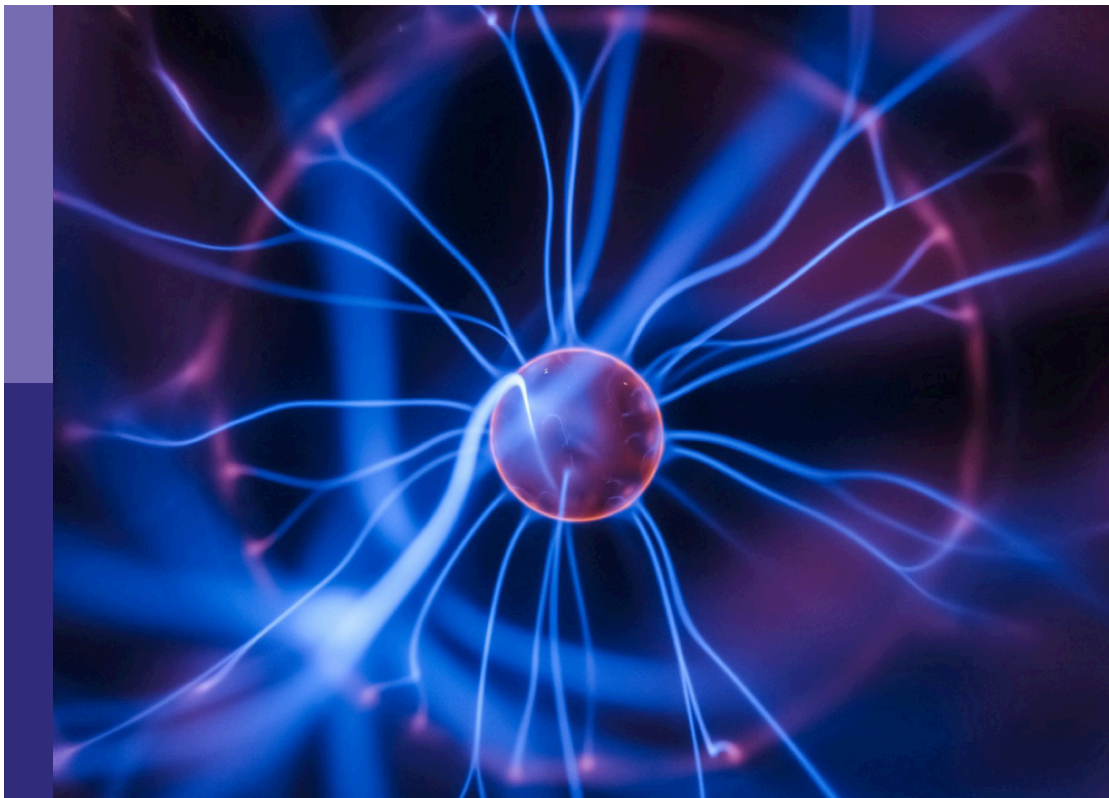
frontiers | Research Topics

Interdisciplinary techniques in biomedical photonics

Edited by
Wei Gong, Xunbin Wei, Liwei Liu, Zhiwei Huang, Junjie Yao and Minbiao Ji

Published in
Frontiers in Physics



FRONTIERS EBOOK COPYRIGHT STATEMENT

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

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

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

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

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

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

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

ISSN 1664-8714
ISBN 978-2-83252-107-6
DOI 10.3389/978-2-83252-107-6

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

Interdisciplinary techniques in biomedical photonics

Topic editors

Wei Gong — Zhejiang University, China
Xunbin Wei — Peking University, China
Liwei Liu — Shenzhen University, China
Zhiwei Huang — National University of Singapore, Singapore
Junjie Yao — Duke University, United States
Minbiao Ji — Fudan University, China

Citation

Gong, W., Wei, X., Liu, L., Huang, Z., Yao, J., Ji, M., eds. (2023). *Interdisciplinary techniques in biomedical photonics*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-83252-107-6

Table of contents

04 Real-Time Reconstruction of the Complex Field of Phase Objects Based on Off-Axis Interferometry
Xiang Li, Guanyuan Qin, Wenhui Yu, Huixian Li, Rui Hu, Junle Qu and Liwei Liu

11 Three-Dimensional Virtual Optical Clearing With Cycle-Consistent Generative Adversarial Network
Jiajia Chen, Zhenhong Du and Ke Si

22 Automatic quantifying and monitoring follow-ups for implantable collamer lens implantation using AS-OCT images
Yiming Sun, Jinhao Li, Peifang Xu, Pengjie Chen, Yaqi Wang, Shaodan Hu, Gangyong Jia, Shuai Wang and Juan Ye

31 Portable flexible probe for detecting blood supply status in clinical surgery
Jie Xia, Jiadong Pan, Ruijian Yan, Fan Zhang, Luxi Zhang, Gang Feng, Jack Luo, Shurong Dong and Yueming Wang

41 Color fundus photograph registration based on feature and intensity for longitudinal evaluation of diabetic retinopathy progression
Jingxin Zhou, Kai Jin, Renshu Gu, Yan Yan, Yueyu Zhang, Yiming Sun and Juan Ye

51 FOCMS: An ultrafast optical clearing method with quantified analysis
Xiaobin Xu, Yameng Zheng, Xiao Xiao, Zizheng Wang, Jie Lv, Yongjie Wang, Jianmin Zhang, Shumin Duan and Ke Si

61 Three-dimensional natural color imaging based on focus level correlation algorithm using structured illumination microscopy
Mengrui Wang, Tianyu Zhao, Zhaojun Wang, Kun Feng, Jingrong Ren, Yansheng Liang, Shaowei Wang and Ming Lei

69 Automatic classification of gastric lesions in gastroscopic images using a lightweight deep learning model with attention mechanism and cost-sensitive learning
Lingxiao Wang, Yingyun Yang, Jingyang Li, Wei Tian, Kun He, Tianming Xu, Zhaohui Fang, Aiming Yang and Ting Li

83 Design of an ellipsoidal mirror assisted oblique plane microscopy
Yuzhe Liu, Qiushu Chen, Yiyan Fei, Lan Mi and Jiong Ma

90 MSDenoiser: Muti-step adaptive denoising framework for super-resolution image from single molecule localization microscopy
Qianghui Feng, Qihang Song, Meng Yan, Zhen Li Huang and Zhengxia Wang

frontiers | Frontiers in Physics

ORIGINAL RESEARCH
published: 29 June 2022
doi: 10.3389/fphy.2022.943016

Real-Time Reconstruction of the Complex Field of Phase Objects Based on Off-Axis Interferometry

*Xiang Li†, Guanyuan Qin†, Wenhui Yu, Huixian Li, Rui Hu, Junle Qu and Liwei Liu**

Key Laboratory of Optoelectronic Devices and Systems of Guangdong Province and Ministry of Education, College of Physics and Optoelectronic Engineering, Shenzhen University, Shenzhen, China

Quantitative phase imaging (QPI) can acquire dynamic data from living cells without the need for physical contact. We presented a real-time and stable dynamic imaging system for recording complex fields of transparent samples by using Fourier transform based on off-axis interferometry. We calculated and removed the system phase without sample to obtain the real phase of the sample, so as to ensure that the system has the ability to accurately measure the phase. The temporal and spatial phase sensitivity of the system was evaluated. Benefit from the ability to record the dynamic phase and phase profile of a specimen, a standard sample (polystyrene microspheres) is investigated to demonstrate the efficiency of this imaging system and we have observed the variation of erythrocyte membrane during Red Blood Cells (RBCs) spontaneous hemolysis with different mediums. Experimental results indicate that the phase of non-anticoagulant RBC changed apparently than anticoagulant RBC and the system could be applied to real-time noninvasive and label-free identification of living cells.

Keywords: real-time, complex field, living cells, off-axis interferometry, dynamic imaging

OPEN ACCESS

Edited by:
Baoli Yao,
Xian Institute of Optics and Precision Mechanics, (CAS), China

Reviewed by:
Rongli Guo,
Xi'an Technological University, China
Jianglei Di,
Guangdong University of Technology, China

****Correspondence:***
Liwei Liu
liulw@szu.edu.cn

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

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

Received: *13 May 2022*
Accepted: *10 June 2022*
Published: *29 June 2022*

Citation:
Li X, Qin G, Yu W, Li H, Hu R, Qu J and Liu L (2022) Real-Time Reconstruction of the Complex Field of Phase Objects Based on Off-Axis Interferometry. Front. Phys. 10:943016. doi: 10.3389/fphy.2022.943016

1 INTRODUCTION

Most living cells are almost transparent when illuminated by visible light, essentially acting as phase objects. Same techniques such as phase-contrast microscopy and differential interference difference microscopy can carry out microscopic imaging of transparent samples, consequently revealing the structural details of biological systems [1, 2]. In spite of this, the information of the illumination field obtained by these techniques is only qualitative, and it is difficult to describe the morphology of the sample quantitatively. Both non-interference and interference methods have been widely used in quantitative phase imaging of biological samples. For example, the microscopy based on intensity transfer equations can realize phase imaging of biological samples through a series of numerical operations [3–6]. However, it is limited by the complexity of the calculation process and the long time required. The advantages of digital holographic microscopy (DHM) are rapid, non-destructive, and high-resolution which is widely used in the study of cell structural characteristics, cell deformation, cell dynamics, etc., [7–10]. Meanwhile, it can also be combined with other technologies [11–13] to form a multi-mode microscopic imaging technology. The acquisition rate of this technique is limited only by CCD and has the ability to measure the morphological characteristics of living cells in real-time [14, 15]. However, it should not be ignored that the real-time monitoring quality and the longest observation time of QPI will be limited by the overall stability of the imaging system [16, 17]. In off-axis DHM, the low contrast of interference fringes usually reduces the phase sensitivity of the system. In addition, camera dark noise, read noise and

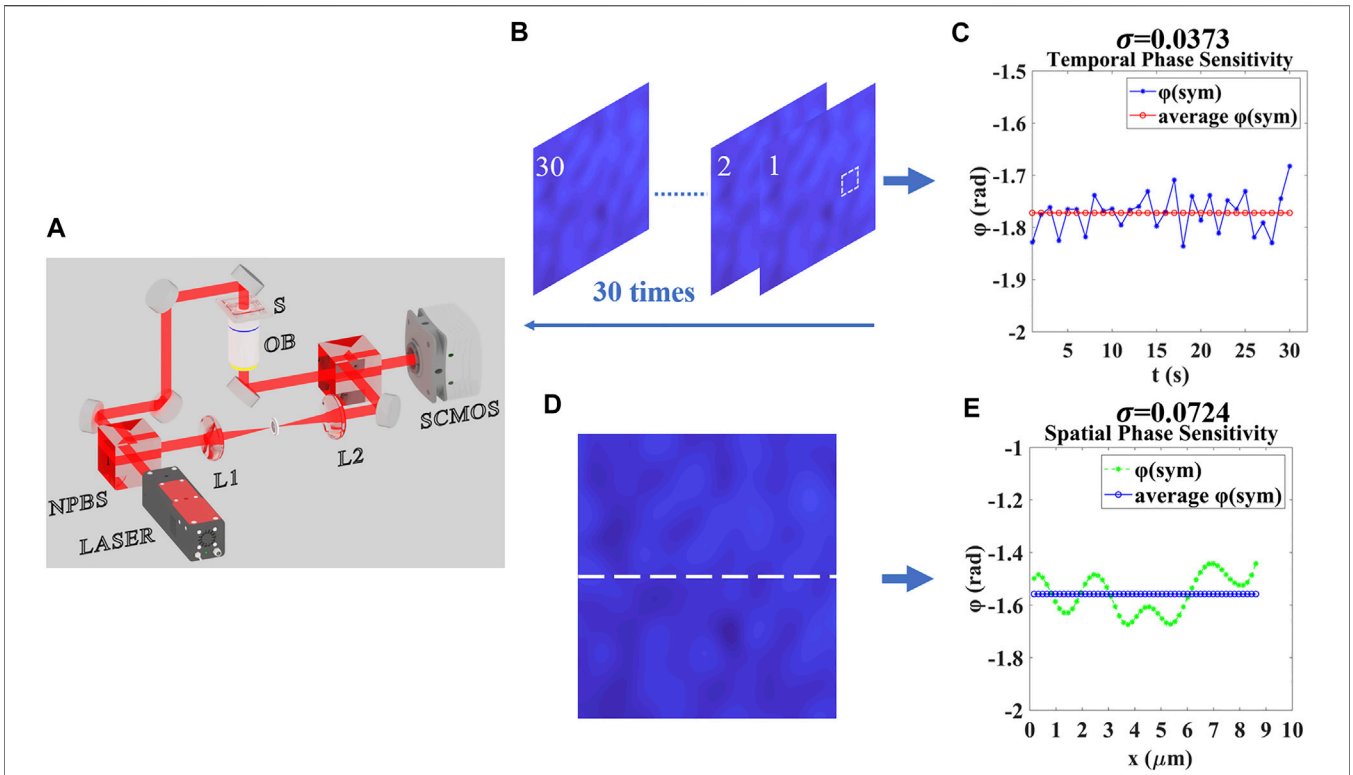


FIGURE 1 | Working principle and spatiotemporal calibration of the system. **(A)** Schematic of the experimental setup. NPBS, non-polarized beam splitter; SCMOS, scientific complementary metal-oxide-semiconductor camera; OB, objective lens; S, stage. **(B)** 30 s phase measurement of the system without sample. **(C)** Temporal Phase Sensitivity. **(D)** phase measurement of the system without sample along the white line. **(E)** Spatial Phase Sensitivity. φ(sym): phase of the system without sample, σ: Standard deviation of the phase without sample.

other instrumental parameters may affect the measurement sensitivity of the system [18]. The spatial sensitivity of the system is easily affected by speckle noise factors such as scattering field of impurities on optical elements and random interference noise patterns generated by specular reflection of various surfaces in the system, or environmental factors such as mechanical vibration and air density fluctuations. Spatial light interference microscopy (SLIM) is considered as a method to reduce speckle noise inherent in laser light source [19]. In conclusion, it is significant to check and reduce the influence of noise and environmental factors on measurement sensitivity during post-processing. In the process of off-axis spatial filtering, the Fourier transform involved will also bring unnecessary noise to the image. Therefore, in order to ensure the spatial phase sensitivity of QPI, it is particularly important to subtract the background phase.

In this paper, an off-axis real-time digital holographic microscopy system based on Mach- Zehnder interferometer was designed, to solve the above problems, we acquired the background phase firstly by dealing with the interference fringe without a sample then the phase caused only by the sample can be computed by subtracting the background phase. Consequently, the dynamic imaging of the phase only caused by the sample could be achieved. To verify the feasibility of the imaging system, we have demonstrated the experiments on polystyrene microspheres and red blood cells [20–23].

2 SYSTEM AND METHODS

2.1 Experimental Setup

A typical setup of off-axis interferometry is depicted in **Figure 1A**. A continuous wave (CW) laser (MRL-III-650L, Changchun new industry), which was used for the imaging Mach-Zehnder interferometer, was steered to the first non-polarized beam splitter (NPBS), after which the beam was separated to perform off-axis interferometry.

The sample was placed on a three-axis displacement table for wide-field illumination. An objective (Daheng Optics, GCO-213 40x, NA = 0.60) imaged the sample to a scientific complementary metal-oxide-semiconductor (sCMOS) camera which was positioned at the imaging plane of the objective, where an exact (magnified) replica of the sample field can be formed. The acquisition rate of the sCMOS that we used (PCO.Panda.4.2, Germany) is 48 frames/s when acquiring at the full resolution of 2048 × 2048 pixels. To produce a clean reference beam, a 20 μm pinhole was placed within the reference path at the common focus of a pair of lenses (L1 and L2)

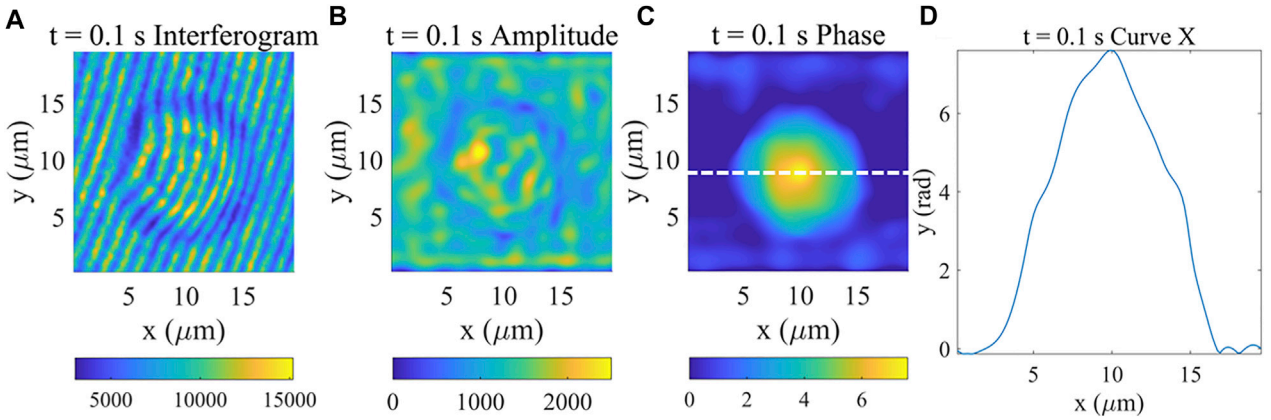


FIGURE 2 | Complex field and Phase measurement results of polystyrene microspheres. **(A)** Interferogram image at 0.1 s. **(B)** Amplitude image at 0.1 s. **(C)** Phase image at 0.1 s. **(D)** Phase profile of horizontal direction curves at the location indicated by the white dashed line shown in **(C)**.

performing spatial filtering. Finally, the reference field was slightly tilted relative to the sample beam and interfered with the sample beam to form uniform phase modulation fringes in sCMOS. The standard Fourier transform (FT) algorithm was adopted to reconstruct the amplitude and phase [24]. Specifically, the camera respectively recorded a hologram with and without a sample and performed a fast two-dimensional Fourier transform (By selecting the higher-order information in the spectrum to fundamental and performing the inverse Fourier transform), by subtracting and unwrapping the phase information, the phase directly related to the sample can be obtained. Off-axis interferogram was recorded after optimized the fringe contrast with an exposure time of 10 ms. The temporal and spatial phase sensitivity was evaluated as shown in **Figure 1B,D** (a 30 s continuous and a series of different points measurement of the phase without samples). The absolute value of the phase is not meaningful, but the relative change of the phase is meaningful. The standard deviation of points was selected to demonstrate the time-space domain phase fluctuation, the smaller the value, the more stable the phase, which shows great temporal-spatial stability as shown in **Figure 1C,E**.

2.2 Principle

As for the spatial coherence imaging system, the intensity comes out at sCMOS has the form:

$$I = |U_0|^2 + |U_1(x,y)|^2 + 2|U_0||U_1(x,y)|\cos\left(\varphi(x,y) + 2\pi f_x x + 2\pi f_y y + \varphi_n\right) \tag{1}$$

$|U_0|^2$ and $|U_1(x,y)|^2$ represent the irradiance distribution of reference and the irradiance distribution of sample respectively, $\varphi(x,y)$ represents the optical delay caused by the sample, which is the amount of interest in the experiments, f_x and f_y, respectively, represent the spatial frequencies of fringes with X and Y direction, and φ_n is the additional phase modulation introduced by the environment noise. For easy description, we denote $2\pi f_x x + 2\pi f_y y$ as φ_{sym}. By Fourier high-pass filtering, the interference term $U(x,y)$ can be isolated:

$$U(x,y) = 2|U_0||U_1(x,y)|\cos\left(\varphi(x,y) + \varphi_{sym} + \varphi_n\right) \tag{2}$$

By applying Euler's formula, j represents an imaginary unit:

$$U(x,y) = |U_0||U_1(x,y)|\left(e^{j\left(\varphi(x,y)+\varphi_{sym}+\varphi_n\right)} + e^{-j\left(\varphi(x,y)+\varphi_{sym}+\varphi_n\right)}\right) \tag{3}$$

In the spectrum, the interference term $U(x,y)$ is divided into two parts as follow, which distribute along the center fundamental frequency signal symmetrically and contain the same high-frequency information, $\mathcal{F}$ is the Fourier operator, $u(k_x,k_y)$ is the Fourier transform of $U(x,y)$.

$$\begin{aligned} u\left(k_x,k_y\right)_{+1} &= \left|U_0\right|\mathcal{F}\left(\left|U_1(x,y)\right|e^{j\left(\varphi(x,y)+\varphi_{sym}+\varphi_n\right)}\right) \\ &= \left|U_0\right|u_{\left(k_x-2\pi f_x x, k_y-2\pi f_y y\right)} \end{aligned} \tag{4}$$

$$\begin{aligned} u\left(k_x,k_y\right)_{-1} &= \left|U_0\right|\mathcal{F}\left(\left|U_1(x,y)\right|e^{-j\left(\varphi(x,y)+\varphi_{sym}+\varphi_n\right)}\right) \\ &= \left|U_0\right|u_{\left(k_x+2\pi f_x x, k_y+2\pi f_y y\right)} \end{aligned} \tag{5}$$

By taking any term $(u(k_x,k_y)_{+1}$ or $u(k_x,k_y)_{-1})$ in the spectrum as higher-order information and return it to the fundamental frequency:

$$u\left(k_x,k_y\right)_{\pm 1} = \left|U_0\right|\mathcal{F}\left(\left|U_1(x,y)\right|e^{\pm j\left(\varphi(x,y)+\varphi_n\right)}\right) \tag{6}$$

The term φ_{sym} introduced by moving reference beam is eliminated.

