted PAT/US imaging (Agrawal et al., 2019; Hariri et al., 2020; Agrawal et al., 2021b). DL would substantially impact the advancement of modern PAT/US imaging processing methods. Future scope of this work involves 3D simulations and validation studies of different organs to simulate real optical and acoustic heterogeneity, artifacts, shadow effects, and systemic noise. Thus, AI algorithms can be combined to detect the invasion depth and boundary of tumors more precisely in the coming future; 8) At last, FDA cleared PAT/US devices have a greater potential to be a quicker way for clinical application and translation.

We hope this review can be helpful for researchers who wish to learn more about PAT/US dual-modal detections and to use PAT/US dual-modal imaging in their clinical applications.

Author contributions

YW and HJ contributed to the conception of this review; DG, JZ, XL, DW, and TL collated the literature and helped manuscript preparation; TL and LL helped collating the pictures in this review; YW, DG, and JZ wrote the manuscript; HJ and SJ helped perform the analysis with constructive discussions.

Funding

This work was supported by grants from the program of Chengdu Fifth people's hospital Fund (No. KYJJ 2021-29), the Xinglin Scholars research program (No. 20210067), the Chengdu Medical Research Project (No. 2022055), the Doctoral Innovative Talents Program of Chongqing University of Posts and Telecommunications (No. BYJS202117), Chongqing Education Commission, Youth Fund (No. KJQN202000607) and Chongqing post doctoral research project (special funding project, No. 2021XM3040).

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

Agrawal, S., Dangi, A., Kothapalli, S. R., Ghouth, S. B., Albahrani, H., and Frings, N. (2019). Optimal design of combined ultrasound and multispectral photoacoustic deep tissue imaging devices using hybrid simulation platform. *Photons Plus Ultrasound* 7, 0277–786X. doi:10.1117/12.2510869

Agrawal, S., Johnstonbaugh, K., Clark, J. Y., Raman, J. D., Wang, X., and Kothapalli, S. R. (2020). Design, development, and multi-characterization of an integrated clinical transrectal ultrasound and photoacoustic device for human prostate imaging. *Diagn. (Basel)* 10, 566. doi:10.3390/diagnostics10080566

Agrawal, S., Singh, M. K. A., Johnstonbaugh, K., Han, D. C., Pameijer, C. A., and Kothapalli, S-R. (2021a). Photoacoustic imaging of human vasculature using LED versus laser illumination: A comparison study on tissue phantoms and *in vivo* humans. *Sensors* 21, 424. doi:10.3390/s21020424

Agrawal, S., Suresh, T., Garikipati, A., Dangi, A., and Kothapalli, S. R. (2021b). Modeling combined ultrasound and photoacoustic imaging: Simulations aiding device development and artificial intelligence. *Photoacoustics* 24, 100304. doi:10.1016/j.pacs.2021.100304

Aguirre, A., Guo, P., Gamelin, H., Yan, S., Sanders, M. M., Brewer, M., et al. (2009). Coregistered three-dimensional ultrasound and photoacoustic imaging system for ovarian tissue characterization. *J. Biomed. Opt.* 14, 054014. doi:10.1117/1.3233916

Ahn, H. S., Kim, H. J., and Welch, H. G. (2014). Korea's thyroid -cancer "epidemic"— Screening and overdiagnosis. *N. Engl. J. Med.* 371, 1765–1767. doi:10.1056/NEJMp1409841

Ahn, S., Kang, J., Kim, P., Lee, G., Jeong, E., and Jung, W. "Smartphone-based portable ultrasound imaging system: Prototype implementation and evaluation," in Proceedings of the 2015 IEEE International Ultrasonics Symposium (IUS)Taipei, Taiwan, October 2015 (IEEE). doi:10.1109/ULTSYM.2015.0517

Anas, E. M. A., Zhang, H. K., Kang, J., and Boctoret, E. (2018). Enabling fast and high quality LED photoacoustic imaging: A recurrent neural networks based approach. *Biomed. Opt. Express* 9, 3852–3866. doi:10.1364/BOE.9.003852

Aragon, S. S., Oliver, M. R., Madariaga, A., Tabuenca, M. J., Martinez, M., Galindo, A., et al. (2022). Accuracy and limitations of sentinel lymph node biopsy after Neoadjuvant Chemotherapy in breast cancer patients with positive nodes. *Breast J.* 2022, 1507881. doi:10.1155/2022/1507881

Banaka, I., Kaltsas, G., Antoniou, S., Kanakis, G., Zilos, A., Baltas, C. S., et al. (2011). Prognostic value of vascularity index for the diagnosis of autoimmune thyroid disease. *JBR-BTR.* 4, 185–190. doi:10.5334/jbr-btr.582

Bayer, C. L ., Wlodarczyk, B. J ., Finnell, R. H., and Emelianovl, S. Y. (2017). Ultrasound-guided spectral photoacoustic imaging of hemoglobin oxygenation during development. *Biomed. Opt. Express* 8, 757–763. doi:10.1364/BOE.8.000757

Bayer, C. L., Joshi, P. P., and Emelianov, S. Y. (2013). Photoacoustic imaging: A potential tool to detect early indicators of metastasis. *Expert Rev. Med. Devices* 1, 125–134. doi:10.1586/erd.12.62

Beard, P. (2011). Biomedical photoacoustic imaging. *Interface Focus* 1, 602–631. doi:10.1098/rsfs.2011.0028

Becker, A., Masthoff, M., Claussen, J., Ford, S. J., Roll, W., Burg, M., et al. (2018). Multispectral optoacoustic tomography of the human breast: Characterisation of healthy tissue and malignant lesions using a hybrid ultrasound-optoacoustic approach. *Eur. Radiol.* 28, 602–609. doi:10.1007/s00330-017-5002-x

Bene, I., Ciurea, A. I., Ciortea, C. A., Ştefan, P. A., Ciule, L. D., et al. (2022). Radiomic signatures Derived from hybrid contrast-enhanced ultrasound images (CEUS) for the assessment of Histological characteristics of breast cancer: A pilot study. *Cancers* 14, 3905. doi:10.3390/cancers14163905

Bouchard, R., Sahin, O., and Emelianov, S. (2014). Ultrasound-guided photoacoustic imaging: Current state and future development. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 61, 450–466. doi:10.1109/TUFFC.2014.2930

Center, M. M., Jemal, A., Lortet-Tieulent, J., Ward, E., Ferlay, J., Brawley, O., et al. (2012). International variation in prostate cancer incidence and mortality rates. *Eur. Urol.* 61, 1079–1092. doi:10.1016/j.eururo.2012.02.054

Chitgupi, U., and Lovell, J. F. (2018). Naphthalocyanines as contrast agents for photoacoustic and multimodal imaging. *Biomed. Eng. Lett.* 8, 215–221. doi:10.1007/s13534-018-0059-2

Choi, W., Park, E. Y., Jeon, S., and Kim, C. (2018). Clinical photoacoustic imaging platforms. *Biomed. Eng. Lett.* 8, 139–155. doi:10.1007/s13534-018-0062-7

Choi, W., Park, E. Y., Jeon, S., Yang, Y., Park, B., Ahn, J., et al. (2022). Three-dimensional Multistructural quantitative photoacoustic and US imaging of human feet *in vivo*. *Radiology* 303, 467–473. doi:10.1148/radiol.211029

Cohen, J. L., Cheirif, J., Segar, D. S., Gillam, L. D., Gottdiener, J. S., Hausnerova, E., et al. (1998). Improved left ventricular endocardial border delineation and opacification with OPTISON (FS069), a new echocardiographic contrast agent. Results of a phase III Multicenter Trial. *J. Am. Coll. Cardiol.* 32, 746–752. doi:10.1016/s0735-1097(98)00311-8

Crisan, M., Crisan, D., Sannino, G., Lupsor, M., Badea, R., and Amzica, F. (2013). Ultrasonographic staging of cutaneous malignant tumors: An ultrasonographic depth index. *Arch. Dermatol. Res.* 305, 305–313. doi:10.1007/s00403-013-1321-1

Dadkhah, A., and Jiao, S. (2021). Integrating photoacoustic microscopy with other imaging technologies for multimodal imaging. *Exp. Biol. Med.* 246, 771–777. doi:10.1177/1535370220977176

Dahlstrand, U., Sheikh, R., Merdasa, A., Chakari, R., Malmsj, M., Cinthio, M., et al. (2020). Photoacoustic imaging for three-dimensional visualization and delineation of basal cell carcinoma in patients. *Photoacoustics* 18, 100187. doi:10.1016/j.pacs.2020.100187

Dai, X., Xi, l., Duan, C., Yang, H., Xie, H., and Jiang, H. (2015). Miniature probe integrating optical-resolution photoacoustic microscopy, optical coherence tomography, and ultrasound imaging: Proof-of-concept. *Opt. Lett.* 40, 2921–2924. doi:10.1364/OL.40.002921

Dai, X., Yang, H., and Jiang, H. (2017a). *In vivo* photoacoustic imaging of vasculature with a low-cost miniature light emitting diode excitation. *Opt. Lett.* 42, 1456–1459. doi:10.1364/OL.42.001456

Dai, X., Yang, H., Shan, T., Xie, H., Berceli, S. A., and Jiang, H. (2017b). Miniature endoscope for multimodal imaging. *ACS Photonics* 1, 174–180. doi:10.1021/acsphotonics.6b00852

Das, D., Sharma, A., Rajendran, P., and Pramanik, M. (2009). Another decade of photoacoustic imaging. *Phys. Med. Biol.* 66, 05TR01. doi:10.1088/1361-6560/abd669

Datta, S ., Coussios, C. C ., Ammi, A. Y., Mast, T. D., de Courten-Myers, G. M., and Holland, C. K. (2008). Ultrasound-enhanced thrombolysis using Definity as a cavitation nucleation agent. *Ultrasound Med. Biol.* 34, 1421–1433. doi:10.1016/j.ultrasmedbio.2008.01.016

Daoudi, K., Kersten, B. E., van den Ende, C.H.M., van den Hoogen, F.H.J., Vonk, M.C., de Korte, C.L., et al. (2021). Photoacoustic and high-frequency ultrasound, imaging of systemic sclerosis patients. *Arthritis Res Ther.* 23, 22. doi:10.1186/s13075-020-02400-y

Deán-Ben, X. L., and Razansky, D. (2021). Optoacoustic imaging of the skin. *Exp. Dermatol.* 30, 1598–1609. doi:10.1111/exd.14386

Dima, A., and Ntziachristos, V. (2016). *In-vivo* handheld optoacoustic tomography of the human thyroid. *Photoacoustics* 4 (2), 65–69. doi:10.1016/j.pacs.2016.05.003

Dong, B., Sun, C., and Zhang, H. F. (2017). Optical detection of ultrasound in photoacoustic imaging. *IEEE Trans. Biomed. Eng.* 64, 4–15. doi:10.1109/TBME.2016.2605451

Duan, F., Nie, L., Zhang, J., Wu, Z., Chen, R., et al. (2020). *In vivo* photoacoustic imaging Dynamically monitors the structural and functional changes of Ischemic Stroke at a very early stage. *Theranostics* 10 (2), 816–828. doi:10.7150/thno.38554

Dumani, D. S., Sun, I. C., and Emelianov, S. Y. (2019). Ultrasound-guided immunofunctional photoacoustic imaging for diagnosis of lymph node metastases. *Nanoscale* 11, 11649–11659. doi:10.1039/c9nr02920f

Ermilov, S. A., Khamapirad, T., Conjusteau, A., Leonard, M. H., Lacewell, R., Mehta, K., et al. (2009). Laser optoacoustic imaging system for detection of breast cancer. *J. Biomed. Opt.* 14, 024007. doi:10.1117/1.3086616

Farnia, P., Najafzadeh, E., Hariri, A., Lavasani, S. N., Makkiabadi, B., Ahmadian, A., et al. (2020). Dictionary learning technique enhances signal in LED-based photoacoustic imaging. *Biomed. Opt. Express* 11, 2533–2547. doi:10.1364/BOE.387364

Feng, T., Zhu, Y., Liu, C., Du, S., Ta, D., Cheng, H., et al. (2020). Ultrasound-guided detection and Segmentation of photoacoustic signals from bone tissue *in vivo*. *Appl. Sci.* 11, 19. doi:10.3390/app11010019

Galanzha, E. I., Menyaev, Y. A., Yadem, A. C., Sarimollaoglu, M., Juratli, M. A., Nedosekin, D. A., et al. (2019). *In vivo* liquid biopsy using Cytophone platform for photoacoustic detection of circulating tumor cells in patients with melanoma. *Sci. Transl. Med.* 11, eaat5857. doi:10.1126/scitranslmed.aat5857

Garcia-Uribe, A., Erpelding, T. N., Krumholz, A., Ke, H., Maslov, K., AppletonC, et al. (2015). Dual-modality photoacoustic and ultrasound imaging system for noninvasive sentinel lymph node detection in patients with breast cancer. *Sci. Rep.* 5, 15748. doi:10.1038/srep15748

Gerling, M., Zhao, Y., Nania, S., Norberg, K. J., Verbeke, C. S., Englert, B., et al. (2014). Real-time assessment of tissue hypoxia *in vivo* with combined photoacoustics and high-frequency ultrasound. *Theranostics* 4, 604–613. doi:10.7150/thno.7996

Gharib, H., Papini, E., Garber, J. R., Duick, D. S., Harrell, R. M., Hegedus, L., et al. (2016). AACE/ACE/AME Task Force on thyroid nodules, "American association of clinical Endocrinologists, American College of Endocrinology, and Associazione Medici Endocrinologi medical Guidelines for clinical practice for the diagnosis and management of thyroid nodules-2016 update. *Endocr. Pract.* 22, 622–639. doi:10.4158/EP161208.GLGL

Ghoncheh, M., Pournamdar, Z., and Salehiniya, H. (2016). Incidence and mortality and Epidemiology of breast cancer in the World. *Asian pac. J. Cancer Prev.* 17, 43–46. doi:10.7314/apjcp.2016.17.s3.43

Goh, Y., Balasundaram, G., Moothanchery, M., Attia, A., Li, X., Lim, H. Q., et al. (2019). Ultrasound guided optoacoustic tomography in assessment of tumor margins for Lumpectomies. *Transl. Oncol.* 13, 254–261. doi:10.1016/j.tranon.2019.11.005

Gonzalez, E. A., Jain, A., and Bell, M. A. L. (2021). Combined ultrasound and photoacoustic image guidance of spinal pedicle cannulation demonstrated with Intact *ex vivo* specimens. *IEEE Trans. Biomed. Eng.* 68, 2479–2489. doi:10.1109/TBME.2020.3046370

Gröhl, J., Schellenberg, M., Dreher, K., and Maier-Hein, L. (2021). Deep learning for biomedical photoacoustic imaging: A review. *Photoacoustics* 22, 100241. doi:10.1016/j.pacs.2021.100241

Guggenheim, J. A., Allen, T. J., Plumb, A., Zhang, E., Rodriguez-Justo, M., Punwani, S., et al. (2015). Photoacoustic imaging of human lymph nodes with endogenous lipid and hemoglobin contrast. *J. Biomed. Opt.* 20, 050504. doi:10.1117/1.JBO.20.5.050504

Halpern, E. J., Mccue, P. A., Aksnes, A. K., Hagen, E. K., Frauscher, F., and Gomellal, L. G. (2002). Contrast-enhanced US of the prostate with Sonazoid: Comparison with whole-mount prostatectomy specimens in 12 patients. *Radiology* 222, 361–366. doi:10.1148/radiol.2222010582

Han, S., Lee, H., Kim, C., and Kim, J. (2022). Review on multispectral photoacoustic analysis of cancer: Thyroid and breast. *Metabolites* 12, 382. doi:10.3390/metabo12050382

Hariri, A., Alipour, K., Mantri, Y., Schulze, J. P., and Jokerst, J. V. (2020). Deep learning improves contrast in low-fluence photoacoustic imaging. *Biomed. Opt. Express* 11, 3360–3373. doi:10.1364/BOE.395683

Hartman, R. K., Hallam, K. A., Donnelly, E. M., and Emelianov, S. Y. (2019). Photoacoustic imaging of gold nanorods in the brain delivered via microbubble-assisted focused ultrasound: A tool for *in vivo* molecular neuroimaging. *Laser Phys. Lett.* 16, 025603. doi:10.1088/1612-202X/aaf89e

Heibel, H. D., Hooey, L., and Cockerell, C. J. (2020). A review of noninvasive techniques for skin cancer detection in dermatology. *Am. J. Clin. Dermatol.* 21, 513–524. doi:10.1007/s40257-020-00517-z

Heijblom, M., Piras, D., Xia, W., van Hespen, J. C. G., Klaase, J. M., van Leeuwen, T. G., et al. (2012). Visualizing breast cancer using the Twente photoacoustic mammoscope: What do we learn from twelve new patient measurements? *Opt. Express* 20, 11582–11597. doi:10.1364/OE.20.011582

Hindelang, B., Aguirre, J., Schwarz, M., Berezhnoi, A., Eyerich, K., Ntziachristos, V., et al. (2019). Non-invasive imaging in dermatology and the unique potential of raster-scan optoacoustic mesoscopy. *J. Eur. Acad. Dermatol. Venereol.* 33, 1051–1061. doi:10.1111/jdv.15342

Jennings, L. E., and Long, N. J. (2009). 'Two is better than one'--probes for dual-modality molecular imaging. *Chem. Commun.* 28, 3511–3524. doi:10.1039/b821903f

Jeon, M., and Kim, C. (2014). Photoacoustic and ultrasound imaging with nanosized contrast agents. *Nanotechnol. Biomed. Imaging Diagnostics* 10, 293–323. doi:10.1002/9781118873151.ch10

Jiang, H. (2015). Photoacoustic tomography. *IEEE Trans. Med. Imaging* 34 (12), 2645. doi:10.1109/TMI.2015.2499098

Jo, J., Tian, C., Xu, G., Sarazin, J., Schiopu, E., Gandikota, G., et al. (2018). Photoacoustic tomography for human musculoskeletal imaging and inflammatory arthritis detection. *Photoacoustics* 12, 82–89. doi:10.1016/j.pacs.2018.07.004

Jo, J., Xu, G., Cao, M., Marquardt, A., Francis, S., Gandikota, G., et al. (2017). A functional study of human inflammatory arthritis using photoacoustic imaging. *Sci. Rep.* 7, 15026. doi:10.1038/s41598-017-15147-5

Jose, J., Manohar, S., Kolkman, R. G., Steenbergen, W., and van Leeuwen, T. G. (2009). Imaging of tumor vasculature using Twente photoacoustic systems. *J. Biophot.* 2, 701–717. doi:10.1002/jbio.200910025

Kang, J., Chang, J. H., Kim, S. M., Lee, H. J., Kim, H., Wilson, B. C., et al. (2017). Real-time sentinel lymph node biopsy guidance using combined ultrasound, photoacoustic, fluorescence imaging: *In vivo* proof-of-principle and validation with nodal obstruction. *Sci. Rep.* 7, 45008. doi:10.1038/srep45008

Kang, J., Koehler, R. C., Graham, E. M., and Boctor, E. M. (2022). Photoacoustic assessment of the fetal brain and placenta as a method of non-invasive antepartum and intrapartum monitoring. *Exp. Neurol.* 347, 113898. doi:10.1016/j.expneurol.2021.113898

Karlas, A., Kallmayer, M., Bariotakis, M., Fasoula, N. A., Liapis, E., Hyafil, F., et al. (2021a). Multispectral optoacoustic tomography of lipid and hemoglobin contrast in human carotid atherosclerosis. *Photoacoustics* 23, 100283. doi:10.1016/j.pacs.2021.100283

Karlas, A., Masthoff, M., Kallmayer, M., Helfen, A., Bariotakis, M., Fasoula, N. A., et al. (2021b). Multispectral optoacoustic tomography of peripheral arterial disease based on muscle hemoglobin gradients—A pilot clinical study. *Ann. Transl. Med.* 9, 36. doi:10.21037/atm-20-3321

Karmacharya, M. B., Sultan, L. R., and Sehgal, C. M. (2021). Photoacoustic monitoring of oxygenation changes induced by therapeutic ultrasound in Murine Hepatocellular carcinoma. *Sci. Rep.* 11 (1), 4100. doi:10.1038/s41598-021-83439-y

Karpiouk, A. B ., Wang, B., Amirian, J., Smalling, R. W., and Emelianov, S. Y. (2012). Feasibility of *in vivo* intravascular photoacoustic imaging using integrated ultrasound and photoacoustic imaging catheter. *J. Biomed. Opt.* 17, 096008. doi:10.1117/1.JBO.17.9.096008

Kelly, C., Refaee, A., and Salcudean, S. E. (2020). Integrating photoacoustic tomography into a multimodal automated breast ultrasound scanner. *J. Biomed. Opt.* 25, 116010. doi:10.1117/1.JBO.25.11.116010

Kim, C., Erpelding, T. N., Jankovic, L., Pashley, M. D., and Wang, L. V. (2010b). Deeply penetrating *in vivo* photoacoustic imaging using a clinical ultrasound array system. *Biomed. Opt. Express* 1, 278–284. doi:10.1364/BOE.1.000278

Kim, C., Qin, R., Xu, J. S., Wang, L. V., and Xu, R. (2010a). Multifunctional microbubbles and nanobubbles for photoacoustic and ultrasound imaging. *J. Biomed. Opt.* 15, 010510. doi:10.1117/1.3302808

Kim, C., Song, K. H., Gao, F., and Wang, L. V. (2010c). Sentinel lymph nodes and lymphatic vessels: Noninvasive dual-modality *in vivo* mapping by using indocyanine green in rats—volumetric spectroscopic photoacoustic imaging and planar fluorescence imaging. *Radiology* 255 (2), 442–450. doi:10.1148/radiol.10090281

Kim, H., and Chang, J. H. (2018). Multimodal photoacoustic imaging as a tool for sentinel lymph node identification and biopsy guidance. *Biomed. Eng. Lett.* 8, 183–191. doi:10.1007/s13534-018-0068-1

Kim, J., Lee, D., Jung, U., and Kim, C. (2015). Photoacoustic imaging platforms for multimodal imaging. *Ultrasonography* 34, 88–97. doi:10.14366/usg.14062

Kim, J., Park, B., Ha, J., Steinberg, I., Hooper, S. M., Jeong, C., et al. (2021). Multiparametric photoacoustic analysis of human thyroid cancers *in vivo*. *Cancer Res.* 81, 4849–4860. doi:10.1158/0008-5472.CAN-20-3334

Kim, J., Park, E. Y., Park, B., Choi, W., Lee, K. J., Kimet, C., et al. (2020). Towards clinical photoacoustic and ultrasound imaging: Probe improvement and real-time graphical user interface. *Exp. Biol. Med.* 245, 321–329. doi:10.1177/1535370219889968

Kim, J., Park, S., Jung, Y., Chang, S., Park, J., Zhang, Y., et al. (2016). Programmable real-time clinical photoacoustic and ultrasound imaging system. *Sci. Rep.* 6, 35137. doi:10.1038/srep35137

Knieling, F., Neufert, C., Hartmann, A., Claussen, J., Urich, A., Egger, C., et al. (2017). Multispectral optoacoustic tomography for assessment of Crohn's disease Activity. *N. Engl. J. Med.* 376, 1292–1294. doi:10.1056/NEJMc1612455

Knorring, T. V., and Mogensen, M. (2021). Photoacoustic tomography for assessment and quantification of cutaneous and metastatic malignant melanoma - a systematic review. *Photodiagnosis Photodyn. Ther.* 33, 102095. doi:10.1016/j.pdpdt.2020.102095

Kole, A., Cao, Y., Hui, J., Bolad, I. A., Alloosh, M., Cheng, J-X., et al. (2019). Comparative quantification of arterial lipid by intravascular photoacoustic-ultrasound imaging and near-infrared spectroscopy-intravascular ultrasound. *J. Cardiovasc. Transl. Res.* 12, 211–220. doi:10.1007/s12265-018-9849-2

Kolkman, R. G. M., Bosschaart, N., Kok, B., van Leeuwen, T. G., and Steenbergen, W. (2006). Photoacoustic imaging of valves in superficial veins. *Lasers Surg. Med.* 38, 740–744. doi:10.1002/lsm.20390

Kolkman, R. G. M., Brands, P. J., Steenbergen, W., and van Leeuwen, T. G. (2008). Real-time *in vivo* photoacoustic and ultrasound imaging. *J. Biomed. Opt.* 13, 050510. doi:10.1117/1.3005421

Kothapalli, S. R., Sonn, G. A., Choe, J. W., Nikoozadeh, A., Bhuyan, A., Park, K. K., et al. (2019). Simultaneous transrectal ultrasound and photoacoustic human prostate imaging. *Sci. Transl. Med.* 11, eaav2169. doi:10.1126/scitranslmed.aav2169

Kratkiewicz, K., Manwar, R., Rajabi-Estarabadi, A., Fakhoury, J., Meiliute, J., Daveluy, S., et al. (2019). Photoacous- tic/ultrasound/optical coherence tomography evaluation of melanoma lesion and healthy skin in a Swine model. *Sensors* 19, 2815. doi:10.3390/s19122815

Kruizinga, P., Mastik, F., Koeze, D., Jong, N. D., van der Steen, A. F. W., Soest, G. V., et al. (2013). Ultrasound-guided photoacoustic image reconstruction: Image completion and boundary suppression. *J. Biomed. Opt.* 18, 096017. doi:10.1117/1.JBO.18.9.096017

Laufer, J., Elwell, C., Delpy, D., and Beard, P. (2005). *In vitro* measurements of absolute blood oxygen saturation using pulsed near-infrared photoacoustic spectroscopy: Accuracy and resolution. *Phys. Med. Biol.* 50, 4409–4428. doi:10.1088/0031-9155/50/18/011

Leng, X., Chapman, W., Rao, B., Nandy, S., Chen, R., Rais, R., et al. (2018). Feasibility of co-registered ultrasound and acoustic-resolution photoacoustic imaging of human colorectal cancer. *Biomed. Opt. Express* 9, 5159–5172. doi:10.1364/BOE.9.005159

Levi, J., Sathirachinda, A., and Gambhir, S. S. (2014). A high-affinity, high-stability photoacoustic agent for imaging gastrin-releasing peptide receptor in prostate cancer. *Clin. Cancer Res.* 20, 3721–3729. doi:10.1158/1078-0432.CCR-13-3405

Li, D., Humayun, L., Vienneau, E., Vu, T., and Yao, J. (2021). Seeing through the skin: Photoacoustic tomography of skin vasculature and beyond. *JID Innov.* 1, 100039. doi:10.1016/j.xjidi.2021.100039

Li, M., Tang, Y., and Yao, J. (2018). Photoacoustic tomography of blood oxygenation: A mini review. *Photoacoustics* 10, 65–73. doi:10.1016/j.pacs.2018.05.001

Li, R., Wang, P., Lan, L., Lloyd, F. P., Goergen, C. J., Chen, S. X., et al. (2015). Assessing breast tumor margin by multispectral photoacoustic tomography. *Biomed. Opt. Express* 6 (4), 1273–1281. doi:10.1364/BOE.6.001273

Li, X., Wang, D., Ran, H., Hao, L., Cao, Y., Ao, M., et al. (2018). A preliminary study of photoacoustic/ultrasound dual-mode imaging in melanoma using MAGE-targeted gold nanoparticles. *Biochem. Biophys. Res. Commun.* 502, 255–261. doi:10.1016/j.bbrc.2018.05.155

Li, Y., Lu, G., Chen, J. J., Jing, J. C., Huo, T. C., Chen, R., et al. (2019). PMN-PT/Epoxy 1-3 composite based ultrasonic transducer for dual-modality photoacoustic and ultrasound endoscopy. *Photoacoustics* 15, 100138. doi:10.1016/j.pacs.2019.100138

Liang, B., Wang, S., Shen, F., Liu, Q., Gong, Y., and Yao, J. (2021). Acoustic impact of the human skull on transcranial photoacoustic imaging. *Biomed. Opt. Express* 12, 1512–1528. doi:10.1364/BOE.420084

Lin, L., Hu, P., Shi, J., Appleton, C. M., Maslov, K., Li, L., et al. (2018). Single-breath-hold photoacoustic computed tomography of the breast. *Nat. Commun.* 9, 2352–2359. doi:10.1038/s41467-018-04576-z

Liu, W. Z., and Zhang, H. F. (2016). Photoacoustic imaging of the eye: A mini review. *Photoacoustics* 4, 112–123. doi:10.1016/j.pacs.2016.05.001

Lou, Y., Zhou, W., Matthews, T. P., Appleton, C. M., and Anastasio, M. A. (2017). Generation of anatomically realistic numerical phantoms for photoacoustic and ultrasonic breast imaging. *J. Biomed. Opt.* 22, 41015. doi:10.1117/1.JBO.22.4.041015

Luke, G. P., and Emelianov, S. Y. (2015). Label-free detection of lymph node metastases with US-guided functional photoacoustic imaging. *Radiology* 277, 435–442. doi:10.1148/radiol.2015141909

Luke, G. P., Myers, N. J., Emelianov, S. Y., and Sokolov, K. V. (2014). Sentinel lymph node biopsy revisited: Ultrasound- guided photoacoustic detection of micrometastases using molecularly targeted plasmonic nanosensors. *Cancer Res.* 74, 5397–5408. doi:10.1158/0008-5472.CAN-14-0796

Mallidi, S., Kim, S., Karpiouk, A., Joshi, P. P., Sokolov, K., and Emelianov, S. (2015a). Visualization of molecular composition and functionality of cancer cells using nanoparticle-augmented ultrasound-guided photoacoustics. *Photoacoustics* 3, 26–34. doi:10.1016/j.pacs.2014.12.003

Mallidi, S., Watanabe, K., Timerman, D., Schoenfeld, D., and Hasan, T. (2015b). Prediction of tumor recurrence and therapy monitoring using ultrasound-guided photoacoustic imaging. *Theranostics* 5, 289–301. doi:10.7150/thno.10155

Maneas, E., Aughwane, R., Huynh, N., Xia, W., Ansari, R., Singh, M. K. A., et al. (2020). Photoacoustic imaging of the human placental vasculature. *J. Biophot.* 13, e201900167. doi:10.1002/jbio.201900167

Manohar, S., and Dantuma, M. (2019). Current and future trends in photoacoustic breast imaging. *Photoacoustics* 16, 100134. doi:10.1016/j.pacs.2019.04.004

Manohar, S., and Gambhir, S. S. (2020). Clinical photoacoustic imaging. *Photoacoustics* 19, 100196. doi:10.1016/j.pacs.2020.100196

Masthoff, M., Helfen, A., Claussen, J., Karlas, A., Markwardt, N. A., Ntziachristos, V., et al. (2018). Use of multispectral optoacoustic tomography to diagnose vascular Malformations. *JAMA Dermatol.* 154, 1457–1462. doi:10.1001/jamadermatol.2018.3269

Menke, J. (2015). Photoacoustic breast tomography prototypes with reported human applications. *Eur. Radiol.* 25, 2205–2213. doi:10.1007/s00330-015-3647-x

Mennes, O. A., Netten, J. J., Slart, R. H., and Steenbergen, W. (2018). Novel optical techniques for imaging microcirculation in the diabetic foot. *Curr. Pharm. Des.* 24, 1304–1316. doi:10.2174/1381612824666180302141902

Menzel, J., and Domschke, W. (2000). Gastrointestinal miniprobe sonography: The current status. *Am. J. Gastroenterol.* 95, 605–616. doi:10.1111/j.1572-0241.2000.01832.x

Merčep, E., Deán-Ben, X. L., and Razansky, D. (2018). Imaging of blood flow and oxygen state with a multi-segment optoacoustic ultrasound array. *Photoacoustics* 10, 48–53. doi:10.1016/j.pacs.2018.04.002

Miranda, C., Barkley, J., and Smith, B. (2018). Intrauterine photoacoustic and ultrasound imaging probe. *J. Biomed. Opt.* 23, 1–9. doi:10.1117/1.JBO.23.4.046008

Morel, D. R., Schwieger, I., Hohn, L., Terrettaz, J., Llull, J. B., Cornioley, Y. A., et al. (2000). Human pharmacokinetics and safety evaluation of SonoVue, a new contrast agent for ultrasound imaging. *Invest. Radiol.* 35, 80–85. doi:10.1097/00004424-200001000-00009

Mozaffarzadeh, M., Moore, C., Golmoghani, E. B., Mantri, Y., Hariri, A., Jorns, A., et al. (2021). Motion-compensated noninvasive periodontal health monitoring using handheld and motor-based photoacoustic-ultrasound imaging systems. *Biomed. Opt. Express* 12, 1543–1558. doi:10.1364/BOE.417345

Nagae, K., Asao, Y., Sudo, Y., Murayama, N., Tanaka, Y., Ohira, K., et al. (2018). Real- time 3D photoacoustic visualization system with a wide field of view for imaging human limbs. *F1000Res.* 7, 1813. doi:10.12688/f1000research.16743.2

Nam, S. Y., Ricles, L. M., Suggs, L. J., and Emelianov, S. Y. (2012). *In vivo* ultrasound and photoacoustic monitoring of mesenchymal stem cells labeled with gold nanotracers. *PLoS One* 7, 37267. doi:10.1371/journal.pone.0037267

Needles, A., Heinmiller, A., Sun, J., Theodoropoulos, C., Bates, D., Hirson, D., et al. (2013). Development and initial application of a fully integrated photoacoustic micro-ultrasound system. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 60 (5), 888–897. doi:10.1109/TUFFC.2013.2646

Neuschler, E. I., Butler, R., Young, C. A., Barke, L. D., Bertrand, M. L., Böhm-Vélez, M., et al. (2017). A Pivotal study of optoacoustic imaging to diagnose benign and malignant breast Masses: A new evaluation tool for radiologists. *Radiology* 287, 398–412. doi:10.1148/radiol.2017172228

Neuschmelting, V., Lockau, H., Ntziachristos, V., Grimm, J., and Kircher, M. F. (2016). Lymph node micrometastases and in-transit metastases from melanoma: *In vivo* detection with multispectral optoacoustic imaging in a mouse model. *Radiology* 280, 137–150. doi:10.1148/radiol.2016160191

Niederhauser, J. J., Jaeger, M., Lemor, R., Weber, P., and Frenz, M. (2005). Combined ultrasound and optoacoustic system for real-time high- contrast vascular imaging *in vivo*. *IEEE Trans. Med. Imaging* 24, 436–440. doi:10.1109/tmi.2004.843199

Nyayapathi, N., and Xia, J. (2019). Photoacoustic imaging of breast cancer: A mini review of system design and image features. *J. Biomed. Opt.* 24, 1–13. doi:10.1117/1.JBO.24.12.121911