The phase and amplitude information of the sample could be obtained by inverse Fourier transform of $u(k_x,k_y)_{\pm 1}$ in the spectrum back to the spatial domain:

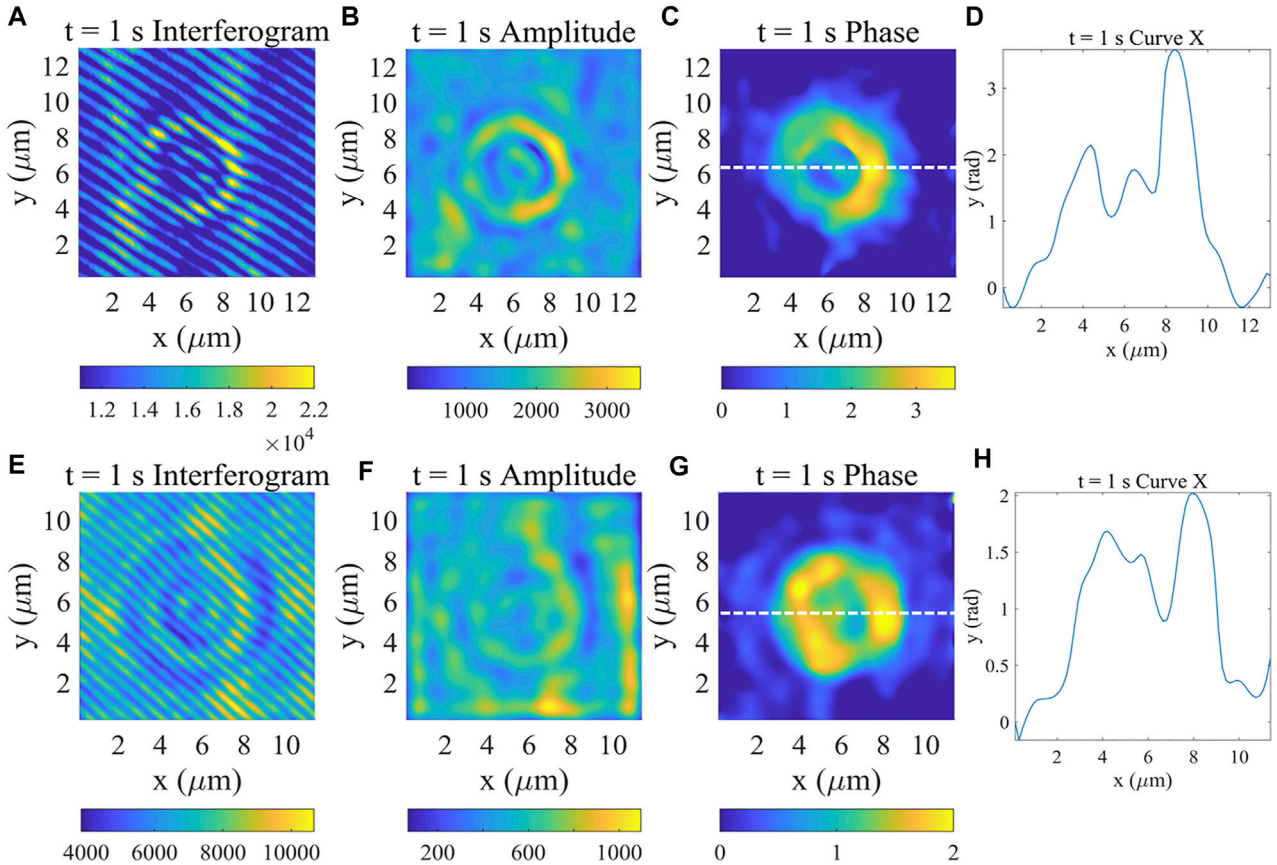


FIGURE 3 | Complex field and Phase measurement results of red blood cells without and with anticoagulant (the results of the first row are without anticoagulant). **(A,E)** Interferogram image at 1 s **(B,F)** Amplitude image at 1 s **(C,G)** Phase image at 1 s **(D,H)** Phase profile of horizontal direction curves at the location indicated by the white dashed line shown in **(C,G)**.

$$\phi(x, y) = \varphi(x, y) + \varphi_n \tag{7}$$

The phase calculated here includes the phase caused by non-samples, which we solve in this paper by the following method. The phase caused by ambient noise or the possible presence of impurity scattering field φ_n in the system can be obtained by performing the same process above on the interferogram without samples. Finally, the real phase value of the target sample can be obtained by calculating the difference between the phase value obtained with and without the sample $\varphi(x, y) = \phi(x, y) - \varphi_n$. It can be seen from **Figure 1** that this method can make the system have good measurement sensitivity.

2.3 Sample Preparation

To demonstrate the phase stabilization ability of the proposed method, we performed QPI of polystyrene sphere as well as the RBC spontaneous hemolysis. For the study, all cell samples were taken from three-month-old mice. After euthanasia on mice (Put the mouse's body straight, lift it up about 30° diagonally, and instantly break the cervical vertebra according to IACUC guidelines), we extracted blood by removing the eyeballs of mice. For comparison, we divide the obtained blood into two parts, one immediately poured into a centrifuge tube with anticoagulant and the other not, centrifuged for 5 min at a speed of 2000 RPM, and then sucked 2 μl red blood cells on a slide for imaging experiments with our experimental system. Animal related experiments were approved by Guangdong Medical Experimental Animal Center (Code: C202110-01).

3 RESULTS

Using quantitative phase imaging techniques for live cell monitoring can better reflect other cell information such as phase, optical thickness, etc. We first achieved the complex field and phase measurement ability of our system using standard samples (polystyrene microspheres) as shown in **Figure 2**. Polystyrene microsphere is a common test target because of its simple and easily identifiable structure. The polystyrene microspheres we use are about 10 microns with refractive index of 1.60. We took a hologram of the sample as shown in **Figure 2A** in Olympus oil medium with approximate refractive index of 1.52. **Figure 2B,C** show the overall amplitude and phase distributions with Fourier transform (FT) algorithm. The phase information shows the spherical structure of the sample, what's more, we plotted the phase profile as shown in **Figure 2D** to verify the effectiveness of our imaging system. (Real-time dynamic process of the polystyrene microspheres See **Supplementary Figure S1**).

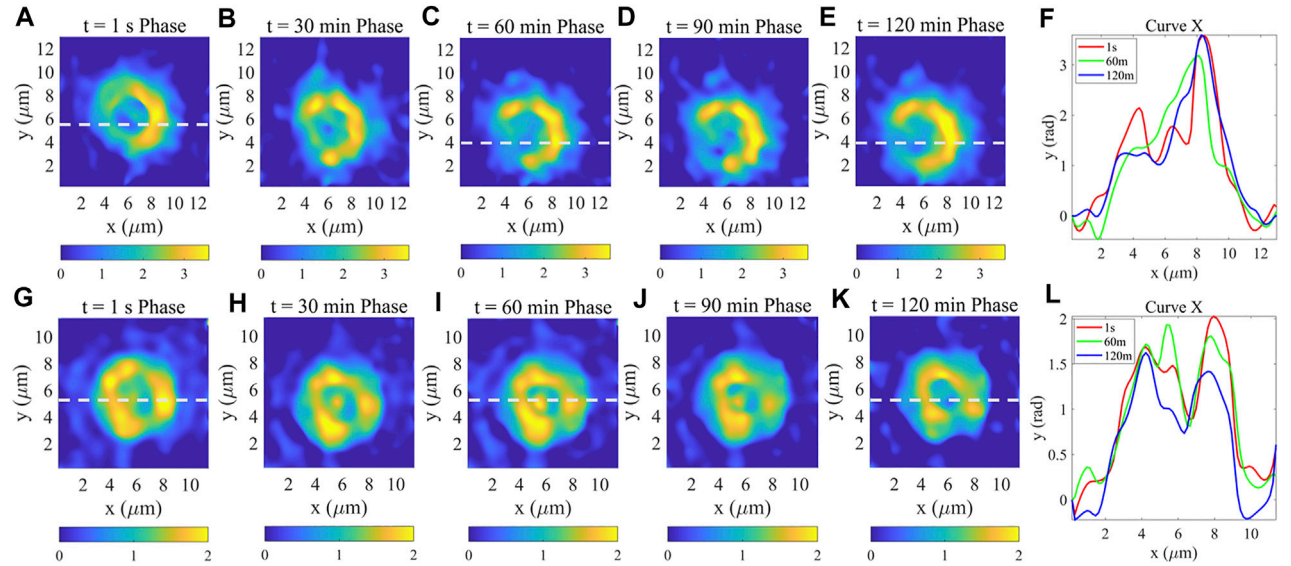


FIGURE 4 | (A–E) are phase measurement results of red blood cells without anticoagulant in 2 hours. **(F)** Phase profile of horizontal direction curves at the location indicated by the white dashed line shown in **(A,C,E)**; **(G–K)** are phase measurement results of red blood cells with anticoagulant in 2 h. **(L)** Phase profile of horizontal direction curves at the location indicated by the white dashed line shown in **(G,I,K)**.

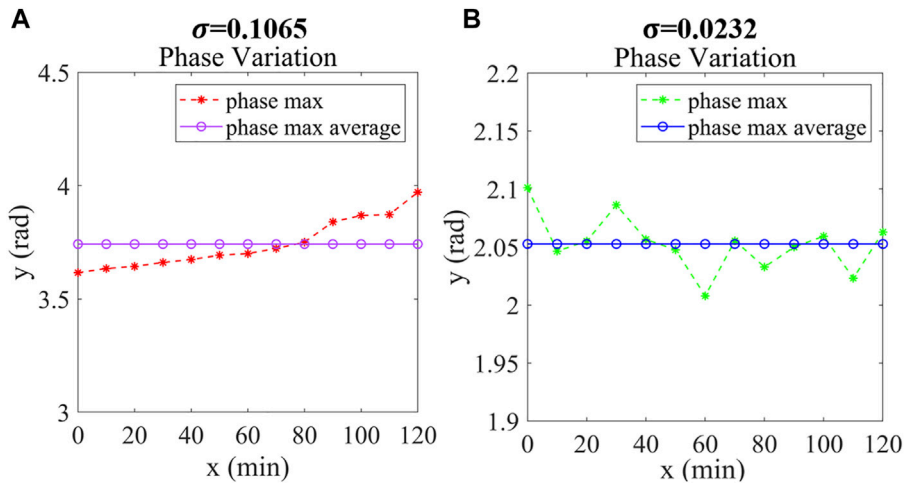


FIGURE 5 | (A) Phase Maximum variation of red blood cells without anticoagulant. **(B)** Phase Maximum variation of red blood cells with anticoagulant. σ: Standard deviation of the RBC Phase Maximum.

Then, we imaged living red blood cells with different mediums. The result of the complex field of the blood cell without anticoagulant and its phase profile at 1s are shown in **Figure 3A–D**, and **Figure 3E–H** shows the results of the red blood cell with anticoagulant. After obtaining the interferogram (**Figure 3A**) with high contrast, the complex field of this sample was obtained by using the algorithm introduced above, as shown in **Figure 3B,C**. To verify the accuracy of the achieved phase, the phase value across the middle white line was plotted (**Figure 3D**) which showed the unique structure of RBCs.

To observe the morphological changes of red blood cells during spontaneous hemolysis, phase measured for 4 consecutive hours, the results of the first 2 h were shown in **Figure 4**. On the one hand, it can be clearly observed that the structure of the non-anticoagulant RBC membrane has undergone significant changes as shown in **Figure 4A–E** which is caused by the variation of its osmotic pressure between internal and external during cell spontaneous hemolysis [7]. Specifically, the RBC (**Figure 4A**) has a complete structure that means it has two peaks which can be seen from its phase curve shown in **Figure 3D**. An hour later, as shown in **Figure 4F**, the RBC only has one peak which indicates both that its cell membrane morphology has changed and it has died. From **Figure 4 C,D,E**, we

can find that the morphology of the RBC is basically unchanged cause of death of the cell. On the other hand, the results in **Figure 4G–K** show that the anticoagulant RBC membrane has not changed significantly and the RBC has two peaks all the time as shown in **Figure 4L**. Comparing **Figure 4A–E** with **Figure 4G–K**, it can be found that due to the addition of anticoagulants to the samples, which enabled the RBC to maintain physiologically active for a long time and with no obvious variation of the RBC phase. The dynamic phase of the RBC was continuously acquired after 2 hours for approximately 120 min with 30 s intervals as shown in **Supplementary Figures S2, S3**.

To show the phase variation of the non-anticoagulant and anticoagulant RBC in 2 hours qualitatively, we continually recorded the maximum phase of the cells with 10 min intervals and calculate its average value as shown in **Figure 5A** and **Figure 5B**. Apparently, the maximum phase of non-anticoagulant RBC trended to change in one direction, but of which the anticoagulant RBC changed little and leveled off at the same time. The standard deviation was used to show the stability of data usually. Compared with the anticoagulant RBC, the standard deviation of the non-anticoagulant RBC phase maximum is bigger which means non-anticoagulant RBC has changed sharply and the addition of anticoagulants has the effect of making the cell morphology last longer.

4 CONCLUSION

In summary, we have measured the complex field of objects using off-axis interferometry. The dynamic phase and phase profile are used to describe the morphology changes of the sample. We have observed the spontaneous hemolysis process of red blood cells with two different mediums in 2 hours (with anticoagulant or not). For the non-anticoagulant RBC, the cell membrane changed significantly during spontaneous hemolysis. We also quantitatively described the spontaneous hemolysis of red blood cells in two ambient fluids by using the standard deviation of the maximum phase. The advantages of our system are that good measurement stability of the system is obtained by subtracting the background phase, the phase and profile information are combined and can be detected in real-time, and the time of one single shoot needed is about 0.1 s. We believe that this work can be applied to the physiological detection of living cells.

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 animal study was reviewed and approved by Animal related experiments were implemented according to the guidance of the Medical Department of Shenzhen University.

AUTHOR CONTRIBUTIONS

XL and GQ conceptualized the study. XL, GQ, and WY performed the analysis of data. XL and GQ performed the data collection and wrote the original draft. XL, GQ, WY, HL, RH, JQ, and LL reviewed and edited the manuscript. LL reviewed and supervised writing the manuscript.

FUNDING

This work was supported in part by The National Natural Science Foundation of China (62175163/61935012/61961136005/61835009/62127819), Shenzhen Key projects JCYJ20200109105404067, Shenzhen International Cooperation Project (GJHZ20190822095420249), and Shenzhen Talent Innovation Projects RCJC20210706091949022.

ACKNOWLEDGMENTS

We thank students in the group for providing mice in the sample preparation procedure in our imaging system.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2022.943016/full#supplementary-material

REFERENCES

1. Zernike F. Phase Contrast, a New Method for the Microscopic Observation of Transparent Objects Part II. *Physica* (1942) 9(10):974–86. doi:10.1016/S0031-8914(42)80079-8
2. Nomarski G. Microinterféromètre différentiel à ondes polarisées. *J Phys Rad* (1955) 16:9S–13S.
3. Li Y, Di J, Ma C, Zhang J, Zhong J, Wang K, et al. Quantitative Phase Microscopy for Cellular Dynamics Based on Transport of Intensity Equation. *Opt Express* (2018) 26(1):586. doi:10.1364/OE.26.000586
4. Greenbaum A, Zhang Y, Feizi A, Chung P-L, Luo W, Kandukuri SR, et al. Wide-Field Computational Imaging of Pathology Slides Using Lens-Free On-Chip Microscopy. *Sci Transl Med* (2014) 6(267):ra175. doi:10.1126/scitranslmed.3009850
5. Zuo C, Chen Q, Yu Y, Asundi A. Transport-of-Intensity Phase Imaging Using Savitzky-Golay Differentiation Filter - Theory and Applications. *Opt Express* (2013) 21(5):5346–62. doi:10.1364/OE.21.005346
6. Gureyev TE, Nugent KA. Rapid Quantitative Phase Imaging Using the Transport of Intensity Equation. *Opt Commun* (1997) 133(1-6):339–46. doi:10.1016/S0030-4018(96)00454-3

7. Popescu G, Ikeda T, Best CA, Badizadegan K, Dasari RR, Feld MS. Erythrocyte Structure and Dynamics Quantified by Hilbert Phase Microscopy. *J Biomed Opt* (2005) 10(6):060503. doi:10.1117/1.2149847
8. Lue N, Popescu G, Ikeda T, Dasari RR, Badizadegan K, Feld MS. Live Cell Refractometry Using Microfluidic Devices. *Opt Lett* (2006) 31(18):2759–61. doi:10.1364/OL.31.002759
9. Popescu G, Ikeda T, Goda K, Best-Popescu CA, Laposata M, Manley S, et al. Optical Measurement of Cell Membrane Tension. *Phys Rev Lett* (2006) 97(21): 218101. doi:10.1103/PhysRevLett.97.218101
10. Masters BR. Quantitative Phase Imaging of Cells and Tissues. *J Biomed Opt* (2012) 17(2):029901. doi:10.1117/1.JBO.17.2.029901
11. Park Y, Popescu G, Badizadegan K, Dasari RR, Feld MS. Diffraction Phase and Fluorescence Microscopy. *Opt Express* (2006) 14(18):8263. doi:10.1364/OE.14.008263
12. Hu C, He S, Lee YJ, He Y, Kong EM, Li H, et al. Live-Dead Assay on Unlabeled Cells Using Phase Imaging with Computational Specificity. *Nat Commun* (2022) 13(1):713. doi:10.1038/s41467-022-28214-x
13. Park Y, Depeursinge C, Popescu G. Quantitative Phase Imaging in Biomedicine. *Nat Photon* (2018) 12(10):578–89. doi:10.1038/s41566-018-0253-x
14. Hoover EE, Squier JA. Advances in Multiphoton Microscopy Technology. *Nat Photon* (2013) 7(2):93–101. doi:10.1038/nphoton.2012.361
15. Pham HV, Bhaduri B, Tangella K, Best-Popescu C, Popescu G. Real Time Blood Testing Using Quantitative Phase Imaging. *PLoS One* (2013) 8(2): e55676. doi:10.1371/journal.pone.0055676
16. Lee K, Kim K, Jung J, Heo J, Cho S, Lee S, et al. Quantitative Phase Imaging Techniques for the Study of Cell Pathophysiology: from Principles to Applications. *Sensors* (2013) 13(4):4170–91. doi:10.3390/s130404170
17. Bhaduri B, Edwards C, Pham H, Zhou R, Nguyen TH, Goddard LL, et al. Diffraction Phase Microscopy: Principles and Applications in Materials and Life Sciences. *Adv Opt Photon* (2014) 6(1):57–119. doi:10.1364/AOP.6.000057
18. Hosseini P, Zhou R, Kim Y-H, Peres C, Diaspro A, Kuang C, et al. Pushing Phase and Amplitude Sensitivity Limits in Interferometric Microscopy. *Opt Lett* (2016) 41(7):1656. doi:10.1364/OL.41.001656
19. Wang Z, Millet L, Mir M, Ding H, Unarunotai S, Rogers J, et al. Spatial Light Interference Microscopy (SLIM). *Opt Express* (2011) 19:1016–26. doi:10.1364/OE.19.001016
20. Memmolo P, Miccio L, Merola F, Mugnano M, Ferraro P. Hydrodynamic Red Blood Cells Deformation by Quantitative Phase Microscopy and Zernike Polynomials. *Front Phys* (2019) 7:111. doi:10.3389/fphy.2019.00111
21. Vasudevan S, Chen GCK, Andika M, Agarwal S, Chen P, Olivo M. Dynamic Quantitative Photothermal Monitoring of Cell Death of Individual Human Red Blood Cells upon Glucose Depletion. *J Biomed Opt* (2010) 15(5):057001. doi:10.1117/1.3484260
22. Memmolo P, Merola F, Miccio L, Mugnano M, Ferraro P. Investigation on Dynamics of Red Blood Cells through Their Behavior as Biophotonic Lenses. *J Biomed Opt* (2016) 21(12):121509. doi:10.1117/1.JBO.21.12.121509
23. Memmolo P, Miccio L, Merola F, Gennari O, Netti PA, Ferraro P. 3D Morphometry of Red Blood Cells by Digital Holography. *Cytometry A* (2014) 85(12):1030–6. doi:10.1002/cyto.a.22570
24. Ikeda T, Popescu G, Dasari RR, Feld MS. Hilbert Phase Microscopy for Investigating Fast Dynamics in Transparent Systems. *Opt Lett* (2005) 30(10):1165–7. doi:10.1364/OL.30.001165

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.

Copyright © 2022 Li, Qin, Yu, Li, Hu, Qu and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

ORIGINAL RESEARCH
published: 19 July 2022
doi: 10.3389/fphy.2022.965095

Three-Dimensional Virtual Optical Clearing With Cycle-Consistent Generative Adversarial Network

Jiajia Chen[1,2,3†], *Zhenhong Du*[1,2†] *and Ke Si*[1,2,3,4*]

[1]State Key Laboratory of Modern Optical Instrumentation, Department of Psychiatry, First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2]College of Optical Science and Engineering, Zhejiang University, Hangzhou, China, [3]Intelligent Optics and Photonics Research Center, Jiaxing Research Institute, Zhejiang University, Jiaxing, China, [4]MOE Frontier Science Center for Brain Science and Brain-Machine Integration, NHC and CAMS Key Laboratory of Medical Neurobiology, School of Brain Science and Brain Medicine, Zhejiang University, Hangzhou, China

High-throughput deep tissue imaging and chemical tissue clearing protocols have brought out great promotion in biological research. However, due to uneven transparency introduced by tissue anisotropy in imperfectly cleared tissues, fluorescence imaging based on direct chemical tissue clearing still encounters great challenges, such as image blurring, low contrast, artifacts and so on. Here we reported a three-dimensional virtual optical clearing method based on unsupervised cycle-consistent generative adversarial network, termed 3D-VoCycleGAN, to digitally improve image quality and tissue transparency of biological samples. We demonstrated the good image deblurring and denoising capability of our method on imperfectly cleared mouse brain and kidney tissues. With 3D-VoCycleGAN prediction, the signal-to-background ratio (SBR) of images in imperfectly cleared brain tissue areas also showed above 40% improvement. Compared to other deconvolution methods, our method could evidently eliminate the tissue opaqueness and restore the image quality of the larger 3D images deep inside the imperfect cleared biological tissues with higher efficiency. And after virtually cleared, the transparency and clearing depth of mouse kidney tissues were increased by up to 30%. To our knowledge, it is the first interdisciplinary application of the CycleGAN deep learning model in the 3D fluorescence imaging and tissue clearing fields, promoting the development of high-throughput volumetric fluorescence imaging and deep learning techniques.

Keywords: optical clearing, deep learning, deep tissue imaging, light-sheet, image processing

OPEN ACCESS

Edited by:
Liwei Liu,
Shenzhen University, China

Reviewed by:
Ming Lei,
Xi'an Jiaotong University, China
Sihua Yang,
South China Normal University, China

****Correspondence:***
Ke Si
kesi@zju.edu.cn

†These authors have contributed equally to this work

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

Received: *09 June 2022*
Accepted: *22 June 2022*
Published: *19 July 2022*

Citation:
Chen J, Du Z and Si K (2022) Three-Dimensional Virtual Optical Clearing With Cycle-Consistent Generative Adversarial Network. Front. Phys. 10:965095. doi: 10.3389/fphy.2022.965095

INTRODUCTION

Fluorescence microscopy has been playing an increasingly indispensable role in depiction of biological microstructures and functions. Up to now, confocal microscopy is still the most extensive and successful commercial fluorescence imaging system [1]. Nevertheless, the tissue anisotropy, the signal attenuation or absorption, the optical aberration of imaging system will all cause severe image blurring and degradation in the practical imaging process, limiting the further development of biological research at micro-scale [2]. On the one hand, the low fluorescence image quality greatly decreases the resolving power and further analysis accuracy of imaging systems for microstructural information. On the other hand, the reduction of fluorescence signal deep inside

biological tissues will influence the imaging depth and imaging speed of thick tissues, restricting the experimental research efficiency for large-scale biological tissues. So far, many researchers have made great efforts to improve the imaging efficiency and image quality from different aspects, including physical, chemical, and digital ways [3–5].

To acquire detailed 3D information physically, a series of advanced optical imaging techniques have been developed for deep tissue imaging in recent years, such as two-photon excitation microscopy (TPEM) [6], fluorescence micro-optical sectioning tomography (fMOST) [7], and light-sheet fluorescence microscopy (LSFM) [8]. Compared with the confocal microscopy, the two-photon absorption effect provides lower background signal level, phototoxicity and photobleaching for biological imaging. Besides, longer wavelength of laser used in TPEF could realize larger penetration depth for fluorescence excitation and detection, improving the 3D imaging capability of fluorescence microscopy. fMOST broke the 3D imaging limitations for brainwide mapping neurite level by using a continuous tissue sectioning microtome and synchronous wide-field detection. And the synchronous tissue sectioning and imaging idea could be also introduced into various conventional imaging systems for high-speed 3D imaging and reconstruction, including serial two-photon tomography [9], automatic serial sectioning polarization sensitive optical coherence tomography [10] and so on. Further, as a rapid, high-resolution imaging technique, LSFM has played an important part in large-scale mesoscopic biological research due to its large field of view and good optical sectioning capability. Especially for millimeter-level biological tissue imaging, LSFM has shown unprecedented imaging speed and throughput, which is at least dozens of times higher than some conventional fluorescence microscopies [11].

Although these microscopic imaging systems have made significant progress in depicting biological microstructures and functions, it is not sufficient for us to improve the image quality and imaging depth only by physical means. It is because the strong scattering and attenuation effect introduced by the biological tissue anisotropy will directly cause severe image degradation and noise, which could not easily be overcome or bypassed by upgrading the optical system. Hence, the chemical tissue clearing techniques were proposed to improve the tissue homogeneity and ensure refractive matching between tissues and surrounding buffers. Especially as a powerful combination with LSFM imaging techniques, various tissue clearing protocols have been developed and modified for larger imaging depth and better imaging quality in 3D tissue imaging [12–14]. For example, CUBIC-series allows whole-brain even whole-body clearing and enables single-cell-resolution visualization and quantification of nucleus and neural activities in centimeter-scale brains [15]. And the previously unknown details and anatomical connections such as non-dividing stem cells near perisinusoidal areas under the fluorescence microscope could also be revealed by using DISCO-series protocols [16]. Besides, an ultrafast optical clearing method (FOCM) was also proposed to clarify 300-um-thick mouse tissue slices in 2 min with low morphological deformation, fluorescent toxicity and easy operation [17]. Nevertheless, in spite of the great clearing effect on tissues of rodent animals, these clearing protocols have not perfectly resolved the compactness and refractoriness of brain tissues, especially the white matter. So far, as the most important partner of high-throughput imaging systems (e.g., LSFM), the development of chemical tissue clearing techniques is still booming for expanding the 3D tissue imaging depth and image quality.

Except for improvements of imaging systems and tissue preparation techniques, another popular approach is image deconvolution. An image restoration process, namely deconvolution, is established for enhancing the tissue details and image quality by typically modeling the image acquisition and degradation process as the summation of image noise and convolution between sample and systematic point spread function [18]. Many classic deconvolution methods, such as Richardson-Lucy deconvolution and Huygens deconvolution have shown great image enhancement performance for different requirements, including resolution improvement, image deblurring and noise suppression [19, 20]. However, inevitably, tissue anisotropy and scattering generally lead to deficiency of some 3D information, getting in the way of accurate acquiring or estimation of systematic point spread function, which is very important for the deconvolution process. Particularly, the fast large-scale tissue imaging process is generally accompanied by unforeseeable uncertainty of point spread function distortion and image degradation. As a kind of emerging state-of-the-art technique, deep learning has gradually shown powerful efficiency and wide feasibility, especially in image super-resolution, image restoration and aberration correction [21–23]. However, almost current deep learning-based image processing methods need a large number of exquisitely prepared paired datasets. Due to the hardware and experimental limitation, it is difficult even impossible to acquire enough high-quality ground-truth in paired datasets for some deep learning models, such as convolutional neural networks and U-Net [24–26]. In recent years, a series of unsupervised deep learning models are proposed to realize feature transformation between two types of data with unpaired datasets [27–29]. For example, CycleGAN model has been widely applied in two-dimensional (2D) medical image processing, which realized good efficiency comparable to supervised deep learning models [30–32]. Nevertheless, CycleGAN was mainly used in non-fluorescent imaging and 2D image processing. The wide application and successful verification in 3D high-throughput fluorescence microscopy has hardly ever been reported before.

Here we report a three-dimensional virtual optical clearing method based on cycle-consistent generative adversarial network, termed 3D-VoCycleGAN, to improve the transparency of imperfect cleared biological tissues and image quality of LSFM images. First of all, we selected the blurred 3D image volumes and clear 3D image volumes from the raw 3D image data with varying transparency and contrast as datasets for further network training. Then we built a CycleGAN deep learning model with two 3D ResUNet-based generators and two 3D PatchGAN-based discriminators to realize fast prediction and transformation from

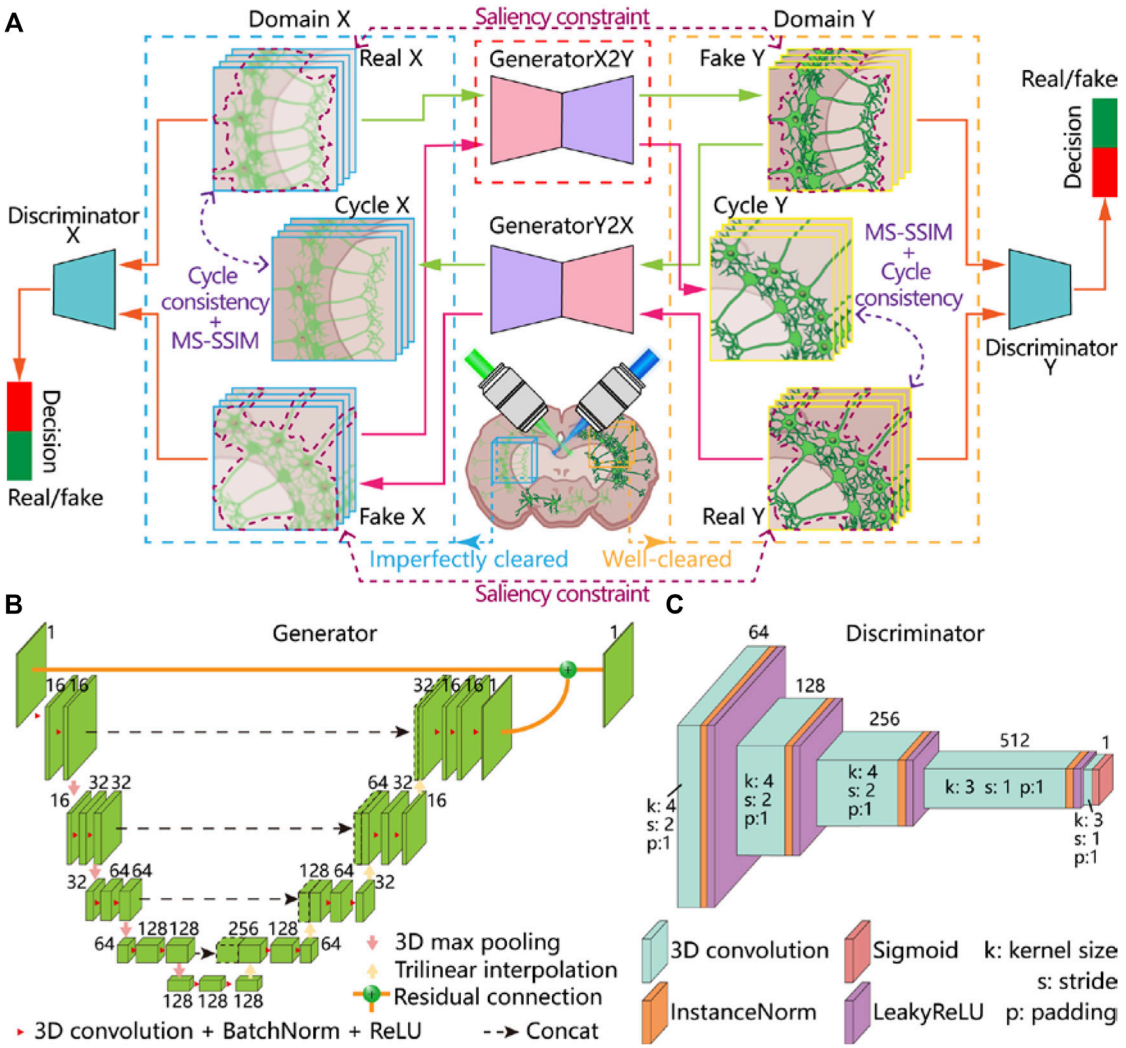


FIGURE 1 | Framework of virtual optical clearing method and 3D-VoCycleGAN architecture. **(A)** The framework and workflow of our virtual optical clearing method. The 3D-VoCycleGAN consists of two generators and two discriminations **(B)** The network structure of generators. 3D ResUNet was used to build the two generators in our network **(C)** The network structure of discriminators. 3D PatchGAN was used to build the two discriminators in our network.

blurred image volumes to clear image volumes. By testing the method on the 3D image data acquired from a custom-built LSFM system, we verified evident improvements with homogeneous digital tissue clearing and good image contrast on imperfect cleared mouse brain tissues and mouse kidney tissues. Besides, compared to other deconvolution methods, our virtual clearing method showed good image restoration and transparency enhancement effect with evident speed advantage, especially for 3D images deep inside biological tissues. To our knowledge, it is the first time that CycleGAN model has been used for enhancing the clearing effect of chemical tissue clearing, and restoring the 3D blurred LSFM images. Our virtual optical clearing method could effectively remedy the insufficiency of chemical tissue clearing and deep tissue imaging techniques, illustrating promising potential in future 3D histology and volumetric fluorescence imaging.

METHODS

The Main Framework of Virtual Optical Clearing

The tissue spatial anisotropy and imperfect tissue clearing effect generally leads to heterogeneous image contrast or image blurring, limiting further image biological structure identification and analysis. We proposed a CycleGAN-based approach to restore image quality with various tissue transparency and contrast in 3D LSFM imaging. The main framework of our virtual optical clearing method is shown in **Figure 1A**. According to different types of tissue properties and clearing effect, we selected several representative 3D image volumes with specific structural information from the raw LSFM data to generate datasets for network training. In the tissues, 3D volumes with high image contrast, low noise and

good transparency were regarded as well-cleared tissue areas. And the cropped 3D image data from the well-cleared tissue areas was performed 3D deconvolution to further suppress background signal out of focus and improve image contrast for generating the target domain image data. At the same time, 3D image volumes with severe blurring and background noise in imperfectly cleared area was selected as the source domain image data. It is noticed that unlike the supervised deep learning-based methods, we need not realize accurate data pre-alignment between the two domains in our method. Hence, various types of biological samples could be used to build datasets for our experiments, including mouse brain tissues and kidney tissues acquired by using our custom-built LSFM system.

In the specific network training process, we defined the source domain and the target domain as domain X and domain Y, respectively. Taking the GPU memory requirements of network training into account, 200 groups of unpaired 3D image volumes with 256 × 256×16 pixel^3 size were randomly cropped from the domain X and domain Y to generate the training datasets. In order to improve the training efficiency and avoid overfitting, image volumes with insufficient foreground were automatically discarded before training [33]. And these image volumes were normalized and then fed into the network for training. To ensure higher generalization capability of our network, during each training process, we performed data augmentation on the input image volumes by introducing a series of random changes, such as flipping, rotation, and so on. As the loss function iteratively minimized, the non-transparent 3D tissue image identified and learned the clearer structural features from the cleared 3D tissue image. After the deep learning network was well optimized, the original large LSFM image data with varying spatial transparency and contrast was sent to the network for implementing fast image prediction and image quality restoration.

3D-VoCycleGAN Architecture

In our 3D-VoCycleGAN, we built the two generators by using ResUNet, which has been proved to have great biological feature extraction capability [34]. The ResUNet used in our generators is a kind of encoder-decoder cascade structure, which is composed of four downsampling blocks and upsampling blocks (**Figure 1B**). In the encoder path, each downsampling block contained a max pooling and two convolution layers, each layer of which comprised 3 × 3×3 kernel followed by a batch-normalization layer and a ReLU activation function. And the encoder encoded the input image stack into multiple feature representations at different levels through max pooling. Then the decoder consisting of four upsampling blocks symmetrical to the encoder was used to decode data information back to the original dimension. Each upsampling block consisted of trilinear interpolation, skip connection, and convolution layers. The skip connection concatenated the high-level features and spatial information between encoder and decoder, thus retaining more details and capturing finer information. Finally, we established a residual connection between the input and the output of the decoder to avoid the gradient vanishing problem and improve the performance of the network [35].

For building the discriminators, we modified the five-layer conventional structure of PatchGAN [31] into a 3D form (**Figure 1C**). In the PatchGAN, conventional operation with a stride of 2 and padding of 1 was used in the first three conventional layers, each layer of which was followed by an instance normalization layer, and a LeakyReLU activation function. And the channel numbers were doubled as the image resolution was halved in the first three conventional each time. And the channel numbers of the five convolutional layers were 64, 128, 256, 512 and 1, respectively. Considering that the size of the input stack was 256 × 256×16, we set the kernel size of the last two convolution layers to 3 × 3×3, making the size of the output patch 32 × 32×2. Finally, the last convolutional layer reduced the channel numbers of the feature map to 1, and a sigmoid activation function was used to normalize the output value into [0, 1].

Our 3D-VoCycleGAN was implemented based on the Pytorch deep learning framework. The learning rate of the Adam optimizer was set to 0.0002 in the first 100 epochs and linearly decayed to 0 in the next 100 epochs. The batch size was set to 1. The overall training and prediction processes based on our method were implemented in a Dell 7,920 workstation equipped with RTX 3090 GPU (24 GB memory).