Oeri, M., Bost, W., Sénégond, N., Tretbar, S., and Fournelle, M. (2017). Hybrid photoacoustic/ultrasound Tomograph for real-time finger imaging. *Ultrasound Med. Biol.* 43, 2200–2212. doi:10.1016/j.ultrasmedbio.2017.05.015

Oraevsky, A. A., Clingman, B., Zalev, J., Stavros, A. T., Yang, W. T., and Parikh, J. R. (2018). Clinical optoacoustic imaging combined with ultrasound for coregistered functional and anatomical mapping of breast tumors. *Photoacoustics* 12, 30–45. doi:10.1016/j.pacs.2018.08.003

Pan, D., Pramanik, M., Senpan, A., Allen, J. S., Zhang, H., Wickline, S. A., et al. (2010). Molecular photoacoustic imaging of angiogenesis with integrin-targeted gold nanobeacons. *FASEB J.* 25, 875–882. doi:10.1096/fj.10-171728

Pang, W. R., Wang, Y. J., Guo, L. L., Wang, B., Lai, P. X., and Xiao, J. Y. (2022). Two-dimensional photoacoustic/ultrasonic endoscopic imaging based on a line-focused transducer. Ultrasonic endoscopic imaging based on a line-focused transducer. *Front. Bioeng. Biotechnol.* 9, 807633. doi:10.3389/fbioe.2021.807633

Paola, V. D., Mazzotta, G., Pignatelli, V., Bufi, E., D'Angelo, A., Conti, M., et al. (2022). Beyond N staging in breast cancer: Importance of MRI and ultrasound-based imaging. *Cancers* 14, 4270. doi:10.3390/cancers14174270

Paproski, R. J., Forbrich, A., Huynh, E., Chen, J., Lewis, J. D., Zheng, G., et al. (2016). Porphyrin nanodroplets: Sub-micrometer ultrasound and photoacoustic contrast imaging agents. *Small* 12, 371–380. doi:10.1002/smll.201502450

Park, B., Bang, C. H., Lee, C., Han, J. H., Choi, W., Kim, J ., et al. (2021). 3D wide-field multispectral photoacoustic imaging of human melanomas *in vivo*: A pilot study. *J. Eur. Acad. Dermatol. Venereol.* 35, 669–676. doi:10.1111/jdv.16985

Park, S., Jang, J., Kim, J., Kim, Y. S., and Kim, C. (2017a). Real-time Triple-modal photoacoustic, ultrasound, and magnetic resonance fusion imaging of humans. *IEEE Trans. Med. Imaging* 36, 1912–1921. doi:10.1109/TMI.2017.2696038

Park, S., Jung, U., Lee, S., Lee, D., and Kim, C. (2017b). Contrast-enhanced dual mode imaging: Photoacoustic imaging plus more. *Biomed. Eng. Lett.* 7, 121–133. doi:10.1007/s13534-016-0006-z

Petri, M., Stoffels, I., Jose, J., Leyh, J., Schulz, A., Dissemond, J., et al. (2016). Photoacoustic imaging of real-time oxygen changes in Chronic Leg Ulcers after Topical application of a Haemoglobin spray: A pilot study. *J. Wound Care* 25 (2), 89–91. doi:10.12968/jowc.2016.25.2.87

Porter, P. L. (2009). Global trends in breast cancer incidence and mortality. *Salud Publica Mex.* 51 (2), s141–s146. doi:10.1590/s0036-36342009000800003

Poudel, J., Lou, Y., and Anastasio, M. A. (2019). A survey of computational frameworks for solving the acoustic inverse problem in three-dimensional photoacoustic computed tomography. *Phys. Med. Biol.* 64, 14TR01. doi:10.1088/1361-6560/ab2017

Raes, F., Sobilo, J., Mée, M. L., Rétif, S., Natkunarajah, S., Lerondelet, S., et al. (2016). High resolution ultrasound and photoacoustic imaging of Orthotopic lung cancer in Mice: New perspectives for Onco-Pharmacology. *PLoS One* 11, 0153532. doi:10.1371/journal.pone.0153532

Regensburger, A. P., Fonteyne, L. M., Jüngert, J., Wagner, A. L., Gerhalter, T., Nagel, A. M., et al. (2019). Detection of collagens by multispectral optoacoustic tomography as an imaging biomarker for Duchenne muscular dystrophy. *Nat. Med.* 25, 1905–1915. doi:10.1038/s41591-019-0669-y

Roll, W., Markwardt, N. A., Masthoff, M., Helfen, A., Claussen, J., Eisenblätter, M., et al. (2019). Multispectral optoacoustic tomography of benign and malignant thyroid Disorders: A pilot study. *J. Nucl. Med.* 60 (10), 1461–1466. doi:10.2967/jnumed.118.222174

Salehi, H. S., Li, H., Merkulov, A., Kumavor, P. D., Vavadi, H., Sander, M., et al. (2016). Coregistered photoacoustic and ultrasound imaging and classification of ovarian cancer: *Ex vivo* and *in vivo* studies. *J. Biomed. Opt.* 21, 046006. doi:10.1117/1.JBO.21.4.046006

Shi, X. F., Ji, B., Kong, Y., Guan, Y., and Ni, R. (2021). Multimodal contrast agents for optoacoustic brain imaging in small animals. *Front. Bioeng. Biotechnol.* 9, 746815. doi:10.3389/fbioe.2021.746815

Shiina, T., Toi, M., and Yagi, T. (2018). Development and clinical translation of photoacoustic mammography. *Biomed. Eng. Lett.* 8, 157–165. doi:10.1007/s13534-018-0070-7

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2021). Cancer Statistics, 2017. *Ca. Cancer J. Clin.* 71, 7–30. doi:10.3322/caac.21387

Sinha, S., Dogra, V. S., Chinni, B. K., and Rao, N. A. (2017). Frequency Domain analysis of Multiwavelength photoacoustic signals for Differentiating Among malignant, benign, and Normal thyroids in an *ex vivo* study with human thyroids. *J. Ultrasound Med.* 36, 2047–2059. doi:10.1002/jum.14259

Song, K. H., Stein, E. W., Margenthaler, J. A., and Wang, L. V. (2008). Noninvasive photoacoustic identification of sentinel lymph nodes containing methylene blue *in vivo* in a rat model. *J. Biomed. Opt.* 13, 054033. doi:10.1117/1.2976427

Steinberg, I., Huland, D. M., Vermesh, O., Frostig, H. E., Tummers, W. S., and Gambhir, S. S. (2019). Photoacoustic clinical imaging. *Photoacoustics* 14, 77–98. doi:10.1016/j.pacs.2019.05.001

Su, J. L., Wang, B., Wilson, K. E., Bayer, C. L., Chen, Y. S., Kim, S., et al. (2010). Advances in clinical and biomedical applications of photoacoustic imaging. *Expert Opin. Med. diagn.* 4, 497–510. doi:10.1517/17530059.2010.529127

Taruttis, A., Wildgruber, M., Kosanke, K., Beziere, N., Licha, K., Haag, R., et al. (2013). Multispectral optoacoustic tomography of Myocardial Infarction. *Photoacoustics* 1 (1), 3–8. doi:10.1016/j.pacs.2012.11.001

Tian, C., Pei, M., Shen, K., Liu, S., Hu, Z., and Feng, T. (2020). Impact of system factors on the performance of photoacoustic tomography scanners. *Phys. Rev. Appl.* 13, 014001. doi:10.1103/PhysRevApplied.13.014001

Upputuri, P. K., and Pramanik, M. (2017). Recent advances toward preclinical and clinical translation of photoacoustic tomography: A review. *J. Biomed. Opt.* 22, 41006. doi:10.1117/1.JBO.22.4.041006

Vaccarella, S., Franceschi, F., Bray, F., Wild, C. P., Plummer, M., and Maso, L. D. (2016). Worldwide thyroid-cancer epidemic? The increasing impact of overdiagnosis. *N. Engl. J. Med.* 375, 614–617. doi:10.1056/NEJMp1604412

van den Berg, P. J., Daoudi, K., Bernelot Moens, H. J., and Steenbergen, W. (2017). Feasibility of photoacoustic/ultrasound imaging of synovitis in finger joints using a point-of-care system. *Photoacoustics* 8, 8–14. doi:10.1016/j.pacs.2017.08.002

Wang, C., Guo, L., Wang, G., Ye, T., Wang, B., Xiao, J., et al. (2021). *In-vivo* imaging of melanoma with simultaneous dual-wavelength acoustic-resolution-based photoacoustic/ultrasound microscopy. *Appl. Opt.* 60, 3772–3778. doi:10.1364/AO.412609

Wang, L. V. (2008). Prospects of photoacoustic tomography. *Med. Phys.* 35, 5758–5767. doi:10.1118/1.3013698

Wang, S., Lin, J., Wang, T., Chen, X., and Huang, P. (2016). Recent advances in photoacoustic imaging for deep-tissue biomedical applications. *Theranostics* 6, 2394–2413. doi:10.7150/thno.16715

Wang, S. L., Zhang, L. L., Zhao, J. J., He, M., and Zhao, S. L. (2021). A tumor Microenvironment–induced absorption red-Shifted Polymer nanoparticle for simultaneously Activated photoacoustic imaging and Photothermal therapy. *Sci. Adv.* 7 (12), eabe3588. doi:10.1126/sciadv.abe3588

Wang, X., Fowlkes, J. B., Cannata, J. M., Hu, C., and Carson, P. L. (2011). Photoacoustic imaging with a commercial ultrasound system and a custom probe. *Ultrasound Med. Biol.* 37, 484–492. doi:10.1016/j.ultrasmedbio.2010.12.005

Wang, X., Pang, Y., Geng, K., Xie, X., and Wang, L. V. (2003). Noninvasive laser-induced photoacoustic tomography for structural and functional *in vivo* imaging of the brain. *Nat. Biotechnol.* 21 (7), 803–806. doi:10.1038/nbt839

Wang, Y. H., Liao, A. H., Chen, J. H., Chris Wang, C. R., and Li, P. C. (2012). Photoacoustic/ultrasound dual-modality contrast agent and its application to thermotherapy. *J. Biomed. Opt.* 17, 045001. doi:10.1117/1.JBO.17.4.045001

Wang, Y., Yuan, C., Jiang, J., Peng, K., and Wang, B. (2022). Photoacoustic/ultrasound endoscopic imaging reconstruction algorithm based on the Approximate Gaussian acoustic field. *Biosens. (Basel)* 12 (7), 463. doi:10.3390/bios12070463

Wen, Y. T., Wu, D., Zhang, J., Guo, D., Xiong, C. Y., Chi, Z. H., et al. (2022a). Multiple wavelength photoacoustic tomography of human cervical Trachea. *Prog. Biochem. Biophysics* 03, 600–606. doi:10.16476/j.pibb.2021.0131

Wen, Y. T., Wu, D., Zhang, J., Jiang, S. X., Xiong, C. Y., Guo, D., et al. (2022b). Evaluation of Tracheal Stenosis in Rabbits using multispectral optoacoustic tomography. *Front. Bioeng. Biotechnol.* 10, 860305. doi:10.3389/fbioe.2022.860305

Wiacek, A., and Bell, M. A. L. (2021). Photoacoustic-guided surgery from head to toe [Invited]. *Biomed. Opt. Express* 12, 2079–2117. doi:10.1364/BOE.417984

Wu, D., Guo, X. Y., Cui, R. H., Wu, M., Shang, Q., and Jiang, H. (2019a). *In vivo* hemodynamic visualization of Berberine-induced effect on the cerebral Cortex of a mouse by photoacoustic tomography. *Appl. Opt.* 58 (1), 1–8. doi:10.1364/AO.58.000001

Wu, D., Huang, L., Jiang, M. S., Jiang, H. b., et al. (2014). Contrast agents for photoacoustic and thermoacoustic imaging: A review. *Int. J. Mol. Sci.* 15 (12), 23616–23639. doi:10.3390/ijms151223616

Wu, D., Zhang, X., Rong, J., and Jiang, H. (2019b). Photoacoustic molecular imaging using combined Acupuncture and gold nanorods as a composite contrast agent. *J. Innov. Opt. Health Sci.* 12, 1941004. doi:10.1142/S1793545819410049

Wu, M., Awasthi, N., Rad, N. M., Pluim, J. P. W., and Lopata, R. G. P. (2021). Advanced ultrasound and photoacoustic imaging in Cardiology. *Sensors* 21, 7947. doi:10.3390/s21237947

Wu, M., Jansen, K., van der Steen, A. F., and Soest, V. G. (2015). Specific imaging of atherosclerotic plaque lipids with two-wavelength intravascular photoacoustics. *Biomed. Opt. Express* 6, 3276–3286. doi:10.1364/BOE.6.003276

Wu, Y., Zeng, F., Zhao, Y., and Wu, S. (2021). Emerging contrast agents for multispectral optoacoustic imaging and their biomedical applications. *Chem. Soc. Rev.* 14, 7924–7940. doi:10.1039/d1cs00358e

Xavierselvan, M., Singh, M., and Mallidi, S. "Preclinical cancer imaging using a multispectral LED-based photoacoustic and ultrasound imaging system," in SPIE Photonics West, May 2021 (SPIE). doi:10.1117/12.2579263

Xia, W., Maneas, E., Huynh, N. T., Singh, M. K. A., Brown, N. M., Ourselin, S., et al. "Imaging of human peripheral blood vessels during cuff occlusion with a compact LED-based photoacoustic and ultrasound system," in Photons Plus Ultrasound: Imaging and Sensing, 2019 (International Society for Optics and Photonics). doi:10.1117/12.2509666

Xia, W., Maneas, E., Nikitichev, D. I., Mosse, C. A., Santos, G. S. D., Vercauteren, T., et al. (2015). Interventional photoacoustic imaging of the human placenta with ultrasonic tracking for minimally invasive fetal surgeries. *Med. Image Comput. Comput. Assist. Interv.* 9349, 371–378. doi:10.1007/978-3-319-24553-9_46

Xu, G., Rajian, J. R., Girish, G., Kaplan, M. J., Fowlkes, J. B., Carson, P. L., et al. (2013). Photoacoustic and ultrasound dual-modality imaging of human peripheral joints. *J Biomed Opt.* 18 (1), 10502. doi:10.1117/1.JBO.18.1.010502

Xu, J. S., Huang, J., Qin, R., Hinkle, G. H., Povoski, S. P., Martin, E. W., et al. (2010). Synthesizing and binding dual-mode poly (lactic-co-glycolic acid) (PLGA) nanobubbles for cancer targeting and imaging. *Biomaterials* 31, 1716–1722. doi:10.1016/j.biomaterials.2009.11.052

Xu, Q. W., Wang, X., and Jiang, H. B. (2019). Convolutional neural network for breast cancer diagnosis using Diffuse optical tomography. *Vis. Comput. Ind. Biomed. Art.* 2 (1), 1. doi:10.1186/s42492-019-0012-y

Xu, Z., Zhu, Q., and Wang, L. V. (2011). *In vivo* photoacoustic tomography of mouse cerebral edema induced by cold injury. *J. Biomed. Opt.* 16, 066020. doi:10.1117/1.3584847

Yamada, H., Matsumoto, N., Komaki, T., Konishi, H., Kondo, T., Son, A., et al. (2020). Photoacoustic *in vivo* 3D imaging of tumor using a highly tumor-targeting probe under high-threshold Conditions. *Sci. Rep.* 10 (1), 19363. doi:10.1038/s41598-020-76281-1

Yamaleyeva, L. M., Brosnihan, K. B., Smith, L. M., and Sun, Y. (2018). Preclinical ultrasound-guided photoacoustic imaging of the placenta in Normal and Pathologic Pregnancy. *Mol. Imaging* 17, 1536012118802721. doi:10.1177/1536012118802721

Yan, Y., Basij, M., Garg, A., Varrey, A., Alhousseini, A., Hsu, R., et al. (2021). Spectroscopic photoacoustic imaging of cervical tissue composition in excised human samples. *PLoS One* 16, e0247385. doi:10.1371/journal.pone.0247385

Yang, G., Amidi, E., Chapman, W., Nandy, S., Mostafa, A., Abdelal, H., et al. (2019). Co-registered photoacoustic and ultrasound imaging of human colorectal cancer. *J. Biomed. Opt.* 24, 1–13. doi:10.1117/1.JBO.24.12.121913

Yang, H., Qian, W. P., Yang, L. L., Xie, H. K., and Jiang, H. B. (2020). *In vivo* evaluation of a miniaturized fluorescence molecular tomography (FMT) endoscope for breast cancer detection using targeted Nanoprobes. *Int. J. Mol. Sci.* 21 (24), 9389. doi:10.3390/ijms21249389

Yang, J. M., Favazza, C., Chen, R., Yao, J. J., Cai, X., Maslov, K., et al. (2012). Simultaneous functional photoacoustic and ultrasonic endoscopy of internal organs *in vivo*. *Nat. Med.* 18, 1297–1302. doi:10.1038/nm.2823

Yang, J. M., and Ghim, C. M. (2021). Photoacoustic tomography opening new Paradigms in biomedical imaging. *Adv. Exp. Med. Biol.* 1310, 239–341. doi:10.1007/978-981-33-6064-8_11

Yang, M., Jiang, Y., Zhao, L., Su, N., Wei, Y., Wang, M., et al. "Photoacoustic/ultrasound dual imaging of human superficial lesions: An initial clinical study (Conference Presentation)," in Multimodal Biomedical Imaging XIII, 2018, March (SPIE), 104870C. doi:10.1117/12.2286872

Yang, M., Zhao, L., He, X., Su, N., Zhao, C. Y., Tang, H., et al. (2017). Photoacoustic/ultrasound dual imaging of human thyroid cancers: An initial clinical study. *Biomed. Opt. Express* 8, 3449–3457. doi:10.1364/BOE.8.003449

Yao, D. K., Maslov, K., Shung, K. K., Zhou, Q., and Wang, L. V. (2010). *In vivo* label-free photoacoustic microscopy of cell nuclei by excitation of DNA and RNA. *Opt. Lett.* 35, 4139–4141. doi:10.1364/OL.35.004139

Yao, J., and Wang, L. V. (2018). Recent progress in photoacoustic molecular imaging. *Curr. Opin. Chem. Biol.* 45, 104–112. doi:10.1016/j.cbpa.2018.03.016

Zalev, J., Herzog, D., Clingman, B., Miller, T., Kist, K., Dornbluth, N. C., et al. "Clinical feasibility study of combined optoacoustic and ultrasonic imaging modality providing coregistered functional and anatomical maps of breast tumors," in Proceedings of SPIE - The International Society for Optical Engineering, 2012 (SPIE), 858103–858109. doi:10.1117/12.2009112

Zhang, M., Kim, H. S., Jin, T ., Yi, A., and Moon, W. K. (2016). Ultrasound-guided photoacoustic imaging for the selective detection of EGFR-expressing breast cancer and lymph node metastases. *Biomed. Opt. Express* 7, 1920–1931. doi:10.1364/BOE.7.001920

Zhang, Y., and Wang, L. (2020). Video-rate Ring-array ultrasound and photoacoustic tomography. *IEEE Trans. Med. Imaging* 39, 4369–4375. doi:10.1109/TMI.2020.3017815

Zhao, T., Desjardins, A. E., Ourselin, S., Vercauteren, T., and Xia, W. (2019). Minimally invasive photoacoustic imaging: Current status and future perspectives. *Photoacoustics* 16, 100146. doi:10.1016/j.pacs.2019.100146

Zhu, Y., Feng, T., Cheng, Q., Wang, X., Du, S., Sato, N., et al. (2020). Towards clinical translation of LED-based photoacoustic imaging: A review. *Sensors* 20, 2484. doi:10.3390/s20092484

Zhu, Y., Xu, G., Yuan, J., Jo, J., Gandikota, G., Demirci, H., et al. (2018). Light emitting diodes based photoacoustic imaging and potential clinical applications. *Sci. Rep.* 8, 9885. doi:10.1038/s41598-018-28131-4

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 31 October 2022
DOI 10.3389/fphys.2022.1040777

OPEN ACCESS

EDITED BY
Hao Yang,
University of South Florida,
United States

REVIEWED BY
Bo Lan,
Summation Bio Inc, United States
Haiyan Shang,
Chongqing Jiaotong University, China
Fei Yang,
Guangdong University of Technology,
China

*CORRESPONDENCE
Xiuxin Wang,
wangxx@cqupt.edu.cn
Yin Tian,
tianyin@cqupt.edu.cn

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

RECEIVED 09 September 2022
ACCEPTED 17 October 2022
PUBLISHED 31 October 2022

CITATION
Jiang Y, Li H, Pang Y, Ling J, Wang H, Yang Y, Li X, Tian Y and Wang X (2022), Cell image reconstruction using digital holography with an improved GS algorithm.
Front. Physiol. 13:1040777.
doi: 10.3389/fphys.2022.1040777

Cell image reconstruction using digital holography with an improved GS algorithm

Yuhao Jiang[1,2], Hongzhong Li[3], Yu Pang[1], Jiwei Ling[1], Hao Wang[1], Yuling Yang[1], Xinyu Li[1], Yin Tian[1,4]* and Xiuxin Wang[1,5]*

[1]Institute for Advanced Sciences, Chongqing University of Posts and Telecommunications, Chongqing, China, [2]Central Nervous System Drug Key Laboratory of Sichuan Province, Luzhou, China, [3]School of Civil Engineering, Guangdong Communication Polytechnic, Guangzhou, China, [4]Guangyang Bay Laboratory, Chongqing Institute for Brain and Intelligence, Chongqing, China, [5]Key Laboratory of Opto-technology and Intelligent Control, Ministry of Education, Lanzhou Jiaotong University, Lanzhou, China

Digital holography is an effective technology in image reconstruction as amplitude and phase information of cells can be acquired without any staining. In this paper, we propose a holographic technique with an improved Gerchberg-Saxton (GS) algorithm to reconstruct cell imaging based on phase reconstruction information. Comparative experiments are conducted on four specific models to investigate the effectiveness of the proposed method. The morphological parameters (such as shape, volume, and sphericity) of abnormal erythrocytes can be obtained by reconstructing cell hologram of urinary sediment. Notably, abnormal red blood cells can also be detected in mussy circumstances by the proposed method, owing to the significantly biophysical contrast (refractive index distribution and mass density) between two different cells. Therefore, this proposed method has a broad application prospect in cell image reconstruction and cell dynamic detection.

KEYWORDS

computer hologram, cell image, GS algorithm, phase reconstruction, reconstruction

Introduction

In the research field of biomedical and life sciences, researchers expect to capture the cell morphological shapes, dynamic characteristics, physiological parameters, interaction among cells, cell reactions to medicines and drug transport information by visual observation of biological cells in nutrient solution or under natural circumstances, which has an important role in early diagnosis and medicine design (Dixit and Cyr, 2003). Optical microscopy, which is regarded as a conventional medical imaging tool, is capable of magnifying images of minute object specimens with visible light. However, the conventional optical microscope can only examine the strength of the optical wave and achieve a two-dimensional, light intensity distribution map. Optical microscopy has difficulty capturing the three-dimensional spatial information of specimens. In addition, fluorescence microscopy demonstrated great performance in enhancing the contrast in cell images (Otuboah et al., 2019; Wen et al., 2020). Before observation using fluorescence microscopy, stains and dyes, such as rhoda mine, acridine orange, green fluorescent protein or other substances, are usually employed to mark biological cells to enhance the contrast in different biological specimens (Kricka and Fortina, 2009). The staining pretreatment process not only increases the complexity of cell imaging in operation but also greatly impacts cellular activity. Therefore, with the development of biomedical science, conventional cell imaging technologies have difficulty meeting the requirements of the dynamic observation and quantitative analysis of living cells (Ushenko et al., 2018; Borovkova et al., 2019; Ushenko et al., 2019; Peyvasteh et al., 2020).

Optical holography is a technology that adopts a photosensitive medium to record the optical wavefront information and that subsequently superimposes a second wavefront, which is referred to as a diffracted wave, to reconstruct the final image (Su et al., 2003; Wang et al., 2014; Lucchetta et al., 2015; Deng and Li, 2017). Digital holography is established by optical holography, dividing the procedures into recording and reconstruction (Demoli et al., 2019; Xia et al., 2019). Digital holography has drawn much attention from researchers as a cell imaging technique, owing to its advantageous characteristics of low production cost, rapid reconstruction speed, flexible recording and capturing of phase information of optical fields (Wang et al., 2018; Zhong et al., 2020). Digital holography utilizes photoelectric sensors (i.e., CCD or COMS) instead of a photosensitive medium to record the wavefront information, which is processed by computers. The light transmitting process is simulated by a computer based on diffraction propagation theory to quantitatively calculate the amplitude and phase distribution of the recorded object. Although existing optical instruments easily capture the light intensity data of high-frequency optical waves, the phase information is hard to collect or even lost in the course of recording. Notably, the optical wave phase has important morphological information with respect to the specimens and may contain up to 80% information of the imaged specimens. A phase retrieval algorithm has been reported to promote the speed of reconstructing images of X-rays, and this method is responsible for phase holograms in both the visible band and the non-visible band (Hennelly and Sheridan, 2003). Digital holographic techniques have been reported for dynamic estimation by analyzing the specimen shape and deformation analysis (Cui et al., 2010; Jin et al., 2017). Recently, digital holographic techniques also had been adopted for specific neural activity patterns (Adesnik and Abdeladim, 2021), monitoring cell characteristic variations in their density (Delikoyun et al., 2021), and cells migrating in a 3D environment (Hellesvik et al., 2020).

In this paper, we proposed a holographic technique to reconstruct cell imaging based on phase recovery information using an improved GS algorithm. The improved GS algorithm could optimize the initial phase based on local convergence theorem, and accomplish the reconstructed holographic image from contrapuntal phase information. We built four specific models to investigate the effectiveness of the proposed method. Comparative experiment results demonstrated that the proposed method is able to reconstruct phase images with a high signal-to-noise ratio, in which the edge and whole structural information has been reserved. The proposed method is utilized to detect abnormal red blood cells in urine sediment. Specifically, the improved GS algorithm enhances the quality of the phase-reconstructed holographic cell image and successfully detects abnormal red blood cells with distinct edges from the mussy circumstance. This method has great potential for practical biomedical applications in monitoring cell activities.

Methods

Digital holographic instruments

In this section, Figure 1 shows the general setup of the holographic imaging measurement system for holographic information acquisition and object reconstruction, and the optical holography adopts optical interference and diffraction theories to establish the light path. The system is composed of a laser source, beamsplitter cube, lens, reflector and beam receiver CCD. The laser beam emitted by the laser source is split into two orthogonal laser beams *via* the beamsplitter cube, and the laser beams are expanded by a beam expander. The outgoing beams from the beam expander are then collimated by a convex lens, and the beams are regarded as plane waves. One incident plane wave is projected onto the sample object, and the other plane wave goes through the acoustic-optic modulator as the reference wave. The frequency of outgoing waves from the acoustic-optic modulator has changed. Then, the light wave projected by the sample object and reference wave transmit through different optical paths and then become incident on the CCD sensor, on which optical interference simultaneously occurs. The CCD sensor acquires the holographic data, which are saved by the computer. The digital holographic technique eliminates the laser source and complex light path structure with respect to optical holography. The digital holographic technique saves the interference pattern generated on the conventional sensor plane of optical holography as a digital image for simulation (Wang et al., 2019). By computing the complex amplitude map of the optical wave, the phase information or amplitude information can be coded, and the encoding function is adopted to depict the diffracted propagation of the object projected wave from the holographic plane to the object plane. The digital holographic

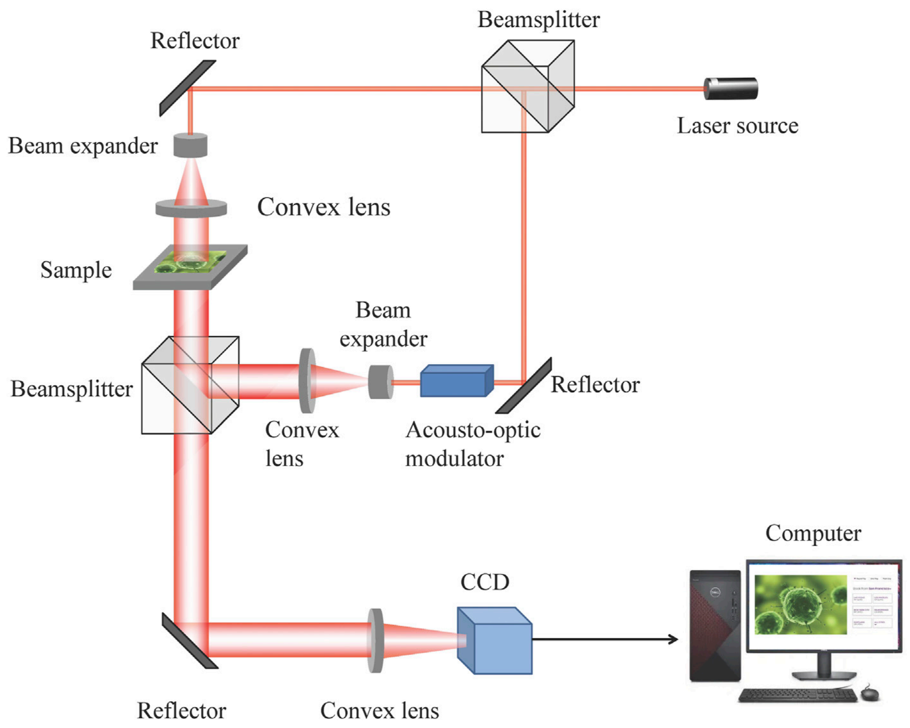


FIGURE 1
General setup of the holographic imaging measurement system.

map is reconstructed by computer holography (Schnars and Jüptner, 2002).

Holographic imaging algorithm

The holographic image is reconstructed in the steps of computer holography. First, we conduct a discrete sampling operation on the optical images to obtain a discrete digital image. A random initial phase distribution of the input optical field corresponding to the amplitude of the sampled image is generated for iterative process. Second, we conduct Fresnel or Fourier transformation on the discrete optical field image. Third, we use an optimal encoding scheme to transform the amplitude and phase information into the optimal transmittance expression function. Last, the holographic image is reconstructed through iterative calculations based on the encoding function.

During computer holographic image reconstruction, we simulate the holographic imaging process and optical diffraction theory by computer calculation based on scalar diffraction theory, which is responsible for describing optical propagation. In computer holography, z defines the distance between the diffractive plane and the observed plane, and $U(x, y, 0)$ and $U(x, y, z)$ demonstrate the complex amplitudes of the optical wave on the diffractive plane and observed plane, respectively. The Fourier spectra regarding the complex amplitudes of the optical wave, diffractive plane and observed plane are $F_0(f_x, f_y)$ and $F_z(f_x, f_y)$. Given that the spectrum of the optical wave on the observed plane is known, the inverse Fourier transformation is conducted to calculate the amplitude on the observed plane, and the calculating function can be expressed as follows:

$$U(x,y,z) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} F_z(f_x,f_y)e^{j2\pi(f_x x+f_y y)}df_x df_y \quad (1)$$

where (f_x, f_y) demonstrates spatial frequency of simple harmonic plane wave and $e^{j2\pi(f_x x+f_y y)}$ demonstrates complex amplitude distribution of simple harmonic plane wave in (x, y) coordinates. Next, the correlation between $F_0(f_x, f_y)$ and $F_z(f_x, f_y)$ is calculated by substituting Eq. 1 into the Helmholtz equation (Liu and Scott, 1987), which is expressed as follows:

$$F_z(f_x,f_y) = F_0(f_x,f_y)e^{j\frac{2\pi}{\lambda}z\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}} \quad (2)$$

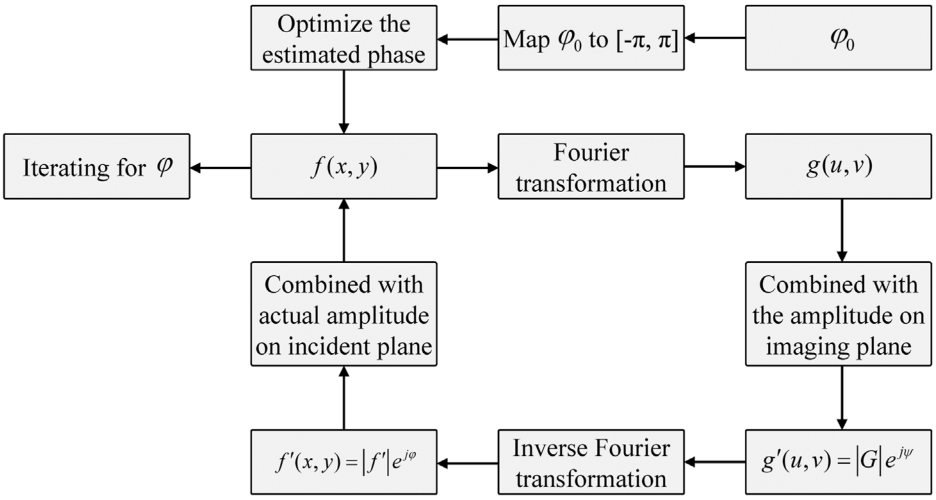


FIGURE 2
Schematic of the improved algorithm calculating steps.

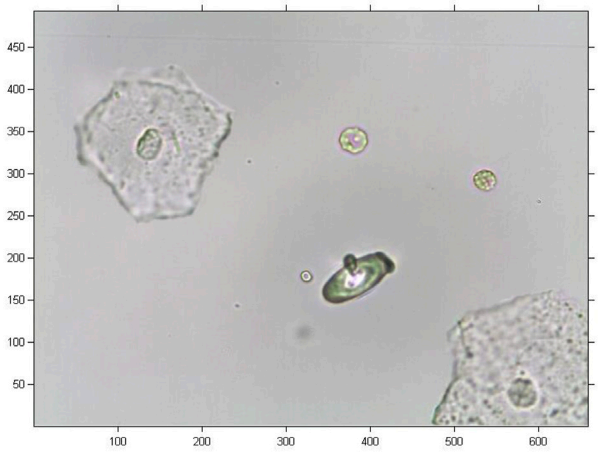


FIGURE 3
Original urine sediment image.