Loss Function

The deep learning model of 3D-VoCycleGAN contained two generators G_{XY} (from domain X to Y), G_{YX} (from domain Y to X) and two discriminators D_X, D_Y. The two generators G_{XY}, G_{YX} aimed to realize transformation from 3D image volumes of domain X to 3D image volumes of domain Y and inverse image transformation from domain Y to domain X, respectively. In each domain, an image volume could be the output of the generator or come from the original training data. Hence, two discriminators were used to judge which situation the image volumes belong to. In a word, the goal of the generator was synthesizing images and fooling the discriminator, while the discriminator tried to accurately winnow truth from falsehood. The standard GAN loss [36] was given as follows:

$$l_{GAN}(G_{XY}) = \mathbf{E}_{y\sim p_{\text{data}}(y)}[\log D_Y(y)] + \mathbf{E}_{x\sim p_{\text{data}}(x)} [\log(1 - D_Y(G_{XY}(x))], \quad (1)$$

$$l_{GAN}(G_{YX}) = \mathbf{E}_{x\sim p_{\text{data}}(x)}[\log D_X(x)] + \mathbf{E}_{y\sim p_{\text{data}}(y)} [\log(1 - D_X(G_{YX}(y))], \quad (2)$$

where x denotes the blurred stack in domain X, y denotes the clear stack in domain Y. $\mathbb{E}_{x\sim p_{\text{data}}(x)}$, $\mathbb{E}_{y\sim p_{\text{data}}(y)}$ were the expectation operators. However, standard GAN loss usually suffers from mode collapse and vanishing gradients, which severely degrades network performance and stability. Here we used LS-GAN loss [37] as objective function:

$$l_{GAN}(G_{XY}) = \frac{1}{2}\mathbf{E}_{x\sim p_{data}(x)}[(D_Y(G_{XY}(x)) - 1)^2], \quad (3)$$

$$l_{GAN}(D_Y) = \frac{1}{2}\mathbf{E}_{y\sim p_{\text{data}}(y)}[(D_Y(y) - 1)^2] + \frac{1}{2}\mathbf{E}_{x\sim p_{data}(x)} [(D_Y(G_{XY}(x)))^2], \quad (4)$$

$$l_{GAN}(G_{YX}) = \frac{1}{2}\mathbb{E}_{y\sim p_{data}(y)}\left[(D_X(G_{YX}(y)) - 1)^2\right], \quad (5)$$

$$l_{GAN}(D_X) = \frac{1}{2}\mathbb{E}_{x\sim p_{\text{data}}(x)}\left[(D_X(x) - 1)^2\right] + \frac{1}{2}\mathbb{E}_{y\sim p_{data}(y)}\left[(D_X(G_{YX}(y)))^2\right], \quad (6)$$

The original cycle consistency loss calculated the L_1 loss between the original image and reconstructed image. We incorporated a 3D multi-scale structural similarity index metric (MS-SSIM) to construct a detail preserving transformation [38]. The cycle consistency loss was defined as follows:

$$l_{\text{cycle}}(G_{XY}, G_{YX}) = \mathbb{E}_{x\sim p_{\text{data}}(x)}\left[\|G_{YX}(G_{XY}(x)) - x\|_1\right] + \mathbb{E}_{y\sim p_{\text{data}}(y)}\left[\|G_{XY}(G_{YX}(y)) - y\|_1\right] + \eta \cdot l_{MS\text{-}SSIM}, \quad (7)$$

$$l_{MS\text{-}SSIM}(G_{XY}, G_{YX}) = 1 - \text{MS-SSIM}(x, G_{YX}(G_{XY}(x))) + 1 - \text{MS-SSIM}(y, G_{XY}(G_{YX}(y))), \quad (8)$$

where η represents the weight factor for the $l_{MS-SSIM}$ term. And it was set to 0.2 in all experiments. An additional identity loss was used to constrain the generator to build an identity mapping when well-cleared volumes of the target domain were provided.

$$l_{identity}(G_{XY}, G_{YX}) = \mathbb{E}_{x\sim p_{\text{data}}(x)}\left[\|G_{YX}(x) - x\|_1\right] + \mathbb{E}_{y\sim p_{\text{data}}(y)}\left[\|G_{XY}(y) - y\|_1\right], \quad (9)$$

Saliency constraint has been proved to be effective for content preservation on 2D microscopic images [39]. And we added it to our 3D-VoCycleGAN for maintaining fine structures and information. By appropriately setting the threshold value, we could separate the foreground with detailed structure from the 3D samples. Therefore, the two generators focused on the essential structures of foreground at the early training stage and avoided appearance of artifacts in background with the aid of saliency constraint. The saliency constraint loss can be written as:

$$l_{saliency}(G_{XY}, G_{YX}) = \mathbb{E}_{x\sim p_{\text{data}}(x)}\left[\|\mathsf{T}_\alpha(x) - \mathsf{T}_\beta(G_{XY}(x))\|_1\right] + \mathbb{E}_{y\sim p_{\text{data}}(y)}\left[\|\mathsf{T}_\beta(y) - \mathsf{T}_\alpha(G_{YX}(y))\|_1\right], \quad (10)$$

where $\mathsf{T}_\alpha(x) = \text{sigmoid}[100(x-\alpha)]$ and $\mathsf{T}_\beta(x) = \text{sigmoid}[100(x-\beta)]$ are binary functions parameterized by threshold α and β.

The total loss for generators could be expressed by:

$$L_{\text{Generator}} = l_{GAN}(G_{XY}) + l_{GAN}(G_{YX}) + \lambda \cdot l_{identity}(G_{XY}, G_{YX}) + \rho \cdot l_{saliency}(G_{XY}, G_{YX}) + \mu \cdot l_{\text{cycle}}(G_{XY}, G_{YX}), \quad (11)$$

where the parameters λ, μ, ρ denotes the weight factor for the identity loss, saliency constraint loss and cycle loss, respectively. We assigned $\lambda = 5$, $\mu = 10$, respectively. And we set $\rho = 10 \times e^{-0.1\times n}$, where n denotes training epoch number, making the network pay more attention to the transformation of important details and ignoring the interference from the background at the early stage of training.

Sample Preparation and 3D Image Data Acquisition

The sample used in this study contained mouse brain and mouse kidney tissues. For verifying the performance of our method on tissue slices, we prepared Thy1-GFP mouse brains and mouse kidney tissues, which were sectioned into 300-μm and 200-μm thick slices, respectively. The mouse kidney tissue slices were stained with DRAQ5 for labelling the cell nucleus before cleared by FOCM reagents. The FOCM reagents were prepared as 30% (wt/vol) urea (Vetec), 20% (wt/vol) D-sorbitol (Vetec), and 5% (wt/vol) glycerol dissolved in DMSO. When preparing the reagent, urea and D-sorbitol were dissolved in DMSO and stirred at room temperature overnight. After complete dissolution, glycerol was added and stirred further. The reagents should be stored at room temperature and shaken gently before using. Before imaging experiments, the well-stained mouse brain and kidney tissue slices were incubated in FOCM reagents for several minutes.

For verifying the performance of our method on deep tissues with millimeter-thickness, we prepared stereoscopic mouse brain and kidney tissue blocks labelled with Alexa Fluor 647 anti-mouse CD31 antibody (CD31-AF647, BioLegend) by caudal vein injection. The Alexa Fluor 647 anti-mouse CD31 antibody (20 mg) was then diluted in sterile saline (total volume of 150 ml). After the injection, mice were placed in a warm cage for 30 min prior to perfusion. Then mice were rapidly anesthetized with chloral hydrate [5%, wt/vol, 0.1 ml/10 g, intraperitoneal (i.p.)] and transcardially perfused with ice-cold 0.01 M phosphate buffered saline (PBS, Coolaber) and paraformaldehyde (PFA, 4% in PBS wt/vol, Saiguo Biotechnology Co., Ltd). Mouse brains and kidneys were collected and incubated in the same PFA solution at 4°C for 24–48 h for uniform fixation. After fixation, mouse brains and kidneys were washed in 0.01 M PBS at room temperature (20–25°C) for 6–12 h. The mouse brains and kidneys were clarified by the CUBIC-L/R+ protocol [xxx]. CUBIC-L [10 wt% of N-butyldiethanolamine (Vetec) and 10 wt% of Triton X-100 (Sigma) in water] and CUBIC-R+ [45 wt% of antipyrine (Vetec) and 30 wt% of nicotinamide (Vetec) in water, buffered with 0.5% (v/w) N-butyldiethanolamine (pH ~ 10)] was prepared for tissue clearing. Mouse organs and tissues were incubated in CUBIC-L for 7 d at 37°C with gentle shaking followed by PBS washing at room temperature. After PBS clearing, the mouse brains and kidneys was incubated and stored in CUBIC-R+ at room temperature.

For generating 3D datasets and verifying the application of our method, we acquired the experimental data of mouse organs and tissues via a custom-built LSFM system. For exciting fluorescence signal of mouse tissues labeled with Thy1-GFP, DRAQ5, and CD31-AF647, two semiconductor lasers (OBIS 488LS/637LX nm, Coherent) were aligned and expanded by a pair of achromatic lenses with 30 and 250 mm focal length, respectively. Like the classical selective plane illumination microscopy, we used a cylindrical lens with 100 mm focal length and a low-NA objective (Olympus ×4/NA 0.1/WD 18.5 mm) with long working distance to generate a thin illumination sheet. And a mechanical slit (VA100 C/M, Thorlabs) was set to 1 mm for controlling the thickness of the illumination sheet. The fluorescence signal was collected by a

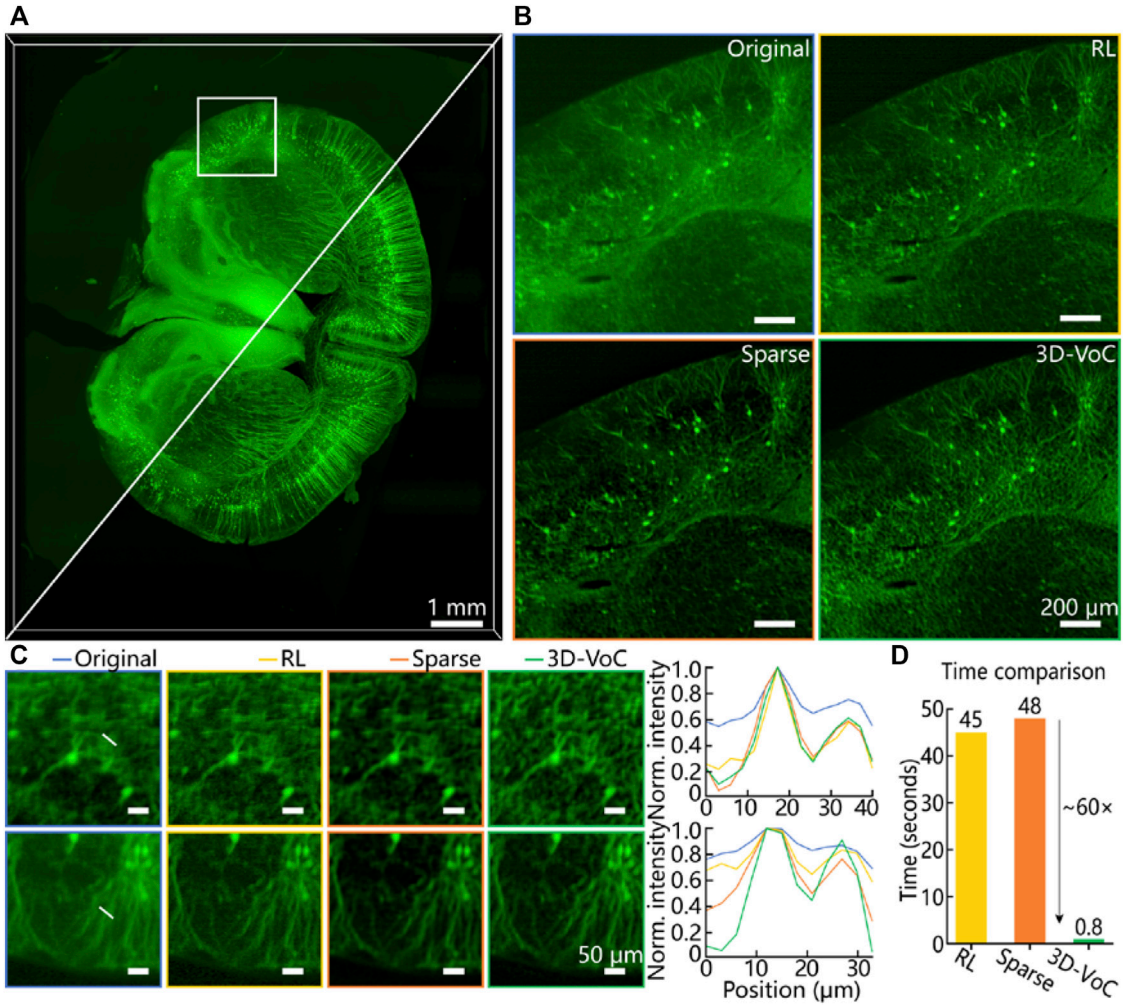


FIGURE 2 | The image enhancement performance of 3D-VoCycleGAN. **(A)** The comparison results before and after using 3D-VoCyleGAN **(B)** The image quality enhancement results with Richardson-Lucy deconvolution, sparse deconvolution, and 3D-VoCycleGAN method. The size of cropped image volume was limited by the GPU memory requirements of Richardson-Lucy deconvolution and sparse deconvolution. **(C)** The quantitative evaluation of the image quality enhancement performance. The plotting curves of two line profiles showed the signal level and SBR of nerve fiber in the brain slice **(D)** The time consumption of different image processing methods for the image volume. RL: Richardson-Lucy deconvolution, Sparse: sparse deconvolution, 3D-VoC: 3D-VoCycleGAN.

tube lens (ITL200, Thorlabs), detection objective (Nikon ×4/NA 0.2/WD 20 mm), a multi-channel emission filter (#87-247, Edmund Optics) and a sCMOS camera (ORCA-Flash 4.0 V3, Hamamatsu). And the biological samples were loaded on a 3D motorized stage (KMT50SE/M-3D, Thorlabs) and performed scanned imaging. In this system, we used the external trigger mode with a synchronous signal to ensure high-frame image acquisition. During the LSFM imaging process, each frame was captured with a constant acquisition interval of 2 μm and total exposure time of 10 ms. The corresponding camera acquisition speed could be up to 100 fps. Due to the rapid imaging advantage, the image acquisition process of a whole mouse brain tissue slice could be finished in 4 min. All images were transformed and stored as 16-bit Multi-TIFF format for post-processing.

RESULTS

Image Enhancement Performance of 3D-VoCycleGAN for Mouse Brain Slices

Due to spatial anisotropy and structural complexity of biological tissues, chemical tissue clearing still could not realize perfect tissue transparency and image contrast. As for a pre-cleared tissue, the clearing extent and spatial image degradation is random across the whole 3D volume. Here we first demonstrated our 3D-VoCycleGAN on Thy1-GFP mouse brain slices to digitally improve the optical clearing extent and image quality. Although the whole image showed the distribution of fluorescence signal in the brain slice, some details such as nerve fiber were still blurred as shown in **Figure 2A**. By using our 3D-VoCycleGAN, the image background was greatly suppressed.

Besides, the blurred nerve fiber in the original image became more distinguishable after enhancement. Further, we cropped a region-of-interest (ROI) from the brain slice to compare the virtual optical clearing efficiency and image enhancing performance with other deconvolution methods. Here we performed sparse deconvolution [40], Richardson-Lucy deconvolution and 3D-VoCycleGAN on a 500 × 500×32 pixel^3 image volume. According to the results in **Figure 2B**, three methods all realized image quality improvement to different extents. Then we plotted the profiles of two lines across the nerve fiber to quantitatively evaluate their differences of performance. Although Richardson-Lucy also enhanced the image resolution to some extent, the image background noise suppression capability was not as good as the other methods. Our 3D-VoCycleGAN and sparse deconvolution showed better image contrast and SBR improvements, which were both increased by above 40% (**Figure 2C**). It is worth noting that although sparse deconvolution showed very powerful image quality enhancing capability comparable to our 3D-VoCycleGAN, the processing time consumption for the image volume with same size is larger than our method (**Figure 2D**). It usually takes a considerable amount of time to finish the image processing by using Richardson-Lucy deconvolution. And sparse deconvolution could shorten the processing time with the aid of GPU acceleration. Nevertheless, due to GPU memory limitation, sparse deconvolution could not realize image processing of large 3D data in one time. Although sparse deconvolution could process the 3D image data volume by volume after splitting the large 3D image data into several image stacks, the final stitched 3D image will show uneven brightness and background since these sub-volumes exist sparsity differences. As for this image volume, the processing time consumption of 3D-VoCycleGAN, sparse deconvolution, and Richardson-Lucy deconvolution were 0.8, 48 and 45s, respectively. Thereinto, we performed 50 iterations of the Richardson-Lucy deconvolution by using the DeconvolutionLab2 plugin in ImageJ/Fiji. Then, with 3D-VoCycleGAN, image processing of the whole brain slice with 3,100 × 3,500×180 pixel^3 volume could be finished in only 141s. Hence, our 3D-VoCycleGAN could realize great image quality enhancement with short time consumption, showing the great image processing efficiency in brain slice imaging.

Information Restoration of Images Deep Inside Kidney Tissue Slices

Tissue slice imaging with chemical optical clearing protocols are widely used in biological research. Although many chemical tissue clearing methods have promoted biological structure and function research, tissue scattering and refractive index mismatching between multiple media in the imaging system usually influence the image quality and microstructure analysis under the slice surface. Our 3D-VoCycleGAN could contribute to the information restoration of images deep inside tissue slices. As shown in **Figure 3A**, labelled cell nucleus of the mouse kidney tissue slice showed the distribution of glomeruli. However, the tissue transparency was not enough to distinguish the detailed nucleus or glomeruli. Especially, due to the uneven thickness and structural anisotropy in axial direction, the clearing reagents could not perfectly make the kidney tissue slice transparent. For example, although the kidney tissue slice could be optical cleared by the clearing reagents, the images deep inside the tissue slice were still blurred due to inevitable light scattering or attenuation. The fluorescence signal was nearly overwhelmed in the strong background signal except the area from the tissue surface to 115-μm depth (**Figure 3B**). By using our 3D-VoCycleGAN, the fluorescence signal in deep tissue could be quickly recovered. The axial images in **Figures 3B,C** showed the profile of glomeruli and tubules across about 150-μm depth, which improved the clearing depth of mouse kidney tissues by up to 30%. Besides, we could also compare the 2D images in different depths. When the imaging depth was 100 μm, the original image could show distinguished cell nucleus with faint noise. Our virtual optical clearing method could improve the image contrast with details maintained as shown in **Figure 3D**. When the imaging depth was 140 μm, the morphology of cell nucleus and glomeruli were severely blurred (**Figure 3E**). With 3D-VoCycleGAN, the final image quality was evidently improved, where some information of nucleus and glomeruli was recovered. Hence, to some extent, we could realize structural information restoration of imperfectly cleared tissue slices in deep depth by virtual optical clearing technique.

Virtual Optical Clearing of Large-Scale 3D Tissues

Except for tissue slice imaging, high-throughput imaging for large-scale 3D tissues also plays an increasingly important part in biological research, especially in digital organ mapping and brain network reconstruction. Here we demonstrated the 3D image enhancement capability for stereoscopic mouse tissue with above 1 × 1 × 1 mm^3 volume by using our virtual clearing method. To further depict the glomeruli distribution or morphology of kidney tissue, we imaged and reconstructed a mouse kidney tissue block with a custom-bulit LSFM system. As shown in **Figure 4A**, the original 3D image volume has evident space-variant opaqueness and structure blurring. In particular, the structural contours of glomeruli were gradually blurred as the axial depth increases, which was marked by the white arrow in **Figure 4A**. The connections between glomeruli and arteries were also very important for supporting some kidney functions. However, in the imperfectly cleared kidney tissue, a branch of the kidney artery was also overwhelmed in the background noise (yellow arrow in **Figure 4A**). By using our virtual optical clearing method, the glomeruli and artery branch were all evidently recovered and distinguished (**Figure 4B**). And the anisotropic tissue transparency was eliminated after virtual optical clearing. From the depth color-coded z-projections, we could see the whole image contrast and fluorescence signal intensity in various depths were enhanced (**Figures 4C,D**).

In order to quantify the virtual optical clearing effect, we selected two 2D images from two different z-depths. As shown in **Figure 4E**, the original 2D image had low image quality, including strong noise, blurring, and structure missing (two blue arrows). After virtual optical clearing, the details and

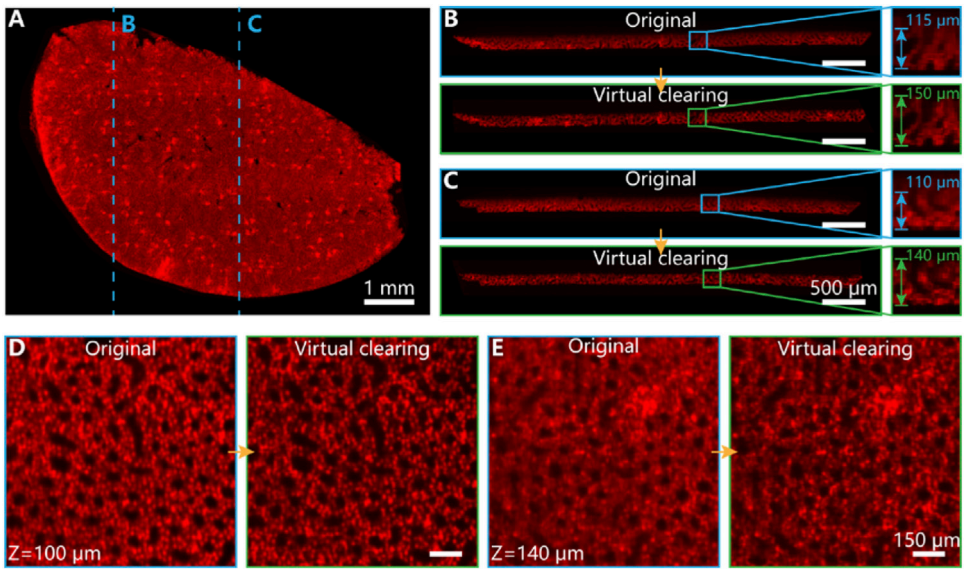


FIGURE 3 | The image information restoration effect in a mouse kidney tissue slice. **(A)** Fluorescence image of a mouse kidney slice labelled with cell nucleus **(B,C)** The comparison of clearing extent between original and fluorescence image processed by 3D-VoCycleGAN. **(D,E)** The image improvements in different imaging depths. The image of z = 100 μm was given in **(D)** and image of z = 140 μm was given in **(E)**.

contours of glomeruli were enhanced greatly, which might be meaningful to the morphological analysis of kidney microstructures. Although the original 2D image in **Figure 4F** had lower noise and blurring extent than the original image in **Figure 4E** because of shallow imaging depth, the vascular walls of two artery branches were still difficult to be distinguished directly (two green arrows). Our virtual optical clearing successfully recovered the details of two vascular walls. Further, we evaluated the fluorescence signal intensity and noise level in the two depths by plotting four lines. As shown in **Figure 4G**, images processed by our method had more distinguishable details about glomeruli and artery, especially the contour of glomeruli and vascular walls. Besides, our method improved the SBR of images by above 25%. Meanwhile, we tested our method on 3D images of brain vessels, which showed complex structures and dense distribution in the mouse brain. For deep brain tissue 3D imaging, tissue anisotropy and imperfect tissue clearing will introduce strong scattering and attenuation of fluorescence signal, resulting in vessel blurring and artifacts as shown in **Figure 4H**. According to the comparison results of depth color-coded z-projections, we could see the definition and sharpness of vessels were greatly improved in various depths by using our virtual optical clearing method. As shown in **Figures 4I,J**, the vessel artifacts were suppressed after virtual clearing. Similar to the results of kidney tissue, the original vessel contours existed severe blurring. Our virtual optical clearing method could effectively restore the information and improve the image quality, indicating its powerful image enhancement capability in 3D fluorescence imaging.

DISCUSSION

High-throughput 3D fluorescence imaging and tissue clearing techniques are playing an increasingly important role in biological research. However, due to the tissue anisotropy and structural complexity, chemical tissue clearing techniques sometimes could not imperfectly clear the whole tissue, resulting in image quality degradation, such as image blurring, background noise, artifacts and so on. A series of methods including physical, chemical and digital ways have been proposed to improve the fluorescence image quality in recent years. Here, we presented an unsupervised deep learning-based image processing method, called 3D-VoCycleGAN to realize further virtual optical clearing of imperfectly cleared tissues. By making full use of the tissue anisotropy and space-variant clearing extent, we built a virtual optical clearing method to enhance the clearing effect of chemical tissue clearing techniques. In our virtual optical clearing, we established a CycleGAN architecture which consists of two pairs of 3D image generators and discriminators to realize the transformation and evaluation from low-quality LSFM images to high-quality LSFM images. Needless of accurate data pre-alignment between the source domain and target domain, our method need not accurate data pre-alignment between the source domain and target domain, which greatly improves the image processing efficiency. Compared with other image enhancement methods, our method showed more powerful image enhancement effect and faster processing speed. With the 3D-VoCycleGAN, we could restore more detailed mouse tissue structural information with high SBR and image contrast, such as distinguished nerve fibers, somas, glomeruli, and vessels.

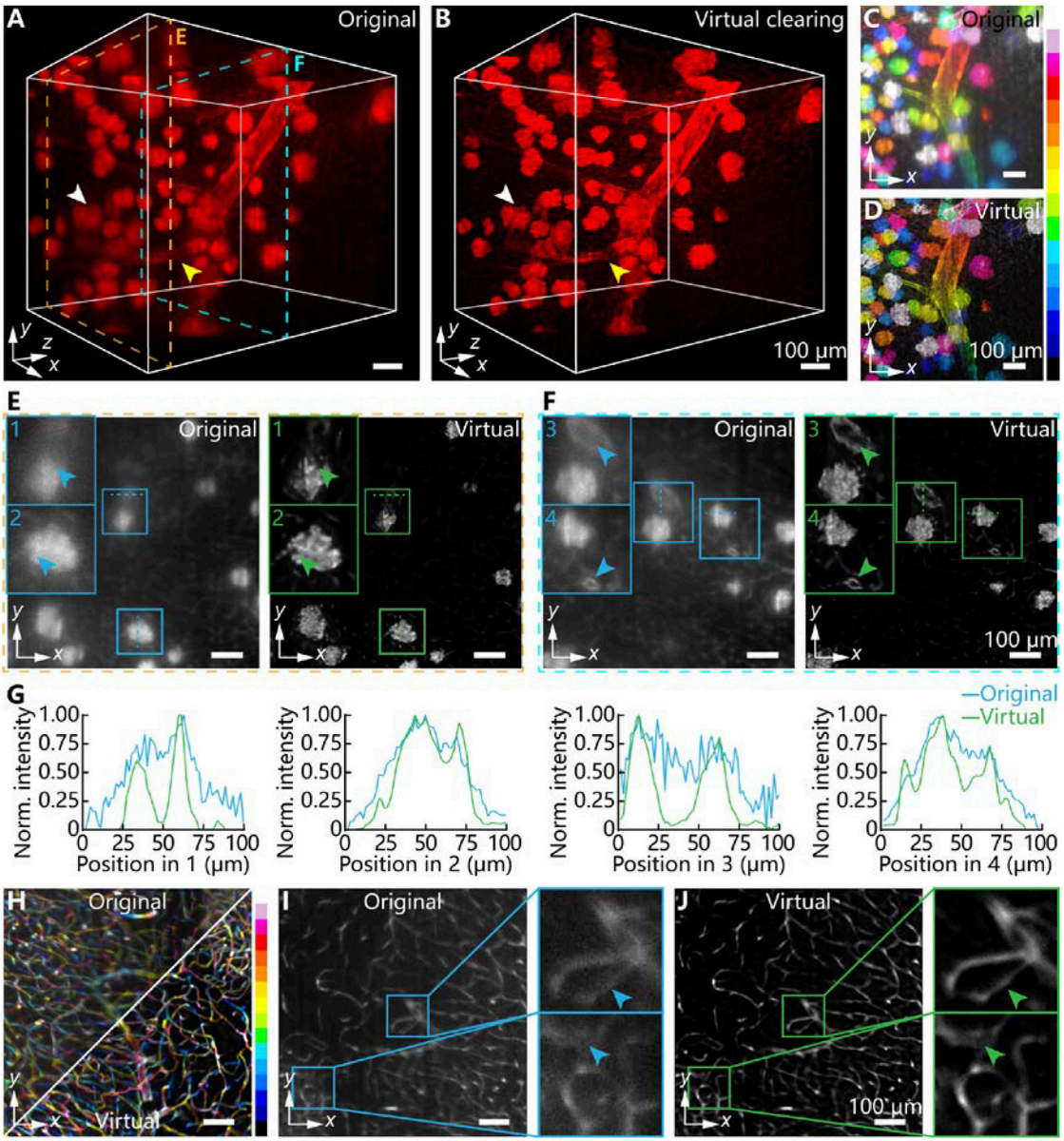


FIGURE 4 | The virtual optical clearing effect for millimeter-thickness mouse tissues. **(A)** 3D fluorescence image of imperfectly cleared mouse kidney tissue. The two dashed boxes represents 2D sections from two imaging depths **(B)** 3D fluorescence image of virtually optically cleared mouse kidney tissue by using 3D-VoCycleGAN. The white and yellow arrows represents the glomerulus and artery branch, respectively. **(C,D)** The depth color-coded image of image volume in **(A)** and **(B)**, respectively. **(E)** The comparison results between original image and virtually cleared image for shallow depth shown in **(A)**. The magnifications of insets were ×2. **(F)** The comparison results between original image and virtually cleared image for deep depth shown in **(A)**. The magnifications of insets were ×1.5. The blue and green arrows in **(E)** and **(F)** represented the structure details of glomeruli and artery branch, respectively. **(G)** The quantitative evaluation of tissue structural details. The plotting curves of four groups of line profiles represented the glomeruli and artery branch **(H)** The comparison results between original and virtually cleared mouse brain vessels by using depth color-coded z-projections. **(I)** The original 2D image showing blurred brain vessels **(J)** The 2D image corresponding to image **(I)** after virtual optical clearing. The blue and green arrows represented the brain vessel details. The magnifications of insets were ×3.

Particularly, it is the first time that the CycleGAN deep learning model has been used for enhancing the clearing effect of chemical tissue clearing, and restoring the 3D blurred LSFM images. Furthermore, except for LSFM, our 3D-VoCycleGAN could also be used to process 3D images captured by other 3D fluorescence imaging systems such as confocal, two-photon microscopy and so on. From the aspect of deep learning network architecture, the 3D operations such as 3D max pooling and 3D convolution in our 3D-VoCycleGAN are not only limited to specific

fluorescence imaging systems. From the aspect of dataset preparation, the sparse deconvolution method which we used to pre-process the 3D cleared tissue images was proposed to improve the image quality of structured illumination microscopy and proved to be available in various 3D fluorescence imaging systems such as confocal, two-photon, expansion microscopy and so on [40]. Hence, the whole image processing flow of our 3D-VoCycleGAN could be transferred to various 3D fluorescence imaging systems by using related datasets. In summary, our study promoted the combination and application of digital image processing, chemical tissue clearing and 3D fluorescence imaging techniques, showing the promising development of interdisciplinary technology in future high-throughput 3D biomedical imaging and biological research.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

All mouse experiments followed the guidelines for the Care and Use of Laboratory Animals of Zhejiang University approved by the Committee of Laboratory Animal Center of Zhejiang University.

AUTHOR CONTRIBUTIONS

JC and ZD conceived the idea. KS supervised the project. JC and ZD performed the experiments and image processing. JC, ZD, and KS wrote the paper. All the authors contributed to the discussion on the results for this manuscript.

FUNDING

This work was supported in part by the Key R&D Program of Zhejiang Province (2021C03001), National Natural Science Foundation of China (61735016), CAMS Innovation Fund for Medical Sciences (2019-I2M-5-057), Fundamental Research Funds for the Central Universities, Alibaba Cloud.

REFERENCES

1. Wilson T, Sheppard C. *Theory and Practice of Scanning Optical Microscopy*. London: Academic Press (1984).
2. Feuchtinger A, Walch A, Dobosz M. Deep Tissue Imaging: A Review from a Preclinical Cancer Research Perspective. *Histochem Cel Biol* (2016) 146: 781–806. doi:10.1007/s00418-016-1495-7
3. Ntziachristos V. Going Deeper Than Microscopy: The Optical Imaging Frontier in Biology. *Nat Methods* (2010) 7:603–14. doi:10.1038/nmeth.1483
4. Yu H, Rahim NAA. *Imaging in Cellular and Tissue Engineering*. 1st ed. Boca Raton: CRC Press (2013).
5. Yoon S, Kim M, Jang M, Choi Y, Choi W, Kang S, et al. Deep Optical Imaging within Complex Scattering media. *Nat Rev Phys* (2020) 2:141–58. doi:10.1038/s42254-019-0143-2
6. Helmchen F, Denk W. Deep Tissue Two-Photon Microscopy. *Nat Methods* (2005) 2:932–40. doi:10.1038/nmeth818
7. Li A, Gong H, Zhang B, Wang Q, Yan C, Wu J, et al. Micro-Optical Sectioning Tomography to Obtain a High-Resolution Atlas of the Mouse Brain. *Science* (2010) 330:1404–8. doi:10.1126/science.1191776
8. Dodt H-U, Leischner U, Schierloh A, Jährling N, Mauch CP, Deininger K, et al. Ultramicroscopy: Three-Dimensional Visualization of Neuronal Networks in the Whole Mouse Brain. *Nat Methods* (2007) 4:331–6. doi:10.1038/nmeth1036
9. Economo MN, Clack NG, Lavis LD, Gerfen CR, Svoboda K, Myers EW, et al. A Platform for Brain-wide Imaging and Reconstruction of Individual Neurons. *eLife* (2016) 5:e10566. doi:10.7554/eLife.10566
10. Wang H, Magnain C, Wang R, Dubb J, Varjabedian A, Tirrell LS, et al. as-PSOCT: Volumetric Microscopic Imaging of Human Brain Architecture and Connectivity. *NeuroImage* (2018) 165:56–68. doi:10.1016/j.neuroimage.2017.10.012
11. Stelzer EHK, Strobl F, Chang B-J, Preusser F, Preibisch S, McDole K, et al. Light Sheet Fluorescence Microscopy. *Nat Rev Methods Primers* (2021) 1:73. doi:10.1038/s43586-021-00069-4
12. Richardson DS, Guan W, Matsumoto K, Pan C, Chung K, Ertürk A, et al. Tissue Clearing. *Nat Rev Methods Primers* (2021) 1:84. doi:10.1038/s43586-021-00080-9
13. Brenna C, Simioni C, Varano G, Conti I, Costanzi E, Melloni M, et al. Optical Tissue Clearing Associated with 3D Imaging: Application in Preclinical and Clinical Studies. *Histochem Cel Biol* (2022) 157:497–511. doi:10.1007/s00418-022-02081-5
14. Yu T, Zhu J, Li D, Zhu D. Physical and Chemical Mechanisms of Tissue Optical Clearing. *iScience* (2021) 24:102178. doi:10.1016/j.isci.2021.102178
15. Matsumoto K, Mitani TT, Horiguchi SA, Kaneshiro J, Murakami TC, Mano T, et al. Advanced CUBIC Tissue Clearing for Whole-Organ Cell Profiling. *Nat Protoc* (2019) 14:3506–37. doi:10.1038/s41596-019-0240-9
16. Molbay M, Kolabas ZI, Todorov MI, Ohn T-L, Ertürk A. A Guidebook for DISCO Tissue Clearing. *Mol Syst Biol* (2021) 17:e9807. doi:10.15252/msb.20209807
17. Zhu X, Huang L, Zheng Y, Song Y, Xu Q, Wang J, et al. Ultrafast Optical Clearing Method for Three-Dimensional Imaging with Cellular Resolution. *Proc Natl Acad Sci U.S.A* (2019) 116:11480–9. doi:10.1073/pnas.1819583116
18. Cannell MB, McMorland A, Soeller C. Image Enhancement by Deconvolution. In: JB Pawley, editor. *Handbook of Biological Confocal Microscopy*. Boston, MA: Springer US (2006). p. 488–500. doi:10.1007/978-0-387-45524-2_25
19. Richardson WH. Bayesian-Based Iterative Method of Image Restoration*. *J Opt Soc Am* (1972) 62:55–9. doi:10.1364/josa.62.000055
20. Lucy LB. An Iterative Technique for the Rectification of Observed Distributions. *Astronomical J* (1974) 79:745. doi:10.1086/111605
21. Belthangady C, Royer LA. Applications, Promises, and Pitfalls of Deep Learning for Fluorescence Image Reconstruction. *Nat Methods* (2019) 16: 1215–25. doi:10.1038/s41592-019-0458-z
22. Wang H, Rivenson Y, Jin Y, Wei Z, Gao R, Günaydın H, et al. Deep Learning Enables Cross-Modality Super-resolution in Fluorescence Microscopy. *Nat Methods* (2019) 16:103–10. doi:10.1038/s41592-018-0239-0
23. Zhang B, Zhu J, Si K, Gong W. Deep Learning Assisted Zonal Adaptive Aberration Correction. *Front Phys* (2021) 8. doi:10.3389/fphy.2020.621966
24. Schmidhuber J. Deep Learning in Neural Networks: An Overview. *Neural Networks* (2015) 61:85–117. doi:10.1016/j.neunet.2014.09.003
25. Jin KH, McCann MT, Froustey E, Unser M. Deep Convolutional Neural Network for Inverse Problems in Imaging. *IEEE Trans Image Process* (2017) 26:4509–22. doi:10.1109/tip.2017.2713099
26. Salim UT, Ali F, Dawwd SA, System D. U-net Convolutional Networks Performance Based on Software-Hardware Cooperation Parameters: A Review. *Int J Comput Digital Syst* (2021) 11. doi:10.12785/ijcds/110180
27. Karhunen J, Raiko T, Cho K. Unsupervised Deep Learning. In: E Bingham, S Kaski, J Laaksonen, J Lampinen, editors. *Advances in Independent*

Component Analysis and Learning Machines. Academic Press (2015). p. 125–42. doi:10.1016/b978-0-12-802806-3.00007-5
28. Li Z, Zhang T, Wan P, Zhang D. SEGAN: Structure-Enhanced Generative Adversarial Network for Compressed Sensing MRI Reconstruction. *Proceedings of the AAAI Conference on Artificial Intelligence* (2019) 33: 1012–9. doi:10.1609/aaai.v33i01.33011012
29. Lv J, Wang C, Yang G. PIC-GAN: A Parallel Imaging Coupled Generative Adversarial Network for Accelerated Multi-Channel MRI Reconstruction. *Diagnostics* (2021) 11:61. doi:10.3390/diagnostics11010061
30. Quan TM, Nguyen-Duc T, Jeong W-K. Compressed Sensing MRI Reconstruction Using a Generative Adversarial Network with a Cyclic Loss. *IEEE Trans Med Imaging* (2018) 37:1488–97. doi:10.1109/tmi.2018.2820120
31. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired Image-To-Image Translation Using Cycle-Consistent Adversarial Networks. In: IEEE International Conference on Computer Vision (ICCV); 22-29 October, 2017; Venice, Italy (2017). p. 2242–51. doi:10.1109/ICCV.2017.244
32. Li M, Huang H, Ma L, Liu W, Zhang T, Jiang Y. Unsupervised Image-To-Image Translation with Stacked Cycle-Consistent Adversarial Networks. In: Proceedings of the European conference on computer vision (ECCV); 8-14 September, 2018; Munich, Germany (2018). p. 184–99. doi:10.1007/978-3-030-01240-3_12
33. Chen J, Sasaki H, Lai H, Su Y, Liu J, Wu Y, et al. Three-dimensional Residual Channel Attention Networks Denoise and Sharpen Fluorescence Microscopy Image Volumes. *Nat Methods* (2021) 18:678–87. doi:10.1038/s41592-021-01155-x
34. Weigert M, Schmidt U, Boothe T, Müller A, Dibrov A, Jain A, et al. Content-aware Image Restoration: Pushing the Limits of Fluorescence Microscopy. *Nat Methods* (2018) 15:1090–7. doi:10.1038/s41592-018-0216-7
35. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 27-30 June 2016; Las Vegas, NV, USA (2016). p. 770–8. doi:10.1109/cvpr.2016.90
36. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets. *Adv Neural Inf Process Syst* (2014) 2:2672–80. doi:10.5555/2969033.2969125
37. Mao X, Li Q, Xie H, Lau RY, Wang Z, Paul Smolley S. Least Squares Generative Adversarial Networks. In: Proceedings of the IEEE international conference on computer vision; 22-29 October, 2017; Venice, Italy (2017). 2794–802. doi:10.1109/iccv.2017.304
38. Wang Z, Simoncelli EP, Bovik AC. Multiscale Structural Similarity for Image Quality Assessment. In: The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers; November 09-12, 2003; Pacific Grove,CA (2003). p. 1398–402. doi:10.1109/ACSSC.2003.1292216
39. Li X, Zhang G, Qiao H, Bao F, Deng Y, Wu J, et al. Unsupervised Content-Preserving Transformation for Optical Microscopy. *Light Sci Appl* (2021) 10: 44. doi:10.1038/s41377-021-00484-y
40. Zhao W, Zhao S, Li L, Huang X, Xing S, Zhang Y, et al. Sparse Deconvolution Improves the Resolution of Live-Cell Super-resolution Fluorescence Microscopy. *Nat Biotechnol* (2021) 40:606–17. doi:10.1038/s41587-021-01092-2

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.

Copyright © 2022 Chen, Du and Si. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 30 August 2022
DOI 10.3389/fphy.2022.969683

Check for updates

OPEN ACCESS

EDITED BY
Minbiao Ji,
Fudan University, China

REVIEWED BY
Yu Zhao,
Technical University of Munich, Germany
Jiancheng Yang,
Swiss Federal Institute of Technology Lausanne, Switzerland

*CORRESPONDENCE
Juan Ye,
yejuan@zju.edu.cn
Shuai Wang,
shuaiwang.tai@gmail.com

†The first two authors contributed equally to this work.

SPECIALTY SECTION
This article was submitted to Optics and Photonics, a section of the journal Frontiers in Physics

RECEIVED 15 June 2022
ACCEPTED 15 August 2022
PUBLISHED 30 August 2022

CITATION
Sun Y, Li J, Xu P, Chen P, Wang Y, Hu S, Jia G, Wang S and Ye J (2022), Automatic quantifying and monitoring follow-ups for implantable collamer lens implantation using AS-OCT images. *Front. Phys.* 10:969683. doi: 10.3389/fphy.2022.969683

COPYRIGHT
© 2022 Sun, Li, Xu, Chen, Wang, Hu, Jia, Wang and Ye. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Automatic quantifying and monitoring follow-ups for implantable collamer lens implantation using AS-OCT images

Yiming Sun[1†], Jinhao Li[2†], Peifang Xu[1], Pengjie Chen[1], Yaqi Wang[3], Shaodan Hu[1], Gangyong Jia[4], Shuai Wang[2*] and Juan Ye[1*]

[1]Department of Ophthalmology, The Second Affiliated Hospital of Zhejiang University, School of Medicine, Hangzhou, Zhejiang, China, [2]Intellegent Information Processing Research Institute, School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai, Shandong, China, [3]College of Media Engineering, Communication University of Zhejiang, Hangzhou, China, [4]College of Computer Science, Hangzhou Dianzi University, Hangzhou, China

Purpose: To develop a deep learning method to automatically monitor the implantable collamer lens (ICL) position and quantify subtle alterations in the anterior chamber using anterior segment optical coherence tomography (AS-OCT) images for high myopia patients with ICL implantation.