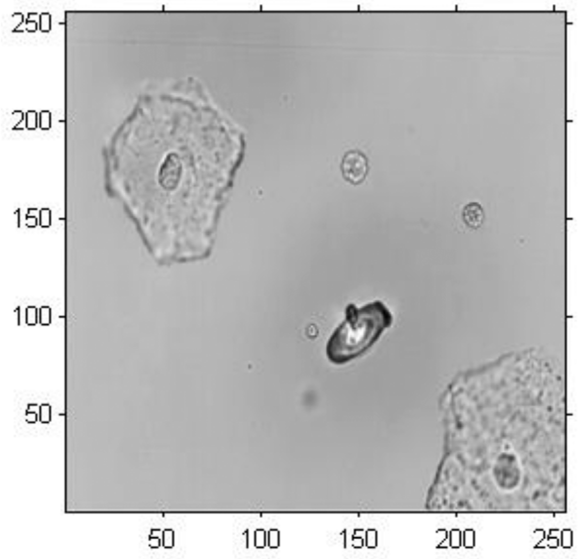


FIGURE 4
Pretreatment of image of urine sediment.

Then, we obtain the following equation.

$$U(x,y,z) = IFFT\langle FFT\{U(x,y,0)\}e^{j\frac{2\pi}{\lambda}z\sqrt{1-(\lambda f_x)^2-(\lambda f_y)^2}}\rangle \quad (3)$$

where FFT demonstrates the Fourier transformation and IFFT demonstrates the inverse Fourier transformation.

The phase reconstruction algorithm is divided into two categories: The first interference phase detection category primarily includes the holographic technique and shear interference method, and the second non-interference detection category primarily includes the iterative method and transport-of-intensity equation method (Teague, 1983; Deutsch et al., 2008). The iterative method is based on the GS algorithm, reported by Gerchberg and Saxton (Gerchberg and Saxton, 1972), which is a typical algorithm to overcome the phase reconstruction problem. This algorithm adopts an iterative method for phase reconstruction by evaluating the relationship between the optical intensity distribution on the incident plane and that on the imaging plane. The GS method has a fast rate of convergence and high accuracy. In this section, we propose an improved GS algorithm, and Figure 2 demonstrates a

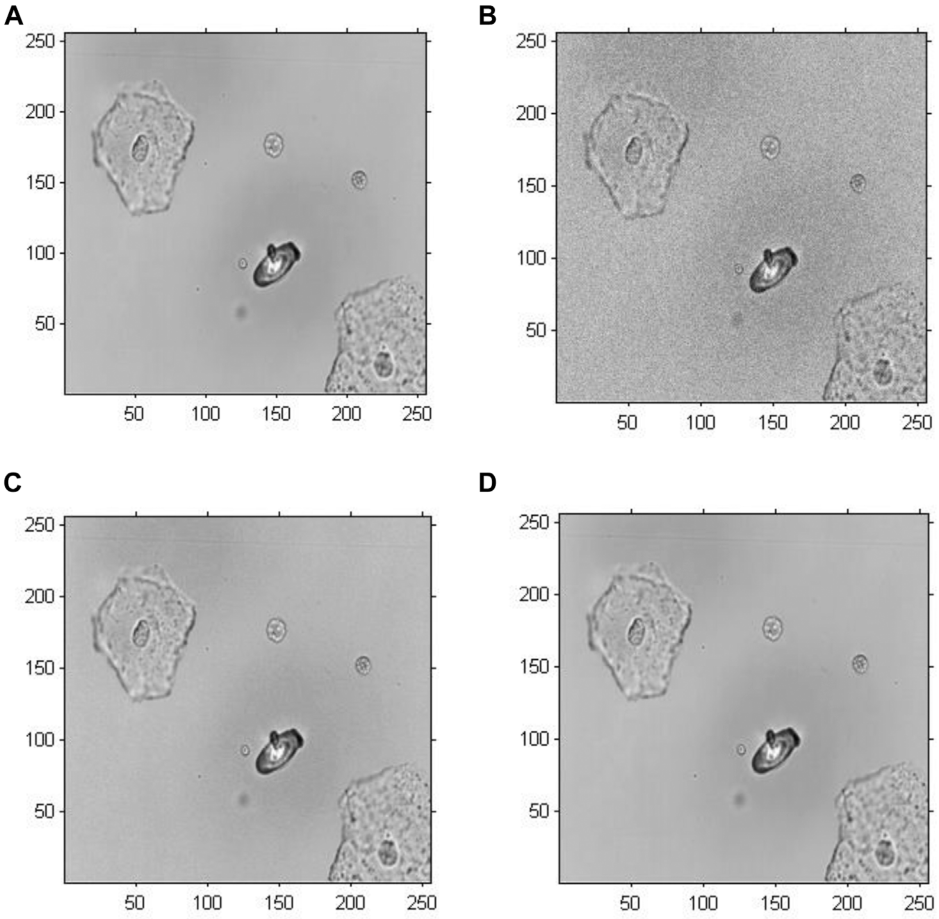


FIGURE 5
Amplitude reconstruction-based holographic image for Model 1. **(A)** Original greyscale image, **(B)** 20 iterative reconstruction images of amplitude information, **(C)** 200 iterative reconstruction images of amplitude information, and **(D)** 2000 iterative reconstruction images of amplitude information.

schematic of holographic imaging with the improved GS algorithm. We normalize the phase and then set a threshold value of the processed phase to achieve a better reconstruction performance. Here, the optical field display on the incident plane using the wavefront coding technique is expressed as:

$$f(x,y) = |f(x,y)|e^{i\varphi(x,y)} \tag{4}$$

The optical field on the imaging plane using the wavefront coding technique is expressed as:

$$g(u,v) = |g(u,v)|e^{i\psi(u,v)} \tag{5}$$

where $f(x,y)$ and $g(u,v)$ demonstrate the amplitude of the optical field on the input plane and that on the imaging plane, respectively. $\varphi(x,y)$ and $\psi(u,v)$ demonstrate the phase on the input plane and imaging plane, respectively. The correlation between $f(x,y)$ and $g(u,v)$ can be expressed as follows:

$$f(x,y) = FFT\{g(u,v)\} \tag{6}$$

A random generating phase of the optical field on the incident plane is input into the iterative operation as the initial phase, and the phase is continuously updated by the following iterations. Then, we normalize the phase and multiply the normalized phase by a factor value to optimize the initial phase. We combine $g(u,v)$ with the amplitude on the imaging plane to obtain $g'(u,v)$. Inverse Fourier transformation is conducted on $g'(u,v)$ to obtain $f'(x,y)$. The above steps are repeated to obtain the final phase.

Comparative experimental models

In our study, an improved GS algorithm was adopted to reconstruct the wavefront phase map of urine sediment. Comparative experiments were conducted to validate the

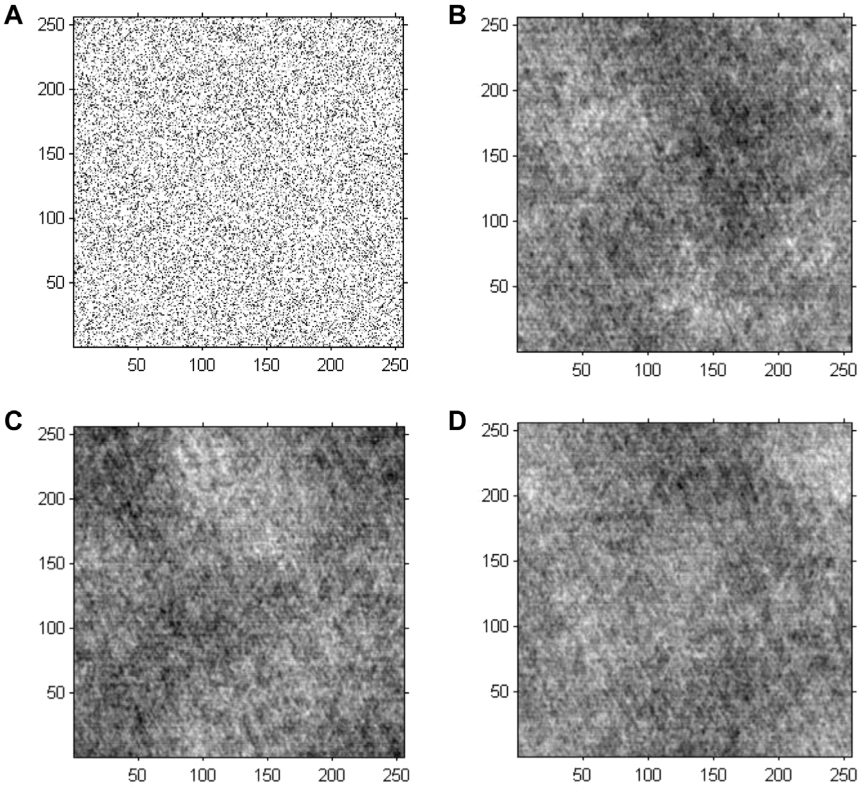


FIGURE 6
Phase reconstruction-based holographic image for Model 2 using the conventional GS algorithm. **(A)** Initial random phase, **(B)** 20 iterative reconstruction images of phase information, **(C)** 200 iterative reconstruction images of phase information, and **(D)** 2000 iterative reconstruction images of phase information.

effectiveness of the proposed method. The performance of computer holography was analysed based on the phase and amplitude information of the cell reconstructed holographic image.

The urine sediment image was adopted as the experimental object. Computer holographic GS modelling is performed as follows: First, we perform preprocessing on an original RGB cell image, including downsampling the image with M × N pixels to an image with 256 × 256 pixels and conducting a greyscale operation. Second, the amplitude in each pixel was multiplied by a randomly generated phase value. Last, the iteration operation using the proposed method was conducted to obtain the reconstructed holographic cell image.

In this study, four models were built to investigate the effectiveness of the improved GS algorithm in cell holographic imaging. For Model 1, the original GS algorithm was adopted to obtain the amplitude information of the image, and the number of iterations was set to 20, 200, and 2,000. Then, we compared the amplitude of the holographic images with different iterations. For Model 2, the GS algorithm was adopted to obtain the phase information of the cell image. Neither Model 1 nor Model 2 was performed with phase mapping values. For Model 3, normalization processing and phase mapping were conducted on the amplitude and phase in Model 2, including transforming the image into greyscale from 0 to 1 and restricting the initial phase map value in the range of $-\pi$ to π. This strategy could avoid the impact of phase deviation on the estimated accuracy. Then, we transformed the input image into an image with a complex format, and a random function was employed to generate a pair of images. The estimated phase was continuously calculated to approximate the actual phase in the iterative loop until the estimated phase was similar to the actual phase. The output estimated phase was mapped from $[-\pi, \pi]$ to $[0, 1]$. The number of iterations was set to 20, 200 and 2,000 for comparison. Since the holographic algorithm takes the negative direction of the gradient of high-dimensional function as the maximum value, when the initial phase is not selected properly or the number of iterations is insufficient, the result will be very different from the optimal value. Thus for Model 4, β factor was introduced to restrict the initial value selection, and the phase matrix was normalized and then multiplied by the β factor value. The number of iterations was also set to 20, 200 and 2,000 for a horizontal comparison.

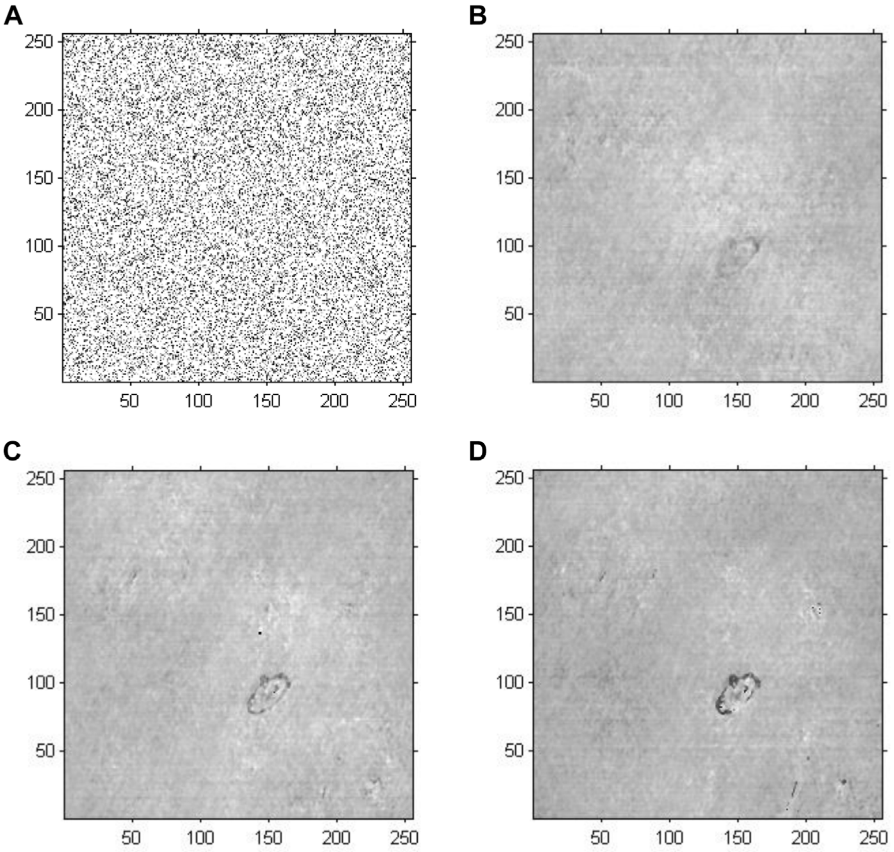


FIGURE 7
Phase reconstruction-based holographic image for Model 3 using the improved GS algorithm with the optimal initial phase. **(A)** Optimized initial phase. **(B)** Twenty iterative reconstruction images of phase information. **(C)** 200 iterative reconstruction images of phase information. **(D)** 2000 iterative reconstruction images of phase information.

To quantitatively estimate the impact of the phase reconstruction, the root mean squared error (RMSE) is employed as an indicator to assess the results. The indicator of the RMSE is expressed as follows:

$$\mathbf{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i')^2} \tag{7}$$

where n is the number of iterations, y_i is the calculated phase of each iteration, and y_i' is the phase distribution of the last iteration.

Results

In this study, the original image was a coloured image of urine sediment with 493 × 659 pixels, as shown in Figure 3. Note that there are several epithelial cells and abnormal red blood cells in this image. There is a protruding cellular bleb on the membrane of the abnormal red blood cell. In order to extract the amplitude and phase information of the cell image, we conducted cropping and graying arithmetic on the original cell image. After the preprocessing the original image, a greyscale image with 256 × 256 pixels was obtained, as shown in Figure 4.

Figure 5 demonstrates the amplitude reconstruction-based holographic image for Model 1, which adopted the original GS algorithm to extract amplitude information with the number of iterations set to 20, 200 and 2,000. According to the visual comparison, there is a slight difference among the amplitude reconstruction-based holographic reconstructed images with different numbers of iterations. The reconstructed cell images only revealed a difference in brightness. Therefore, the amplitude information showed minimal impact on the cell holographic imaging.

The GS algorithm with a random initial phase was conducted on the preprocessed greyscale image to investigate the performance of cell holographic imaging. The phase reconstruction-based cell holographic images for Model 2 are shown in Figure 6. The comparative results demonstrated that the phase image would be fuzzy and chaotic without phase normalization, leading to different results for each calculation.

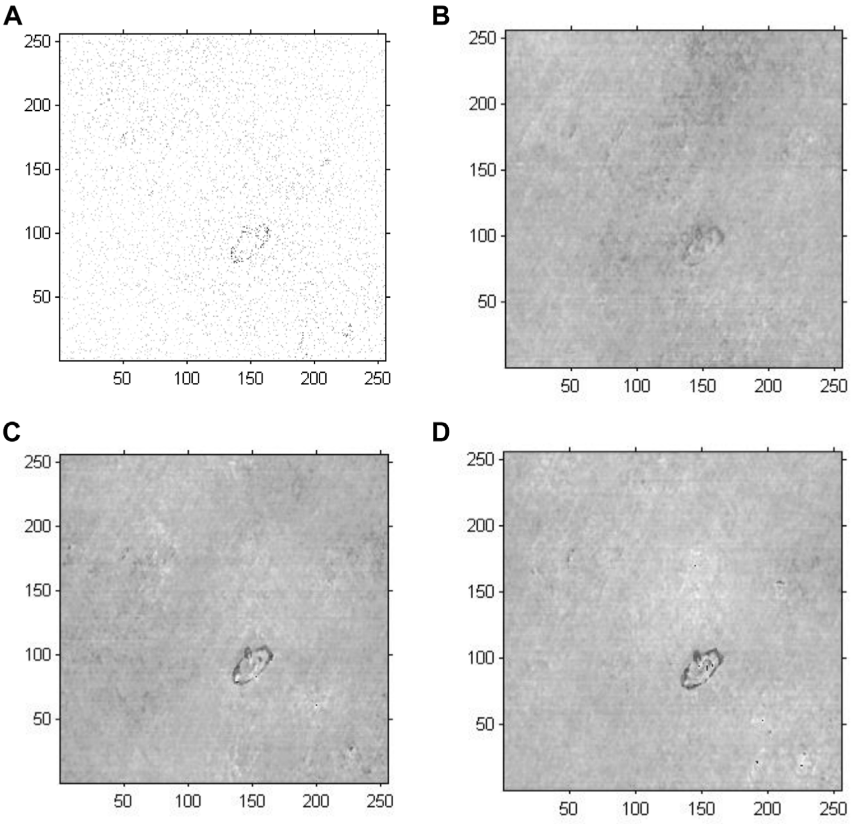


FIGURE 8
Phase reconstruction-based holographic image for Model 4 using the improved GS algorithm. **(A)** Optimized initial phase, **(B)** 20 iterative reconstruction images of phase information, **(C)** 200 iterative reconstruction images of phase information, and **(D)** 2000 iterative reconstruction images of phase information.

Figure 7 demonstrates the phase reconstruction-based holographic image for Model 3, which adopted an improved GS algorithm with a random initial phase to extract phase information with the number of iterations set to 20, 200 and 2,000. According to the cell reconstructed images with different iterations, the comparative results suggested that as the number of iterations increased, the edges and the entire structure of abnormal red blood cells gradually became clearer, and the sharpness of the images increased. Compared with the results of Model 2, the abnormal blood red blood cells were successfully detected, and other impure cells were removed. Specifically, compared with the results of Model 1, the phase information had a more important role in cell holographic imaging than amplitude information, which only depicted the light and dark distribution.

We conducted holographic imaging using an improved GS algorithm with an optimal initial phase to investigate the effectiveness of the proposed method for cell holographic imaging. Figure 8 shows the phase reconstruction-based cell holographic images for Model 4 with the initial phase limited by a certain threshold value and the number of iterations set to 20, 200 and 2,000. A comparison of the cell holographic image with different iterations reveals that the edges and the entire structure of abnormal blood red cells became clearer as the number of iterations increased. A comparison of the image results of Model 3 and Model 4 reveal that the abnormal blood red cell of Model 4 was clearer than the cell image of Model 3, which suggested that the cell holographic image using the proposed method showed effective performance in holographic reconstruction and cell detection.

The root mean squared error (RMSE) was calculated to quantitatively demonstrate the impact of phase reconstruction. We take the logarithm of the RMSE for easy comparison. The RMSE results support the above conclusion that the proposed method is suitable for cell holographic reconstruction.

Model 2 is the unmapped phase original GS algorithm; the phase peaks are disordered and small. Model 3 is the phase of the mapped original GS algorithm; the phase peak at the red blood

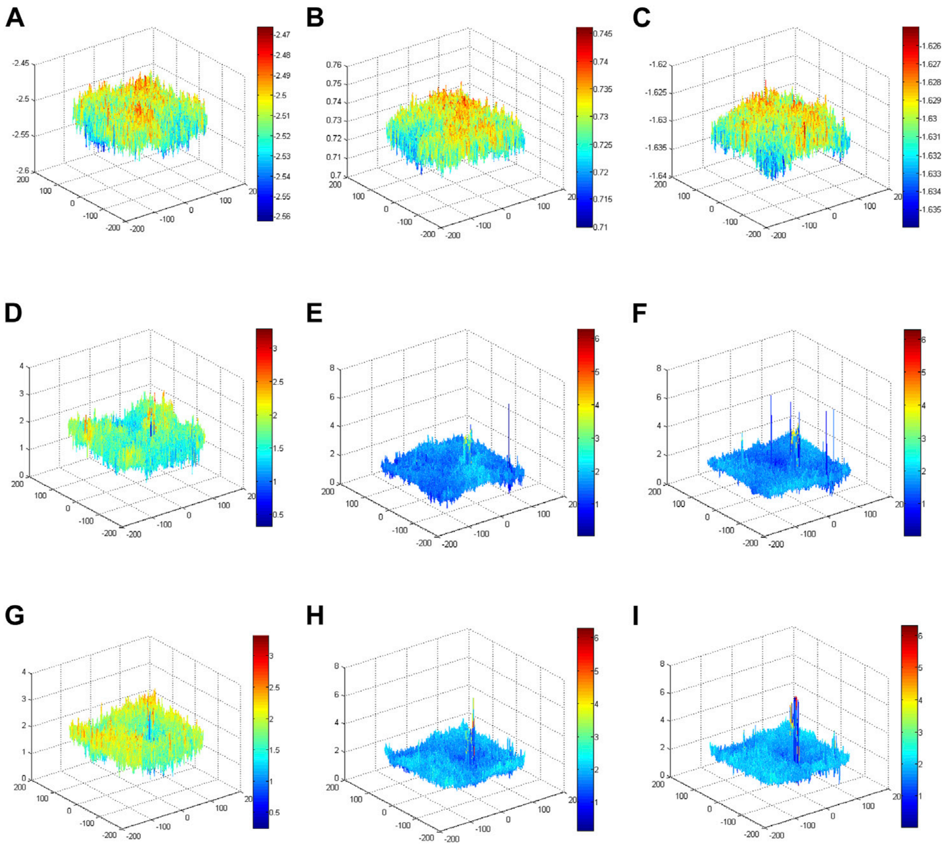


FIGURE 9
Phase energy map using the improved GS algorithm for different models. **(A)** Phase energy map with 20 iterations for Model 2. **(B)** Phase energy map with 200 iterations for Model 2. **(C)** Phase energy map with 2,000 iterations for Model 2. **(D)** Phase energy map with 20 iterations for Model 3. **(E)** Phase energy map with 200 iterations for Model 3. **(F)** Phase energy map with 2,000 iterations for Model 3. **(G)** Phase energy map with 20 iterations for Model 4. **(H)** Phase energy map with 200 iterations for Model 4. **(I)** Phase energy map with 2,000 iterations for Model 4.

cell in the target image is more obvious than that in Model 2. Model 4 is the phase of the mapped improved GS algorithm. Compared with Model 3, the phase peak at the position of the red blood cell is slightly higher, and the performance of the phase imaging image quality is slightly outstanding. Each model ultimately presents a phase energy diagram with 20, 200 and 2,000 iterations.

To verify the enhancement effects of the improved GS algorithm, the phase distributions of different models were conducted using corresponding methods with different iterations, as shown in Figure 9. The phase results of Model 2 revealed that the phase distributions were scrambled and that the peak value was small. Notably, the phase peaks concentrated on the position of the blood red cell in Model 3 and Model 4. More specifically, the phase peaks of Model 4 at the position of the blood red cell were more concentrated and higher than the phase peaks in Model 3. These comparative results validated that the imaging quality using the proposed method was outstanding.

Discussions

In this paper, we proposed an improved GS algorithm for cell reconstructed imaging, which demonstrated the capability of detecting normal red blood cells in urinary mussy circumstances. The proposed method adopted phase mapping and optimization to overcome the random and chaotic defects during conventional phase reconstruction processing.

The holographic image based on amplitude reconstruction results demonstrated that cell edges and detailed information is difficult to extract from the amplitude information. The comparative experimental results suggested that phase information was more important for cell holographic imaging than amplitude information. The amplitude information only demonstrates the light and shade distribution of the image, but does not associate with the content of cell characteristic properties, while the phase information is related to cell characteristic properties, such as edge and structure. According to the results of Model 3 and Model 4, when the phase was mapped in the range of $-\pi$ to π, the impact of the phase deviation on the estimation accuracy could be avoided so that an accurate phase on the imaging plane was obtained. Phase normalization could render subsequent data processing more convenient and accelerate the convergence rate. As the greyscale value has a range, the real part value and imaginary part value of the Fourier transformation were unlimited. The mapping method was adopted to write the two matrices (amplitude and phase) into an image. The optimal initial phase could promote the accuracy of phase reconstruction with a small number of iterations. In addition, the correlation coefficient between the complex amplitude function and the initial input complex amplitude function after inverse transformation reduction was taken as a threshold. Setting constraints on iteration could decrease the iteration times of the GS algorithm and avoid the waste of computing resources.

Cells can be considered phase objects. As many cells have a low absorption or scattering rate of light, cell phase changes are closely associated with the structure and other morphological parameters. Therefore, digital holography has attracted more attention in the field of biological cell imaging as it is capable of quantitatively analysing biological cells and non-destructively and dynamically observing living cells. Specifically, Kemper and Von Bally (2008) applied digital holography technology in biomedical applications and carried out three-dimensional reconstruction of human liver tumour cells and red blood cells. Potcoava and Kim (2008) conducted research on the thickness of blood vessels and retina using digital holographic technology. Wang et al. (2016) promoted the resolution of digital holographic imaging by combining microspheres and image planes. Sandoz et al. (2019) adopted a digital holographic system to estimate the biochemical change in lipid droplets, which are responsible for lipid storage and metabolism. Research on variations in the physicochemical characteristics of the nucleolus in HeLa cells using a digital holographic system has been reported (Kim et al., 2019). Additionally, a report associated with the observation of living diatoms in seawater using a digital holographic system was presented (Umemura et al., 2020). In addition to quantitatively imaging single cells, many studies have referred to cell pathology monitoring using digital holographic techniques (Charrière et al., 2006; Sung et al., 2009; Hsu et al., 2012).

Conclusion

In this paper, digital holographic technique could analyse cell physiological property without staining and dyeing, which is more suitable for the reconstruction of cell phase information, and makes up for the shortage that the traditional optical microscope can only provide amplitude images but is difficult to accurately distinguish the cell edges. An improved Gerchberg-Saxton algorithm was adopted to reconstruct the cell holographic image in light of the phase reconstruction information. Comparative experiments were carried out on four specific models with different phase constraint contidions, and the cell holographic image results clearly verified and revealed that this proposed improved GS method achieved better performance than the conventional GS method on the basis of phase reconstruction data. As the phase image reveals not only the spatial information but also the temporal statistics during the transmission of various frequency waves, the accumulation of phase information is able to depict the accurate shape parameters of the recorded object. Then, validation experiments were conducted on urine sediment images, and the reconstructed holographic image results demonstrated that this improved GS method was able to extract abnormal red blood cells with sharp edges and filter out other impure objects. In conclusion, the reconstructed holographic phase image achieved by the proposed method had a high signal-to-noise ratio, and the method was specific to discriminate the morphological parameters. Thus, these results demonstrated the possibility of this proposed holographic method in practical biomedical applications in cell activities monitoring.

Data availability statement

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

Author contributions

XW: experiments design and conception. JL, HW, and XL: performed experiments and data analysis. YJ, HL, YP, YT, and YY: contributed to data analysis and wrote the draft of the manuscript. All the authors edited the manuscript.

Funding

This work was sponsored in part by the National Natural Science Foundation of China (Grant No. 61605021); in part by Science and Technology Research Program of Chongqing Municipal Education Commission (No. KJQN202000624) in part by Opening Foundation of Key Laboratory of Opto-technology and Intelligent Control (Lanzhou Jiao tong University), The Ministry of Education (KFKT 2020-08); in part by China Postdoctoral Science Foundation (No. 2021MD703941); in part by special support for Chongqing post-doctoral research project (2021XM 2051); in part by Natural Science Foundation of Chongqing (Nos. cstc2019jscx-msxmX0079 and cstc2019jcyj-msxmX0275); in part by Project of Central Nervous System Drug Key Laboratory of Sichuan Province (200028-01SZ); in part by the Doctoral Foundation of Chongqing University of Posts and Telecommunications (A2020-19, A2018-89, X2020-53, A2022-255).

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

Adesnik, H., and Abdeladim, L. (2021). Probing neural codes with two-photon holographic optogenetics. *Nat. Neurosci.* 24, 1356–1366. doi:10.1038/s41593-021-00902-9

Borovkova, M., Trifonyuk, L., Ushenko, V., Dubolazov, O., Vanchulyak, O., Bodnar, G., et al. (2019). Mueller-matrix-based polarization imaging and quantitative assessment of optically anisotropic polycrystalline networks. *PloS One* 14, e0214494. doi:10.1371/journal.pone.0214494

Charriere, F., Pavillon, N., Colomb, T., Depeursinge, C., Heger, T. J., Mitchell, E. A., et al. (2006). Living specimen tomography by digital holographic microscopy: Morphometry of testate amoeba. *Opt. Express* 14, 7005–7013. doi:10.1364/OE.14.007005

Cui, H., Wang, D., Wang, Y., Liu, C., Zhao, J., and Li, Y. (2010). "Automatic procedure for non-coplanar aberration compensation in lensless Fourier transform digital holography," in Proceedings of the 5th international Symposium on advanced optical Manufacturing and testing technologies: Optical Test and measurement Technology and equipment: SPIE, October 2010, 213–219. doi:10.1117/12.865688

Delikoyun, K., Yaman, S., Yilmaz, E., Sarigil, O., Anil-Inevi, M., Telli, K., et al. (2021). HologLev: A hybrid magnetic levitation platform integrated with lensless holographic microscopy for density-based cell analysis. *ACS Sens.* 6, 2191–2201. doi:10.1021/acssensors.0c02587

Demoli, N., Gladić, J., Lovrić, D., and Abramović, D. (2019). Digital holography using LCOS microdisplay as input three-dimensional object. *Optik* 194, 162877. doi:10.1016/j.ijleo.2019.05.083

Deng, Z. L., and Li, G. (2017). Metasurface optical holography. *Mater. Today Phys.* 3, 16–32. doi:10.1016/j.mtphys.2017.11.001

Deutsch, B., Hillenbrand, R., and Novotny, L. (2008). Near-field amplitude and phase recovery using phase-shifting interferometry. *Opt. Express* 16, 494–501. doi:10.1364/OE.16.000494

Dixit, R., and Cyr, R. (2003). Cell damage and reactive oxygen species production induced by fluorescence microscopy: Effect on mitosis and guidelines for non-invasive fluorescence microscopy. *Plant J.* 36, 280–290. doi:10.1046/j.1365-313X.2003.01868.x

Gerchberg, R. W., and Saxton, W. O. (1972). A practical algorithm for the determination of plane from image and diffraction pictures. *Optik* 35, 237–246.