Methods: In this study, 798 AS-OCT images of 203 patients undergoing ICL implantation at our eye center from April 2017 to June 2021 were involved. A deep learning system was developed to first isolate the corneoscleral, ICL, and lens, and then quantify clinical important parameters in AS-OCT images (central corneal thickness, anterior chamber depth, and lens vault).

Results: The deep learning system was able to accurately isolate the corneoscleral, ICL, and lens with the Dice coefficient ranging from 0.911 to 0.960, and all the F1 scores >0.900. The relative error between automated measurements and the ground truth for 95% (188 images out of 198) of LVs was within 10%. Intraclass correlation coefficients (ICCs) of the machine-ground truth measurements ranged from 0.928 to 0.995. The deep learning method also showed better repeatability than human graders.

Conclusion: The deep learning method provides reliable detection and quantification of AS-OCT scans for postoperative ICL implantation, which can simplify and optimize the management of clinical outcomes of ICL implantations. Also, this is a step towards an objective measurement of the postoperative vault, making the data more comparable and repeatable to each other.

KEYWORDS

ICL (implantable collamer lens), anterior segment optical coherence tomography (AS-OCT), follow-up, high myopia, vault, deep learning, artificial neural network

Introduction

High myopia has become a major public health issue regarding its increasing prevalence around the world, with 10% of the world's population estimated to be affected by 2050 [1, 2]. Nevertheless, for the correction of high myopia, current mainstream laser-assisted refractive surgery can be risky due to the thinning of corneal as well as structural alterations in corneal biomolecules [3]. In recent years, the phakic intraocular lens has been widely accepted as an option for high myopia patients with its wide refractive correction range and preservation of accommodation. One of the most worldwide-used phakic intraocular lens types is the posterior chamber phakic intraocular lens (EVO ICL; STAAR Surgical), which involves placing an intraocular lens inside the eye without manipulating the lens itself [4, 5]. Since the ICL is implanted in the posterior chamber, it is crucial to monitor physiological changes in the eye that may lead to adverse postoperative events. For example, the inappropriate distance between the posterior ICL surface and the anterior crystalline lens (lens vault, LV) can lead to the risk of specific complications, such as anterior subcapsular (ASC) cataracts, and considerable endothelial cell loss [5, 6]. Therefore, the management of the postoperative follow-ups is essential to the long-term success of ICL implantation.

The development of the anterior segment optical coherence tomography (AS-OCT) enables the acquisition and visualization of high-resolution images of the anterior segment structures [7, 8]. With its non-invasive character, the device has been widely used in post-operative follow-up for ICL implantation. Nevertheless, current technology typically requires manual identification and measurement of the structures, which would not be clinically viable to manually label each parameter individually in crowded ophthalmology clinics. Hence, an objective method is required to automatically identify and measure each scan.

Deep learning, a subfield of artificial intelligence (AI), has proven to be effective for automatically analyzing ocular images, including AS-OCT images [9–14]. However, there is no attempt to automatically analyze AS-OCT images following ICL implantations in patients, to whom an appropriate method to manage the follow-ups can prevent major postoperative complications. Herein, this study aims to develop a fully automatic method based on deep learning to monitor the ICL position and identify subtle alterations in the anterior chamber for patients receiving ICL surgery, which could promptly evaluate postoperative risks and discover adverse events.

Methods

Subjects

This work included 203 patients undergoing posterior chamber phakic intraocular lens (EVO ICL; STAAR Surgical) implantation between April 2017 and June 2021 at the Eye Center, the Second Affiliated Hospital of Zhejiang University, College of Medicine, China. The surgeries were performed by senior surgeons. Patients with a history of cataracts, glaucoma, uveitis, or ocular surgery that could affect structures in AS-OCT were excluded. The postoperative scans were obtained from the swept-source Casia SS-1000 AS-OCT (Tomey Corporation, Nagoya, Japan).

The Ethics Committee of the Second Affiliated Hospital of Zhejiang University, College of Medicine, approved this study. All methods adhered to the tenets of the Declaration of Helsinki.

Deep learning system development

To fully automatically obtain the values in the AS-OCT images following ICL implantation in clinical practices, the deep learning system consists of two approaches: the detection approach and the quantification approach (Figure 1). We developed the system using 598 images (75%) for training, and the remaining 200 images for testing.

The detection approach involved automated recognition of the corneoscleral, ICL, natural lens, and angle recess points (Figure 1A): Each AS-OCT image was manually segmented into 5 parts (the corneosclera, the ICL, the natural lens, and the angle recess points for both left and right) to guide the training of an improved U-Net network. The U-shape network was composed of an encoder, decoder, and skip connection. The encoder was made up of four downsampling blocks, and each downsampling block consisted of two CBRs followed by a pooling layer, while CBR is referred to as conv 3×3 layer + batch normalization layer + ReLU layer. The decoder was composed of four upsampling blocks, each containing an upsampling layer followed by two CBRs. Skip connection was employed between each downsampling block and upsampling block at the same level to copy the features of the encoder into the decoder for feature fusion. It is worth noting that two CBRs were used for communication between the last downsampling block and the first upsampling block, instead of using a skip connection. The last block of the decoder used two conv 1×1+softmax layers to output a four-channel region segmentation map and a three-channel angle recess points positioning region map respectively. The purpose of

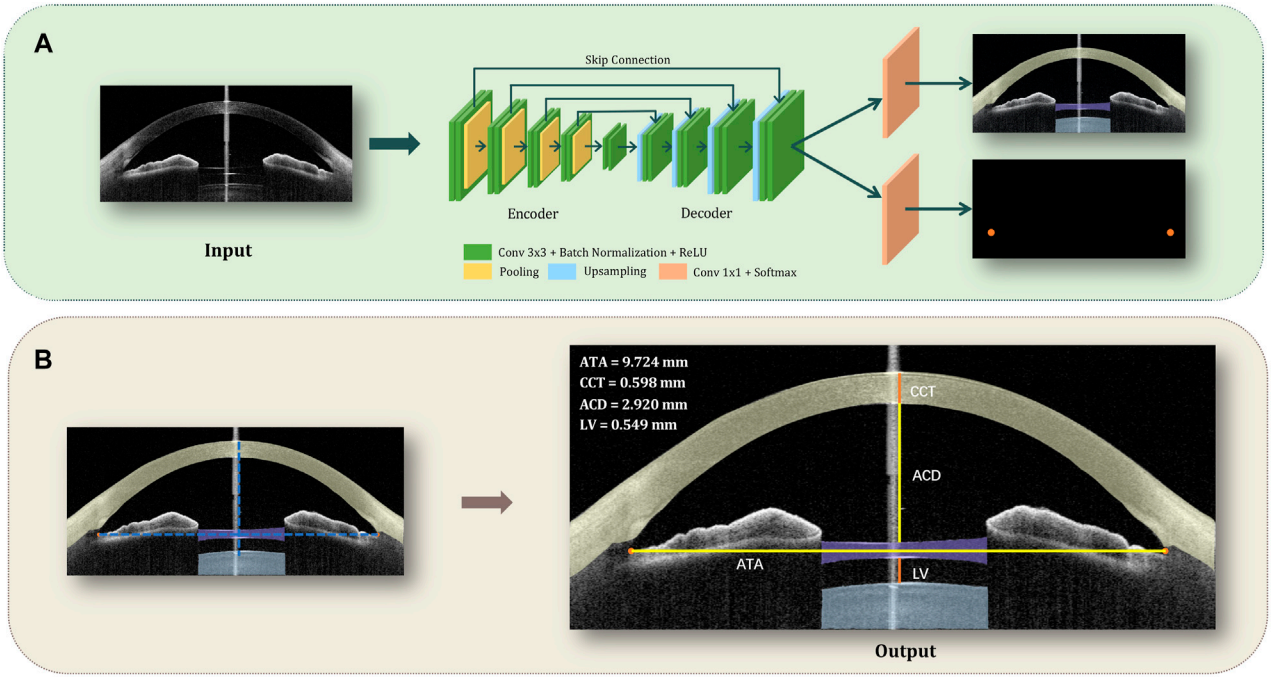


FIGURE 1
Flowchart of the deep learning system, which is able to monitor the ICL position and recognize subtle alterations in the anterior chamber using AS-OCT images for patients with ICL. **(A)**. The detection approach: each AS-OCT image was manually segmented into 5 parts (the corneosclera, ICL, lens, and angle recess points for both left and right) and put into an improved U-Net architecture for the training process. Then, the test images went through the detection approach to generate the segmentation map. **(B)**. The quantification approach: the centroid of the maximum connected area of the angle recess area was calculated and defined as the final coordinate of the points. Then we connected the left and the right angle recess, and make the perpendicular bisector of ATA, which represents the line of axis oculi. Afterward, the intersections of the contour and the axis oculi were connected to calculate CCT, ACD, and LV respectively.

segmenting the angle recess points was to help create the line of axis oculi in the quantification approach. And to improve the segmentation robustness of the angle recess points, regions with a radius of 20 pixels centered on the left and right angle recess points are used for segmentation, instead of using individual points.

Then, the quantification approach automatically obtained central corneal thickness (CCT), anterior chamber depth (ACD), and LV (Figure 1B): To automatically obtain the values, we calculated the centroid of the maximum connected area of the angle recess area, which was the final location coordinate of the point. Then we connected the left and right angle recess points (angle recess to angle recess, ATA), and make the perpendicular bisector of ATA, which represents the line of axis oculi. Afterward, the intersections of the contour and the axis oculi were connected to calculate CCT, ACD, and LV respectively.

For network implementation, we used the Pytorch platform with an Nvidia GeForce RTX 3090 GPU. During model optimization, the number of training epochs was set to 150, and the batch size was set to 1. Dice loss was implemented with a learning rate of 0.00003. We applied an RMSprop optimizer with a weight decay of 1×10^{-8}, and a momentum of 0.9. In terms of the learning scheduler, we used the StepLR scheduler with the period of learning rate decay of 10, and the multiplicative factor of learning rate decay was set to 0.5.

Performance of the deep learning system

To assess the performance of the system, we evaluated the segmentation performance and the quantification performance respectively.

The segmentation performances were assessed using the Dice coefficient, which indicates the similarity between the manual and automated segmentation. Besides, we used the error of X-coordinate, Y-coordinate and absolute values to assess the performance of the two angle recess points segmentation, and precision, recall, F1 score and mean IoU to evaluate the performance of the corneosclera, ICL, and the lens segmentation. For comparison, the ResNet-18, state-of-the-art ReLayNet [15], and DeepLabel V3+ (commonly used for OCT segmentation) [16] models were also implemented.

To assess the accuracy, reproducibility, and repeatability of the measurement, we performed various tests involving the

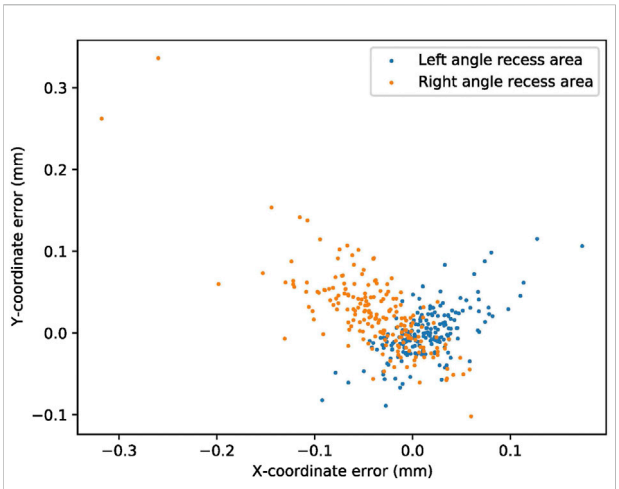


FIGURE 2
The segmentation performance of the angle recess area. The blue and orange dots represent the X-coordinate and Y-coordinate errors of the angle recess area, respectively.

following measurements: G-the ground truth, D-the deep learning system, A-expert s, B-expert c. The ground truth measurements were obtained by manually labeling the edges of the corneosclera, ICL, natural lens, and angle recess points in the test sets and connecting the intersections of the contour and the axis oculi to calculate CCT, ACD, and LV respectively. The relative error, defined as the ratio of the absolute error (D vs. G) and the ground truth value, was used to evaluate the accuracy of the measurements. Bland-Altman plots were used to visualize the distribution of discrepancy between the measurements of ground truth and the deep learning system. The intraclass correlation coefficient (ICC) was used to indicate the degree of agreement and correlation between individual measurements (G vs. D; G vs. A; G vs. B). ICCs of 0.41–0.60, 0.60–0.80, and 0.80–1.00 were taken as moderate, substantial, and excellent agreement, respectively [17]. The root mean square (RMS) difference was calculated to assess human-ground truth differences and machine-ground truth differences, as well as the repeatability of different methods.

Results

Patient characteristics

In total, 798 AS-OCT images from 203 patients (406 eyes) collected from April 2017 to June 2021 were included for analysis after 5 images were excluded (due to poor quality). We used 598 images from 102 patients (204 eyes) for training and validation (training:validation = 7:1) and 200 images from 101 patients (202 eyes) for testing. In the test set, 2 test images were further excluded in the measurement step due to invalid segmentation. The patients were in stable recovery with a mean follow-up period of 130 days. The mean and standard deviation of participant age was *28.6 ± 6.4* years (ranging from 19 to 52 years). There were 67 (33.0%) male participants and 136 (67.0%) females. The flowchart of the deep learning system and an example of the performance are shown in Figure 1.

Segmentation performance

To create the line of axis oculi in the quantification approach, we segmented the angle recess area in the first step. The segmentation performance of the angle recess area was evaluated using the error of X-coordinate, Y-coordinate, absolute value, and Dice coefficient. The system was able to locate the angle recess area accurately. The mean-variance of the left angle recess area was $0.013 \pm 0.034\,mm$ for X-coordinate and$0.001 \pm 0.031\,mm$ for Y-coordinate, and $0.037 \pm 0.030\,mm$ for absolute values (Figure 2). The Dice score was 0.865. The mean-variance of the right angle recess area was $-0.038 \pm 0.050\,mm$ for X-coordinate, $-0.023 \pm 0.049\,mm$ for Y-coordinate, and $0.063 \pm 0.054\,mm$ for absolute values (Figure 2). The Dice score was 0.865 for the left angle recess and 0.788 for the right. Compared with the traditional ResNet-19 model, our method exhibited a much higher performance (Supplementary Table S1).

To obtain the edges of the corneoscleral, ICL, and natural lens, these structures were also separated in the first step. Table 1 lists the segmentation performance of the corneosclera, ICL, and natural lens (using the Dice score, mean IoU, precision, recall, and F1 score), which indicated that the network possessed the ability to accurately identify the structures. The Dice score of these structures ranged from 0.911 to 0.960, and the mean IoU ranged from 0.868 to 0.923. The precision, recall, and F1 score of these structures ranged from 0.945 to 0.971, 0.894 to 0.976, and 0.926 to 0.960 respectively. Compared with previous models, the numerical results showed that our U-Net-based method outperformed the Deep Label V3+ and ReLayNet models (Supplementary Table S2), especially in recognizing the corneoscleral.

Measurement performance

Based on the performance of the segmentation step, we developed an automatic method to quantify these essential anterior segment parameters (ATA, CCT, ACD, and LV). Table 2 lists the outcomes of the automated measurements and ground truth.

We evaluated the accuracy of automated measurements using relative errors between the output measurements and the ground truth, as shown in Figure 3. The relative errors

TABLE 1 The segmentation performance of the structures in AS-OCT following ICL implantation.

	Dice	MeanIoU	Precision	Recall	F1 score
Two points					
Left angle recess	0.865	—	—	—	—
Right angle recess	0.788	—	—	—	—
Three planes					
Corneosclera	0.960	0.923	0.945	0.976	0.960
ICL	0.928	0.873	0.971	0.894	0.931
Lens	0.911	0.868	0.953	0.901	0.926

TABLE 2 The outcomes of the automated measurement and ground truth.

	The automated measurement *Mean* ± *SD* (mm)	The ground truth *Mean* ± *SD* (*mm*)
ATA	9.499 ± 0.329	9.551 ± 0.324
CCT	0.639 ± 0.050	0.600 ± 0.048
ACD	2.887 ± 0.283	2.876 ± 0.284
LV	0.726 ± 0.266	0.729 ± 0.275

ATA, angle recess to angle recess; CCT, central cornea thickness; ACD, anterior chamber distance; LV, lens vault; SD, standard deviation.

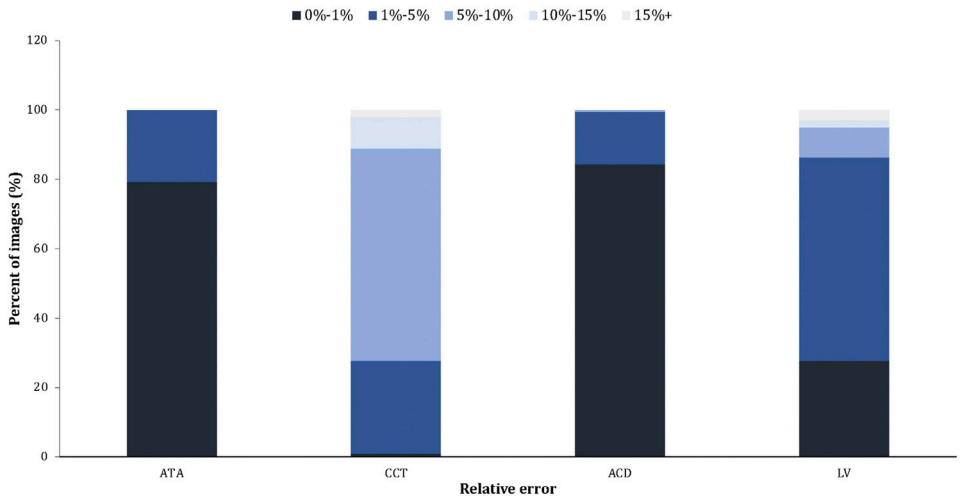


FIGURE 3
The relative error of the automated measurement (Machine vs. Ground truth). The relative error = the absolute error/the ground truth × 100%.

indicated the high accuracy of the deep learning measurement. More specifically, for all measurements of ATAs and ACDs, 89% of CCTs and 95% of LVs, the relative error between automated measurements and ground truth was within 10% compared with the ground truth.

We also assessed the degree of agreement and correlation between measurements using ICCs. The ICCs between the deep learning method and the ground truth of ATA, CCT, ACD, and LV ranged from 0.928 to 0.995, indicating excellent agreement between the automated method and the ground truth. The Bland-Altman plots (Figure 4) also confirmed the excellent agreement and acceptable limits of agreement between the automated method and the ground truth, with the bias ranging from −0.05 to 0.01 mm. Furthermore, the RMS difference between ground truth and various methods of measurement (Figure 5) showed that human-ground truth

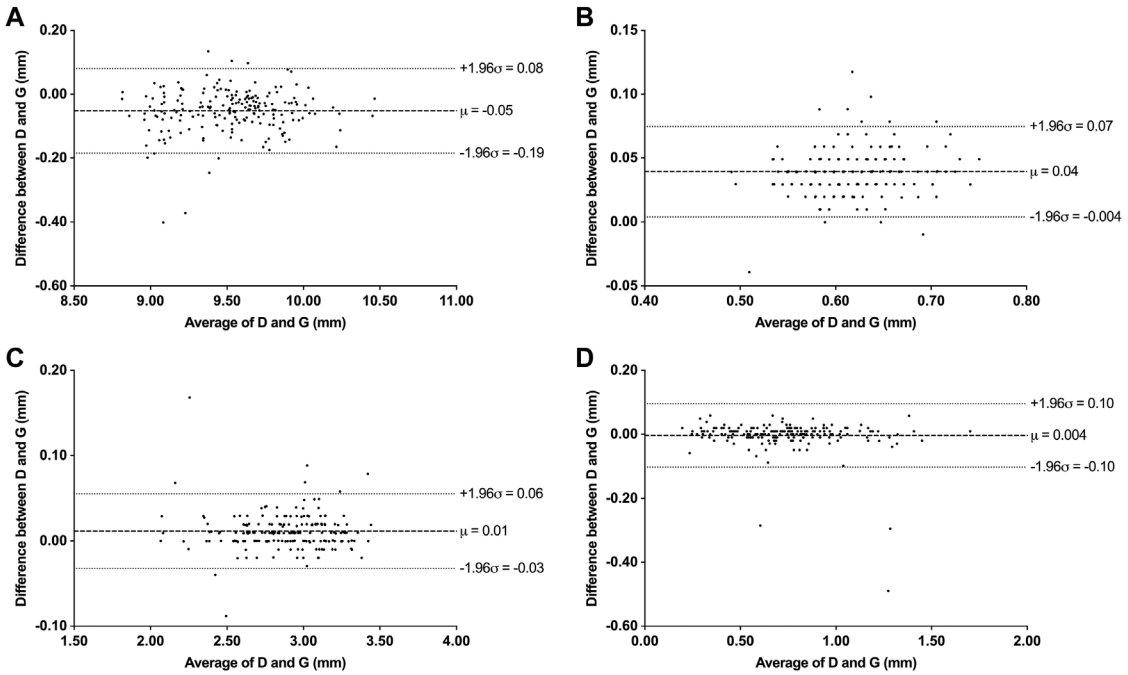


FIGURE 4
Bland-Altman plots between deep learning system **(D)** and ground truth **(G)** measurements. **(A)** Bland-Altman plots for ATA. **(B)** Bland-Altman plots for CCT. **(C)** Bland-Altman plots for ACD. **(D)** Bland-Altman plots for LV.