Hellesvik, M., Øye, H., and Aksnes, H. (2020). Exploiting the potential of commercial digital holographic microscopy by combining it with 3D matrix cell culture assays. *Sci. Rep.* 10, 14680–14714. doi:10.1038/s41598-020-71538-1

Hennelly, B., and Sheridan, J. T. (2003). Fractional fourier transform-based image encryption: Phase retrieval algorithm. *Opt. Commun.* 226, 61–80. doi:10.1016/j.optcom.2003.08.030

Hsu, W. C., Su, J. W., Chang, C. C., and Sung, K. B. (2012). "Investigating the backscattering characteristics of individual normal and cancerous cells based on experimentally determined three-dimensional refractive index distributions," in Proceedings of the Optics in health Care and biomedical optics V: SPIE, December 2012, 248–254. doi:10.1117/12.999804

Jin, D., Zhou, R., Yaqoob, Z., and So, P. T. (2017). Tomographic phase microscopy: Principles and applications in bioimaging [invited]. *J. Opt. Soc. Am. B* 34, B64–B77. doi:10.1364/JOSAB.34.000B64

Kemper, B., and Von Bally, G. (2008). Digital holographic microscopy for live cell applications and technical inspection. *Appl. Opt.* 47, A52–A61. doi:10.1364/AO.47.000A52

Kim, T. K., Lee, B. W., Fujii, F., Kim, J. K., and Pack, C. G. (2019). Physicochemical properties of nucleoli in live cells analyzed by label-free optical diffraction tomography. *Cells* 8, 699. doi:10.3390/cells8070699

Kricka, L. J., and Fortina, P. (2009). Analytical ancestry:"Firsts" in fluorescent labeling of nucleosides, nucleotides, and nucleic acids. *Clin. Chem.* 55, 670–683. doi:10.1373/clinchem.2008.116152

Liu, G., and Scott, P. D. (1987). Phase retrieval and twin-image elimination for in-line Fresnel holograms. *J. Opt. Soc. Am. A* 4, 159–165. doi:10.1364/JOSAA.4.000159

Lucchetta, D., Spegni, P., Di Donato, A., Simoni, F., and Castagna, R. (2015). Hybrid surface-relief/volume one dimensional holographic gratings. *Opt. Mater.* 42, 366–369. doi:10.1016/j.optmat.2015.01.028

Otuboah, F. Y., Jihong, Z., Tianyun, Z., and Cheng, C. (2019). Design of a reduced objective Lens fluorescence dPCR Gene chip detection system with high-throughput and large field of view. *Optik* 179, 1071–1083. doi:10.1016/j.ijleo.2018.11.052

Peyvasteh, M., Dubolazov, A., Popov, A., Ushenko, A., Ushenko, Y., and Meglinski, I. (2020). Two-point Stokes vector diagnostic approach for characterization of optically anisotropic biological tissues. *J. Phys. D. Appl. Phys.* 53, 395401. doi:10.1088/1361-6463/ab9571

Potcoava, M., and Kim, M. (2008). Optical tomography for biomedical applications by digital interference holography. *Meas. Sci. Technol.* 19, 074010. doi:10.1088/0957-0233/19/7/074010

Sandoz, P. A., Tremblay, C., Van Der Goot, F. G., and Frechin, M. (2019). Image-based analysis of living mammalian cells using label-free 3D refractive index maps reveals new organelle dynamics and dry mass flux. *PLoS Biol.* 17, e3000553. doi:10.1371/journal.pbio.3000553

Schnars, U., and Jüptner, W. P. (2002). Digital recording and numerical reconstruction of holograms. *Meas. Sci. Technol.* 13, R85–R101. doi:10.1088/0957-0233/13/9/201

Su, W. C., Chen, Y. W., Ouyang, Y., Sun, C. C., and Wang, B. (2003). Optical identification using a random phase mask. *Opt. Commun.* 219, 117–123. doi:10.1016/S0030-4018(03)01308-7

Sung, Y., Choi, W., Fang-Yen, C., Badizadegan, K., Dasari, R. R., and Feld, M. S. (2009). Optical diffraction tomography for high resolution live cell imaging. *Opt. Express* 17, 266–277. doi:10.1364/OE.17.000266

Teague, M. R. (1983). Deterministic phase retrieval: A green's function solution. *J. Opt. Soc. Am.* 73, 1434–1441. doi:10.1364/JOSA.73.001434

Umemura, K., Matsukawa, Y., Ide, Y., and Mayama, S. (2020). Corrigendum to "Label-free imaging and analysis of subcellular parts of a living diatom cylindrotheca sp. using optical diffraction tomography" [MethodsX: 7 (2020) 100889]. *MethodsX* 7, 100923. doi:10.1016/j.mex.2020.100889

Ushenko, V., Sdobnov, A., Syvokorovskaya, A., Dubolazov, A., Vanchulyak, O., Ushenko, A., et al. (2018). 3D Mueller-matrix diffusive tomography of polycrystalline blood films for cancer diagnosis. *Photonics* 5, 54. doi:10.3390/photonics5040054

Ushenko, V., Sdobnov, A. Y., Mishalov, W., Dubolazov, A., Olar, O., Bachinskyi, V., et al. (2019). Biomedical applications of Jones-matrix tomography to polycrystalline films of biological fluids. *J. Innov. Opt. Health Sci.* 12, 1950017. doi:10.1142/S1793545819500172

Wang, X., Jiang, Y., Li, Z., Wang, W., and Li, Z. (2019). Sensitivity characteristics of microfiber fabry-perot interferometric photoacoustic sensors. *J. Light. Technol.* 37 (17), 4229–4235. doi:10.1109/JLT.2019.2922223

Wang, X., Jin, L., Li, J., Ran, Y., and Guan, B. O. (2014). Microfiber interferometric acoustic transducers. *Opt. Express* 22 (7), 8126–8135. doi:10.1364/oe.22.008126

Wang, X. X., Li, Z. Y., Tian, Y., Wang, W., Pang, Y., and Tam, K. Y. (2018). Two dimensional photoacoustic imaging using microfiber interferometric acoustic transducers. *Opt. Commun.* 419, 41–46. doi:10.1016/j.optcom.2018.02.048

Wang, Y., Guo, S., Wang, D., Lin, Q., Rong, L., and Zhao, J. (2016). Resolution enhancement phase-contrast imaging by microsphere digital holography. *Opt. Commun.* 366, 81–87. doi:10.1016/j.optcom.2015.12.031

Wen, C., Kong, Y., Zhao, J., Li, Y., Yu, M., Zeng, S., et al. (2020). Efficacy of the photon-initiated photoacoustic streaming combined with different solutions on he root canals. *Enterococcus faecalis in tMicrosc. Res. Tech.* 83, 647–657. doi:10.1002/jemt.23455

Xia, P., Wang, Q., Ri, S., and Tsuda, H. (2019). Calibrated phase-shifting digital holography based on space-division multiplexing. *Opt. Lasers Eng.* 123, 8–13. doi:10.1016/j.optlaseng.2019.06.022

Zhong, Z., Zhao, H., Shan, M., Liu, B., Lu, W., and Zhang, Y. (2020). Off-axis digital holographic microscopy with divided aperture. *Opt. Lasers Eng.* 127, 105954. doi:10.1016/j.optlaseng.2019.105954

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 04 November 2022
DOI 10.3389/fphys.2022.1066999

OPEN ACCESS

EDITED BY
Huan Qin,
South China Normal University, China

REVIEWED BY
Liye Zhang,
Shandong University of Technology, China
Min Li,
Nanjing University of Science and Technology, China
Fangmin Sun,
Shenzhen Institutes of Advanced Technology (CAS), China

*CORRESPONDENCE
Guo Zhang,
zhangguo@swmu.edu.cn

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

RECEIVED 11 October 2022
ACCEPTED 25 October 2022
PUBLISHED 04 November 2022

CITATION
Sun W, Pang Y and Zhang G (2022), CCT: Lightweight compact convolutional transformer for lung disease CT image classification.
Front. Physiol. 13:1066999.
doi: 10.3389/fphys.2022.1066999

CCT: Lightweight compact convolutional transformer for lung disease CT image classification

Weiwei Sun[1], Yu Pang[1] and Guo Zhang[1,2]*

[1]College of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China, [2]School of Medical Information and Engineering, Southwest Medical University, Luzhou, China

Computed tomography (CT) imaging results are an important criterion for the diagnosis of lung disease. CT images can clearly show the characteristics of lung lesions. Early and accurate detection of lung diseases helps clinicians to improve patient care effectively. Therefore, in this study, we used a lightweight compact convolutional transformer (CCT) to build a prediction model for lung disease classification using chest CT images. We added a position offset term and changed the attention mechanism of the transformer encoder to an axial attention mechanism module. As a result, the classification performance of the model was improved in terms of height and width. We show that the model effectively classifies COVID-19, community pneumonia, and normal conditions on the CC-CCII dataset. The proposed model outperforms other comparable models in the test set, achieving an accuracy of 98.5% and a sensitivity of 98.6%. The results show that our method achieves a larger field of perception on CT images, which positively affects the classification of CT images. Thus, the method can provide adequate assistance to clinicians.

KEYWORDS

axial attention, compact convolutional transformer, COVID-19, positional bias term, image classification

1 Introduction

According to real-time statistics from the World Health Organization (WHO) and Hopkins University, as of 1 August 2022, there were an estimated 570 million confirmed COVID-19 cases worldwide, with −6.4 million deaths (Dong et al., 2020; Zhu et al., 2020). With an increasing number of new cases recorded worldwide, COVID-19 has considerably impacted industries. Additionally, people's everyday lives have been seriously affected. Therefore, the primary means of prevention and detection entail controlling the spread of COVID-19. In clinical settings, nasopharyngeal and oropharyngeal swabs are the main screening methods for COVID-19 (Xu et al., 2020). However, many circumstances might cause a false negative test result (Bai et al., 2020). For example, at the initial stage, when the virus enters the human body, the amount of virus present in the human body is within an undetectable level. And different sampling times and locations may yield insufficient viral amounts in the samples. In addition, the laboratory equipment and the testing capabilities are poor, and a quality management system has not been established. Thus, these restrictions increase the risk of COVID-19 transmission and cause patients to receive delayed treatment or a wrong diagnosis.

The advantages of computed tomography (CT) are noninvasiveness, high resolution, and timeliness, which help diagnose COVID-19. CT expedites the diagnostic processes and is an effective supplement to nucleic acid detection. CT images can clearly find lesions, observe their size, shape, texture, and other characteristics, and accurately segment them (Bernheim et al., 2020; Rubin et al., 2020; Wong et al., 2020). Analyzing the degree of pulmonary involvement and the severity of infection helps support the follow-up clinical treatment of patients. However, community pneumonia (CP) is also associated with cough, sputum, malaise, and fever (Afshar et al., 2020; Zhang et al., 2020a; Brunese et al., 2020; Han et al., 2020; Mahmud et al., 2020; Oh et al., 2020; Ozturk et al., 2020; Calderon-Ramirez et al., 2021; Ozyurt et al., 2021), and CT images of community pneumonia are very similar to COVID-19. This not only makes it more difficult to read the images (Shi et al., 2020) but also greatly increases the workload of the doctors. Further, manually labeling the infected area is time-consuming, and the accuracy is subject to the doctor's subjectivity.

Deep learning (Li et al., 2009; Li et al. 2010; Li et al. 2015; Ardakani et al., 2020) has demonstrated excellent capabilities in auxiliary lung diagnosis recently. It can automatically mine high-dimensional features related to clinical outcomes from CT images. The deep learning-based COVID-19 image classification model has successfully assisted in patient disease diagnosis (Esteva et al., 2017; Litjens et al., 2017; Ardila et al., 2019; Esteva et al., 2019; Qin et al., 2019; Topol 2019; Mei et al., 2020; Sun et al., 2022). An automatic and accurate method for COVID-19 detection based on the ResNet50 model was proposed (Li et al., 2020). And 4,356 chest CT images of 3,322 patients were used to distinguish between COVID-19, CP, and non-pneumonia. The sensitivity, specificity, and area under the curve (AUC) scores of the model were 90%, 96%, and 0.96, respectively. A method for COVID-19 detection based on the DenseNet201 depth transfer model was proposed (Jaiswal et al., 2020). The model was trained using the Image Net dataset and was 96.3% accurate in classifying and recognizing chest CT images. Further, Wu et al. integrated COVID-19 classification and lesion segmentation into the COVID-CS network, and the two tasks shared the same backbone network (Wu et al., 2021). The classification test set obtained an average sensitivity and specificity of 95.0% and 93.0%, respectively. Some researchers built Dense Net-121 to identify the CT images of COVID-19 in a comparative experiment to achieve self-supervised learning and an accuracy of 85.5% (Chen et al., 2021).

However, the classification of COVID-19 still has the following problems. At present, many algorithms (Li et al., 2020; Wang et al., 2020; Hassani et al., 2021) can be used to partially solve the problem of scarce COVID-19 data. But most methods are difficult to accurately capture the essential feature space of various categories of data in a small amount of image data. And, most of the existing algorithms have poor classification performance for common pneumonia and COVID-19, which seriously affects the overall classification performance of the algorithms. It will hurt the subsequent research and eventually make the algorithms difficult to be applied in the clinic.

Therefore, to increase the recognition ability of the model for common pneumonia and COVID-19, and further improve the accuracy and efficiency of COVID-19 image recognition, we employ a novel method to solve the above problems in the CT image classification of COVID-19. A new sequence pooling approach and convolution are proposed herein, i.e., a smaller and more compact transformer based on CCT suitable for datasets lacking pneumonia images. First, the self-attention mechanism in CCT is decomposed into two one-dimensional (1D) self-attention mechanisms: height axis and width axis (Ho et al., 2019; Huang et al., 2019). Subsequently, while the axial attention mechanism replaces the original self-attention mechanism, location coding is added to obtain a larger receptive field. Finally, the position offset item is added to the position-coding to obtain the dependence of the precise position information during training. Herein, the addition of the axial attention mechanism considerably improved the accuracy of COVID-19 detection on chest CT images, achieving better performance results for both COVID-19 model accuracy and other pneumonia screenings. The main innovations herein are as follows. 1. A new sequence pooling strategy and convolution are proposed along with a smaller and more compact transformer based on CCT; this transformer is suitable for datasets lacking pneumonia images. 2. We improved the self-attention mechanism of the transformer encoder to an axial attention mechanism and added a position offset term. The long-range location dependencies of accurate location information are obtained during training to improve the model's classification performance. 3. Compared to the Vision Transformer (ViT) structure and the traditional Convolutional Neural Network (CNN), the performance on the small COVID-19 dataset is stronger.

2 Materials and methods

Our proposed sequence pooling method and convolution module of the CCT model can reduce the class token and embedding requirements. The convolution module can be adapted to the small COVID-19 dataset. The model belongs to the lightweight transformer structure and comprises a convolution module, embedding, transformer encoder,

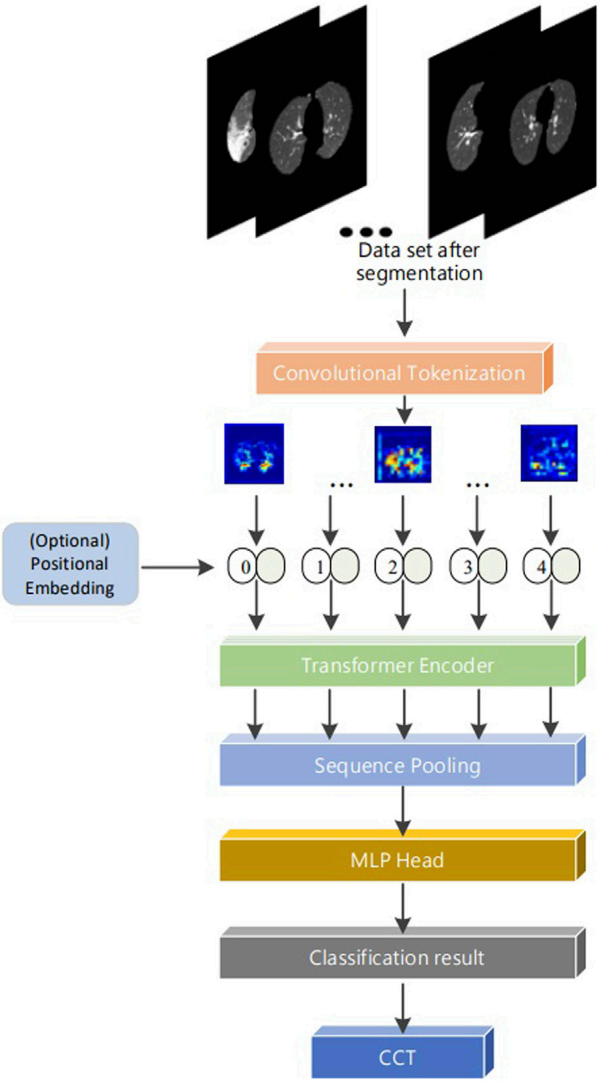


FIGURE 1
Our compact convolutional transformer models.

sequence pooling, and multilayer perceptron (MLP) (Ramchoun et al., 2017) head (Figure 1).

2.1 Improved compact convolutional transformer model structure

We propose a patching method based on small-scale convolutional modules in the CCT to completely preserve local information. This method does not affect how the transformer encoder calculates patch interactions. First, after the input feature vector of the convolution module is normalized, the convolution operation and the ReLU function are used for feature extraction. Second, down-sampling through max-pooling extracts essential information. Third, the residual structure of ResNet50 is employed as an additional feature extraction to prevent the transformer structure data from being unable to be trained during the backpropagation process. Finally, the output vector processed by the convolution module meets the input dimension requirements of the embedding layer. Subsequently, the 3D vector is down-sampled, and the ReLU activation layer is performed. After convolution and flattening operations, the vector dimension of the same size as a position embedding layer of the improved model is obtained (Figure 2).

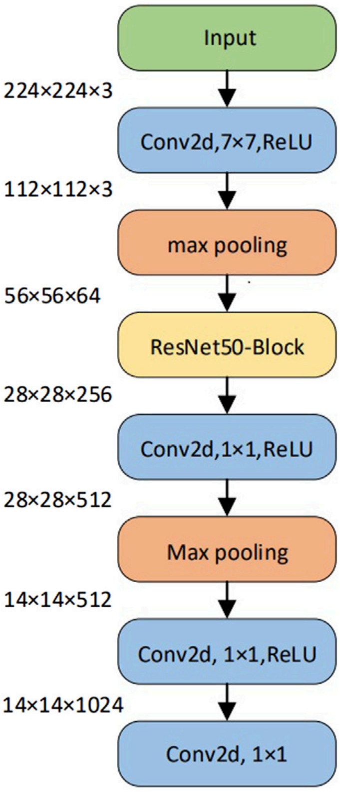


FIGURE 2
CCT diagram of convolution module.

The CCT can adapt to training with smaller datasets by adjusting the size of patches. The CCT introduces a patching method based on convolution. The relationship between patches can be encoded while restraining the local information. This method can effectively tokenize and maintain the local spatial relationship, thereby eliminating dependence on the class token and providing greater model flexibility.

2.2 Transformer encoder

The transformer encoder of the CCT was consistent with that of the ViT. Multiple encoders exist in the model, with no weight

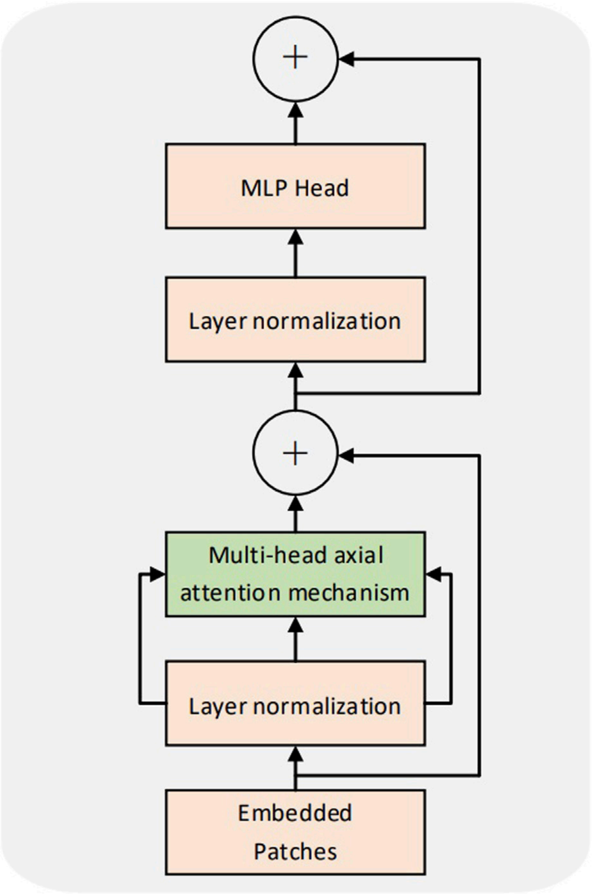


FIGURE 3
Encoder structure drawing.

sharing among them. Figure 3 illustrates the structure of the encoder. Each coding layer comprises two sublayers: multihead self-attention (MSA) and MLP. Each sublayer is preceded by layer normalization. The input sequence was set to x; the output y of a single coding layer was obtained. The formula is as follows:

$$x'_{l-1} = x_{l-1} + W - MSA(LN(x_{l-1})) \tag{1}$$

$$\ell \in 1, 2, \ldots, L \tag{2}$$

$$x_\ell = x'_\ell + MLP(LN(x_{\ell-1})) \tag{3}$$

In Eqs 2, 3, a structure similar to the residual network (He et al., 2016) is laid out. This design retains more information, reduces information loss, and can use a more significant number of encoders for training. L denotes the number of encoders.

The transformer can establish distance dependence on the target while extracting more powerful features by multifocusing on the global content information. The self-attention mechanism in the encoder, given a height of *h*, a width of *w*, and a channel of input embedded patches $X \in \mathbb{R}^{h \times w \times c_{in}}$, and an output formula $y_o \in \mathbb{R}^{h \times w \times c_{out}}$ with position $o = \{i, j \mid i \in \{1, \ldots, h\}, j \in \{1, \ldots, w\}\}$ is defined as follows:

$$q = W_Q x \tag{4}$$

$$k = W_K x \tag{5}$$

$$v = W_V x \tag{6}$$

$$y_o = \sum_p soft\max\left(q_o^T k_p\right) v_p \tag{7}$$

The *q*, *k*, and *v* vectors in Eqs 4–6 are the query, key, and value, respectively. W_Q, W_K and $W_V \in R^{(c_{in} \times c_{outn})}$ are the weight matrices learned during training. In Eq. 6, v is multiplied by the input xi and the trainable matrix WV to obtain the input eigenvector. The dot product of q and k is used to calculate the weight of *v*. In Eq. 7, p = (w, h), q and k are normalized by SoftMax and multiplied by v to obtain the attention value. In contrast to convolution, the self-attention mechanism may obtain nonlocal information from the entire feature map. However, the calculation of this attention value comes at a cost. Applying the self-attention mechanism to the visual model architecture becomes impossible as the feature map increases. Additionally, the self-attention layer does not use any position information when calculating the nonlocal context. However, the position information is vital for obtaining the structure and shape of the target in the visual model.

Based on the abovementioned reasons, the axial attention mechanism is divided into two 1D self-attention mechanisms: the height and width axes. Additionally, a position code was added to the query mechanism. The structural diagram is shown in Figure 4. The axial attention mechanism can also match the original self-attention mechanism dimensions. The width and height dimensions are considered to reduce the number of calculations and improve the calculation efficiency. The position offset terms are set while collating the attention value to make it more responsive to the position information. This bias term is usually called relative position coding and can be learned through training.

The attention model of Ramachandran et al. uses relative position coding for queries only. This study combines the axial attention mechanism and position coding to apply them to all queries, keys, and values. For any given input feature map *x*, an axial position-sensitive attention mechanism with position encoding along the width-axis, the equation is as follows:

$$y_{ij} = \sum_{p=1}^{p=1} soft\max\left(q_{ij}^T k_{iw} + q_{ij}^T r_{iw}^q + k_{ij}^T r_{iw}^k\right)\left(v_{iw} + r_{iw}^v\right) \tag{8}$$

where $r^q, r^k, r^v \in R^{W \times W}$, $\mathrm{r}_{\mathrm{iw}}^{\mathrm{q}}$, $\mathrm{r}_{\mathrm{iw}}^{\mathrm{k}}$, and $\mathrm{r}_{\mathrm{iw}}^{\mathrm{v}}$ are learnable vectors representing the position codes of queries, keys, and values. For example, the attention mechanisms of the height and width axes have the same definition. One axial attention layer spreads information on a specific axis, and both axial attention layers use an MSA mechanism. After the position offset term is an introduction to the axial attention mechanism Compared to the original self-

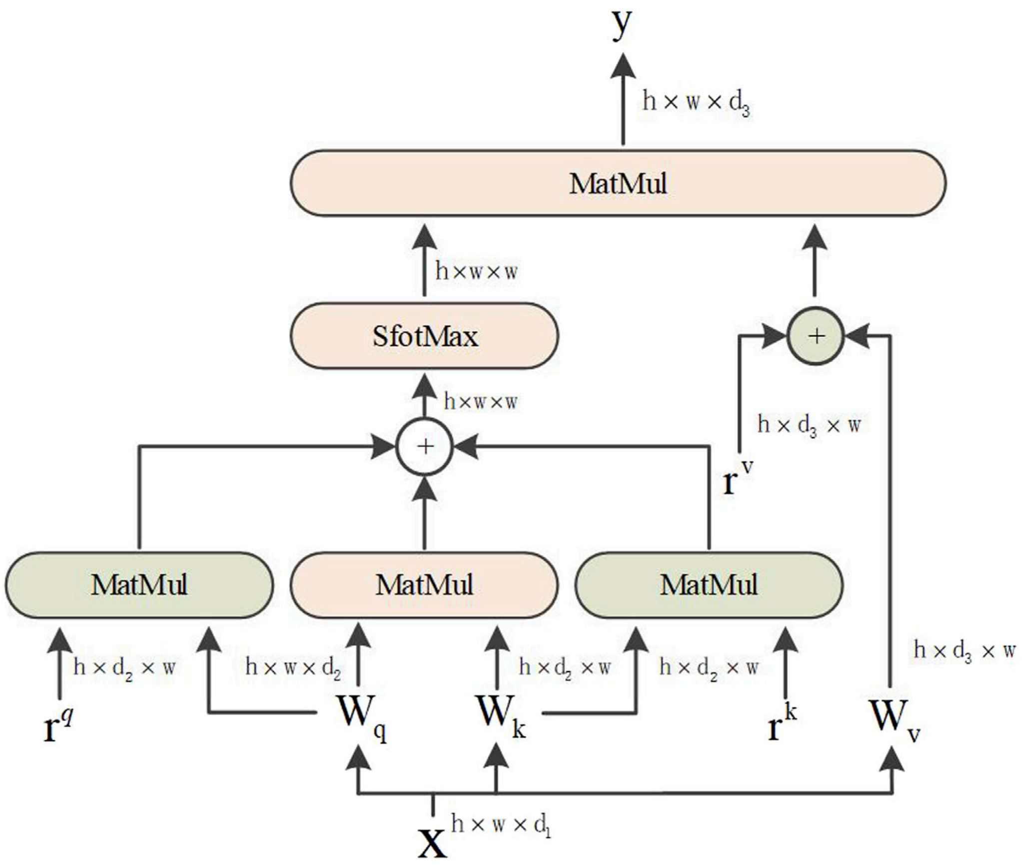


FIGURE 4
Axial attention.

attention mechanism, the global receptive field acquisition feature can be realized, thus reducing the computational complexity.

2.3 Serial pooling

Herein, the feature vector classification results are output using sequence pooling rather than class tokens (Devlin et al., 2018). For the L-layer transformer encoder, the output results are collected in sequence. The model is compact as the data sequence includes information and category information for different parts of the input image, thereby compacting the model. Sequence pooling outputs the sequential embedding of the latent space generated by the transformer encoder to correlate the input data better. The output feature mapping is defined as $T\colon \mathbb{R}^{b\times n\times d} \mapsto \mathbb{R}^{b\times d}$, and the equation is given as follows:

$$X_L = f\,(X_0) \tag{9}$$

where X_L or f (X_0) is the L-layer Transformer encoder, b is the batch size, n is the sequence length, d is the embedding dimension, and $(X_L) \in \mathbb{R}^{d\times 1}$. Using the SoftMax activation function, the equation is given as follows:

$$X_L' = soft\max\left(g\,(X_L)^T\right) \tag{10}$$

As $(X_L) \in \mathbb{R}^{d\times 1}$, we get:

$$Z = X_L' X_L = soft\max\left(g\,(X_L)^T\right) \times X_L \tag{11}$$

where $z \in \mathbb{R}^{b\times 1\times d}$ merge the second dimension to get $z \in \mathbb{R}^{b\times d}$. This output can then be used to obtain the result through a linear classifier.

2.4 Datasets

2.4.1 Lung data COVID-19 CT-CCII

We used the classification dataset from the China Consortium for Chest CT Imaging Research (CC-CCII) (Zhang et al., 2020b; Zhou et al., 2021). Informed consent from the patients was obtained, reviewed, and approved by the Medical Ethics Committee. The dataset comprises

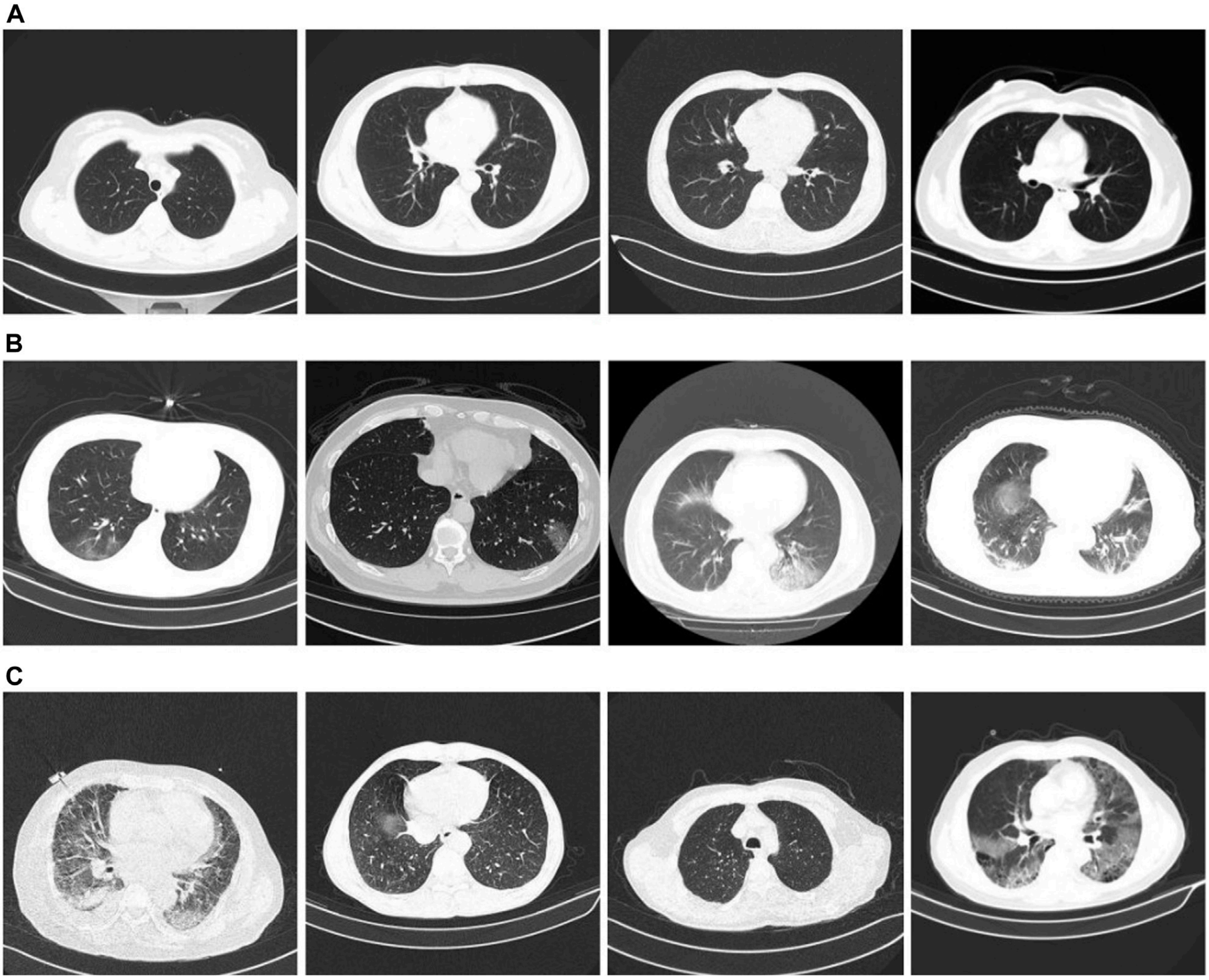


FIGURE 5
CC-CCII chest CT images. **(A)** Normal conditions; **(B)** COVID-19; **(C)** CP.

6752 CT scans of 4,154 participants. For our training test, we used 5985 CT scans. Among them, the training set is 3,017, and the test set is 2,968. The training and test set distributions were consistent, and the ratio of COVID-19, community pneumonia, and normal in the dataset was 1:1:1. The image size is three-channel, 512 × 512 × 3. Figure 5 presents an example of the dataset. Figures 5A–C show CT images of a typical patient, a patient with COVID-19, a patient with community pneumonia (mainly bacteria, viruses, chlamydia, and other microorganisms causing pneumonia), respectively.

2.4.2 Dataset partitioning

To divide the dataset, the K-Fold cross-validation method was employed. First, the dataset was divided into K sets, and each fold training used K-1 sets as the training set to train the model (K = 10). The remaining set was used as a validation set to test the performance evaluation of each folded training model, and the content of each validation set remained unrepeated. The data augmentation methods of random rotation, horizontal flipping, and contrast adjustment were used in training pre-processing to improve the model's generalization ability.

2.4.3 Experimental environment

Ubuntu18.04 was used as the operating system platform, with Intel(R) Core (TM) i5-6500 CPU, Nvidia GeForce GTX 1080ti GPU, with 11 GB of video memory and 16 GB of RAM.

The model performance can be improved, and the training time can be reduced with proper parameter configuration. Stochastic gradient descent was used to train the optimizer, and exponential decay was used to adjust the learning rate. The initial learning rate is 0.001. Additionally, 10-fold cross-validation was used for training with 100 epochs per fold. The

TABLE 1 Training parameter settings.

Type	Setting
Batch size	16
Learning rate	0.001
Optimizer	SGD
Epoch	100
Ubuntu 18.04	PyToch1.6.0

details of the experimental training parameters are shown in Table 1.

2.4.4 Evaluation indicators

To analyze the classification performance of the trained model for COVID-19, CP, and normal, three performance metrics were used: accuracy (Acc), sensitivity (Sens), and AUC of the receiver operating characteristic (ROC). The mPA is the average sum of each category's pixel accuracies. The formulas are as follows:

$$Acc = \frac{TP + TN}{TP + FN + FP + TN} \tag{12}$$

$$Sens = \frac{TP}{TP + FN} \tag{13}$$

$$mPA = \frac{1}{k+1}\sum_{i=0}^{k}\frac{TP}{TP + FN} \tag{14}$$

where, TP represents the number of positive examples that are predicted to be positive examples; FP represents the number of negatives predicted as positives; FN represents the number of positive classes predicted to be negative; TN represents the number of negative classes predicted to be negative; k is the number of categories.

Assuming that the ROC curve is formed by continuous links of points whose coordinates are $\{(x_1, y_1), (x_1, y_2), \ldots, (x_m, y_m)\}$. The AUC formula is as follows:

$$AUC = \frac{1}{2}\sum_{i=1}^{m-1}(x_{i+1} - x_i) \times (y_i + y_{i+1}) \tag{15}$$

$$FLOPs = (2 \times I - 1) \times O \tag{16}$$

where, where I and O represent the input and output neuron numbers, respectively.

3 Results

3.1 Ablation experiment

COVID-19 pneumonia and other pneumonia lesions exhibit the same characteristics of being in the lung area. However, the chest CT scans contain other interfering areas. To ensure that the lung area was unaffected by the interference area, preprocessing was performed during classification training to segment the lung area from the chest CT image. Next, ablation experiments were conducted to verify the segmentation effect of the new, improved model. The segmentation test results are presented in Figure 6. Figure 6A shows the CT images before segmentation, and Figure 6B shows the CT images after segmentation in Figure 6A from left to right. The results show that the newly proposed method can segment tiny lesion details, achieving the highest segmentation performance.

We compared our model to other models to more accurately evaluate its performance. The results are presented in Table 2. First, the convolutional neural network was used to extract enough local information after preprocessing the image features through the convolution module during input. Next, the vector was input into the improved transformer structure, and the initial self-attention mechanism was replaced with the axial attention mechanism. Further, a position offset term was added to improve the model performance. Compared with the CCT model, the accuracy and sensitivity of our improved model are increased by 1.7% and 2.3%, respectively, and the number of floating-point operations (FLOPs) is less than the model calculation amount of the CCT model. Concerning the recognition speed of a single image, the lightweight CCT single image recognition speed is the fastest, only 0.014 s. This is faster than all other models, and its recognition accuracy has not dropped. The comparative results show that our proposed improved method achieves the best results in screening COVID-19 and CP.

In the medical image application of the transformer, the input patch size parameter setting affects the model performance. The self-attention mechanism in the transformer structure has the advantage of obtaining global contextual connections. The matrices of different models were used to evaluate the performance. A total of 2968 CT images were tested. The confusion matrix in Figure 7 shows the difference between the actual and predicted values. The horizontal axis represents the model prediction results, corresponding to the number of predictions of different categories. The vertical axis represents the ground-truth labels (normal, COVID-19, and CP). A 3 × 3 matrix was used to compute the TP, FP, and FN values of the multiclassification task. The numbers on the blue back-ground are the number of correct predictions by the model. The values in the other regions correspond to the values at which the model predicted incorrectly, and the confusion matrix clearly shows the number of types of model mispredictions. The results show that the discrepancies between the chest CT images taken under normal conditions and during pneumonia have different presentation characteristics, leading to differences. Thus, it is easier to make sound judgments about the model. However, a small number of patients with mild COVID-19 or CP are mistaken for normal owing to a lack of apparent symptoms on chest CT images. Each model showed varying degrees of misidentification, misidentifying both COVID-19 and CP as

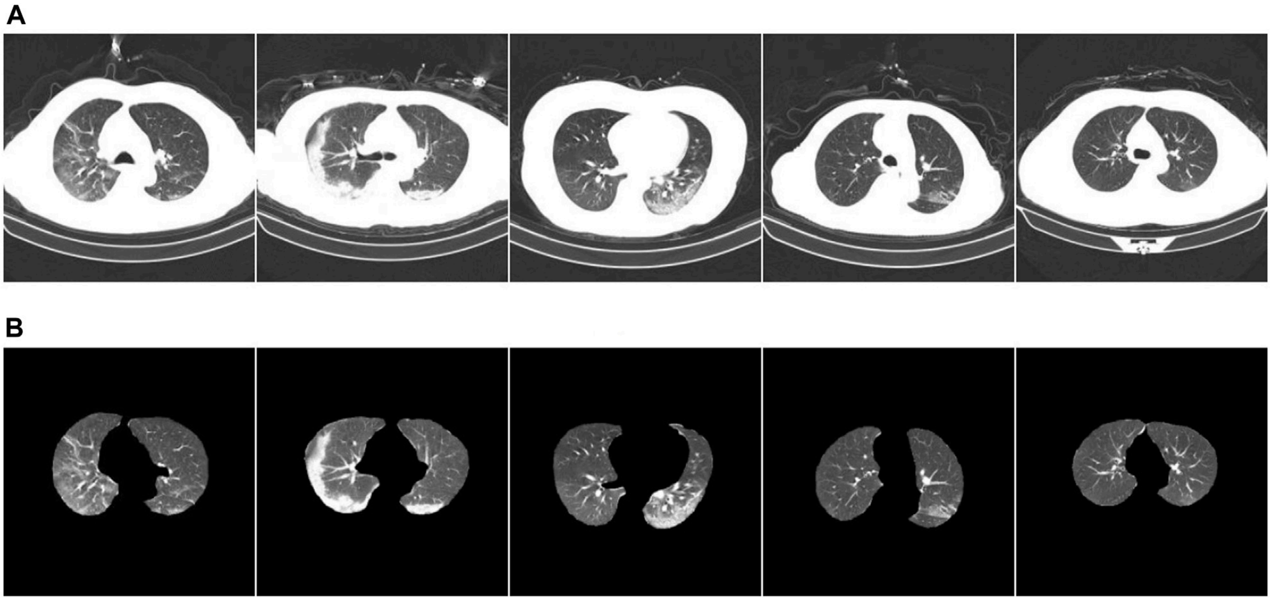


FIGURE 6
Example of ablation experiment comparison. **(A)** The CT image of the COVID-19 patient; **(B)** the result.

TABLE 2 Performance comparison of different models.

Modle	Acc/%	Sens/%	AUC	FLOPs (G)	Time (s)
Efficientnet-b7 Tan and le., 2019	88.4	88.3	0.972	1.02	0.023
Mobilenet-v3 Howard et al., 2019	97.8	97.6	0.997	0.33	0.019
ViT (Nielsen et al., 2015)	95.7	95.6	0.992	0.73	0.017
CCT (Esteva et al., 2019)	96.8	96.3	0.993	1.03	0.015
Ours	98.5	98.6	0.999	0.91	0.014

usual. This misidentification is due to certain similarities between chest CT images of COVID-19 patients and other pneumonia patients, such as ground-glass opacity and lung parenchyma features. Among them, Mobilenet-v3 and our model have fewer misidentifications. The test results in Figure 7E show 1034 COVID-19 CT images. Five and seven CT images were wrongly identified as normal and CT, respectively. The misdiagnosis rate is lower than in other compared models. Our improved model achieved the highest accuracy and the lowest misdiagnosis rate.

3.2 Real dataset model performance comparison

(1) COVID-CT dataset

We investigated the performance of different models on the COVID-CT dataset. Yang et al. collected 349 COVID-19 and 397 normal chest CT images in the COVID-CT dataset for 216 patients (Yang et al., 2020). However, some image data in this dataset were marked or missing. Image quality may have some impact on the training of the model. He et al., 2020 used contrastive self-supervised learning for training and achieved a model accuracy of 86%. Shalbaf et al. used 15 CNN benchmark models for fine-tuning training with the best accuracy and sensitivity of 84.7% and 82.2%, respectively (Gifani et al., 2021). Table 3 shows the comparison between our method and the methods above. The findings demonstrate that their training programs have engaged in significant workloads and relatively complex data preprocessing. However, our improved method achieves the best performance results in the COVID-CT dataset.

(2) SARS-CoV-2 CT-scan dataset

The SARS-CoV-2 CT-scan dataset comprises 2,482 chest CT images, including 1252 COVID-19 and 1,230 non-COVID-

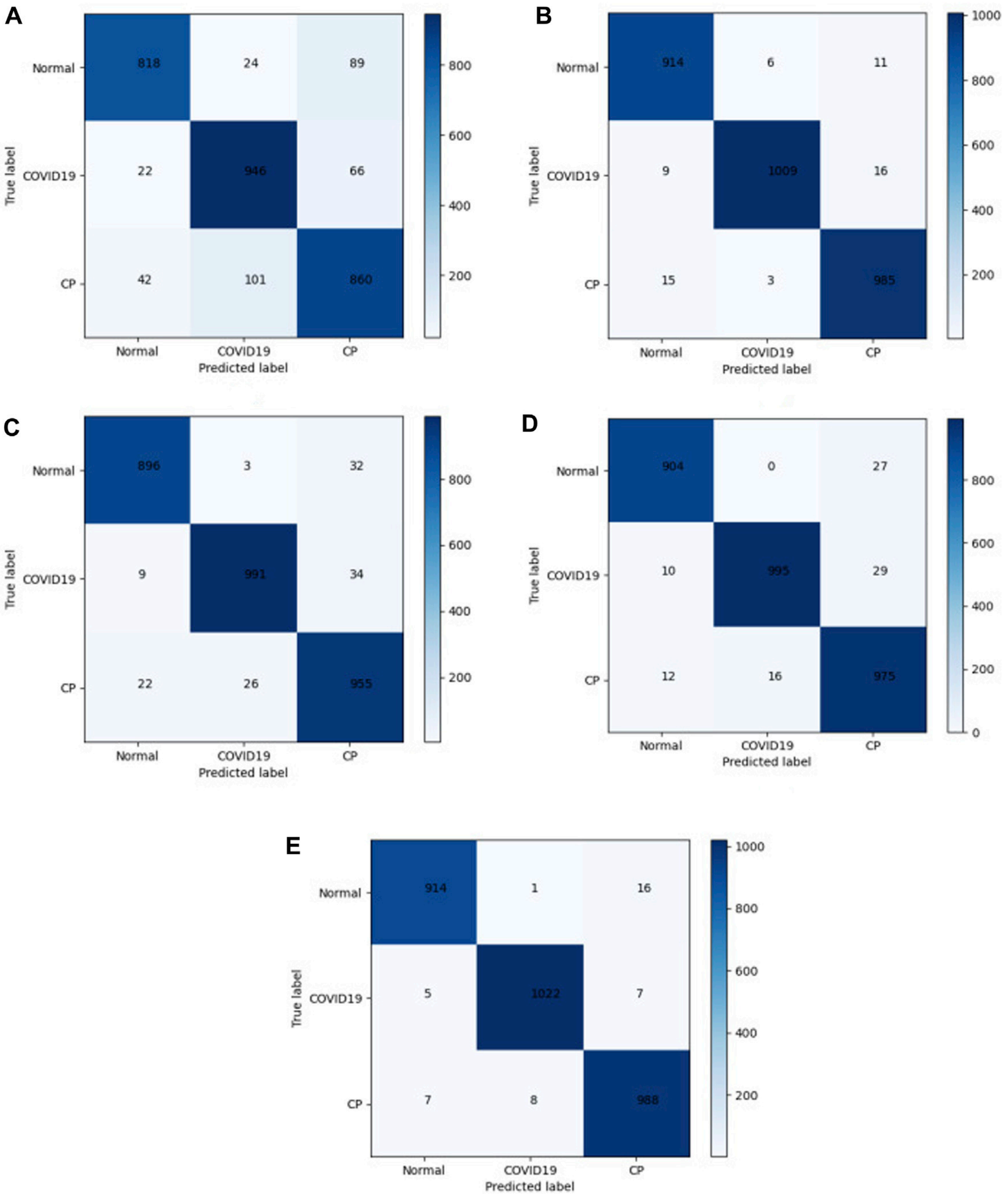


FIGURE 7
Confusion matrix. **(A)** Efficientnet-b7; **(B)** Mobilenet-v3; **(C)** ViT; **(D)** CCT; **(E)** Ours.

19 CT images. Soares et al. proposed the xDNN model and divided the dataset into training and test sets in a 4:1 ratio (Soares et al., 2020). After training, the accuracy and sensitivity rates of the model were 97.38% and 95.53%, respectively, and the essential auxiliary diagnosis ability was realized. Panwar et al. proposed an improved VGG model and used the dataset for training and testing, and the final sensitivity was 94.04% (Panwar et al., 2020). The comparison results between our method and the above methods are presented in Table 4. The results show that our improved method achieves the best performance results, with accuracy and sensitivity values of 98.01% and 98.23%, respectively.

TABLE 3 Model comparison of COVID-19 CT dataset.

Model	Acc/%	Sens/%
He et al., 2020	86	—
Gifani et al., 2021	84.7	82.2
Ours	87.3	86.7

TABLE 4 Model comparison of SARS-CoV-2 CT-scan dataset.

Model	Acc/%	Sens/%	Specificity/%
Soares et al., 2020	97.38	95.53	—
Panwar et al., (2020)	—	94.04	95.86
Ours	98.01	98.23	98.62

TABLE 5 Subjective quality evaluation of a scoring method.

Score	Features of the restored image
0	Severely distorted images
1	Images with severe distortion in some areas
2	Slightly distorted images
3	Difficult to spot distorted images
4	Images with better visual effects
5	Very sharp images

3.3 Subjective evaluation

The classification performance of our models was assessed using a more specific evaluation. Ten lead physicians with over 5 years of clinical experience in radiology were invited to perform independent image analysis (sharpness, resolution, invariance, and acceptability). The scoring method of subjective evaluation is presented in Table 5. One hundred CT images of the lesion area were randomly selected, and 10 sets of test samples were constructed equally.

The subjective quality evaluation results of clinicians are shown in Table 6. The results show that our proposed lightweight CCT achieves the best subjective quality evaluations regarding sharpness, resolution, invariance, and acceptability. This is thanks to our improved ViT as a network framework, using an attention mechanism to compute from image height and width separately, adding a position offset term to improve the model classification performance, and our proposed method has the best performance in maintaining edge and texture feature classification.

TABLE 6 Subjective quality evaluation of different algorithms.

Method	Sharpness	Resolution	Invariance	Acceptability
Efficientnet-b7	3.4 ± 0.35	3.6 ± 0.18	0.5 ± 0.41	3.8 ± 0.54
Mobilenet-v3	3.6 ± 0.72	3.9 ± 0.26	0.6 ± 0.55	3.9 ± 0.18
ViT	3.6 ± 0.39	4.1 ± 0.51	0.6 ± 0.89	4.1 ± 0.36
CCT	3.8 ± 0.65	4.2 ± 0.13	0.7 ± 0.21	4.1 ± 0.68
Proposed	3.9 ± 0.74	4.3 ± 0.29	0.7 ± 0.96	4.2 ± 0.71

4 Discussion

The automatic classification and recognition of chest CT images were improved by improving the CCT model. The self-attention mechanism of the encoder was enhanced to a position-sensitive axial attention mechanism. Meanwhile, the previous architecture was expanded by adding position offset terms to the self-attention mechanism to improve the classification ability of the ability.

Some interference areas were observed in the lung CT images of the patients. Therefore, to keep the model from becoming infected, when the data from the literature were employed simultaneously, sufficient feature extraction of the model was achieved by horizontal and vertical flipping, small angular rotations, and normalized data amplification. Further, it improved the generalization ability of the model to prevent overfitting.

This study CCT employs a new sequence pooling policy, convolution, and smaller, more compact transformers than ViT. Additionally, it compensates for the lack of medical image datasets by eliminating class tokens and positional embedding requirements. However, when the input dimension is large, the model operation cost increases considerably, and global pooling does not use location information when extracting feature information, possibly leading to information loss. Therefore, the self-attention mechanism in the encoder was improved to an axial attention mechanism. The self-attention mechanism was divided into two 1D self-attention mechanisms, the high and wide axes, which were calculated from the two dimensions of the width and height axis. Consequently, the number of calculations and computational efficiency were improved. Additionally, the position deviation was attached to the query, key, and value; an accurate deviation was used to obtain the position information, ensuring that more spatial structural information could be obtained.

According to the results in Tables 2–4, adding the axial attention mechanism considerably improved the accuracy of

COVID-19 detection in chest CT images. In small datasets, the performance was better than that of the standard transformer structural network and comparable to that of the traditional CNN. Although the transformer framework classification model may be suitable for small datasets by changing the patch size or encoder structure, some problems remain. For example, a maximum of three categories of models were trained; however, more categories could be used. As lung CT images of patients with mild COVID-19 symptoms are very similar to normal lung CT images, some of the discriminating errors from the lung CT images of patients with mild symptoms were present when the test set was used to validate the model. Consequently, datasets can be added later to improve the model performance. Although deep learning can represent a predictable information relationship, which has good prospects for medical applications, it is challenging in the context of data differences and other factors in medical images.

5 Conclusion

Although transformers are generally considered to be suitable only for large-scale or medium-scale training, this study shows that our proposed lightweight CCT classification recognition model works successfully on small data regimes and outperforms larger convolutional models. The performance obtained using the proposed model on the small COVID-19 dataset outperforms the standard ViT structured network and is comparable to the performance of traditional CNNs with significantly reduced computational cost and memory constraints. Experiments show that adding a position offset term by using the axial attention mechanism as a Transformer encoder to compute from the image height and width, respectively, can effectively improve the model classification performance. Our proposed classification method achieves the best performance with 98.5% accuracy and 98.6% sensitivity. The subjective quality assessment by physicians is optimal proving that our method is more suitable for clinical practice. Future studies can utilize a lightweight, compact method for initial screening and segmentation network to segment focal features of COVID-19 from chest CT images. We wish to implement a user interface system for digital image processing using a GUI. The main contents include the design of histogram grayscale transformation, edge detection, smooth filtering, and threshold segmentation for lung CT The main contents include the design of histogram grayscale transformation, edge detection, smooth filtering, and threshold segmentation for lung CT images.

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

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

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

Funding

This work was supported by the Doctoral Innovative Talents Project of Chongqing University of Posts and Telecommunications (BYJS202107), Postdoctoral Science Foundation of Chongqing (cstc2021jcyj-bsh0218); Special financial aid to post-doctor research fellow of Chongqing (2011010006445227); The National Natural Science Foundation of China (U21A20447 and 61971079); The Basic Research and Frontier Exploration Project of Chongqing (cstc2019jcyjmsxmX0666); Chongqing technological innovation and application development project (cstc2021jscx-gksbx0051); The Science and Technology Research Program of Chongqing Municipal Education Commission (KJZD-k202000604); the Innovative Group Project of the National Natural Science Foundation of Chongqing (cstc2020jcyj-cxttX0002); the Regional Creative Cooperation Program of Sichuan (2020YFQ0025).

Acknowledgments

We thank the College of Optoelectronic Engineering of Chongqing University of Posts and Telecommunications for their assistance in the research.

Conflict of interest

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

Publisher's note

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

References

Afshar, P., Heidarian, S., Naderkhani, F., Oikonomou, A., Plataniotis, K. N., and Mohammadi, A. (2020). Covid-caps: A capsule network-based framework for identification of Covid-19 cases from x-ray images. *Pattern Recognit. Lett.* 138, 638–643. doi:10.1016/j.patrec.2020.09.010

Ardakani, A. A., Kanafi, A. R., Acharya, U. R., Khadem, N., and Mohammadi, A. (2020). Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks. *Comput. Biol. Med.* 121, 103795. doi:10.1016/j.compbiomed.2020.103795

Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., et al. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat. Med.* 25, 954–961. doi:10.1038/s41591-019-0447-x

Bai, H. X., Hsieh, B., Xiong, Z., Halsey, K., Choi, J. W., Tran, T. M. L., et al. (2020). Performance of radiologists in differentiating Covid-19 from non-Covid-19 viral pneumonia at chest CT. *Radiology* 296, E46–E54. doi:10.1148/radiol.2020200823

Bernheim, A., Mei, X., Huang, M., Yang, Y., Fayad, Z. A., Zhang, N., et al. (2020). Chest CT findings in coronavirus disease-19 (Covid-19): Relationship to duration of infection. *Radiology* 295, 200463. doi:10.1148/radiol.2020200463

Brunese, L., Mercaldo, F., Reginelli, A., and Santone, A. (2020). Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. *Comput. Methods Programs Biomed.* 196, 105608. doi:10.1016/j.cmpb.2020.105608

Calderon-Ramirez, S., Yang, S., Moemeni, A., Elizondo, D., Colreavy-Donnelly, S., Chavarría-Estrada, L. F., et al. (2021). Correcting data imbalance for semi-supervised Covid-19 detection using x-ray chest images. *Appl. Soft Comput.* 111, 107692. doi:10.1016/j.asoc.2021.107692

Chen, X., Yao, L., Zhou, T., Dong, J., and Zhang, Y. (2021). Momentum contrastive learning for few-shot COVID-19 diagnosis from chest CT images. *Pattern Recognit.* 113, 107826. doi:10.1016/j.patcog.2021.107826

Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. (2018). Bert: Pretraining of deep bidirectional transformers for language understanding. *Comput. Lang.* arXiv: 1810.04805 doi:10.48550/arXiv.1810.04805

Dong, E., Du, H., and Gardner, L. M. (2020). An interactive web-based dashboard to track Covid-19 in real time. *Lancet. Infect. Dis.* 20, 533–534. doi:10.1016/S1473-3099(20)30120-1

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542, 115–118. doi:10.1038/nature21056

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., et al. (2019). A guide to deep learning in healthcare. *Nat. Med.* 25, 24–29. doi:10.1038/s41591-018-0316-z

Gifani, P., Shalbaf, A., and Vafaeezadeh, M. (2021). Automated detection of COVID-19 using ensemble of transfer learning with deep convolutional neural network based on CT scans. *Int. J. Comput. Assist. Radiol. Surg.* 16, 115–123. doi:10.1007/s11548-020-02286-w

Han, Z., Wei, B., Hong, Y., Li, T., Cong, J., Zhu, X., et al. (2020). Accurate screening of COVID-19 using attention based deep 3D multiple instance learning. *IEEE Trans. Med. Imaging* 39, 2584–2594. doi:10.1109/TMI.2020.2996256

Hassani, A., Walton, S., Shah, N., Abuduweili, A., Li, J., and Shi, H. (2021). Escaping the big data paradigm with compact transformers. arXiv:2104.05704 doi:10.48550/arXiv.2104.05704

He, K., Zhang, X., Ren, S., and Sun, J. (2016). “Deep residual learning for image recognition,” in Proceedings of the Conference on Computer Vision and Pattern Recognition Las Vegas, NV: IEEE, 770–778. doi:10.1109/CVPR.2016.90

He, X., Yang, X., Zhang, S., Zhao, J., Zhang, Y., Xing, E., et al. (2020). *Sample-efficient deep learning for COVID-19 diagnosis based on CT scans*. medrxiv. doi:10.1101/2020.04.13.20063941

Ho, J., Kalchbrenner, N., Weissenborn, D., and Salimans, T. (2019). Axial attention in multidimensional transformers. arXiv:1912.12180 doi:10.48550/arXiv.1912.12180

Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., et al. (2019). “Searching for mobilenetv3,” in Proceedings of the International Conference on Computer Vision, Seoul, Korea: IEEE, 1314–1324. doi:10.1109/ICCV.2019.00140

Huang, Z., Wang, X., Huang, L., Huang, C., Wei, Y., Liu, W., et al. (2019). “Ccnet: Criss-cross attention for semantic segmentation,” in Proceedings of the International Conference on Computer Vision, Seoul, South Korea: IEEE, 603–612. doi:10.1109/ICCV.2019.00069

Jaiswal, A., Gianchandani, N., Singh, D., Kumar, V., and Kaur, M. (2020). Classification of the Covid-19 infected patients using Dense-Net201 based deep transfer learning. *J. Biomol. Struct. Dyn.* 39, 5682–5689. doi:10.1080/07391102.2020.1788642

Li, T., Mei, T., Yan, S., Kweon, I. S., and Lee, C. (2009). “Contextual decomposition of multi-label images,” in Proceedings of the Conference on Computer Vision and Pattern Recognition, Miami, FL: IEEE, 2270–2277. doi:10.1109/CVPR.2009.5206706

Li, T., Yan, S., Mei, T., Hua, X. S., and Kweon, I. S. (2010). Image decomposition with multilabel context: Algorithms and applications. *IEEE Trans. Image Process.* 20, 2301–2314. doi:10.1109/TIP.2010.2103081

Li, T., Ni, B., Xu, M., Wang, M., Gao, Q., and Yan, S. (2015). Data-driven affective filtering for images and videos. *IEEE Trans. Cybern.* 45, 2336–2349. doi:10.1109/TCYB.2015.2392156

Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., et al. (2020). Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: Evaluation of the diagnostic accuracy. *Radiology* 296, E65–E71. doi:10.1148/radiol.2020200905

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., et al. (2017). A survey on deep learning in medical image analysis. *Med. Image Anal.* 42, 60–88. doi:10.1016/j.media.2017.07.005

Mahmud, T., Rahman, M. A., and Fattah, S. A. (2020). CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Comput. Biol. Med.* 122, 103869. doi:10.1016/j.compbiomed.2020.103869

Mei, X., Lee, H. C., Diao, K. Y., Huang, M., Lin, B., Liu, C., et al. (2020). Artificial intelligence-enabled rapid diagnosis of patients with Covid-19. *Nat. Med.* 26, 1224–1228. doi:10.1038/s41591-020-0931-3

Nielsen, M. A. (2015). Neural networks and deep learning. *Determination Press.* http://neuralnetworksanddeeplearning.com/

Oh, Y., Park, S., and Ye, J. C. (2020). Deep learning COVID-19 features on CXR using limited training data sets. *IEEE Trans. Med. Imaging* 39, 2688–2700. doi:10.1109/TMI.2020.2993291

Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., Acharya, U. R., et al. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput. Biol. Med.* 121, 103792. doi:10.1016/j.compbiomed.2020.103792

Ozyurt, F., Tuncer, T., and Subasi, A. (2021). An automated COVID-19 detection based on fused dynamic exemplar pyramid feature extraction and hybrid feature selection using deep learning. *Comput. Biol. Med.* 132, 104356. doi:10.1016/j.compbiomed.2021.104356

Panwar, H., Gupta, P. K., Siddiqui, M. K., Morales-Menendez, R., Bhardwaj, P., and Singh, V. (2020). A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. *Chaos Solit. Fractals* 140, 110190. doi:10.1016/j.chaos.2020.110190

Qin, Z. Z., Sander, M. S., Rai, B., Titahong, C. N., Sudrungrot, S., Laah, S. N., et al. (2019). Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. *Sci. Rep.* 9, 15000–15010. doi:10.1038/s41598-019-51503-3

Ramchoun, H., Idrissi, M. J., Ghanou, Y., and Ettaouil, M. (2017). “Multilayer Perceptron: Architecture Optimization and training with mixed activation functions,” in Proceedings of the 2nd international Conference on Big Data, Cloud and Applications, Tetouan Morocco, March 2017. IEEE.

Rubin, G. D., Ryerson, C. J., Haramati, L. B., Sverzellati, N., Kanne, J. P., Raoof, S., et al. (2020). The role of chest imaging in patient management during the covid-19 pandemic: A multinational consensus statement from the fleischner society. *Radiology* 296, 172–180. doi:10.1148/radiol.2020201365

Sun, W., Chen, J., Yan, L., Lin, J., Pang, Y., and Zhang, G. (2022). Covid-19 CT image segmentation method based on Swin transformer. *Front. Physiol.* 13, 981463. doi:10.3389/fphys.2022.981463

Shi, H., Han, X., Jiang, N., Cao, Y., Alwalid, O., Gu, J., et al. (2020). Radiological findings from 81 patients with covid-19 pneumonia in wuhan, China: A descriptive study. *Lancet. Infect. Dis.* 20, 425–434. doi:10.1016/S1473-3099(20)30086-4

Soares, E., Angelov, P., Biaso, S., Froes, M. H., and Abe, D. K. (2020). *SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification*. MedRxiv. doi:10.1101/2020.04.24.20078584

Tan, M., and Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. *ICML* 97, 6105–6114. doi:10.48550/arXiv.1905.11946

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nat. Med.* 25, 44–56. doi:10.1038/s41591-018-0300-7

Wang, L., Lin, Z. Q., and Wong, A. (2020). A tailored deep convolutional neural network design for detection of Covid-19 cases from chest x-ray images. *Sci. Rep.* 10, 19549. doi:10.1038/s41598-020-76550-z

Wong, H. Y. F., Lam, H. Y. S., Fong, A. H. T., Leung, S. T., Chin, T. W. Y., Lo, C. S. Y., et al. (2020). Frequency and distribution of chest radiographic findings in patients positive for Covid-19. *Radiology* 296, E72–E78. doi:10.1148/radiol.2020201160

Wu, Y. H., Gao, S. H., Mei, J., Xu, J., Fan, D. P., Zhang, R. G., et al. (2021). Jcs: An explainable Covid-19 diagnosis system by joint classification and segmentation. *IEEE Trans. Image Process.* 30, 3113–3126. doi:10.1109/TIP.2021.3058783

Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., et al. (2020). A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering* 6, 1122–1129. doi:10.1016/j.eng.2020.04.010

Yang, X., He, X., Zhao, J., Zhang, Y., Zhang, S., and Xie, P. (2020). COVID-CT-dataset: A CT scan dataset about COVID-19. arXiv:2003.13865 doi:10.48550/arXiv.2003.13865

Zhang, K., Liu, X., Shen, J., Li, Z., Sang, Y., Wu, X., et al. (2020a). Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of Covid-19 pneumonia using computed tomography. *Cell* 181, 1360–1433. doi:10.1016/j.cell.2020.08.029

Zhang, J., Xie, Y., Pang, G., Liao, Z., Verjans, J., Li, W., et al. (2020b). Viral pneumonia screening on chest X-rays using confidence-aware anomaly detection. *IEEE Trans. Med. Imaging* 40, 879–890. doi:10.1109/TMI.2020.3040950

Zhou, T., Canu, S., and Ruan, S. (2021). Automatic Covid-19 CT segmentation using U-Net integrated spatial and channel attention mechanism. *Int. J. Imaging Syst. Technol.* 31, 16–27. doi:10.1002/ima.22527

Zhu, N., Zhang, D., Wang, W., Li, X., Yang, B., Song, J., et al. (2020). A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* 382, 727–733. doi:10.1056/NEJMoa2001017

frontiers | Frontiers in Physiology

TYPE Original Research
PUBLISHED 16 November 2022
DOI 10.3389/fphys.2022.1067948

OPEN ACCESS

EDITED BY
Hao Yang,
University of South Florida,
United States

REVIEWED BY
Xiong Wang,
ShanghaiTech University, China
Shuangli Liu,
Southwest University of Science and
Technology, China

*CORRESPONDENCE
Yan Luo,
yanluo@scu.edu.cn

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

RECEIVED 12 October 2022
ACCEPTED 02 November 2022
PUBLISHED 16 November 2022

CITATION
Li J, Wu Z, Peng C, Song L and Luo Y
(2022), Microwave-induced
thermoacoustic imaging for the early
detection of canine
intracerebral hemorrhage.
Front. Physiol. 13:1067948.
doi: 10.3389/fphys.2022.1067948

Microwave-induced thermoacoustic imaging for the early detection of canine intracerebral hemorrhage

Jiawu Li[1], Zhenru Wu[2], Chihan Peng[1], Ling Song[1] and Yan Luo[1]*

[1]Department of Ultrasound, West China Hospital of Sichuan University, Chengdu, China, [2]Institute of Clinical Pathology, Key Laboratory of Transplant Engineering and Immunology, NHC, West China Hospital, Sichuan University, Chengdu, China

Purpose: This study aimed to investigate the feasibility and validation of microwave-induced thermoacoustic imaging (TAI) for the early detection of canine intracerebral hemorrhage.

Methods: A TAI system was used to record the thermoacoustic signal (TAS) of canine intracerebral hemorrhage in the study. First, the difference in TAS between deionized water, fresh *ex vivo* porcine blood and brain tissue was explored. Second, the canine hemorrhagic stroke model was established, and canine brain ultrasound examination and TAI examination were performed before modeling and at 0.5 h, 1 h, 2 h, 3 h, 4 h, 4.5 h, 5 h and 6 h after modeling. Finally, pathology and ultrasound were used as the reference diagnoses to verify the accuracy of the thermoacoustic imaging data.

Results: The results showed that significant differences were observed in TASs among deionized water, fresh *ex vivo* porcine blood and brain tissue. The intensity of the thermoacoustic signal of blood was significantly higher than that of *ex vivo* porcine brain tissue and deionized water. The intracerebral hemorrhage model of five beagles was successfully established. Hematomas presented hyperintensity in TAI. Considering ultrasound and pathology as reference diagnoses, TAI can be used to visualize canine intracerebral hemorrhage at 0.5 h, 1 h, 2 h, 3 h, 4 h, 4.5 h, 5 h and 6 h after modeling.

Conclusion: This is the first experimental study to explore the use of TAI in the detection of intracerebral hemorrhage in large live animals (canine). The results indicated that TAI could detect canine intracerebral hemorrhage in the early stage and has the potential to be a rapid and noninvasive method for the detection of intracerebral hemorrhage in humans.

KEYWORDS

microwave-induced thermoacoustic imaging, canine, intracerebral hemorrhage, early detection, thermoacoustic signal

Introduction

Stroke is the second leading cause of death and the third leading cause of disability worldwide, including hemorrhagic and ischemic strokes, and its incidence is increasing, especially in developing countries (Kyu et al., 2018; Feigin et al., 2021). Because the two types of stroke have different causes, their treatments are also completely distinct. Thrombolysis is the main treatment for ischemic stroke, and the "effective treatment window" is very short (within 4.5 h), while thrombolysis is strictly prohibited for hemorrhagic stroke (Powers et al., 2019; Berge et al., 2021). However, the proportion of patients undergoing thrombolytic therapy for acute ischemic stroke is still very low. The principal reason is that it is difficult to effectively distinguish ischemic stroke from hemorrhagic stroke within 3–4.5 h of onset, which leads to the inability to carry out timely thrombolytic therapy for patients with ischemic stroke (Wang et al., 2011; El Khoury et al., 2012). Therefore, it is important to quickly and accurately distinguish hemorrhagic stroke from ischemic stroke.

At present, CT scans are the initial choice for the imaging diagnosis of acute stroke. MRI is more sensitive than CT in the diagnosis of ischemic foci or hematoma within 6 h after the onset of stroke (Chalela et al., 2007), but it is not recommended to use MRI for the early diagnosis of suspected stroke (Powers et al., 2018). However, CT and MRI are difficult to use for detecting intracranial hemorrhage in prehospital conditions or even at the scene of onset due to the large equipment needed. Researchers have been pursuing a portable, prehospital accessible method to accurately detect intracerebral bleeding (Mobashsher et al., 2016; Candefjord et al., 2017; Ljungqvist et al., 2017) to reduce the waiting time for preoperative examination and improve the prognosis of patients with hemorrhagic stroke.

Previous studies have shown that there is an obvious contrast between the conductivity of blood and white matter and gray matter (Peyman et al., 2007), in which the conductivity of gray matter, white matter and blood at a frequency of 3.0 GHz are 2.2189 S/m, 1.5106 S/m and 3.0498 S/m, respectively. The conductivity of blood is significantly higher than that of white matter and gray matter. Of note, in recent years, some studies (Persson et al., 2014; Candefjord et al., 2017; Ljungqvist et al., 2017) have used microwave imaging (MI) based on conductivity to detect acute stroke, and the results showed that brain MI could distinguish hemorrhagic stroke from ischemic stroke. However, the spatial resolution of the MI system is poor. It is difficult to meet the clinical need because the size and location of intracerebral hematoma cannot be provided. Further improvement of the spatial resolution of microwave imaging may make the evaluation of intracerebral hematoma more accurate.

Microwave-induced thermoacoustic imaging (TAI) is a novel noninvasive imaging modality that receives ultrasonic signals generated by the absorption of microwaves in different biological tissues and reconstructs the image to reflect the dielectric properties of biological tissues (Huang et al., 2012; Zheng et al., 2018; Sun et al., 2021; Lin, 2022). It has the advantages of the high contrast of microwave imaging and the high resolution of ultrasound imaging. Currently, TAI technology has attracted increasing attention based on its unique advantages and researchers are actively developing its huge potential role in biomedical diagnosis and even treatment. Differences in conductivity and relative permittivity between different tissues or normal tissues and lesions allow TAI to distinguish them from each other. The image resolution of TAI is determined by the detected ultrasonic signal. Compared with MI, TAI has higher spatial resolution due to its resolution is sub-millimeter (Liu et al., 2022). Up to now, TAI studies in the biomedical field have mainly involved the detection of breast cancer and prostate cancer, joint related disease and brain imaging, etc., especially in the early detection of breast cancer has been extensively studied (Liu et al., 2018; Li et al., 2019; Yuan et al., 2019; Li et al., 2021; Zhang et al., 2022). To date, there are few studies on TAI in brain disease. Xu et al. (Yuan and Wang, 2006) applied TAI to imaging the rhesus monkey brain through the intact skull, and the results showed that the brain parenchyma could be clearly visualized. Huang et al. (2017) made the first preliminary attempt to detect the hemorrhage phantom with the self-built TAI system in the world and successfully imitated a simulated hemorrhage phantom beneath an isolated human skull. The study of Yan et al. (2019) also indicated that thermoacoustic tomography can take images through the adult human skull. Zhao et al. (2017) showed that thermoacoustic tomography could visualize numerous important brain anatomical structures in rats, and they further attempted to detect germinal matrix hemorrhage in neonatal mice (Zhao et al., 2020). The results showed that thermoacoustic tomography can accurately detect a hematoma region at different depths in the neonatal mouse brain. However, to our knowledge, there is no relevant report on the study of thermoacoustic imaging in detecting intracerebral hemorrhage in large animals or humans. Therefore, the main purpose of this study was to investigate the feasibility and validation of thermoacoustic imaging for the rapid detection of canine intracerebral hemorrhage.