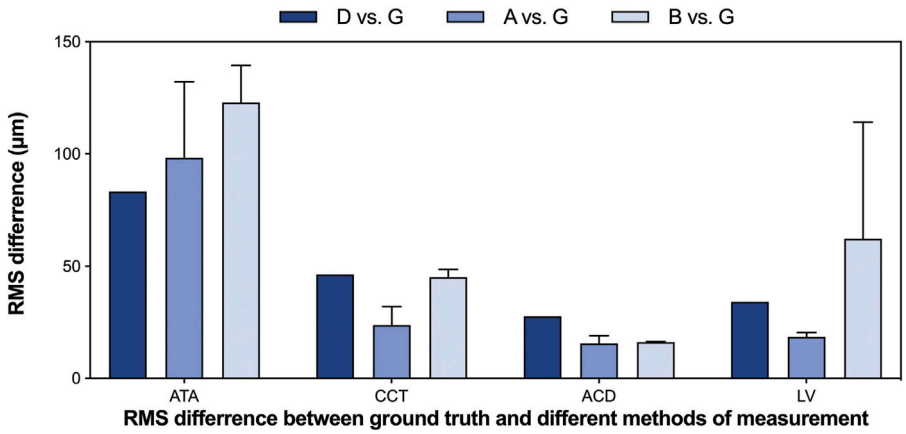


FIGURE 5
The RMS difference between ground truth and different methods of measurement.

differences were not significantly different from machine-ground truth differences in most cases. In addition, the error bars of RMS differences between A, B, and G measurements showed that the deep learning method (without error bars) possessed better repeatability than human experts.

Discussion

Postoperative follow-ups should be effectively managed to prevent major complications for patients undergoing ICL implantation. With the development of the AS-OCT, subtle changes in the anterior chamber could be discovered and

quantified during the follow-ups. In this study, we presented a deep learning method to automatically monitor the position of ICL and promptly identify minor changes in the anterior chamber, including detecting the main structures (the angle recess area, corneoscleral, ICL, and natural lens) and quantifying the anterior chamber parameters (ATA, CCT, ACD, and vault) in AS-OCT images for postoperative ICL patients. The method was based on the U-Net architecture and achieved human expert-level performance.

U-Net, comprising an encoder and a decoder network connected by skip connections, has been showing great promises in segmenting medical images [18–23], including assisting in clinical follow-ups [24–26]. In this study, we developed an improved U-Net architecture for the segmentation module, which includes the segmentation of angle recess points and other structures. Traditionally, in localizing key points of ophthalmic images, the convolutional neural networks (CNN) regression method, such as ResNet18, is usually applied to output the coordinates of the target point [13]. Nevertheless, in our multi-task application scenario, using one network for each task would increase the computation and training costs, and tasks cannot promote each other's performance through interaction. In this study, a more simplified and efficient method was introduced. The method regarded the localization task as finding the segmentation map of the target point regions and the centroids of their largest connected region, to obtain the coordinates. In this way, the two tasks can be completed with only an improved U-Net, without integrating multiple networks, enhancing the simplicity and versatility of the model. Compared with previous models, our method showed a higher accuracy with an absolute error of 0.037 mm for the left angle recess area and 0.063 mm for the right angle recess area (0.487, 0.389 for the Resnet-18 model). In addition, the angle recess points are the key anatomic landmark in the next quantification process, whose performance can be affected by the radius of the points. Therefore, it is crucial to select a proper value for the radius. When trained on the same training data, we found that the performance peaked at 20 pixels (Supplementary Figure S1). For segmentation of the other structures, besides a high Dice coefficient, the module also achieved a high F1 score and mIoU, showing its great potential in accurately isolating AS-OCT structures. For comparison, this method also outperformed the RelayNet and DeepLabel V3+ models in identifying different structures, especially for corneoscleral (Dice coefficient 0.925 vs. 0.888 vs. 0.960). Furthermore, ICCs between automated values and the ground truth were relatively high, making it a reliable method to assist in follow-ups in daily practice.

It is an important mitigation of vault-related adverse events to closely observe ICLs with insufficient or excessive vault [5]. Previous studies showed the risk of cataract formation increases when the vault is low ($< 250\,\mu m$), while the risk of angle closure, pupillary block, or pigment dispersion glaucoma increases when the vault is high ($> 750\,\mu m$) [27]. However, vault varied broadly across studies [28], which ranges from $0\,\mu m$ at two years to $1180\,\mu m$ at one year of follow-up [29]. This may be due to several factors, for example, different standards of various operators when measuring, the size of the ICL [30], the follow-up time when the vault was measured [31–33], the rise of the crystalline lens [34, 35], and the age of the patient [36]. Therefore, an objective method to automatically obtain the values can mitigate the subjectivity of the measurement process, which may also contribute to a better understanding of the postoperative vault. Our results for the ICL segmentation showed a high Dice score, mean IoU, precision, recall, and F1 score (all > 0.85), indicating that our method can accurately detect the ICL. Also, 95% of the relative error of LV was within 10% and the ICCs between different measurements of LV showed excellent agreements (all > 0.90), which illustrated that our deep learning method can obtain the values with high reliability. In the future, with the proposal of our deep learning method, experts are able to take full advantage of the follow-up data and enhance the management of ICL implantation, which also possesses the potential to optimize the formula used for preoperative lens sizing.

Besides the vault-related adverse event, the endothelial cell loss and other subtle changes in the anterior chamber should also be aware. The loss of endothelial cells varied across studies, while it is considered that the largest loss occurs during the early postoperative period, and the surgical procedure is the main cause of the loss [28]. Our deep learning system can separate the corneosclera accurately (with the Dice score, mean IoU, precision, recall and F1 score > 0.85) and automatically obtain the CCT with excellent agreement with the ground truth (89% of relative errors within 10%; ICCs > 0.90). In addition, age-related alterations of the anterior chamber could affect ICL position over the years. For example, there is an age-related increase in ciliary muscle anteroposterior thickness, which might affect the position of ICL [37]. Thus, we also obtained ACD and ATA to quantify the changes in the postoperative anterior chamber with great reliability (all relative errors within 10%; all ICCs > 0.95).

There are also some limitations to our study. First, our study only included a relatively small data set with a specific population (Chinese), which would benefit from external validation of other ethnic groups. Second, the AS-OCT images were obtained from a single type of equipment (Casia SS-1000 AS-OCT). This should be further investigated if there is any difference among measurements of various types of equipment. Finally, to better monitor the postoperative risks, the method could be further developed into a web-based or app-based dataset, which can also record other information during the follow-ups. Above all, there has been an acceleration of adopting new models of healthcare delivery following the rapid changes to healthcare systems during COVID-19 [38].

In summary, we developed a deep learning method to manage the follow-ups after ICL implantations, which can monitor the position of ICL and identify the subtle changes in the anterior chamber with high performance in both the

segmentation and measurement process. This method could assess the postoperative risk and discover the complications timely, which can assist patients and ophthalmologists in daily practice. Also, it is a relatively objective approach to obtain the measurements in AS-OCT images, which can make the data between different studies more comparable and repeatable to each other, including eliminating the deviation caused by the image rotation and personal equation.

Data availability statement

The datasets presented in this article are not readily available because they contain identifying patient information. Requests to access the datasets should be directed to the Second Affiliated Hospital of Zhejiang University, School of Medicine.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

YS, JL, PX, YW, SH, GJ, SW, and JY contributed to conception and design of the study. YS collected and prepared the dataset. JL programmed the deep learning system. YS and PC performed manual measurements. YS, JL, and PX performed the statistical analysis. YS wrote the first draft of the manuscript. PX contributed to the main revision of the first draft. All authors contributed to manuscript revision, read, and approved the submitted version. JY and SW supervised the study.

Funding

1) National Natural Science Foundation Regional Innovation and Development Joint Fund (U20A20386), 2) National key research and development program of China (2019YFC0118400), 3) Key research and development program of Zhejiang Province (2019C03020), 4) National Natural Science Foundation of China (81870635), 5) Clinical Medical Research Center for Eye Diseases of Zhejiang Province (2021E50007), 6) Natural Science Foundation of Shandong Province (2022HWYQ-041).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2022.969683/full#supplementary-material

References

1. Holden BA, Fricke TR, Wilson DA, Jong M, Naidoo KS, Sankaridurg P, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology* (2016) 123(5):1036–42. doi:10.1016/j.ophtha.2016.01.006

2. Morgan IG, Ohno-Matsui K, Saw SM. *Lancet* (2012) 379(9827):1739–48. doi:10.1016/S0140-6736(12)60272-4

3. Brenner LF, Alio JL, Vega-Estrada A, Baviera J, Beltran J, Cobo-Soriano R. Clinical grading of post-LASIK ectasia related to visual limitation and predictive factors for vision loss. *J Cataract Refract Surg* (2012) 38(10):1817–26. doi:10.1016/j.jcrs.2012.05.041

4. Kim TI, Alio Del Barrio JL, Wilkins M, Cochener B, Ang M. Refractive surgery. *The Lancet* (2019) 393(10185):2085–98. doi:10.1016/s0140-6736(18)33209-4

5. Packer M. Meta-analysis and review: Effectiveness, safety, and central port design of the intraocular collamer lens. *Clin Ophthalmol* (2016) 10:1059–77. doi:10.2147/opth.s111620

6. Yang W, Zhao J, Sun L, Zhao J, Niu L, Wang X, et al. Four-year observation of the changes in corneal endothelium cell density and correlated factors after Implantable Collamer Lens V4c implantation. *Br J Ophthalmol* (2021) 105(5):625–30. doi:10.1136/bjophthalmol-2020-316144

7. Ang M, Baskaran M, Werkmeister RM, Chua J, Schmidl D, Aranha Dos Santos V, et al. Anterior segment optical coherence tomography. *Prog Retin Eye Res* (2018) 66:132–56. doi:10.1016/j.preteyeres.2018.04.002

8. Thomas R. Anterior segment optical coherence tomography. *Ophthalmology* (2007) 114(12):2362–3. doi:10.1016/j.ophtha.2007.05.050

9. Ting DSW, Lin H, Ruamviboonsuk P, Wong TY, Sim DA. Artificial intelligence, the internet of things, and virtual clinics: Ophthalmology at the digital translation forefront. *Lancet Digit Health* (2020) 2(1):e8–e9. doi:10.1016/s2589-7500(19)30217-1

10. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* (2016) 316(22):2402–10. doi:10.1001/jama.2016.17216

11. Ting DSW, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye Diseases using retinal images from multiethnic populations with diabetes. *JAMA* (2017) 318(22):2211–23. doi:10.1001/jama.2017.18152

12. Fu H, Baskaran M, Xu Y, Lin S, Wong DWK, Liu J, et al. A deep learning system for automated angle-closure detection in anterior segment optical coherence

tomography images. *Am J Ophthalmol* (2019) 203:37–45. doi:10.1016/j.ajo.2019.02.028

13. Xu BY, Chiang M, Pardeshi AA, Moghimi S, Varma R. Deep neural network for scleral spur detection in anterior segment OCT images: The Chinese American eye study. *Transl Vis Sci Technol* (2020) 9(2):18. doi:10.1167/tvst.9.2.18

14. Pham TH, Devalla SK, Ang A, Soh ZD, Thiery AH, Boote C, et al. Deep learning algorithms to isolate and quantify the structures of the anterior segment in optical coherence tomography images. *Br J Ophthalmol* (2021) 105(9):1231–7. doi:10.1136/bjophthalmol-2019-315723

15. Roy AG, Conjeti S, Karri SPK, Sheet D, Katouzian A, Wachinger C, et al. ReLayNet: Retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomed Opt Express* (2017) 8(8):3627–42. doi:10.1364/boe.8.003627

16. L-C Chen, Y Zhu, G Papandreou, F Schroff, H Adam, 2018. *Encoder-decoder with atrous separable convolution for semantic image Segmentation2018*. Cham: Springer International Publishing.

17. Chen CC, Barnhart HX. Assessing agreement with intraclass correlation coefficient and concordance correlation coefficient for data with repeated measures. *Comput Stat Data Anal* (2013) 60:132–45. doi:10.1016/j.csda.2012.11.004

18. Ibtehaz N, Rahman MS. MultiResUNet : Rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Netw* (2020) 121:74–87. doi:10.1016/j.neunet.2019.08.025

19. Sirinukunwattana K, Pluim JPW, Chen H, Qi X, Heng PA, Guo YB, et al. Gland segmentation in colon histology images: The glas challenge contest. *Med Image Anal* (2017) 35:489–502. doi:10.1016/j.media.2016.08.008

20. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. *Med Image Comput Computer-Assisted Intervention—MICCAI* (2015) 9351:234–41.

21. Yu LQ, Yang X, Chen H, Qin J, Heng PA. Volumetric ConvNets with mixed residual connections for automated prostate segmentation from 3D MR images. In: Thirty-First Aaai Conference on Artificial Intelligence; February 4 –9, 2017; San Francisco (2017). p. 66–72.

22. Christ PF, Mohamed Ezzeldin A, Ettlinger F, Tatavarty S, Bickel M, Bilic P, et al. Automatic liver and lesion segmentation in CT using cascaded fully convolutional neural networks and 3D conditional random fields. International conference on medical image computing and computer-assisted intervention: September 18th to 22nd 2022;Singapore 2016. p. (pp. 415–23).

23. Wang S, He KL, Nie D, Zhou SH, Gao YZ, Shen DG. CT male pelvic organ segmentation using fully convolutional networks with boundary sensitive representation. *Med Image Anal* (2019) 54:168–78. doi:10.1016/j.media.2019.03.003

24. Y Li, J Yang, Y Xu, J Xu, X Ye, G Tao, et al. 2020. *Learning tumor growth via follow-up volume prediction for lung Nodules2020*. Cham: Springer International Publishing.

25. Tao G, Zhu L, Chen Q, Yin L, Li Y, Yang J, et al. Prediction of future imagery of lung nodule as growth modeling with follow-up computed tomography scans using deep learning: A retrospective cohort study. *Transl Lung Cancer Res* (2022) 11(2):250–62. doi:10.21037/tlcr-22-59

26. Zhang Y, Zhang X, Ji Z, Niu S, Leng T, Rubin DL, et al. An integrated time adaptive geographic atrophy prediction model for SD-OCT images. *Med Image Anal* (2021) 68:101893. doi:10.1016/j.media.2020.101893

27. Fernandes P, Gonzalez-Meijome JM, Madrid-Costa D, Ferrer-Blasco T, Jorge J, Montes-Mico R. Implantable collamer posterior chamber intraocular lenses: A review of potential complications. *J Refract Surg* (2011) 27(10):765–76. doi:10.3928/1081597x-20110617-01

28. Montes-Mico R, Ruiz-Mesa R, Rodriguez-Prats JL, Tana-Rivero P. Posterior-chamber phakic implantable collamer lenses with a central port: A review. *Acta Ophthalmol* (2021) 99(3):e288–e301. doi:10.1111/aos.14599

29. Rodriguez-Una I, Rodriguez-Calvo PP, Fernandez-Vega Cueto L, Lisa C, Fernandez-Vega Cueto A, Alfonso JF. Intraocular pressure after implantation of a phakic collamer intraocular lens with a central hole. *J Refract Surg* (2017) 33(4):244–9. doi:10.3928/1081597x-20170110-01

30. Nam SW, Lim DH, Hyun J, Chung ES, Chung TY. Buffering zone of implantable Collamer lens sizing in V4c. *BMC Ophthalmol* (2017) 17(1):260. doi:10.1186/s12886-017-0663-4

31. Fernandez-Vigo JI, Macarro-Merino A, Fernandez-Vigo C, Fernandez-Vigo JA, De-Pablo-Gomez-de-Liano L, Fernandez-Perez C, et al. Impacts of implantable collamer lens V4c placement on angle measurements made by optical coherence tomography: Two-year follow-up. *Am J Ophthalmol* (2017) 181:37–45. doi:10.1016/j.ajo.2017.06.018

32. Yan Z, Miao H, Zhao F, Wang X, Chen X, Li M, et al. Two-year outcomes of visian implantable collamer lens with a central hole for correcting high myopia. *J Ophthalmol* (2018) 2018:1–9. doi:10.1155/2018/8678352

33. Alfonso JF, Fernandez-Vega-Cueto L, Alfonso-Bartolozzi B, Montes-Mico R, Fernandez-Vega L. Five-year follow-up of correction of myopia: Posterior chamber phakic intraocular lens with a central port design. *J Refract Surg* (2019) 35(3):169–76. doi:10.3928/1081597x-20190118-01

34. Gonzalez-Lopez F, Mompean B, Bilbao-Calabuig R, Vila-Arteaga J, Beltran J, Baviera J. Dynamic assessment of light-induced vaulting changes of implantable collamer lens with central port by swept-source OCT: Pilot study. *Transl Vis Sci Technol* (2018) 7(3):4. doi:10.1167/tvst.7.3.4

35. Gonzalez-Lopez F, Bilbao-Calabuig R, Mompean B, Luezas J, Ortega-Usobiaga J, Druchkiv V. Determining the potential role of crystalline lens rise in vaulting in posterior chamber phakic collamer lens surgery for correction of myopia. *J Refract Surg* (2019) 35(3):177–83. doi:10.3928/1081597x-20190204-01

36. Alfonso JF, Fernandez-Vega L, Lisa C, Fernandes P, Jorge J, Montes Mico R. Central vault after phakic intraocular lens implantation: Correlation with anterior chamber depth, white-to-white distance, spherical equivalent, and patient age. *J Cataract Refract Surg* (2012) 38(1):46–53. doi:10.1016/j.jcrs.2011.07.035

37. Strenk SA, Strenk LM, Guo S. Magnetic resonance imaging of the anteroposterior position and thickness of the aging, accommodating, phakic, and pseudophakic ciliary muscle. *J Cataract Refract Surg* (2010) 36(2):235–41. doi:10.1016/j.jcrs.2009.08.029

38. Sifan Zheng JC, Chen J, Joshi S, Sun Z. The barriers to global eye care equity and the role of digital innovations. *Adv Ophthalmol Pract Res* (2021) 1(2):100021. doi:10.1016/j.aopr.2021.100021

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 20 September 2022
DOI 10.3389/fphy.2022.972916

Check for updates

OPEN ACCESS

EDITED BY
Xunbin Wei,
Peking University, China

REVIEWED BY
Xuan Weipeng,
Hangzhou Dianzi University, China
Richard Yongqing Fu,
Northumbria University,
United Kingdom
Fanli Meng,
Northeastern University, China

*CORRESPONDENCE
Ruijian Yan,
yanruijian@zju.edu.cn

SPECIALTY SECTION
This article was submitted to Optics and Photonics, a section of the journal Frontiers in Physics

RECEIVED 19 June 2022
ACCEPTED 29 August 2022
PUBLISHED 20 September 2022

CITATION
Xia J, Pan J, Yan R, Zhang F, Zhang L, Feng G, Luo J, Dong S and Wang Y (2022), Portable flexible probe for detecting blood supply status in clinical surgery.
Front. Phys. 10:972916.
doi: 10.3389/fphy.2022.972916

COPYRIGHT
© 2022 Xia, Pan, Yan, Zhang, Zhang, Feng, Luo, Dong and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Portable flexible probe for detecting blood supply status in clinical surgery

Jie Xia[1,4], Jiadong Pan[1,4], Ruijian Yan[3]*, Fan Zhang[4], Luxi Zhang[1,4], Gang Feng[3], Jack Luo[1], Shurong Dong[1,2] and Yueming Wang[2]

[1]Key Laboratory of Advanced Micro/Nano Electronic Devices and Smart Systems of Zhejiang, College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou, China, [2]Frontier Center of Brain Science and Brain-machine Integration, Cancer Center, Zhejiang University, Hangzhou, China, [3]Department of Orthopedic Surgery, 2nd Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang, China, [4]Key Laboratory of Biomedical Engineering of Education Ministry, Qiushi Academy for Advanced Studies, Zhejiang University, Hangzhou, China

In the medical field, it is important to monitor and evaluate the blood supply status of organs and tissues during the clinical surgery. However, this largely depends on the surgeon's experience and naked eye, which is easy to misjudge due to the interference of blood stains or other factors. A portable and flexible photoplethysmographic (PPG) detection probe is developed in this paper. And a new evaluation methodology of blood supply status is proposed based on this probe. Three typical indicators based on PPG is proposed to comprehensively evaluate the blood supply status, which are the blood oxygen saturation and its pulsation, differential characteristics of different lights, and time-frequency energy spectral characteristic. The probe and its evaluation methodology are verified using the brain of rats as a model.