Methods

The animal study was reviewed and approved by the Animal Ethics Committee of West China Hospital of Sichuan University. All applicate regulations concerning the ethical use of animals were strictly followed during the whole experiment. In this study, we first explored whether there was a difference in

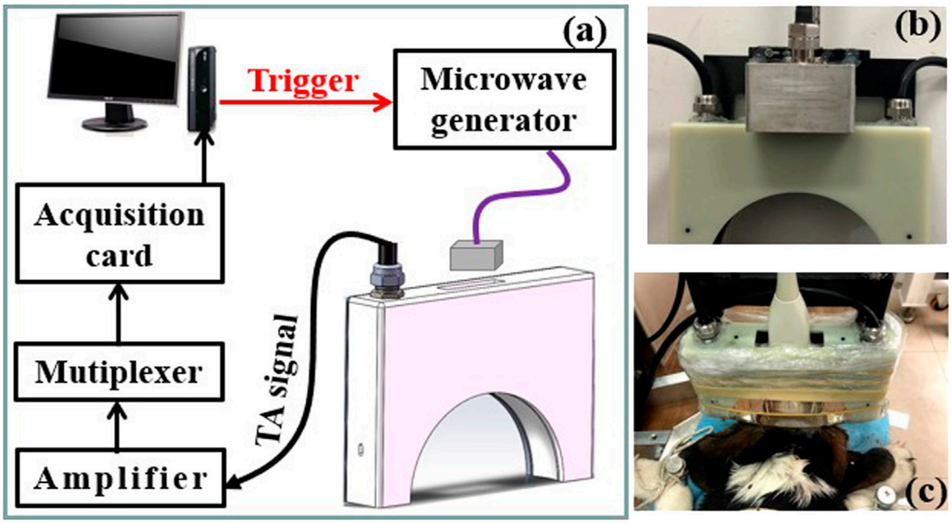


FIGURE 1
(A) Schematic of the thermoacoustic imaging (TAI) system. **(B,C)** Photograph of the antenna and hollow concave transducer array, the dog brain to be imaged, respectively.

thermoacoustic signals among deionized water, fresh *ex vivo* porcine blood and brain tissue. Second, a canine hemorrhagic stroke model was established, and canine brain ultrasound examination and TAI examination were performed before modeling and at 0.5 h, 1 h, 2 h, 3 h, 4 h, 4.5 h, 5 h and 6 h after modeling. Finally, the dogs were sacrificed by air embolization under anesthesia, and their brains were removed and prepared for pathological examination to further verify the accuracy of thermoacoustic imaging.

Microwave-induced thermoacoustic imaging system

A schematic of the TAI system utilized in this study is shown in Figure 1. A custom-designed miniaturized microwave generator (peak power: 60 kW, pulse width: 70–600 ns and repetition rate: 100 Hz) (Huang et al., 2021) coupled with a handheld dipole antenna (aperture size: 60 × 60 mm^2) (Huang et al., 2019) *via* a semirigid coaxial cable (1.5 m long with 1.2 dB insertion loss) was used to evoke thermoacoustic signals. The TA signals were detected by a 128-element hollow concave transducer array (Wang et al., 2021), amplified and averaged 50 times to achieve a good signal-and-noise ratio, and finally recorded by a 64-channel acquisition system with two 32-channel data acquisition cards (5752B, NI. Inc., United States). A back-projection algorithm in MATLAB was used for TA image reconstruction (Hoelen and de Mul., 2000). A B-mode ultrasound imaging platform (iNSIGHT-37°C, SASET. Inc., China) was also used for imaging the dog brains.

Thermoacoustic signal of deionized water, fresh *ex vivo* porcine blood and *ex vivo* porcine brain by thermoacoustic imaging

Considering the accessibility of *ex vivo* animal experimental materials, porcine blood and brain tissue were used in this part of the *in vitro* experiment. Fresh *ex vivo* porcine blood and brain tissue were obtained from a local slaughterhouse and were wrapped in aluminum foil to keep them fresh before experiments. In the experiment, deionized water, *ex vivo* porcine blood and brain tissue were placed into transparent plastic tubes with a diameter of 7 mm, and both sides of the plastic tube were sealed with solid glue. Thermoacoustic imaging was performed by placing plastic tubes equipped with the three components side by side at the same level. In addition, we further explored the TAI of *ex vivo* porcine blood at 1 h, 2 h, 3 h, 4.5 h and 6 h by placing it into a plastic thin-walled tubes with an internal diameter of 7 mm.

Establishment of a canine intracerebral hemorrhage model and detection of canine intracerebral hemorrhage by microwave thermoacoustic imaging

Healthy beagles (n = 5, male; weight: 5–7 kg) used in the experiment were obtained from the Sichuan Institute of Musk Deer Breeding. All dogs were fasted the night before the experiment. The intracerebral hemorrhage model was established by canine autologous blood injection. The process of the intracerebral hemorrhage model was as follows: A beagle was anesthetized by

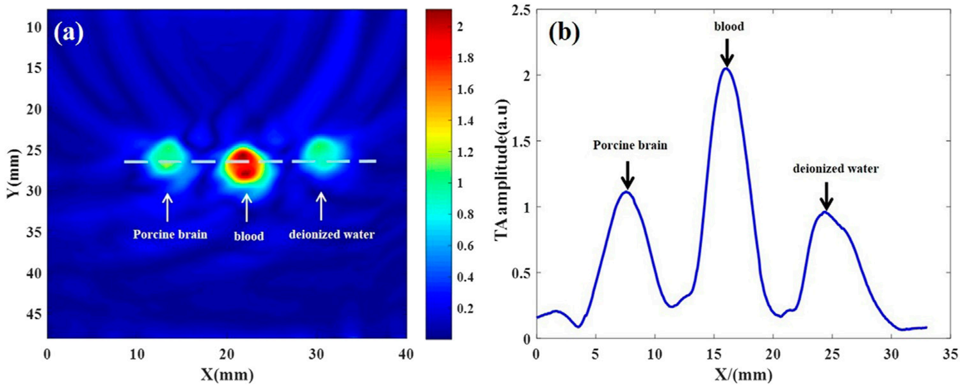


FIGURE 2
Thermoacoustic imaging of fresh *ex vivo* porcine brain, blood and deionized water. **(A)** Thermoacoustic image of *ex vivo* porcine brain, blood and deionized water. **(B)** The one-dimensional distribution curve of relative thermoacoustic signal amplitude along the white dashed line shown in **(A)**.

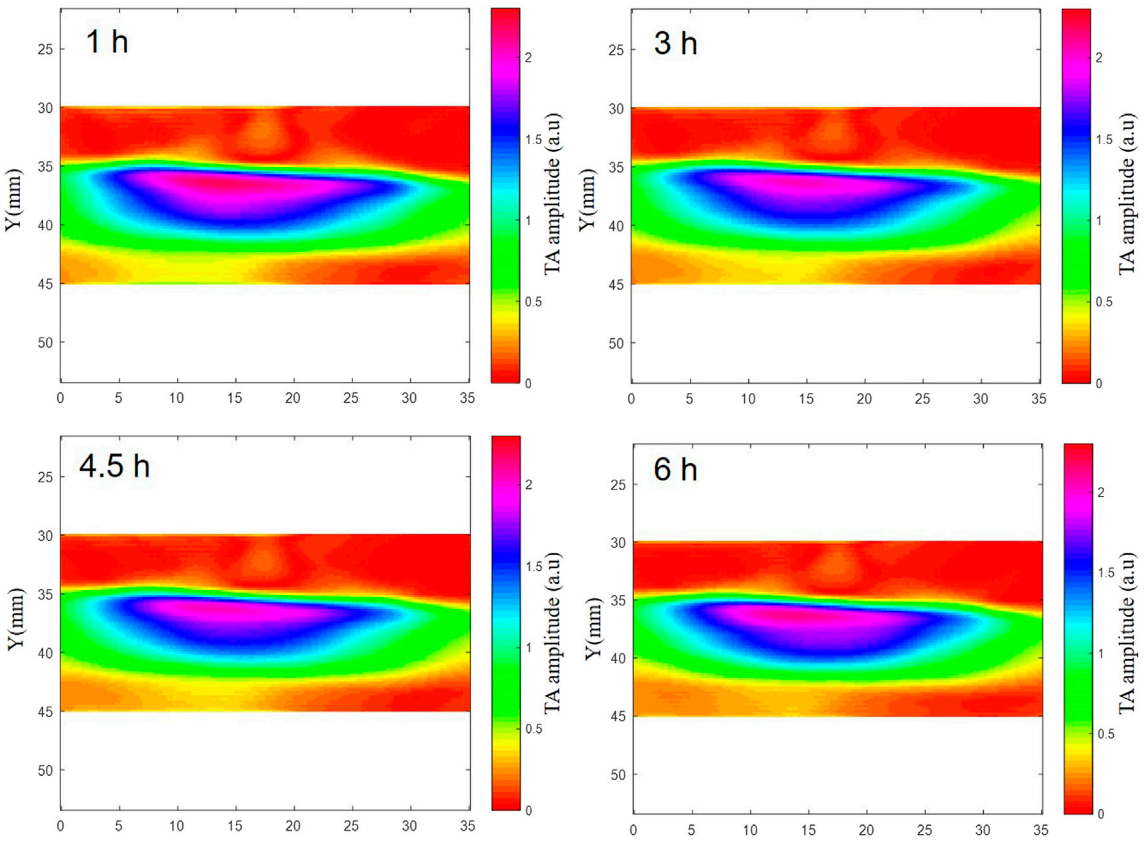


FIGURE 3
The TAI of fresh *ex vivo* porcine blood at 1 h, 2 h, 3 h, 4.5 h and 6 h. The results showed that there was no significant change in the thermoacoustic signals of fresh *ex vivo* porcine blood at 1 h, 2 h, 3 h, 4.5 h and 6 h.

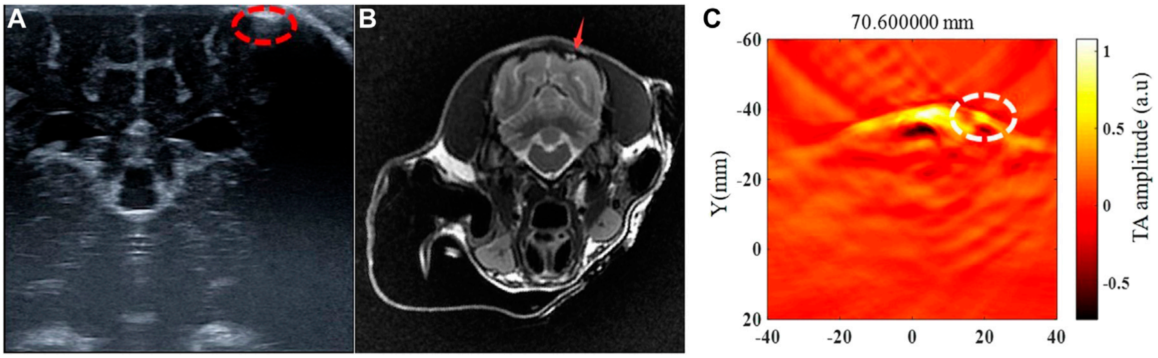


FIGURE 4
Ultrasonography, MRI and thermoacoustic imaging after 6 h of the canine intracerebral hemorrhage modeling. This is a case of canine intracerebral hemorrhage model with blood injected into cerebral sub-cortex. **(A)** Ultrasound showed the hemorrhage area as hyperechoic located at the left cerebral sub-cortex (red circle). **(B)** T2WI of MRI presented hemorrhage area as hyper-intensity located at the left cerebral sub-cortex (red arrow). **(C)** Thermoacoustic imaging showed a hyper-intensity of the hemorrhage area at the same location as presented at ultrasound and MRI (white circle).

intraperitoneal injection of 3% pentobarbital sodium at a dose of 1 ml/kg (body weight) + 2 ml. The dog was fixed on the operating table in a prone position, and a clean, thick towel was used to keep it warm. The head skin was sterilized after hair removal with a pet shaving device. A scalpel was used to cut the scalp in a 1 cm circle along the upper margin of the canine superciliary arch to the front of both ears, separated the temporalis muscle layer by layer, and scraped the periosteum. A craniotomy drill was used to drill a hole 5 mm from the left or right side of the midline of the top of the head (3 left side and 2 right side). During the drilling process, attention was given to gentle movements, and the wound was washed with 0.9% normal saline to reduce the temperature of the grinding area. The whole grinding time was approximately 50 min. During the operation and experiment, the state of the dog was noted at all times. If the dog trembled, 0.1 ml sumianxin was injected into the abdominal cavity for auxiliary anesthesia. After successful drilling, once the dog was fully hemostatic, the skin was sutured. The skulls of all dogs were intact at the time of TAI imaging, except for the small cranial boreholes.

Before the intracerebral hemorrhage model was made in each dog, ultrasound examination and TAI imaging were performed, and then blood injection was carried out to generate the intracerebral hemorrhage model. Five milliliters of blood were taken from the femoral artery of the dog and placed in an anticoagulant tube. Two milliliters of blood were drawn with 5 ml syringes. The dura was punctured vertically, the needle was slowly inserted into the intracranial area (the area 1.5–2 cm away from the syringe needle was clamped with hemostatic forceps to ensure a controlled injection depth), and autologous blood was injected slowly. The vital signs of the dog were closely observed during blood injection. After injection, the needle remained in the brain for 5 min and then was removed slowly. TAI imaging and ultrasound examination were performed at 0.5 h, 1 h, 2 h, 3 h, 4 h, 4.5 h, 5 h and 6 h after modeling. The same transducer was used during the process of TAI and ultrasound examination, and the probe position was kept motionless during each TAI and ultrasound examination. After each ultrasound examination, the position of the probe interface was only switched to ensure that the images obtained by ultrasound and TAI examination were from the same plane. All dogs remained deeply anesthetized during the experiment. Finally, the dogs were sacrificed by air embolization under anesthesia, and the brain was removed and prepared for pathological examination.

Results

Comparison of thermoacoustic signals in deionized water, fresh *ex vivo* porcine blood and *ex vivo* porcine brains

TAI was performed on deionized water, isolated blood and isolated brain tissue simultaneously. The results showed that there were significant differences in microwave thermoacoustic signals among the three groups. The intensity of the thermoacoustic signal of blood was significantly higher than that of *ex vivo* porcine brain tissue and deionized water (Figure 2). In addition, there was no significant change in the thermoacoustic signals of fresh *ex vivo* porcine blood at 1 h, 2 h, 3 h, 4.5 h and 6 h (Figure 3).

Thermoacoustic imaging and ultrasonography of canine intracerebral hemorrhage

The intracerebral hemorrhage model of five beagles was successfully established, and no dogs died during the

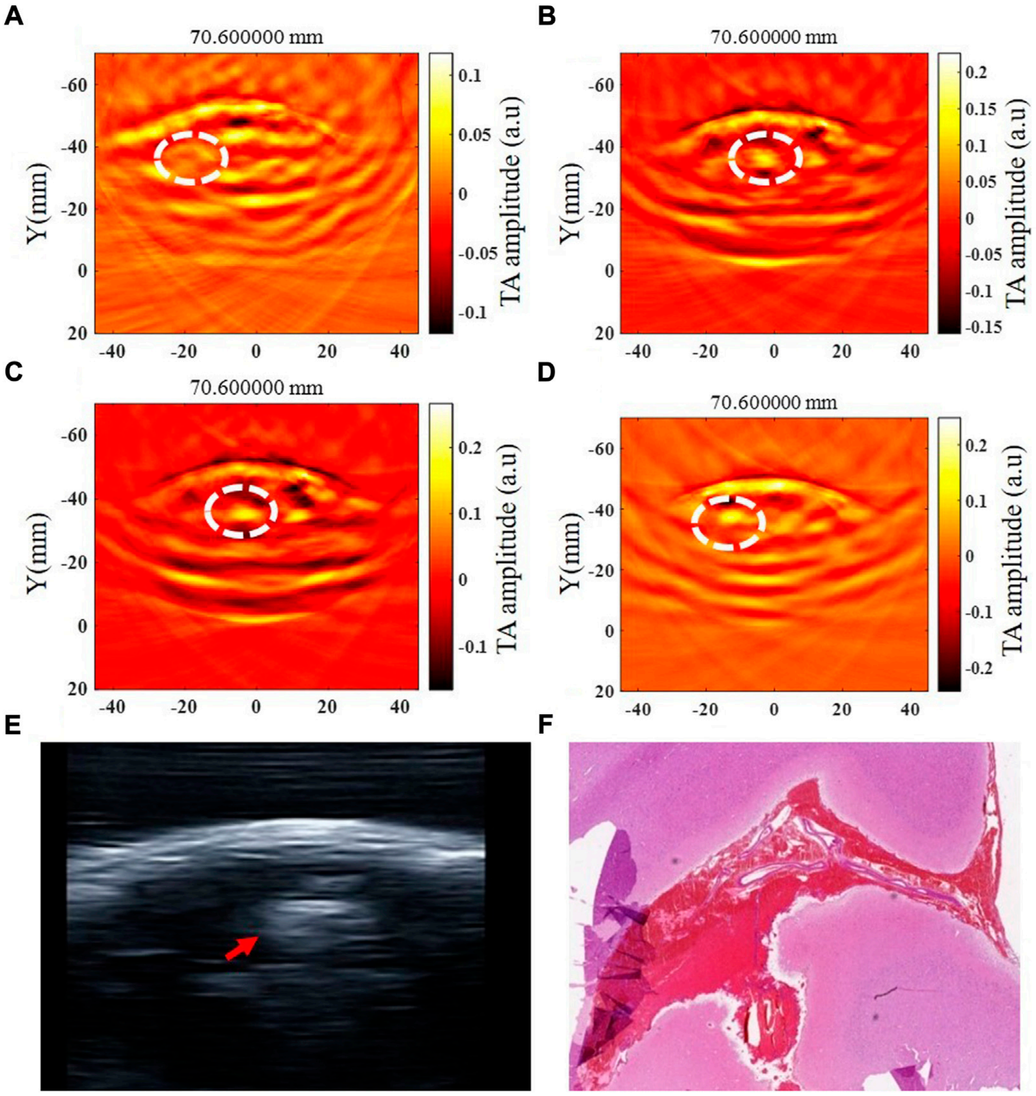


FIGURE 5
Thermoacoustic imaging, ultrasound and pathological examination of a same canine intracerebral hemorrhage. **(A–D)** TAI images at 0.5 h, 1 h, 3 h and 6 h after modeling, respectively (white circle shows the hemorrhagic area). **(E)** Ultrasound image of canine intracerebral hemorrhage (white circle shows the hemorrhagic area). **(F)** Pathological image of canine intracerebral hemorrhage.

experiment. The blood was injected in the right brain parenchyma in two beagles and in the left brain parenchyma in the other three beagles. The injection depth was 1.5 cm–2 cm within the brain parenchyma in four of five beagles, and the other one was injected into the cerebral subcortex due to a thicker scalp. The beagle with the thicker scalp underwent further MRI examination after 6 h of modeling (Figure 4). The time of data collection of TAI for each examination of the canine brain was 1 minute. At 0.5 h, 1 h, 2 h, 3 h, 4 h, 4.5 h, 5 h and 6 h after modeling, TAI and ultrasonography can be used to image intracerebral hemorrhage in the dogs. The hematomas presented hyperintensity in TAI. By comparing ultrasonography and pathology, the size and location of intracerebral hemorrhage in TAI were in line with those of the other groups (Figure 5).

Discussion

In this study, we at first compared the thermoacoustic signal (TAS) in deionized water, fresh *ex vivo* porcine blood and *ex vivo* porcine brain, and the results showed that the intensity of the TAS of blood was significantly higher than that of brain tissue and deionized water, which is related to the apparent difference in conductivity between blood and brain tissues (Andreuccetti et al., 2021). These results provide a solid theoretical basis for further study on thermoacoustic imaging (TAI) in the detection of intracerebral hemorrhage in living animals. Moreover, we demonstrated, for the first time, the ability of TAI to detect intracerebral hemorrhage in large animals by establishing a canine hemorrhagic stroke model with ultrasound and pathological findings as a reference diagnosis.

In previous studies, TAI has demonstrated its potential for brain imaging in experiments on *ex vivo* (Yuan and Wang, 2006; Yan et al., 2019) and phantom (Huang et al., 2017) brain tissues. However, the dielectric properties of brain tissue may change due to a long time *in vitro* or preservation in formalin or mineral oil (Yuan and Wang, 2006), and the experimental results may not necessarily reflect the real status of brain imaging *in vivo*. Similarly, a phantom cannot simulate the complex electromagnetic environment of real living brain tissue. Thus, it is insufficient to prove that TAI is applicable to living tissues differently. Recently, studies (Zhao et al., 2017; Zhao et al., 2020) used TAI for brain imaging in living rats and demonstrated the feasibility of TAI in detecting germinal matrix hemorrhage in neonatal mice *in vivo*. These studies provide good evidence for the application of TAI in brain imaging of living animals. However, the brain structure of humans and rodents has significant discrepancies, and the canine brain is more similar to that of humans (Johnson et al., 20212020). Therefore, we explored the application of TAI in a canine intracerebral hemorrhage model in this study. TAI was performed by injecting blood into the brain parenchyma at different locations in beagles in the study, and the results showed that TAI could visualize intracranial hematomas in all cases. However, the depth of the hematomas in this study was only within 2 cm. Due to the small brain size of dogs, the cerebral cortex below 2 cm reached the lateral ventricle, so the experiment did not explore the modeling of hematoma in the deeper part of the brain, which is a limitation of this study. Of course, the detection of hematoma in the deeper part still needs further study.

In this study, we used a pulsed 3.0 GHz microwave source to transmit microwaves, which can penetrate several centimeters into tissue, for TAI. Thus, the penetration depth and spatial resolution of TAI for brain imaging are determined by the detected thermoacoustic signal. It is well known that skull-induced acoustic attenuation and scattering are the primary factors affecting ultrasound imaging of intracerebral hemorrhage, which also restricts the penetration depth and spatial resolution of TAI in this research. However, for TAI with a pulse width of approximately 600 ns in our experiment, the frequency of the thermoacoustic induced ultrasonic signal is approximately 1 MHz, which can penetrate the dog's skull well and can provide mm-scale spatial resolution. In sum, this study demonstrated the feasibility of TAI in detecting hemorrhagic stroke *in vivo*. While the results obtained are encouraging, improvements are still needed before we can use this type of imaging for human brains. For example, several studies (Venkatasubramanian et al., 2011; Chen et al., 2021) have shown that the brain tissue around hematoma after hemorrhagic stroke will gradually develop edema and peak at 3–4 days, while the main purpose of this study is to explore the detection of intracerebral hemorrhage by TAI in the early stage. Therefore, whether perihematomal edema and the condition of the tissue 6 h after hemorrhage affect the detection of TAI remains to be further studied.

Conclusion

In conclusion, this is the first experimental study to explore the use of TAI in the detection of intracerebral hemorrhage in large live animals (canine). The results indicated that TAI could be used to detect canine intracerebral hemorrhage and has the potential to be a rapid and noninvasive method for the detection of intracerebral hemorrhage in humans. In addition, with the miniaturization of microwave sources and rapid imaging speed, TAI is expected to be used in the future for pre-hospital detection of human intracerebral hemorrhage.

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 animal study was reviewed and approved by Animal Ethics Committee of West China Hospital of Sichuan University.

Author contributions

Study design: All authors. Modelling: JL, CP, and LS. Data Collection: JL, LS, and ZW. Data analysis and interpretation: All authors. Manuscript writing: JL. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by National Natural Science Foundation of China (No. 82001832 and 82071940).

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

Andreuccetti, D., Fossi, R., and Petrucci, C. (2021). *Calculation of the dielectric properties of body tissues in the frequency range 10 Hz-100 GHz florence: IFAC-CNR 1997-2015*. Florence (Italy): Institute for Applied Physics. .

Berge, E., Whiteley, W., Audebert, H., De Marchis, G. M., Fonseca, A. C., and Padiglioni, C., (2021). European Stroke Organisation (ESO) guidelines on intravenous thrombolysis for acute ischaemic stroke. *Eur. Stroke J.* 6, I–LXII. doi:10.1177/2396987321989865

Candefjord, S., Winges, J., Malik, A. A., Yu, Y., Rylander, T., and McKelvey, T., (2017). Microwave technology for detecting traumatic intracranial bleedings: Tests on phantom of subdural hematoma and numerical simulations. *Med. Biol. Eng. Comput.* 55, 1177–1188. doi:10.1007/s11517-016-1578-6

Chalela, J. A., Kidwell, C. S., Nentwich, L. M., Luby, M., Butman, J. A., and Demchuk, A. M., (2007). Magnetic resonance imaging and computed tomography in emergency assessment of patients with suspected acute stroke: A prospective comparison. *Lancet* 369, 293–298. doi:10.1016/S0140-6736(07)60151-2

Chen, Y., Chen, S., Chang, J., Wei, J., Feng, M., and Wang, R. (2021). Perihematomal edema after intracerebral hemorrhage: An update on pathogenesis, risk factors, and therapeutic advances. *Front. Immunol.* 12, 740632. doi:10.3389/fimmu.2021.740632

El Khoury, R., Jung, R., Nanda, A., Sila, C., Abraham, M. G., and Castonguay, A. C., (2012). Overview of key factors in improving access to acute stroke care. *Neurology* 79, S26–S34. doi:10.1212/WNL10.1212/wnl.0b013e3182695a2a

Feigin, V. L., Stark, B. A., Johnson, C. O., Roth, G. A., Bisignano, C., and Abady, G. G., (2021). Global, regional, and national burden of stroke and its risk factors, 1990-2019: A systematic analysis for the global burden of disease study 2019. *Lancet Neurol.* 20, 795–820. doi:10.1016/S1474-4422(21)00252-0

Hoelen, C. G., and de Mul., F. F. (2000). Image reconstruction for photoacoustic scanning of tissue structures. *Appl. Opt.* 39, 5872–5883. doi:10.1364/ao.39.005872

Huang, L., Ge, S., Zheng, Z., and Jiang, H. (2019). Technical Note: Design of a handheld dipole antenna for a compact thermoacoustic imaging system. *Med. Phys.* 46, 851–856. doi:10.1002/mp.13294

Huang, L., Li, T., and Jiang, H. (2017). Technical note: Thermoacoustic imaging of hemorrhagic stroke: A feasibility study with a human skull. *Med. Phys.* 44, 1494–1499. doi:10.1002/mp.12138

Huang, L., Yao, L., Liu, L., Rong, J., and Jiang, H. (2012). Quantitative thermoacoustic tomography: Recovery of conductivity maps of heterogeneous media. *Appl. Phys. Lett.* 101, 244106. doi:10.1063/1.4772484

Huang, L., Zheng, Z., Chi, Z., and Jiang, H. (2021). Technical Note: Compact thermoacoustic imaging system based on a low-cost and miniaturized microwave generator for *in vivo* biomedical imaging. *Med. Phys.* 48, 4242–4248. doi:10.1002/mp.15014

Johnson, P. J., Luh, W. M., Rivard, B. C., Graham, K. L., White, A., and FitzMaurice, M., (2021). Stereotactic cortical atlas of the domestic canine brain. *Sci. Rep.* 1011 (1), 47814911. Erratum in: Sci. doi:10.1038/s41598-020-61665-0

Kyu, H. H., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., and Abbasi, N., (2018). Global, regional, and national disability-adjusted life-years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2017: A systematic analysis for the global burden of disease study 2017. *Lancet* 392, 1859–1922. doi:10.1016/S0140-6736(18)32335-3

Li, J., Wang, B., Zhang, D., Li, C., Zhu, Y., and Zou, Y., (2021). A preclinical system prototype for focused microwave breast hyperthermia guided by compressive thermoacoustic tomography. *IEEE Trans. Biomed. Eng.* 68, 2289–2300. doi:10.1109/TBME.2021.3059869

Li, Y., Tan, Q., Qin, H., and Xing, D. (2019). Defect-rich single-layer MoS2 nanosheets with high dielectric-loss for contrast-enhanced thermoacoustic imaging of breast tumor. *Appl. Phys. Lett.* 115, 073701. doi:10.1063/1.5111892

Lin, J. C. (2022). The microwave auditory effect. *IEEE J. Electromagn. RF Microw. Med. Biol.* 6, 16–28. doi:10.1109/JERM.2021.3062826

Liu, Q., Liang, X., Qi, W., Gong, Y., Jiang, H., and Xi, L. (2022). Biomedical microwave-induced thermoacoustic imaging. *J. Innov. Opt. Health Sci.* 15, 2230007. doi:10.1142/S1793545822300075

Liu, S. L., Zhao, Z. Q., Lu, Y. X., Wang, B. W., Nie, Z., and Liu, Q. H. (2018). Microwave induced thermoacoustic tomography based on probabilistic reconstruction. *Appl. Phys. Lett.* 112, 263701. doi:10.1063/1.5034485

Ljungqvist, J., Candefjord, S., Persson, M., Jönsson, L., Skoglund, T., and Elam, M. (2017). Clinical evaluation of a microwave-based device for detection of traumatic intracranial hemorrhage. *J. Neurotrauma* 34, 2176–2182. doi:10.1089/neu.2016.4869

Mobashsher, A. T., Bialkowski, K. S., Abbosh, A. M., and Crozier, S. (2016). Design and experimental evaluation of a non-invasive microwave head imaging system for intracranial haemorrhage detection. *PLoS One* 11, e0152351. doi:10.1371/journal.pone.0152351

Persson, M., Fhager, A., Trefná, H. D., Yu, Y., McKelvey, T., and Pegenius, G., (2014). Microwave-based stroke diagnosis making global prehospital thrombolytic treatment possible. *IEEE Trans. Biomed. Eng.* 61, 2806–2817. doi:10.1109/TBME.2014.2330554

Peyman, A., Holden, S. J., Watts, S., Perrott, R., and Gabriel, C. (2007). Dielectric properties of porcine cerebrospinal tissues at microwave frequencies: *In vivo*, *in vitro* and systematic variation with age. *Phys. Med. Biol.* 52, 2229–2245. doi:10.1088/0031-9155/52/8/013

Powers, W. J., Rabinstein, A. A., Ackerson, T., Adeoye, O. M., Bambakidis, N. C., and Becker, K., . (2019). Guidelines for the early management of patients with acute ischemic stroke: 2019 update to the 2018 guidelines for the early management of acute ischemic stroke: A guideline for healthcare professionals from the American heart association/American stroke association. *Stroke*. 50, e344-e418. doi:10.1161/STR.0000000000000211

Powers, W. J., Rabinstein, A. A., Ackerson, T., Adeoye, O. M., Bambakidis, N. C., and Becker, K., (2018). 2018 guidelines for the early management of patients with acute ischemic stroke: A guideline for healthcare professionals from the American heart association/American stroke association. *Stroke* 49, e46–e110. doi:10.1161/STR.0000000000000158

Sun, Y., Li, C., Wang, B., and Wang, X. (2021). A low-cost compressive thermoacoustic tomography system for hot and cold foreign bodies detection. *IEEE Sens. J.* 21, 23588–23596. doi:10.1109/JSEN.2021.3109349

Venkatasubramanian, C., Mlynash, M., Finley-Caulfield, A., Eyngorn, I., Kalimuthu, R., and Snider, R. W., (2011). Natural history of perihematomal edema after intracerebral hemorrhage measured by serial magnetic resonance imaging. *Stroke* 42, 73–80. doi:10.1161/STROKEAHA.110.590646

Wang, X., Huang, L., Chi, Z., and Jiang, H. (2021). Integrated thermoacoustic and ultrasound imaging based on the combination of a hollow concave transducer array and a linear transducer array. *Phys. Med. Biol.* 66, 115011. doi:10.1088/1361-6560/abfc91

Wang, Y., Liao, X., Zhao, X., Wang, D. Z., Wang, C., and Nguyen-Huynh, M. N., (2011). Using recombinant tissue plasminogen activator to treat acute ischemic stroke in China: Analysis of the results from the Chinese national stroke registry (CNSR). *Stroke* 42, 1658–1664. doi:10.1161/STROKEAHA.110.604249

Yan, A., Lin, L., Liu, C., Shi, J., Na, S., and Wang, L. V. (2019). Microwave-induced thermoacoustic tomography through an adult human skull. *Med. Phys.* 46, 1793–1797. doi:10.1002/mp.13439

Yuan, C., Qin, B., Qin, H., and Xing, D. (2019). Increasing dielectric loss of a graphene oxide nanoparticle to enhance the microwave thermoacoustic imaging contrast of breast tumor. *Nanoscale* 11, 22222–22229. doi:10.1039/c9nr06549k

Yuan, X., and Wang, L. V. (2006). Rhesus monkey brain imaging through intact skull with thermoacoustic tomography. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* 53, 542–548. doi:10.1109/tuffc.2006.1610562

Zhang, J., Li, C., Jiang, W., Wang, Z., Zhang, L., and Wang, X. (2022). Deep-learning-enabled microwave-induced thermoacoustic tomography based on sparse data for breast cancer detection. *IEEE Trans. Antennas Propag.* 70, 6336–6348. doi:10.1109/TAP.2022.3159680

Zhao, Y., Chi, Z., Huang, L., Zheng, Z., Yang, J., and Jiang, H. (2017). Thermoacoustic tomography of *in vivo* rat brain. *J. Innov. Opt. Heal. Sci.* 10, 1740001-1-1740001-7. doi:10.3233/XST-190599

Zhao, Y., Shan, T., Chi, Z., and Jiang, H. (2020). Thermoacoustic tomography of germinal matrix hemorrhage in neonatal mouse cerebrum. *J. Xray. Sci. Technol.* 28, 83–93. doi:10.3233/XST-190599

Zheng, Z., Huang, L., and Jiang, H. (2018). Label-free thermoacoustic imaging of human blood vessels *in vivo*. *Appl. Phys. Lett.* 113, 253702. doi:10.1063/1.5054652

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 22 November 2022
DOI 10.3389/fphy.2022.1045192

EDITED BY
Hao Yang,
University of South Florida,
United States

REVIEWED BY
Lin Huang,
University of Electronic Science and
Technology of China, China
Jianbo Tang,
Southern University of Science and
Technology, China

*CORRESPONDENCE
Tianqi Shan,
tianqi@cqmu.edu.cn
Yuan Zhao,
yuanzhao@cqmu.edu.cn
Faqi Li,
lifq@cqmu.edu.cn

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

RECEIVED 15 September 2022
ACCEPTED 07 November 2022
PUBLISHED 22 November 2022

CITATION
Yu Q, Liao Y, Liu K, He Z, Zhao Y, Li F and
Shan T (2022), Registration of
photoacoustic tomography vascular
images: Comparison and analysis of
automatic registration approaches.
Front. Phys. 10:1045192.
doi: 10.3389/fphy.2022.1045192

Registration of photoacoustic tomography vascular images: Comparison and analysis of automatic registration approaches

Qinran Yu, Yixing Liao, Kecen Liu, Zhengyan He, Yuan Zhao*,
Faqi Li* and Tianqi Shan*

State Key Laboratory of Ultrasound in Medicine and Engineering, College of Biomedical Engineering,
Chongqing Medical University, Chongqing, China

Image registration is crucial in the clinical application of photoacoustic tomography (PAT) for vascular growth monitoring. Aiming to find an optimized registration scheme for PAT vascular images acquired at different times and with varying imaging conditions, we compared and analyzed different commonly used intensity-based and feature-based automatic registration schemes. To further improve the registration performance, we proposed a new scheme that combines phase correlation with these commonly used intensity-based registration methods and compared their performances. The objective evaluation measures: peak signal-to-noise ratio (PSNR), structural similarity index metric (SSIM), root mean square error (RMSE), and quantitative visual perception (jump percentage P), as well as subjective evaluation using mean opinion score (MOS), were combined to evaluate the registration performance. Results show that the feature-based approaches in this study were not suitable for PAT image registration. And by adding phase correlation as rough registration, the overall registration performance was improved significantly. Among these methods, the proposed scheme of phase correlation combined with mean square error (MSE) similarity measure and regular-step-gradient-descent optimizer provides the best visual effect, accuracy, and efficiency in PAT vascular image registration.