KEYWORDS

optoelectronic sensor, blood supply probe, photoplethysmographic, parathyroidectomy, biomedical photonic device, flexible applications

1 Introduction

In clinical surgery, it is necessary to collect and monitor the blood supply status of organs and tissues. Depend on the blood supply status of organs and tissues, the judgement could be given about its status of normal, abnormal or irreversible inactivation, then the appropriate further surgery can be performed. For instance, the judgment about the blood supply status of parathyroid gland (PG) decides surgery for removing the diseased PG during parathyroidectomy. It is an essential skill for thyroid surgeons to decide to remove or preserve mung bean-sized parathyroid glands through evaluating blood supply status of the thyroid capsule. However, it largely depends on the surgeon's experience due to bloodstains, environmental disturbances and other interferential factors, so the visual observation of the decision is subjective and has a large error. Thus, the rapid and accurate identification of the blood supply status of organs and tissues during the operation has great clinical value. Near-infrared (NIR) [1–3]

photodetectors have been used in more and more medical and health fields. The near-infrared autofluorescence (NIRAF) has been used as a non-invasive, real-time, and automatic live detection method for PG. Under the excitation and irradiation of 785 nm laser, it was found that the fluorescence intensity of the PG in all patients was always greater than the fluorescence intensity of the thyroid and all other tissues of the neck, especially the fluorescence intensity of the PG was 2–11 times higher than that of the thyroid tissue [4]. A study in 2019 further found that NIRAF could accurately distinguish between healthy PG and diseased PG, thereby minimizing hypocalcaemia after thyroidectomy [5]. Another application of organ or tissue blood supply status detection is that for comatose patients, the monitoring of brain activity is required to judge whether the patient is alive or dead [6]. Cerebral blood flow [7, 8] or neural activity [9, 10] can often be used as an auxiliary test to diagnose brain death. Tests used to assess neural activity include EEG and evoked potentials. The main limitations of tests in assessing neural activity are interference from artifacts and the effects of metabolic changes and medications. Cerebral angiography, which is used to assess cerebral blood flow, is recognized as the gold standard method for diagnosing brain death, but its disadvantage is that it requires the patients to leave the intensive care unit and use contrast agents [11]. And Doppler (TCD) ultrasound examination is another method to evaluate blood flow without invasive surgery [7, 8]. However, above methods is not suitable for surgery due to its complex operation and large instrument size. A portable detector to carry out real-time monitoring of blood supply status during surgery is strongly desired by surgeons. In some minimally invasive surgeries, they need a flexible and small tube-like probe to insert into body and approach the organ or tissues to detect the blood supply status *in situ*.

The photoplethysmographic (PPG) and image PPG (iPPG) technology has been used to measure the changes of blood volume [12–14]. The recent publications show the possible application of iPPG for intraoperative monitoring of tissue perfusion during neurosurgery and abdominal surgery [15, 16]. However, in order to capture the changes in blood perfusion, the imaging sensor of the iPPG system has higher requirements on the frame rate, and at the same time, due to its non-contact, it is more affected by movement and ambient light. In addition, the change of PPG of micro organ capillary is too weak to measure it with iPPG [17]. Compared with non-contact iPPG, the contact-based PPG system has better optical accuracy, and the sensitivity of the blood volume change received by the photodiode is higher than that obtained after the secondary conversion of the iPPG imaging sensor. And PPG could be more simple, flexible and integrated.

This paper develop a portable and flexible micro-PPG detection probe for surgeon to monitor and evaluate the blood supply status of organs and tissues during the clinical surgery. A new evaluation methodology of blood supply status is proposed through verification using the brain of rats as a model. The acquisition of PPG signal was extracted and the organ activity was charactered with the sensitive eigenvalue from PPG. Three typical indicators based on PPG is proposed to comprehensively evaluate the blood supply status, which are the Blood oxygen saturation and its pulsation, differential characteristics of different lights, and time-frequency energy spectral characteristic. The micro-probe has been verified using the brain of rats as a model. This technology hopefully provides a powerful technical means for *in-situ* detection of micro-organ's blood supply status in clinical operations.

2 Materials and methods

2.1 Design of portable flexible probe

The working principle of the PPG probe for identifying the blood supply status of organs is described as follows in brief: when a light beam of a certain wavelength illuminates surface of the tissue, the contraction and expansion of blood vessels affect the transmission or reflection of lights every time the pulse beats. When the lights pass through skin and then are reflected back to the photosensitive sensor, the lights will be attenuated to a certain degree. The absorption of lights by muscles, bones, veins and other connecting tissues is similar, but that by arteries is different. Because of the pulsation of blood in the arteries, the absorption of lights will change repeatedly according to the pulsation of the blood. In the cardiovascular system, capillaries are the sites of material exchange with tissues and organs. Since the movement of red blood cells in the capillary microcirculation is a heterogeneous fluid movement, red blood cells pass through the capillaries one by one or one after another, but when the red blood cells exchange substances in the capillary microcirculation and the microcirculation boundary, the interaction of the environment and the interaction between red blood cells, prone to accumulation and aggregation, etc., will affect the blood flow of capillaries. Under normal circumstances, the blood volume is relatively stable due to the non-pulsating nature of the rigid structure of capillaries. However, when organs or tissues are damaged, red blood cells may accumulate and aggregate, and capillary microcirculation is blocked, resulting in corresponding changes in blood volume and changes in blood flow parameters reflected by PPG. Also, the growing transmural pressure of the arteries during the systole compresses the connective tissues of the dermis in a local place, which results in the increasing density of capillaries and changes in blood volume [18–20]. Therefore, PPG can be used to evaluate the blood supply status.

The flexible structure of the probe mainly includes two parts: the first part is the flexible adhesive at the front end of the probe acquisition, mainly using polydimethylsiloxane (PDMS), which is a transparent flexible silicone, non-toxic insulating material

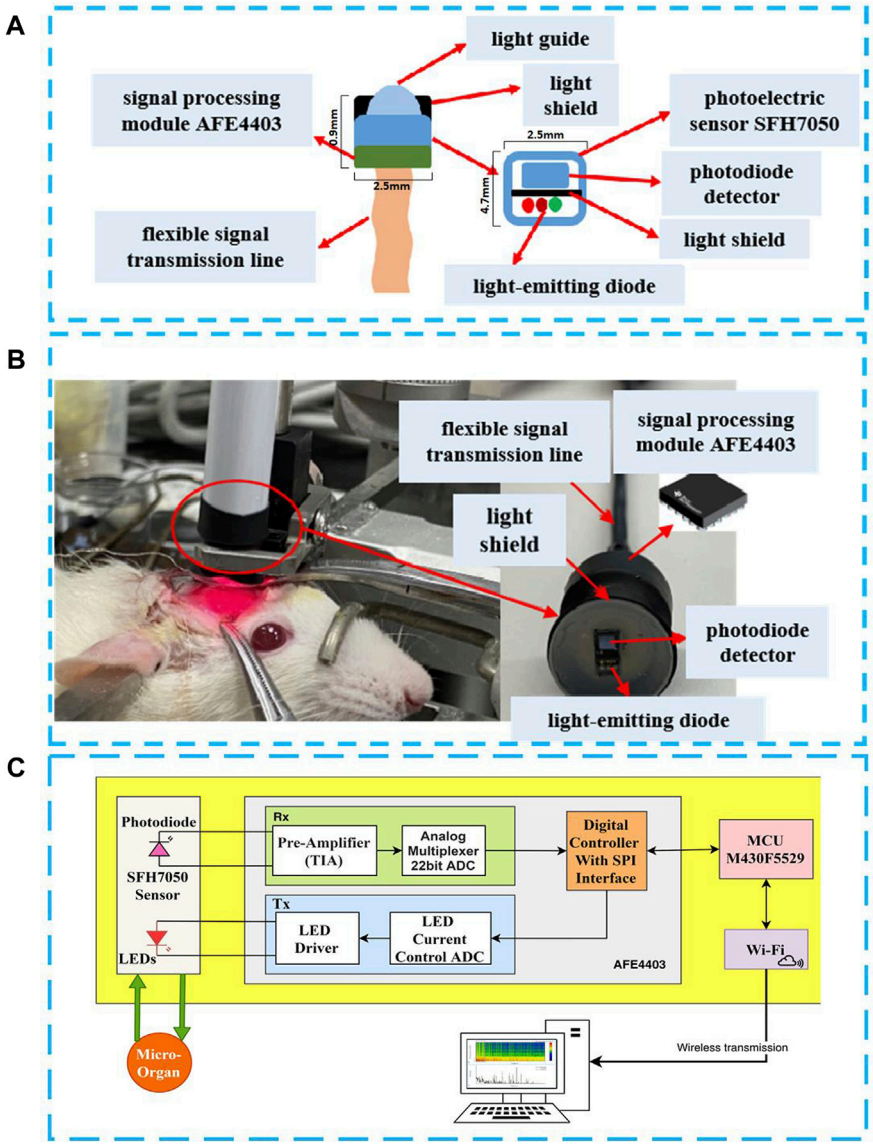


FIGURE 1
Portable and flexible PPG detection probe in this work. **(A)** Probe structure diagram; **(B)** Rat brain *vivo* experiment for verification; **(C)** Probe circuit system diagram system.

with good light transmittance and good biocompatibility, which will not affect the accurate collection of biological probes. The second part is the flexible circuit board structure, which is convenient for the application in the detection of minimally invasive surgery in the body in the future, and can be flexibly bent and moved. Figure 1A is the schematic drawing of the proposed PPG probe in this work.

The photoelectric volume sensor SFH7050 contains three light-emitting diodes (LED) with wavelengths of 532 nm for the green light, 660 nm for the red light and 940 nm for the infrared light, respectively. The PPG probe has a lampshade-like structure as shown in Figure 1B at the front end, and the overall structure of the probe is fixed by a flexible circuit board, which can flexibly adapt to different surgical environments. The head of the PPG probe has lampshade-like structure, similar size with the detected organ or tissue, which can effectively prevent interferences from the ambient light, so as to improve the accuracy and reliability of the detection. In order to make the photoelectric sensor attached closely and comfortable with the organs, a transparent flexible PDMS film was coated inside the light shield. This structure can effectively prevent the interferences of tissue fluid during the detection of organ activity in the body without affecting the light

path. The PDMS film was obtained by coating it on a silicon substrate with a thickness of 100 um by spinning coat PDMS solution at 3,000 rpm for 30 s. Prebaked it at 150°C for 5 min on hotplate, and hard baked at 250°C for 1 h on hotplate. In the end, PDMS film was peeled off the substrate and adhered on the light shield manually. In order to dynamically observe the PPG signal from the organ (rat brain was used for this work) for a long time, we added a casing for fixing. This tubular structure is also suitable to be used as a hand-held or standalone medical device during operation in an open surgical environment. Figure 1C is a photo of the developed flexible PPG probe used in the experiment, and Figure 1C is the functional block diagram of the PPG probe system. The PPG probe system is mainly consisted of a photoelectric sensor SFH7050, a large-area photodetector, an analog front-end AFE4403, which is composed of a low-noise receiver channel with an integrated analog-to-digital converter for PPG signal acquisition, and an ultra-low power microcontroller MSP430 for on-board signal processing. The processed signal is then sent to a PC with WiFi for displaying the results in real-time. The flexible PPG probe is powered by a 4.2 V, 240 mAh lithium battery. A fully charged battery can power the probe for about 16 h.

The mirco-probe mentioned in this paper is mainly used to determine the blood supply status of small organs/tissues, so as to assist surgeons in rapid intervention. For example, in the removal of parathyroid glands, for the judgment of necrotic parathyroid glands, these tiny organs/tissues are often small in size, usually about 6 mm long and 3–4 mm wide, and 1–2 mm anteroposteriorly [21], the size of the rat brain window in this experiment is only 7*6 mm. The size of the PPG sensor in this article is only 4.7 mm × 2.5 mm x 0.9 mm. Therefore, on the one hand, it can be adapted to small organs/tissues, and at the same time, the data collected at a single point is also sufficient for judging the blood supply status.

2.2 Experiments

A health Sprague-Dawley (SD) rat with a weight of 287 g was used for the proof of concept experiment. It was anesthetized with propofol (10 mg/ml solution, 0.012 ml/g) *via* intraperitoneal injection (IP) and set in the stereotaxic frame. A cranial window (size: 5×5 mm^2) was produced into the skull of the rat by drilling, so that the PPG detection probe can be inserted into skull and attached Sulcus gyrus of brain. PPG signal recording was initiated once the cranial window was open, and it was followed with the excessive pentobarbital sodium (150 mg/kg) injection through IP for euthanasia, and observe the response of the brain activity by using the PPG detector.

It should be pointed out that, before euthanasia, the rat was also utilized for other studies to avoid the use of additional animals in our study. All surgical and experimental procedures followed the Guide for The Care and Use of Laboratory Animals (China Ministry of Health) and were approved by the Animal Care Committee of Zhejiang University, China. It should be noted that when SD is injected intraperitoneally or intravenously with sodium pentobarbital ≥100 mg/kg, it will experience euthanasia [22]. Therefore, during the death of SD, the blood circulation in the brain will gradually stop, and the brain gradually loses the arterial blood supply.

2.3 Characteristic indicators and evaluation methodology

In this paper, we provide three characteristic indicators to comprehensively evaluate the blood supply status of organs and tissues organs, which are related with organs and tissues vascular insufficiency or irreversible inactivation. These three indicators are the blood oxygen saturation and its pulsation, differential characteristics of PPG signals of different lights, and time-frequency energy spectral characteristics of PPG, which are explained as follows.

2.3.1 Blood oxygen saturation and pulsation

The blood oxygen saturation [23] represents the ratio of the volume of oxyhemoglobin (HbO_2) bound to oxygen to the volume of all bound hemoglobin (Hb) in the arterial blood, i.e., the amount of oxygen in the blood, usually expressed as a percentage. Lambert Beer's law is the basic law of the spectrophotometry [24], which describes the relationship between the strength of the absorption under a certain wavelength light of a substance, the concentration of the light-absorbing substance, and the thickness of the liquid layer. Derived from Lambert Beer's law, the blood oxygen saturation is expressed as follows:

$$SpO_2 = A - B \cdot R, R = \frac{I_{AC}{}^{\lambda_1} / I_{DC}{}^{\lambda_1}}{I_{AC}{}^{\lambda_2} / I_{DC}{}^{\lambda_2}}, \quad (1)$$

Here, A and B are constants, and can be obtained through experiments. I_{AC} and I_{DC} are the intensity of the reflected lights converted from light 1 (wavelength λ_1) and 2 (wavelength λ_2), respectively. R essentially represents the ratio of the blood absorption rate of the red light to the infrared light per unit time. The value of R could be considered to be a constant because of the relatively invariable blood for a very short period of time. Thus, the spectrum of the PPG obtained in a short period of time could be used to calculate the value of R, which then can be transformed into the ratio of the spectral amplitude of the AC components of the red light and infrared light. Once they are normalized by the respective DC levels (also obtained from the frequency spectrum), it becomes a DC normalized transmittance with time for the tested period. Assuming that the red light and infrared light spectra areY_1 and Y_2 respectively in a short time

interval, the fundamental frequency of the red light PPG wave is ${f_{AC}}^{\lambda_1}$, the DC level frequency is ${f_{DC}}^{\lambda_1}$, the fundamental frequency of the red light PPG wave is ${f_{AC}}^{\lambda_2}$, and the DC level frequency is ${f_{DC}}^{\lambda_2}$, then the blood oxygen saturation can be expressed by Eq. 2:

$$SpO_2 = A - B \cdot \frac{Y_1\left({f_{AC}}^{\lambda_1}\right)/Y_1\left({f_{DC}}^{\lambda_1}\right)}{Y_2\left({f_{AC}}^{\lambda_2}\right)/Y_2\left({f_{DC}}^{\lambda_2}\right)}, \tag{2}$$

As we know, due to the presence of blood supply, PPG will regularly fluctuate with the diastole and contraction of the pulse. After the PPG is converted to the frequency domain, in addition to the DC component absorbed by bones, skin and other tissues, there is also a fundamental frequency near the DC component, which reflects the size of the pulse rate. The pulse rate is also an important physiological parameter of the respiratory cycle [25], which can characterize the activity of the organ. This has been utilized to measure the pulsation as expressed by Eq. 3:

$$PulseRate = \frac{\left({f_{AC}}^{\lambda_1} + {f_{AC}}^{\lambda_2}\right) * 60}{2}, \tag{3}$$

2.3.2 Differential characteristics of PPG of different lights

According to the absorption coefficient curves [26, 27], HbO_2 and Hb have different absorption coefficients for the red and infrared lights. When the red light and infrared light are irradiated on surface of an organ, the reflected light intensities may vary with time, but the difference in the two reflected light intensities remains almost unchanged, i.e. the intensity difference of the two reflected lights is constant for a normal organ. On the other hand, when the organ is malfunction, inactive or dead, there is less or no saturated blood oxygen at all, and the intensity difference of the two reflected lights is different from the normal value, and normally the light intensity difference become stronger as the organ activity becomes weaker. This characteristic difference related to the reflected light intensity of red light and infrared light can be utilized to detect and judge the alive status of an organ during operation, as shown by Eq. 4:

$$Diff = I^{\lambda_1} - I^{\lambda_2}, \tag{4}$$

2.3.3 Time-frequency energy spectral characteristics of PPG

Due to the existence of blood supply, the PPG will beat regularly with the relaxation and contraction of the heart. Therefore, the PPG will have components in its fundamental frequency and harmonic components, and its energy is higher at this time [28]. With the occurrence of organ death, the blood circulation of the organ gradually becomes weak and eventually stops, the PPG no longer presents regular beat function, and the energy gradually decreases when the fundamental frequency and harmonic components are small. Therefore, the instantaneous energy change of the PPG can be observed through the short-time Fourier transform, so as to judge activity of the organ in real-time. Eq. 5 represents the Fourier transform of a signal, and Eq. 6 represents the power spectral density of the signal.

$$Y(t, f) = \int_{-\infty}^{+\infty} w(t-\tau) I_{AC}(\tau) e^{-j2\pi f\tau} d\tau, \tag{5}$$

$$SP_x(t, f) = \left|Y(t, f)\right|^2 = \left|\int_{-\infty}^{+\infty} w(t-\tau) I_{AC}(\tau) e^{-j2\pi f\tau} d\tau\right|^2 \tag{6}$$

2.3.4 Evaluation methodology of blood supply status

In a conclusion, we provide three characteristic indicators to comprehensively evaluate the blood supply status, which are the blood oxygen saturation and its pulsation, differential characteristics of PPG signals of different lights, and time-frequency energy spectral characteristics of PPG [29, 30]. These three indicators can show us the organs and tissues vascular insufficiency or irreversible inactivation according with the whole process of euthanasia of rats, there are actually three stages:

1 Normal blood supply status

$$Diff > 0;$$

Because Hb has a higher absorption coefficient for red light, and HbO_2 has a higher absorption coefficient for infrared light. At this time, the red reflected light detected by our probe is stronger than the infrared reflected light intensity. Therefore, at this time, there is enough HbO_2 in the blood of the organ to absorb a large amount of infrared light, that is, the blood circulation is normal.

2 Abnormal blood supply status but is reversible

If $Diff$ is close to 0, $SpO_2$and $Pulserate$ continue to decrease, and the reflected light intensity of infrared light and red light detected by probe are almost equal, indicating that the content of HbO_2 and Hb in the blood of the organ changes slowly at this time, and blood circulation is blocked, the organs should be treated immediately to restore the blood supply.

$$Diff \to 0; SpO_2 \downarrow; Pulserate \downarrow$$

3 Blood supply stops and is irreversible

At this time, the reflected light intensity of infrared light and