KEYWORDS

image registration, intensity-based registration, photoacoustic tomography, visual perception, photoacoustic imaging, vascular

1 Introduction

Monitoring vascular growth is critical for tumor growth monitoring [1]. Traditional imaging modalities such as magnetic resonance imaging (MRI) [2–4], ultrasound (US) [5–10], computed tomography (CT) [11–14], etc. have been commonly applied to evaluate the structural and functional changes of tumors and surrounding blood vessels in clinical practice, but these methods usually involve contrast agents and suffer from low contrast or resolution. Photoacoustic imaging (PAI), as a new radiation-free and non-ionizing imaging modality, has been developing rapidly in recent years [15]. It utilizes the different light absorptions of tissues to provide unique optical contrasts. The different light absorption coefficients of hemoglobin and various tissue chromophores at different wavelengths give PAI unique advantages in structural and functional vascular imaging [16–22]. Numerous studies have proved the effectiveness of PAI in vasculature imaging [23–27], which can be used to visualize the development of single blood vessels around growing tumors, changes in blood oxygen concentration within tumors, and the depth growth of neovascularization areas [28–33]. Comparing and analyzing the changes of blood vessels around tumors at different stages, including neovascularization and irregular changes in vascular morphology, can offer essential knowledge for monitoring the development and treatment of the disease, and realizing early screening and postoperative evaluation of cancer [34–36].

To observe and compare the morphological and functional changes of vasculatures and tissue structures through the disease progression, images need to be collected at different times. Therefore, to identify the differences in these images, the first problem to be solved is image registration. Image registration as a basic task in medical image analysis is a process of matching the images of the same scene obtained at different times, from different viewpoints, or by different sensors [37]. Single-mode registration is often used to analyze changes in images taken at different imaging periods, such as surgical effect evaluation [38, 39] and tumor growth monitoring [40–42]. MRI usually requires the combination of T1 and T2 sequence images to obtain tissue structure information with different contrast [43]. In photoacoustic imaging, we often need to use different optical wavelengths and sensors with different center frequencies to obtain different structural and functional information of tissues [15]. Image registration is needed before the comparison of tissue changes. Thus, finding a good registration scheme for PAI images taken at different times and with different imaging conditions (e.g., with different optical wavelengths or transducer center frequencies) is necessary for comparative analysis of the changes in target tissues through disease progression.

At present, the registration studies involved in photoacoustic imaging are mostly about real-time registration or PAI image registration with other imaging modalities [44–48], and there is no comparative study on different registration schemes for PAI images taken at different times with different transducers and optical wavelengths. In this study, we tested different automatic registration schemes using four categories of image datasets collected under three optical wavelengths and two transducer center frequencies. In addition, a new scheme was proposed to further improve the registration performances. The results were evaluated and compared by combining subjective and objective evaluation measures. The proposed scheme provides efficient and accurate automatic image registration for PAT vascular imaging, which can be applied to the applications such as vascular change monitoring, early screening, and postoperative treatment evaluation using photoacoustic imaging.

2 Materials and methods

2.1 System overview

A circular-scanning PAT system was used for imaging (See Figure 1). A fast-tuning OPO laser (Beijing ZK Laser Co.,Ltd.; wavelength range: 680–980 nm; repetition frequency: 100 Hz) was used as the excitation source. The light was coupled into a customized optical fiber bundle (CeramOptec GmbH) for light delivery. A digital delay pulse generator (Beijing ZK Laser Co.,Ltd.) sent triggers to laser and data acquisition simultaneously. The ultrasound transducer (ULSO TECH CO., LTD.; center frequency: 5MHz and 7.5 MHz) was mounted on a four-dimensional motion control module (Zolix Instruments Co., Ltd.) which consists of three linear stages and a rotator. The photoacoustic signals were detected by transducers, and then amplified and collected by a customized amplifier and automotive oscilloscopes (Picoscope 5000D). Degassed water was used as the coupling medium for PA waves. The system control interface is developed using Labview to realize the synchronized motor scanning and data acquisition. The image processing module developed using MATLAB includes image reconstrucion, processing, registration, data anaysis and other functions.

2.2 Image processing workflow

The workflow of image processing is shown in Figure 2. Firstly, the images were reconstructed using the delay and sum algorithm (DAS). Furthermore, to reduce the artifacts and background noises, preprocessing using bilateral filter and fuzzy C-means (FCM) was applied to achieve relatively clean background and extract target tissues. Finally, different registration schemes were applied to the PAT images, and the results were evaluated and compared by subjective visual evaluation (MOS) and objective evaluation measures (PSNR, SSIM, RMSE, jump percentage P).

2.3 Image datasets

The image data used in this study were all generated by the self-built PAT system. The backs of the volunteers' hands were imaged. In the experiments, the hands were held firmly on the

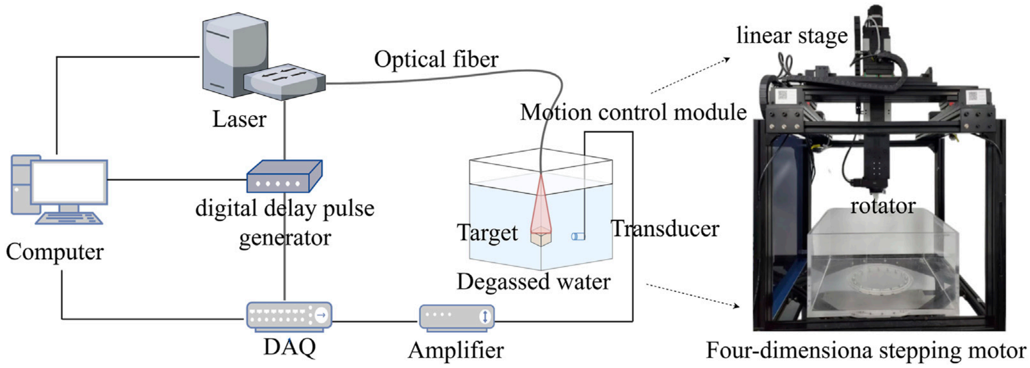


FIGURE 1
Schematics of the experimental system.

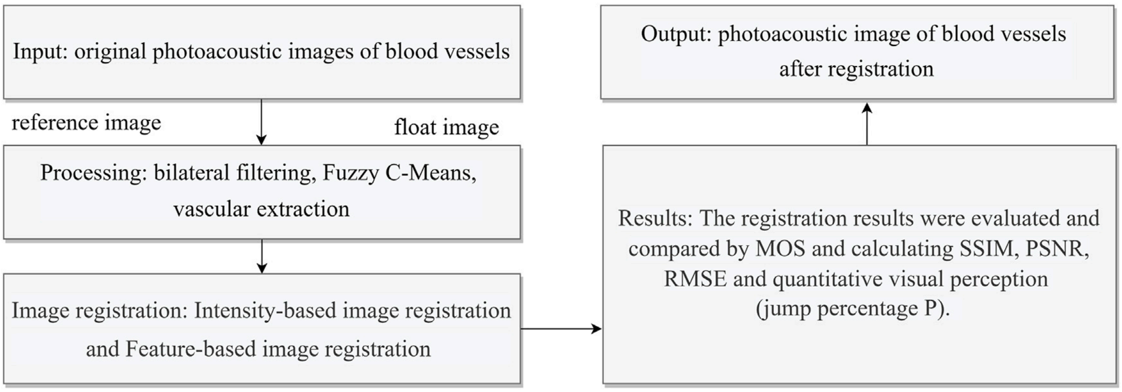


FIGURE 2
Flowchart of image processing.

holder (at a certain angle) in the tank of degassed water and were completely submerged (see Figure 3). The relative angle and distance of the imaging area of the hand back to the transducer was measured and kept consistent for the same subject in different experiments. Laser pulses (under ANSI limit) are diffused and incident from the top so that the light uniformly covers the entire imaging area. The receiving direction of the transducer is perpendicular to the incident direction of the laser, and the target is scanned circularly. The position of laser incidence and the position of transducer reception are controlled and recorded by the four-dimensional motion module to ensure the consistency of the imaging conditions of the same target tissue. During the imaging process, the sampling rate was 62.5MHz, and the scanning step was 1° covering 180°. 5,000 data points were collected at each position and were averaged 20 times. For different sets of data, three different wavelengths (720°nm, 850°nm, and 960 nm) and transducers with different center frequencies (5°MHz and 7.5 MHz) were used for imaging. The experiments were approved by Chongqing medical university.

Since the transducer we used was a flat unfocused single-crystal transducer with a large acceptance angle, and the signals from the same tissue could be received within a certain range. In addition, the blood vessels we imaged were superficial. Lights are uniformly irradiated, and one hand was held on the holder in a certain angle. Therefore, the problem caused by slight changes of projection angles at different times would not occur.

In order to verify the feasibility of the scheme, we carried out phantom experiments (See Figure 4). The silicone tubes (inner diameter 1.5 mm, wall thickness 0.5 mm) and Y-junction were used to simulate the vascular morphology *in vivo*, and Indian ink solution with light absorption coefficient of blood

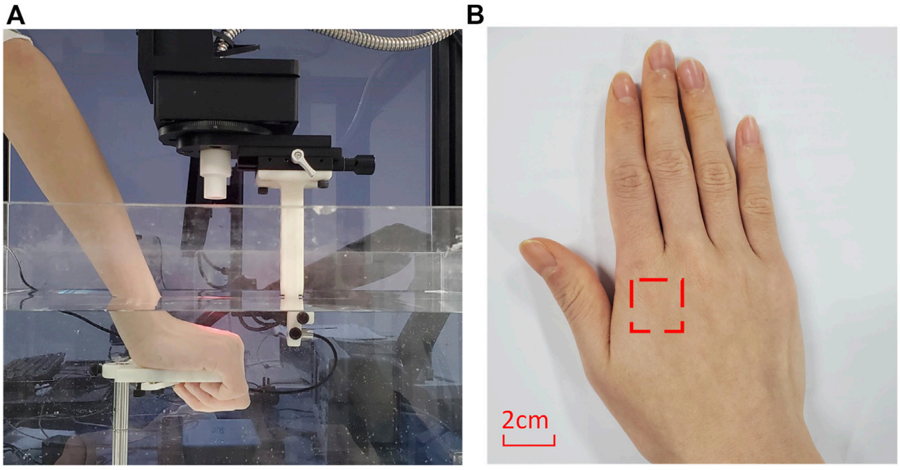


FIGURE 3
(A) is PAT imaging dataset acquisition experiment. **(B)** is a schematic diagram of imaging area of the hand.

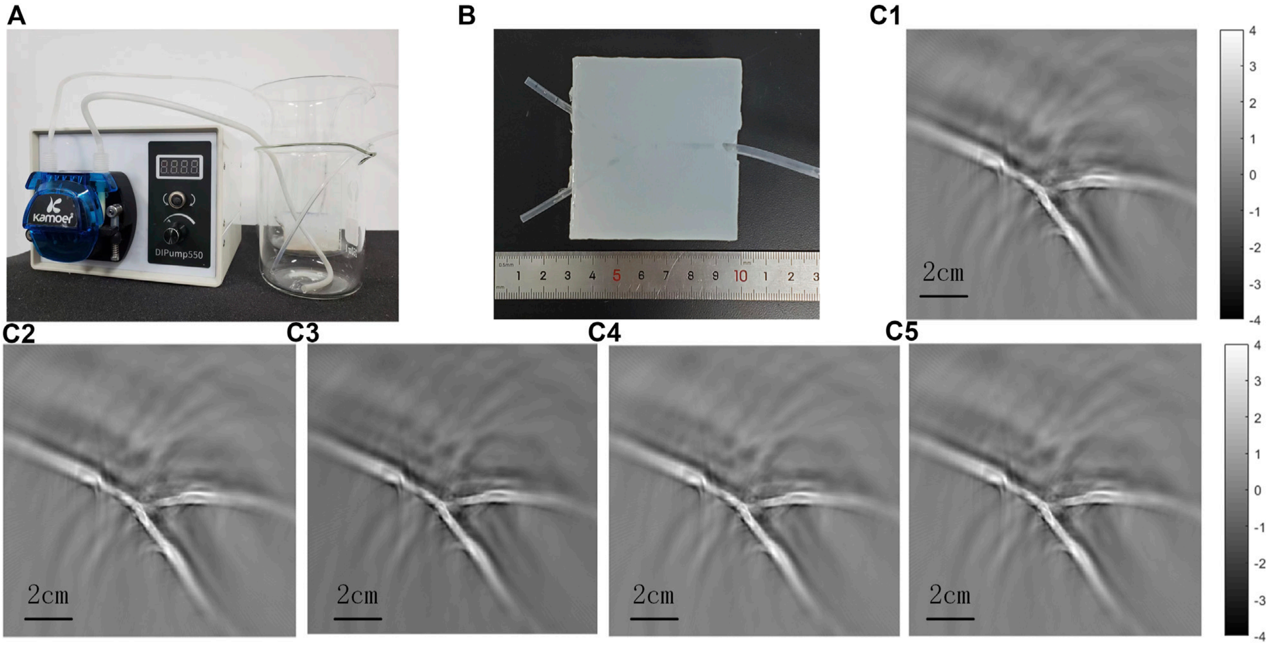


FIGURE 4
Phantom experiments to simulate blood flow imaging. **(A)** is the experimental device of the phantom experiment, which simulates the flow process of blood in the body through a peristaltic pump. **(B)** is the phantom of Y-shaped silicone tube. (c1) -(c5) are the first to fifth layers of the PAT images.

was circulated by a peristaltic pump to simulate the blood flow. Every time the transducer completed a circular scanning, it moved down 1 mm for another scanning, and it moved down five 5 mm in total. The results showed the morphology of the tubes clearly, and there was no significant difference among the five tomographic images. In addition, to verify the change of 2D projection caused by the change of detection angle, we tilted the phantom 15°, which was much larger than the angle difference that could occur in the experiments of the hand back imaging, and then compared the results with that of the phantom placed

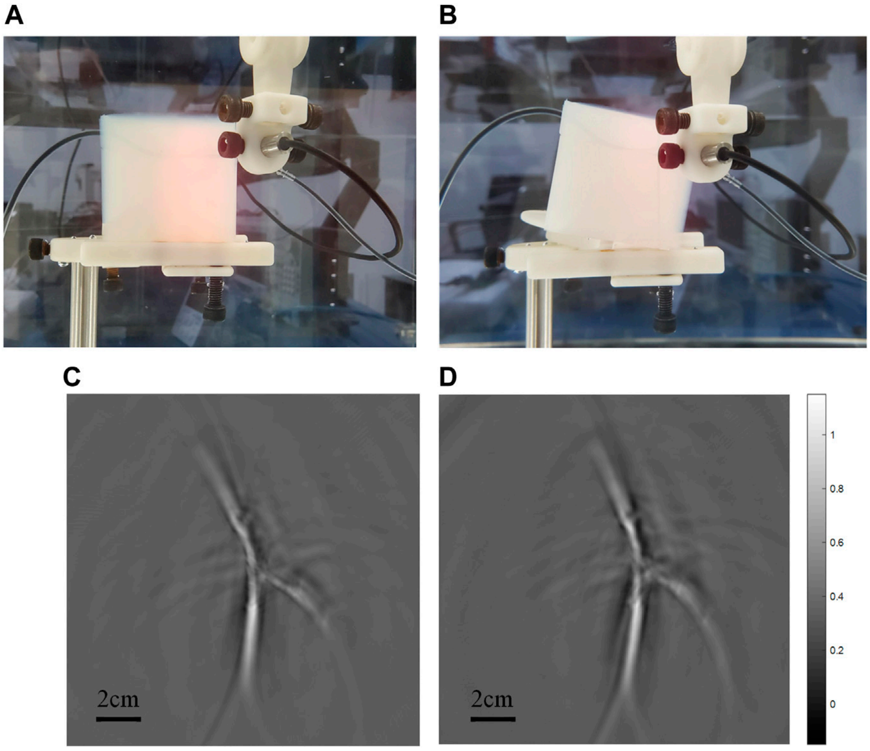


FIGURE 5
(A) is the scanning imaging experiment diagram under horizontal state, and **(B)** is the scanning imaging diagram under phantom tilt of 15°. **(C,D)** are photoacoustic imaging images corresponding to **(A,B)** respectively.

TABLE 1 Overview of hand vascular photoacoustic image datasets.

Number of datasets	Type	Wavelength	Transducer frequency	Imaging parts
Dataset1	Single frequency single wavelength	850 nm	7.5 M	Volunteer 1 right hand back
Dataset2		720 nm	7.5 M	Volunteer 1 left hand back
Dataset3		960 nm	7.5 M	Volunteer 2 right hand back
Dataset4	Single frequency multi-wavelength	850nm/960 nm	7.5 M	Volunteer 1 right hand back
Dataset5		720nm/850 nm	7.5 M	Volunteer 3 left hand back
Dataset6	Multi-frequency single wavelength	850 nm	5M/7.5 M	Volunteer 2 right hand back
Dataset7		720 nm	5M/7.5 M	Volunteer 1 left hand back
Dataset8	Multi-frequency multi-wavelength	720nm/850 nm	5M/7.5 M	Volunteer 1 left hand back
Dataset9		850nm/960 nm	5M/7.5 M	Volunteer 1 right hand back

horizontally (See Figure 5). It is proved that the transducer could receive signals from the same target within a certain range and the change of the results due to the slight differences in the angle can be neglected. Therefore, for superficial blood vessel imaging, the image would not change due to the slight difference of 3D projection angle.

To evaluate the performance and accuracy of different registration methods, the blood vessels of the backs of human hands were imaged. Four categories (nine groups) of datasets collected by transducers with two center frequencies and at three optical wavelengths were used for the registration test (see Table 1). Image registration methods are implemented in

MATLAB, using an Intel Core i7, 10th generation with 1.61 GHz clock speed and 16 GB RAM.

2.4 Preprocessing

The images were acquired by transducers with different center frequencies at different wavelengths and the original photoacoustic images had a low signal-to-background ratio (SBR). Therefore, preprocessing is necessary to bring images taken with different imaging conditions to a similar intensity and improve SBR for subsequent registration. Firstly, normalization was applied to bring the images to similar intensity in order to simplify the computation. Then, bilateral filter was applied to smooth and denoise the background, as well as to preserve the edges [49], and FCM was applied to extract the target tissue [50].

2.5 Image registration

After preprocessing the images, the reference image and floating image were obtained. We tested four commonly used intensity-based (see Figures 7A–D) and two feature-based (SIFT and SURF) registration schemes. In addition, we proposed a new scheme by adding phase correlation as rough registration to improve the performance of these intensity-based methods. The registration schemes were investigated using four types of image datasets (see Table 1) collected at three wavelengths by transducers with two central frequencies.

2.5.1 Registration based on intensity information

The intensity-based registration method uses grayscale information to directly calculate the similarity degree of the image, which has the advantages of simple operation without complex preprocessing and extensive computation. Its registration process is to select the corresponding similarity measure function according to the characteristics of the images, and then by applying a specific search algorithm in the parameter space of the chosen geometric transformation model, the geometric transformation parameter that maximizes the similarity is found. The image registration quality mainly depends on the similarity measure function.

2.5.1.1 Similarity measure function

Mean square error (MSE) and mutual information (MI) are common measure parameters in intensity-based registration [51]. MSE usually represents the deviation between the calculated and the actual values. In image registration, MSE is used to evaluate the accuracy and performance of the algorithm. The smaller MSE represents higher accuracy and better performance of the algorithm. It is defined as Eq. 1:

$$\mathrm{MSE} = \frac{1}{\mathrm{mn}}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}\left[I(i,j) - K(i,j)\right]^2 \tag{1}$$

where I and K are the floating and reference images, respectively. The m, n is the number of samples, that is, the number of pixels corresponding to the floating image I and the reference image K. I, j denotes any pixel. MI is an essential concept in information theory, which describes the correlation between two systems or how much information they contain with each other. In image registration, the MI of the two images reflects the degree of mutual inclusion through their entropy and joint entropy. The mutual information of the two images can be expressed as Eq. 2:

$$MI(X,Y) = H(X) + H(Y) - H(X,Y) \tag{2}$$

$$H(Y) = -\sum_{k=0}^{\mathrm{L}-1} p(r_k)\log p(r_k) \tag{3}$$

$$P(r_k) = \frac{h(r_k)}{n} = \frac{n_k}{n} = \frac{h(r_k)}{\sum_k h(r_k)} \qquad k = 0, 1, \ldots, L{-}1 \tag{4}$$

$$H(X,Y) = -\sum_{xy} p_{xy}(x,y)\log p_{xy}(x,y) \tag{5}$$

H is the entropy of the image. For an image Y, the entropy can be expressed as Eq. 3. Let the gray level value of image Y be $r \in [0, L-1]$, the $\mathrm{r_k}$ represent the gray value of level k. The p $(\mathrm{r_k})$ represents the probability of gray level k, which can be expressed as Eq. 4. The h $(\mathrm{r_k})$ is the histogram discrete function representation of the image Y, $\mathrm{n_k}$ represents the number of pixels whose gray level value is $\mathrm{r_k}$ in image Y. For two images X and Y, the joint information entropy of two images can be expressed as Eq. 5. When the similarity of two images is higher, or the overlap part is larger, the correlation is higher and the joint entropy is smaller, that is, the mutual information is larger.

2.5.1.2 Optimizer

The role of the optimizer is to guide each parameter of the objective function to update the appropriate size in the correct direction during the updating process of the iterative function, so that the updated parameter drives the value of the objective function to approach the global minimum continuously. The regular step gradient descent optimizer (RSGD) and the one-plus-one evolutionary optimizer (OPOE) are commonly used optimizers. The regular step gradient descent optimizer follows the gradient of image similarity measure in the extreme direction [52]. It uses a constant step length along the gradient between computations until the gradient change direction. Thereafter, with each change in the gradient direction, the step size is reduced according to the relaxation factor. The one-plus-one evolutionary optimizer iterates to find a set of parameters that yield the best registration result. It does this by tuning the arguments (the parent arguments) from the last iteration. If the new (child) parameters yield a better result, the new parameter becomes the adjusted new parent parameter. The next iteration will be more aggressive. If the result of the child parameter is not as good as that of the parent parameter, the parent parameter remains, and the next perturbation will be less aggressive [53].

2.5.2 Registration based on feature information

The feature-based registration method extracts the common features from the reference image and the floating image as the registration primitives. It then estimates the geometric transformation model and parameter values between the reference and floating images by establishing the corresponding relationship between the registration primitives. It has the advantages of low computational complexity and strong robustness. It is suitable for registering images with complex geometric deformation, but not for registering images with blurred feature points or a smaller number of features.

2.5.2.1 Scale-invariant feature transform

Scale-invariant feature transform (SIFT) is proposed and further improved by Lowe et al. [54]. It can effectively solve the problem of image scale invariance and rotation invariance, and has good robustness to noises and illumination changes. The main idea of the algorithm is: firstly, the scale space of the image is established; then, the extremum points of the image are searched in the scale space; the feature descriptors are established for the extremum points; the similarity matching is carried out by the feature descriptors; the parameter estimation of the model is transformed; and finally the registration is completed.

2.5.2.2 Speeded up robust features

Speeded up robust features (SURF) is an improvement of SIFT [55]. The main feature of SURF is to use the Hessian determinant value as the feature point and respond to the Harr wavelet transform. It uses the integral graph effectively, and the processing speed is accelerated.

3 Evaluation

After image registration, specific evaluation measures were used to assess the performance of the registration algorithms. The MOS, time consumption, SSIM, PSNR, RMSE, and percentage P of jump in quantitative visual perception were calculated and compared to evaluate the registration performance of different registration methods.

3.1 Subjective visual evaluation

Subjective evaluation is the most common and direct image evaluation method [56]. Image registration must first satisfy the qualitative visual perception. In MOS, image quality is divided into five grades according to its merits, and the best rating is five.

3.2 Objective and quantitative evaluation

When visually challenging to judge, objective parameters are usually applied to evaluate the registration performance quantitatively. In this study, the objective evaluation measures used are computation time, RMSE, SSIM, PSNR, and jump percentage P in quantitative visual perception.

3.2.1 Computation time

Computation time is used to measure the computing speed of image registration. Under the same conditions, the algorithm with less computation time is faster in image registration. There is often a tradeoff between computation time and registration accuracy in practical applications. Therefore, balancing these two factors to optimize quality and efficiency is also the pursuit of image registration algorithms.

3.2.2 Root mean square error

RMSE is usually applied to indicate the deviation between the calculated value and the actual value. The smaller the RMSE, the higher the accuracy and the better the algorithm's performance. It is represented as Eq. 6:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\left\|T\left(p'i,\theta\right)-pi\right\|\right)^{2}} \tag{6}$$

Where p' and p represent the matching points in the image to be registered and the reference image.

3.2.3 Structural similarity index metric

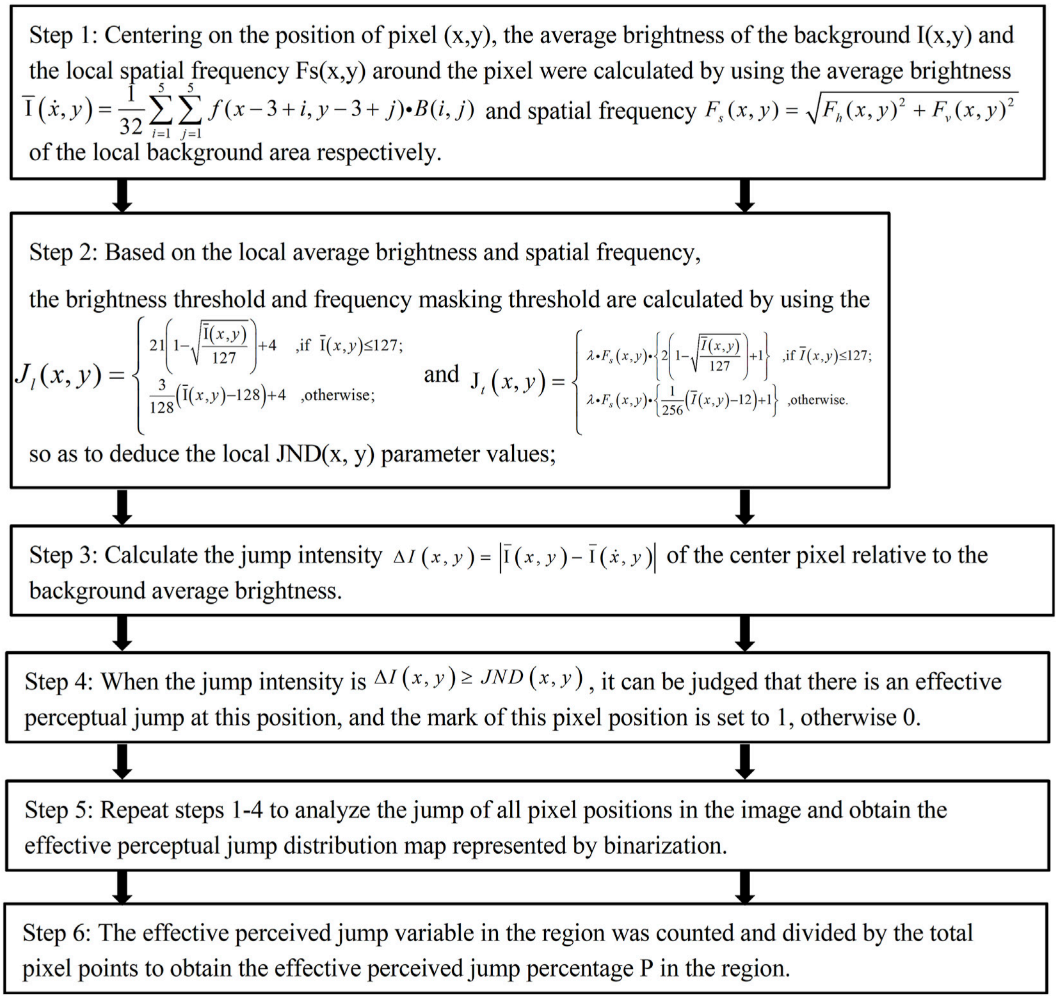
SSIM is a measure of the similarity between two images. SSIM is designed based on the ability of the human visual system to capture the structural features of images. The image is evaluated by brightness, contrast, and structure. The value ranges from 0 to 1, and the higher value represents the higher the similarity between the two images in brightness, contrast, and structure. In practical applications, the Gaussian function, variance, and covariance are generally used to calculate the mean value of images instead of traversing the pixels to achieve higher efficiency. It can be expressed as Eq. 7:

$$SSIM\left(x,y\right)=\frac{\left(2u_{x}u_{y}+C_{1}\right)\left(2\sigma_{xy}+C_{2}\right)}{\left(u_{x}^{2}+u_{y}^{2}+C_{1}\right)\left(\sigma_{x}^{2}+\sigma_{y}^{2}+C_{2}\right)} \tag{7}$$

Where, μ_x is the mean of x, μ_y is the mean of y, σ_x^2 is the variance of x, σ_y^2 is the variance of y, and σ_{xy} is the covariance of x and y.

3.2.4 Peak signal-to-noise ratio

PSNR is used to measure the difference between two images, which can compute the impact of the noises that affect the quality



Step 1: Centering on the position of pixel (x,y), the average brightness of the background I(x,y) and the local spatial frequency Fs(x,y) around the pixel were calculated by using the average brightness $\bar{\mathrm{I}}(\dot{x},y)=\frac{1}{32}\sum_{i=1}^{5}\sum_{j=1}^{5}f(x-3+i,y-3+j)\bullet B(i,j)$ and spatial frequency $F_s(x,y)=\sqrt{F_h(x,y)^2+F_v(x,y)^2}$ of the local background area respectively.

Step 2: Based on the local average brightness and spatial frequency,

the brightness threshold and frequency masking threshold are calculated by using the

$$J_l(x,y)=\begin{cases}21\left(1-\sqrt{\frac{\bar{\mathrm{I}}(x,y)}{127}}\right)+4 & ,\text{if } \bar{\mathrm{I}}(x,y)\le 127;\\ \frac{3}{128}\left(\bar{\mathrm{I}}(x,y)-128\right)+4 & ,\text{otherwise;}\end{cases} \quad \text{and} \quad \mathrm{J}_t(x,y)=\begin{cases}\lambda\bullet F_s(x,y)\bullet\left\{2\left(1-\sqrt{\frac{\bar{I}(x,y)}{127}}\right)+1\right\} & ,\text{if } \bar{I}(x,y)\le 127;\\ \lambda\bullet F_s(x,y)\bullet\left\{\frac{1}{256}\left(\bar{I}(x,y)-12\right)+1\right\} & ,\text{otherwise.}\end{cases}$$

so as to deduce the local JND(x, y) parameter values;

Step 3: Calculate the jump intensity $\Delta I(x,y)=\left|\bar{\mathrm{I}}(x,y)-\bar{\mathrm{I}}(\dot{x},y)\right|$ of the center pixel relative to the background average brightness.

Step 4: When the jump intensity is $\Delta I(x,y)\ge JND(x,y)$, it can be judged that there is an effective perceptual jump at this position, and the mark of this pixel position is set to 1, otherwise 0.

Step 5: Repeat steps 1-4 to analyze the jump of all pixel positions in the image and obtain the effective perceptual jump distribution map represented by binarization.

Step 6: The effective perceived jump variable in the region was counted and divided by the total pixel points to obtain the effective perceived jump percentage P in the region.

FIGURE 6
Quantitative visual perception algorithm.

of its representation between registered and reference images. PSNR is the most commonly used objective evaluation index of images. It is based on the error between corresponding pixels. Thus it is the image quality evaluation based on error sensitivity. The higher value of PSNR indicates better registration quality. Since it doesn't consider the visual characteristics of the human eyes, the evaluation results are sometimes inconsistent with visual perception. Mathematically, it is given as Eq. 8:

$$PSNR(X,Y)=10\log_{10}\left(\frac{\max(Y)}{1/NM\sum_{x,y=1}^{NM}(X-Y)^2}\right) \tag{8}$$

3.2.5 Visual perception

The brightness effect and the spatial frequency masking effect are two essential characteristics of human vision. According to these effects, an important parameter of visual discrimination, just noticeable difference (JND), is derived. This parameter is used to calculate the change of effective pixels number of an image, namely, jump percentage P (higher p-value is better). Thus, it can obtain quantitative evaluation results consistent with human visual perception. The evaluation algorithm is given in Figure 6.

4 Results

The different intensity-based (see Figure 7) and feature-based registrations are evaluated and compared. In our results, the feature-based registration method has a low registration rate (SIFT 33.3% and SURF 11.1%), and poor registration quality. This is due to the insufficient feature points or feature pairs in the PAT images. The results

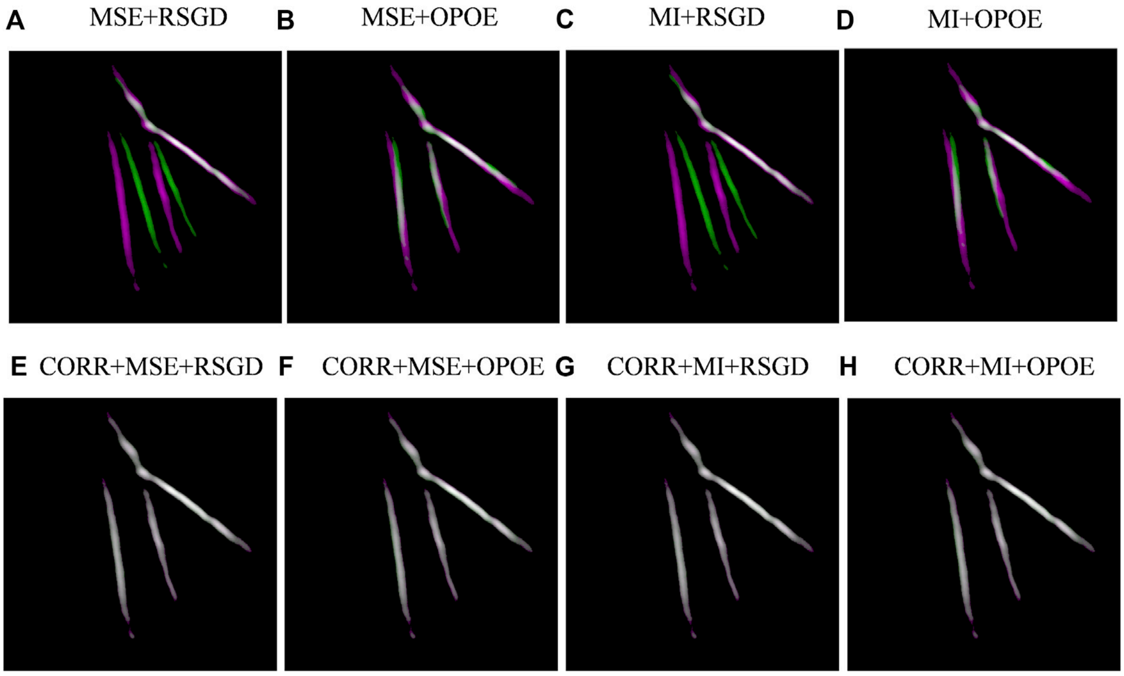


FIGURE 7
Example of image registration for dataset 9. **(A–H)** represents different image registration algorithms.

indicate that the feature-based registration in this study is not suitable for applications like hand blood vessel PAT images. In Figure 7, examples of registered images (dataset 9) of each registration method are shown. Although phase correlation (CORR) cannot obtain high registration quality, it is fast. And by applying it as a rough registration with other methods, the overall registration performance can be significantly improved. Therefore, we took it as a rough registration to combine with other registration schemes and compared the results with and without CORR. In Figure 7, A-H refers to the eight different registration schemes. A–D are the registration results without phase correlation, which are MSE and regular step gradient descent optimizer, MSE and one-plus-one evolutionary optimizer, MI and regular step gradient descent optimizer, MI and one-plus-one evolutionary optimizer. E-F are the registration results of phase correlation combined with the previous methods. In the registration images, magenta and green represent the differences between the two images, and the registered part is white.

4.1 Mean opinion score

Three people who did not participate in the experiment were invited to evaluate the images. A score of 5 refers to the best result, and the percentage of a score of 4 or above in all datasets was used as the evaluation index. A higher percentage indicates a better result. The MOS evaluation results for different registration schemes are shown in Figure 8. Before phase correlation was applied as a rough registration, the scheme of MI with regular step gradient descent optimizer got the best registration result, which was 33.33%. After using phase correlation as rough registration, the overall registration quality was improved. The percentage was increased by 14.81%–48.15%. Overall, the scheme of phase correlation combined with MSE and regular step gradient descent optimizer offers the best registration results (59.26%).

4.2 Computation time

Figure 9 shows the time consumption of the eight registration schemes with different datasets. We repeated the experiments five times for each dataset with each registration method and plotted the averaged time consumption, which excludes the contingency and is representative to a certain extent. As can be seen from Figure 9, similarity measure MSE has a faster registration rate than MI when the optimizer is consistent. When the similarity measure is consistent, one-plus-one evolutionary optimizer takes less time than regular step gradient descent

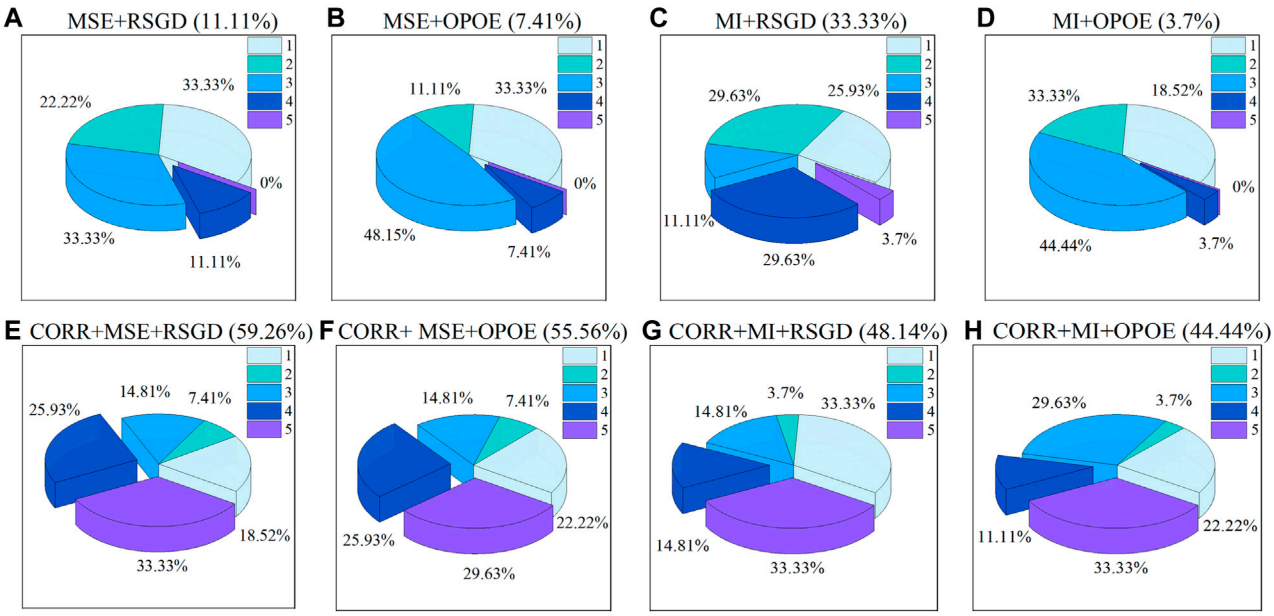


FIGURE 8
MOS evaluation of different registration methods. **(A–H)** represents registration schemes. The proportion of results with a score of four or above is shown in parentheses.

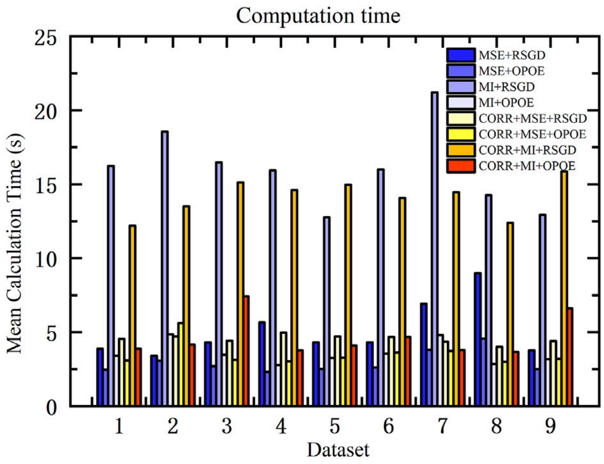


FIGURE 9
Mean computation time of 8 registration schemes in 9 groups of datasets.

optimizer. Overall, whether rough registration is applied or not, the scheme of MI with regular step gradient descent optimizer takes longer. The scheme of MSE with one-plus-one evolutionary optimizer, takes the shortest time. Other schemes take about the same time.

4.3 Objective evaluation measures

The SSIM, RMSE, PSNR, and percentage of visual perception P were calculated (see Figure 10). We compared the performances with and without phase correlation rough registration across all the nine groups of datasets. In addition, we also compared the performances among the eight registration schemes in each type of datasets. Figure 10A shows the SSIM value of registration results using different registration schemes in the nine groups of datasets. After introducing phase correlation as rough registration, the total SSIM value of nine data groups did not change significantly, which means they have similar contrast and structural degradation. Comparing the eight schemes, the scheme of phase correlation combined with MSE and regular step gradient descent optimizer has good SSIM values for each type of datasets. Figure 10B shows the RMSE value. Overall, with rough registration, the registered image has a better RMSE value, indicating better registration image quality. Taken individually, phase correlation combined MSE and regular step gradient descent optimizer had high accuracy in most datasets. Figure 10C represents the PSNR of the datasets. The overall image quality is slightly improved after rough registration. Figure 10D shows the quantified percentage of visual perception. For each type of datasets, phase correlation combined MSE and regular step gradient descent optimizer has a higher value of P, in other words, it is more perceptive to the human eye. As a whole, when the optimizer is consistent, the similarity measure MSE

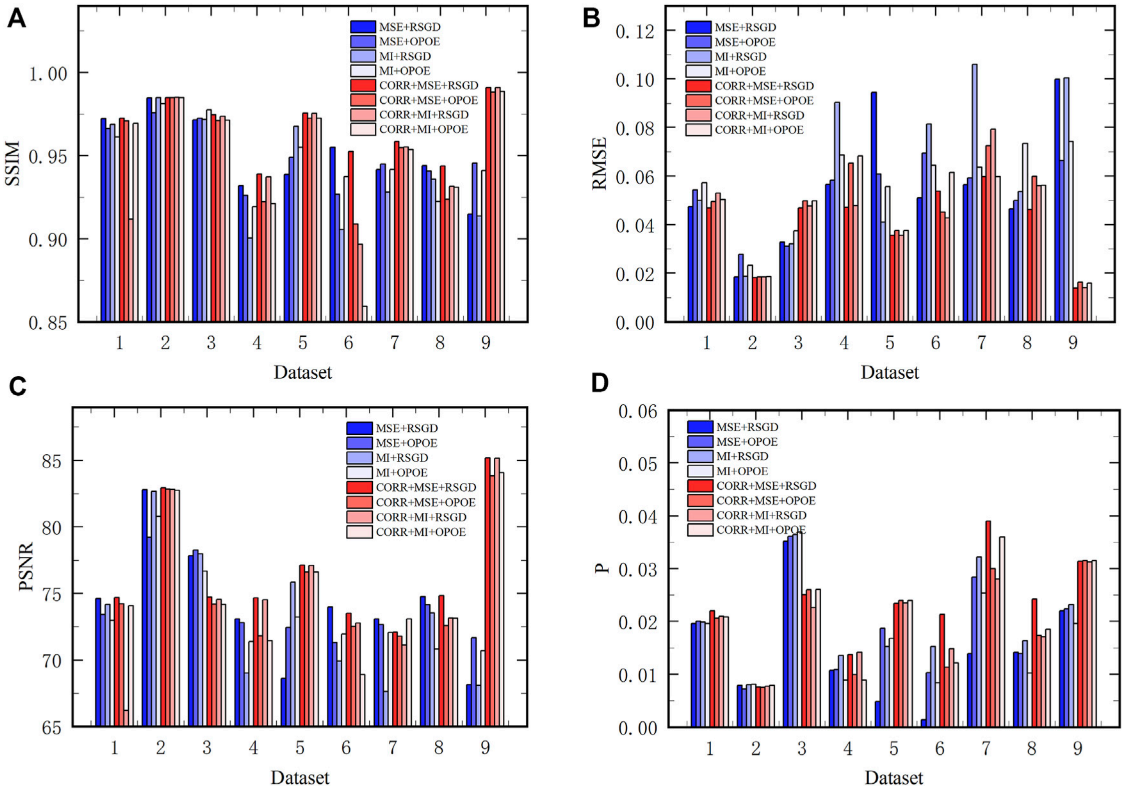


FIGURE 10
The evaluation of 9 groups of datasets using objective measures. **(A–D)** represent SSIM, RMSE, PSNR, and jump percentage P, respectively.

performs better than MI. That is, it has a higher value of SSIM and PSNR, as well as a lower value of RMSE. When the similarity measure is consistent, the regular step gradient descent optimizer has a better performance than the one-plus-one evolutionary optimizer, which is also reflected in a higher value of SSIM and PSNR, and a lower value of RMSE. Moreover, this result remains unchanged after phase correlation is introduced for rough registration.

5 Discussion

5.1 Optical wavelength

Different endogenous contrast agents have different absorption spectra. Hemoglobin is commonly used as an endogenous contrast agent, which is widely used for vascular imaging in the visible and near-infrared spectral ranges. When the wavelength is 720 nm, the absorption coefficient of HbO2 is less than that of HbR. When the wavelength is 850 nm, the absorption coefficient of HbO2 is greater than that of HbR. When the wavelength is 960 nm, more deep information can be provided [57]. In the human body, the oxygen content of arterial blood is higher than that of venous blood [58]. Therefore, when imaging at different wavelengths, veins and arteries can be effectively distinguished. These wavelengths are commonly used in vascular functional imaging [59].

5.2 Image registration

Feature-based methods are widely used in image registration. However, our results show that sufficient feature points or feature pairs cannot be obtained in the PAT images of hand blood vessels, resulting in low registration efficiency. The reasons for this could be that PAT images are lack of rich curve inflection points and local curvature discontinuity points, which do not meet the corner features commonly extracted in SIFT and SURF feature extraction. Therefore, the feature-based registration method is unsuitable for registering PAT images of a small range of hand blood vessels.

Intensity-based registration is the primary registration method analyzed in this study. Different similarity measures and optimization methods turn the registration problem into an optimization problem. The grayscale calculation in a small range provides a fast computation while ensuring high accuracy. In our study, we found that among the commonly used registration methods we tested, the mean square error similarity measurement combined with the regular step gradient descent optimizer has better accuracy, but it is time-consuming. In order to improve the performance, we proposed to apply phase correlation as a rough registration to combine with the previous methods. Although the registration accuracy of phase correlation is low, it has the advantage of fast registration in simple translation. Therefore, we proposed a new scheme that divided the registration into two parts. First, the registration of large positions is completed by phase correlation, and then the registration adjustment of small errors is carried out using MSE and regular step gradient descent optimizer. The results showed a significant improvement in the overall registration performance. In addition, we found that all these registration methods showed a relatively consistent variation in performance on different types of data, which indicates that they can be applied for image registration of all the four types of PAT data. Comparing the eight registration schemes by combining MOS, time consumption, and objective evaluation measures, the proposed scheme of phase correlation combined MSE and regular step gradient descent optimizer has stable and superior performance.

5.3 Evaluation

Image registration must first satisfy the subjective visualization. In this study, subjective evaluations were conducted by different subjects to satisfy the qualitative visual evaluation, as well as to reduce subjectivity. When the differences are hard to identify for the visual assessment, objective evaluation measures are often applied to analyze the registration results quantitatively. The reasons for selecting SSIM, RMSE, and PSNR are as follows:1) by calculating SSIM and PSNR, we can compare the image quality of the registered images relative to reference images, including image contrast, brightness, structural degradation, and unwanted noise; 2) The error between the float image and the reference image can be measured by calculating RMSE.

Since the SSIM, PSNR, and RMSE are calculated based on pixel intensity, it is possible that these objective measures give results different from the subjective visual evaluation. When subjective visual evaluation cannot be made, it is necessary to add the measure of quantitative visual perception to reflect human visual perception. Therefore, both subjective and objective evaluation measures should be considered when judging image registration performance. Meanwhile, computation time is also worth considering. To some extent, there is a tradeoff between the computation time and registration accuracy. Our goal is to achieve high-quality fast registration. Thus, the overall computation time and accuracy also need to be considered comprehensively.

By combining the four objective measures, subjective visual evaluation, and computation time, we evaluated the performances of different combinations of the two most commonly used similarity measures and optimizers for intensity-based registration with four types of PAT datasets. We also combined phase correlation with these registration methods and evaluated the change in registration performance. The results show that by adding phase correlation, the overall performance can be greatly improved, and by combining phase correlation with MSE and regular step gradient descent optimizer, the registration gives better performance in all four categories of datasets. The results were validated using multiple sets of data for each category and were consistent.

6 Conclusion

In this study, intensity-based and feature-based automatic registration methods were investigated in the application of PAT vascular imaging using four types of human hand vascular PAT data. In addition, a new scheme with phase correlation was proposed to improve the performance of the previous registration methods. The feature-based registration methods (SIFT and SURF) did not provide good performance in our application. We had evaluated the performances of the intensity-based schemes by applying subjective visual evaluation and four objective evaluation measures SSIM, RMSE, PSNR, and jump percentage P. In addition, computation time was also considered. We found that by adding phase correlation as a rough registration, the overall registration performance can be significantly improved. From the results, we can conclude that the proposed scheme combining the phase correlation rough registration, mean square error measurement, and regular step gradient descent optimizer gives the best overall performance. This study provides a useful tool of image registration for clinical applications of PAT vascular imaging, such as vascular growth monitoring for early screening and postoperative evaluation of cancers.

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 Medical Research Ethics Committee of Chongqing 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 and approved it for publication.

Funding

This research was supported by the Youth fund project of National Natural Science Foundation of China under grant 62201103,62101083; Project funded by China Postdoctoral Science Foundation 2020M683260; and Natural Science Foundation of Chongqing, China under grant cstc2021jcyj-msxmX0739, cstc2021jcyj-msxmX0104.

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

1. Miller J. C., Pien H. H., Dushyant S., Gregory S. A., Thrall J. H., Imaging angiogenesis: Applications and potential for drug development. *JNCI J Natl Cancer Inst* (2005) 3:172–87. doi:10.1093/jnci/dji023

2. Hectors S. J., Jacobs I., Lok J., Peters J., Bussink J., Hoeben F. J., et al. mproved evaluation of antivascular cancer therapy using constrained tracer-kinetic modeling for multiagent dynamic contrast-enhanced mri. *Cancer Res* (2018) 78(6):1561–70. doi:10.1158/0008-5472.Can-17-2569

3. Vriens D., de Geus-Oei L. F., Heerschap A., van Laarhoven H. W., Oyen W. J., Vascular and metabolic response to bevacizumab-containing regimens in two patients with colorectal liver metastases measured by dynamic contrast-enhanced mri and dynamic 18 F-Fdg-Pet. *Clin Colorectal Cancer* (2011) 10(1): E1–E5. doi:10.3816/ccc.2011.n.010

4. Yang X., Review article quantifying tumor vascular heterogeneity with dynamic contrast-enhanced magnetic resonance imaging: A review *Journal of Biomedicine and Biotechnology* 2011 12 (2013).doi:10.1155/2011/732848

5. Chang Y. C., Huang Y. H., Huang C. S., Chang R. F., Vascular morphology and tortuosity analysis of breast tumor inside and outside contour by 3-D power Doppler ultrasound. *Ultrasound Med Biol* (2012) 38(11):1859–69. doi:10.1016/j.ultrasmedbio.2012.06.010

6. Cuinet C., Sarran F., Uzan-Augui J., Bonastre J., Lassau N., Coiffier B., et al. Selection of an early biomarker for vascular normalization using dynamic contrast-enhanced ultrasonography to predict outcomes of metastatic patients treated with bevacizumab. *Ann Oncol* 2016 27 1922, doi:10.1093/annonc/mdw280

7. Kenneth H., Heidi U., Mark L., Robbin M., Forero-Torres A., *Ultrasound in medicine & biology* (2015).Ultrasound imaging of breast tumor perfusion and neovascular morphology, 41 2292-302. doi:10.1016/j.ultrasmedbio.2015.04.016

8. Lassau N., Bonastre J., Kind M., Lassau N., Bonastre J., Kind M., et al. Validation of dynamic contrast-enhanced ultrasound in predicting outcomes of antiangiogenic therapy for solid tumors: The French multicenter support for innovative and expensive. *Investigative Radiology* (2014) 49:794–800. doi:10.1097/rli.0000000000000085

9. Lassau N., Chami L., Benatsou B., Peronneau P., Roche A., Dynamic contrast-enhanced ultrasonography (Dce-Us) with quantification of tumor perfusion: A new diagnostic tool to evaluate the early effects of antiangiogenic treatment. *Eur Radiol Suppl* (2007) 17(6):89–98. doi:10.1007/s10406-007-0233-6

10. Wu Z., Yang X., Chen L., Wang Z., Shi Y., Mao H., et al. Anti-angiogenic therapy with contrast-enhanced ultrasound in colorectal cancer patients with liver metastasis. *other* (2017)(20) e6731 96.doi:10.1097/MD.0000000000006731

11. Kim E., Stamatelos S., Cebulla J., Bhujwalla Z. M., Popel A. S., Pathak A. P., Multiscale imaging and computational modeling of blood flow in the tumor vasculature. *Ann Biomed Eng* (2012)(11) 40.2425-41. doi:10.1007/s10439-012-0585-5

12. Ma1 E., Ren1 A., Gao1 B., Yang1 M., Zhao1 Q., Wang1 W., et al. Roi for outlining an entire tumor is a reliable approach for quantification of lung cancer tumor vascular parameters using ct perfusion. *OncoTargets & Therapy* (2016).9: 2377-84. doi:10.2147/OTT.S98060

13. Schmitz S., Rommel D., Michoux N., Lhommel R., Hanin F. X., Duprez T., et al. Dynamic contrast-enhanced computed tomography to assess early activity of cetuximab in squamous cell carcinoma of the head and neck. *Radiol Oncol* (2015) 49(1):17–25. doi:10.2478/raon-2014-0030

14. Yao J., Yang Z. G., Chen H. J., Chen T. W., Huang J., Gastric adenocarcinoma: Can perfusion ct help to noninvasively evaluate tumor angiogenesis? *Abdom Imaging* (2011) 36(1):15–21. doi:10.1007/s00261-010-9609-5

15. Vu T, Razansky D, Yao J, Listening to tissues with new light: Recent technological advances in photoacoustic imaging. *J Opt* (2019) 21(10):103001. doi:10.1088/2040-8986/ab3b1a

16. Huang P. H., Tsai J. C., Kuo L. T., Lee C. W., Lai H. S., Tsai L. K., et al. Clinical application of perfusion computed tomography in neurosurgery. *J Neurosurg* (2014) 120(2):473–88. doi:10.3171/2013.10.jns13103

17. Jonathan L., Maxime H., Jean L., Josserand V., Exploration of melanoma metastases in mice brains using endogenous contrast photoacoustic imaging. *Int J Pharmaceutics* (2017) 532:704–9. doi:10.1016/j.ijpharm.2017.08.104

18. Kumar D. M., Christos M., Michael M., Petersen C. R., Moselund P. M., Bang O., High-pulse energy supercontinuum laser for high-resolution spectroscopic photoacoustic imaging of lipids in the 1650-1850 Nm region. *Biomed Opt Express* (2018) 9(4):1762. doi:10.1364/boe.9.001762

19. Neuschmelting V., Lockau H., Ntziachristos V., Grimm J., Kircher M. F., Lymph node micrometastases and in-transit metastases from melanoma: *In vivo* detection with multispectral optoacoustic imaging in a mouse model (2016).*Radiology* 160191 280 doi:10.1148/radiol.2016160191

20. Sangha G. S., Phillips E. H., Goergen C. J., Vivo photoacoustic lipid imaging in mice using the second near-infrared window. *Biomedical Optics Express* (2017). 8: 736–742. doi:10.1364/BOE.8.000736

21. Zhang H. F., Maslov K., Stoica G., Wang L. V., Functional photoacoustic microscopy for high-resolution and noninvasive *in vivo* imaging. *Nat Biotechnol* (2006) 24:848–51. doi:10.1038/nbt1220

22. Liu Y., Zhang L., Li S., Han X., Yuan Z., Imaging molecular signatures for clinical detection of scleroderma in the hand by multispectral photoacoustic elastic tomography. *J Biophotonics* (2018) 11(6):e201700267. doi:10.1002/jbio.201700267

23. Shang S. S., Chen Z. J., Zhao Y., Yang S. H., Xing D., Simultaneous imaging of atherosclerotic plaque composition and structure with dual-mode photoacoustic and optical coherence tomography. *Opt Express* (2017) 25(2):530–9. doi:10.1364/oe.25.000530

24. Brown E., Brunker J., Bohndiek S. E., Photoacoustic imaging as a tool to probe the tumour microenvironment. *Dis Model Mech* (2019) 12(7):dmm039636. doi:10.1242/dmm.039636

25. Brown E. L., Lefebvre T. L., Sweeney P. W., Stolz B. J., Grohl J., Hacker L., et al. Quantification of vascular networks in photoacoustic mesoscopy. *Photoacoustics* (2022) 26:100357. doi:10.1016/j.pacs.2022.100357

26. Haedicke K., Agemy L., Omar M., Berezhnoi A., Roberts S., Longo-Machado C., et al. High-resolution optoacoustic imaging of tissue responses to vascular-

targeted therapies. *Nat Biomed Eng* (2020) 4(3):286–97. doi:10.1038/s41551-020-0527-8

27. Tang J., Dai X., Jiang H., Wearable scanning photoacoustic brain imaging in behaving rats. *J Biophotonics* (2016) 9(6):570–5. doi:10.1002/jbio.201500311

28. Kajita H., Suzuki Y., Sakuma H., Imanishi N., Tsuji T., Jinzaki M., et al. Visualization of lymphatic vessels using photoacoustic imaging. *Keio J Med* (2021) 70(4):82–92. doi:10.2302/kjm.2020-0010-OA

29. Mantri Y., Tsujimoto J., Donovan B., Fernandes C. C., Garimella P. S., Penny W. F., et al. Photoacoustic monitoring of angiogenesis predicts response to therapy in healing wounds. *Wound Repair Regen* (2022) 30(2):258–67. doi:10.1111/wrr.12992

30. Shan T., Zhao Y., Jiang S., Jiang H., *In-vivo* hemodynamic imaging of acute prenatal ethanol exposure in fetal brain by photoacoustic tomography. *J Biophotonics* (2020) 13(5):e201960161. doi:10.1002/jbio.201960161

31. Wang P. P., Chen Z. J., Xing D., Multi-parameter characterization of atherosclerotic plaques based on optical coherence tomography, photoacoustic and viscoelasticity imaging. *Opt Express* (2020) 28(9):13761–74. doi:10.1364/oe.390874

32. Zhang J., Duan F., Liu Y., Nie L., High-resolution photoacoustic tomography for early-stage cancer detection and its clinical translation. *Radiol Imaging Cancer* (2020) 2(3):e190030. doi:10.1148/rycan.2020190030

33. Yang J., Zhang G., Wu M., Shang Q., Huang L., Jiang H., Photoacoustic assessment of hemodynamic changes in foot vessels. *J Biophotonics* (2019) 12(6):e201900004. doi:10.1002/jbio.201900004

34. Lin L., Wang L. H. V., The emerging role of photoacoustic imaging in clinical oncology. *Nat Rev Clin Oncol* (2022) 19(6):365–84. doi:10.1038/s41571-022-00615-3

35. Zhang C., Feng W., Zhao Y. J., Yu T. T., Li P. C., Xu T. H., et al. A large, switchable optical clearing skull window for cerebrovascular imaging. *Theranostics* (2018) 8(10):2696–708. doi:10.7150/thno.23686

36. Yang J., Zhang G., Shang Q., Wu M., Huang L., Jiang H., Detecting hemodynamic changes in the foot vessels of diabetic patients by photoacoustic tomography. *J Biophotonics* (2020) 13(8):e202000011. doi:10.1002/jbio.202000011

37. Zitová B., Flusser J., Image registration methods: A survey. *Image Vis Comput* (2003) 21(11):977–1000. doi:10.1016/s0262-8856(03)00137-9

38. Amelio L., Amelio A., Ct image registration in acute stroke monitoring. Proceedings of the 41th Jubilee International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO) (May 2018). Opatija, Croatia (Piscataway, NJ, United States). doi:10.23919/MIPRO.2018.8400275

39. Kim K. W., Lee J. M., Klotz E., Kim S. J., Kim S. H., Kim J. Y., et al. Safety margin assessment after radiofrequency ablation of the liver using registration of preprocedure and postprocedure ct images. *Am J Roentgenology* (2011) 196(5):565–72. doi:10.2214/ajr.10.5122

40. Demir A., Unal G., Karaman K., editors Anatomical landmark based registration of contrast enhanced T1-weighted mr images, B. Fischer, B. M. Dawant, C. Lorenz, Anatomical landmark based registration of contrast enhanced T1-weighted mr images. *Biomedical Image Registration*. Berlin, Heidelberg: Springer (2010).doi:10.1007/978-3-642-14366-3_9

41. Lin W. -C., Wu C. -C., Huang T. -C., Lin W. -C., Chiu B. Y. -C., Liu R. -S., et al. Red blood cell velocity measurement in rodent tumor model: An *in vivo* microscopic study. *J Med Biol Eng* (2012) 32(2):97–102. doi:10.5405/jmbe.875

42. Matl S., Brosig R., Baust M., Navab N., Demirci S., Vascular image registration techniques: A living review. *Med Image Anal* (2017) 35:1–17. doi:10.1016/j.media.2016.05.005

43. Ringstad G. A., Emblem K. E., Holland D., Dale A. M., Bjornerud A., Hald J. K., Assessment of pituitary adenoma volumetric change using longitudinal mr image registration. *Neuroradiology* (2012) 54(5):435–43. doi:10.1007/s00234-011-0894-7

44. Dumani D. S., Yu A., Thompson W., Brecht H. -P., Ivanov V., Anastasio M. A., et al. editors. Preclinical small animal imaging platform providing Co-registered 3d maps of photoacoustic response and fluorescence. Proccedings of the Conference on photons plus ultrasound - imaging and sensing. San Francisco, CA, USA (March 2019).doi:10.1117/12.2514489

45. Farnia P., Makkiabadi B., Alimohamadi M., Najafzadeh E., Basij M., Yan Y., et al. Photoacoustic-mr image registration based on a Co-sparse analysis model to compensate for brain shift. *Sensors* (2022) 22(6):2399. doi:10.3390/s22062399

46. Gehrung M., Tomaszewski M., McIntyre D., Disselhorst J., Bohndiek S., Co -registration of optoacoustic tomography and magnetic resonance imaging data from murine tumour models. *Photoacoustics* (2020) 18:100147. doi:10.1016/j.pacs.2019.100147

47. Gonzalez E. A., Jain A., Bell M. A. L., Combined ultrasound and photoacoustic image guidance of spinal pedicle cannulation demonstrated with intact *ex vivo* specimens. *IEEE Trans Biomed Eng* (2021) 68(8):2479–89. doi:10.1109/tbme.2020.3046370

48. Ren W., Skulason H., Schlegel F., Rudin M., Klohs J., Ni R., Automated registration of magnetic resonance imaging and optoacoustic tomography data for experimental studies. *Neurophotonics* (2019) 6(2):1. doi:10.1117/1.NPh.6.2.025001

49. Wang W. Q., Wang W. H., Yin J., A bilateral filtering based ringing elimination approach for motion-blurred restoration image. *Curr Opt Photon* (2020) 4(3):200–9. doi:10.3807/copp.2020.4.3.200

50. Ranjbarzadeh R., Saadi S. B., Automated liver and tumor segmentation based on concave and convex points using fuzzy C-means and mean shift clustering. *Measurement* (2020) 150:107086. doi:10.1016/j.measurement.2019.107086

51. Sen A., Anderson B. M., Cazoulat G., McCulloch M. M., Elganainy D., McDonald B. A., et al. Accuracy of deformable image registration techniques for alignment of longitudinal cholangiocarcinoma ct images. *Med Phys* (2020) 47(4):1670–9. doi:10.1002/mp.14029

52. Kamsing P., Torteeka P., Yooyen S., An enhanced learning algorithm with a particle filter-based gradient descent optimizer method. *Neural Comput Appl* (2020) 32(16):12789–800. doi:10.1007/s00521-020-04726-9

53. Styner M., Brechbuhler C., Szekely G., Gerig G., Parametric estimate of intensity inhomogeneities applied to mri. *IEEE Trans Med Imaging* (2000) 19(3):153–65. doi:10.1109/42.845174

54. Wang H., Fan J., Li Y., Research of shoeprint image matching based on sift algorithm. *J Comput Methods Sci Eng* (2016) 16:349–59. doi:10.3233/JCM-160622

55. Zhu H., Jiang Y., Zhang C., Liu S., Research on mosaic method of uav low-altitude remote sensing image based on sift and surf. *J Phys : Conf Ser* (2022) 2203(1):012027. doi:10.1088/1742-6596/2203/1/012027

56. Streijl R. C., Winkler S., Hands D. S., Mean opinion score (mos) revisited: Methods and applications, limitations and alternatives. *Multimedia Syst* (2016) 22(2):213–27. doi:10.1007/s00530-014-0446-1

57. Li M. C., Tang Y. Q., Yao J. J., Photoacoustic tomography of blood oxygenation: A mini review. *Photoacoustics* (2018) 10:65–73. doi:10.1016/j.pacs.2018.05.001

58. Harrop G. A., The oxygen and carbon dioxide content of arterial and of venous blood in normal individuals and in patients with anemia and heart disease. *J Exp Med* (1919) 30(3):241–57. doi:10.1084/jem.30.3.241

59. Yao J., Wang L. V., Sensitivity of photoacoustic microscopy. *Photoacoustics* (2014) 2(2):87–101. doi:10.1016/j.pacs.2014.04.002

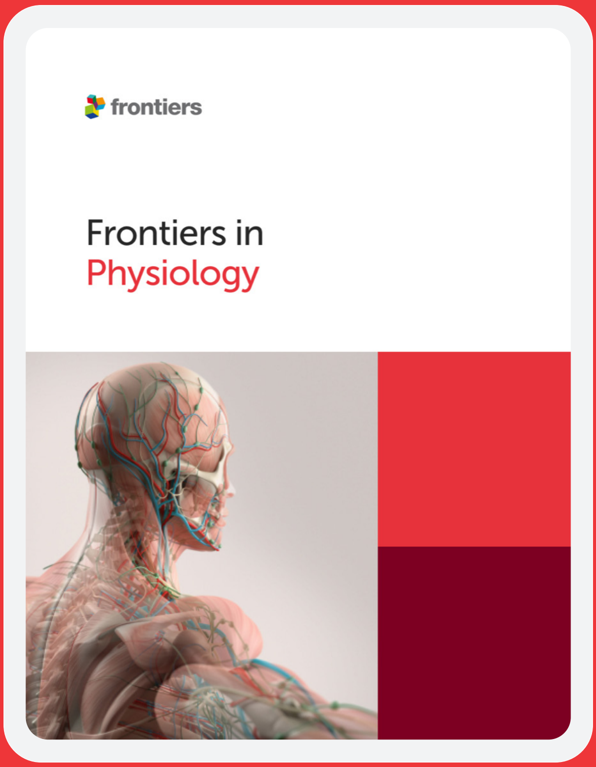
Frontiers in Physiology

Understanding how an organism's components work together to maintain a healthy state

The second most-cited physiology journal, promoting a multidisciplinary approach to the physiology of living systems - from the subcellular and molecular domains to the intact organism and its interaction with the environment.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact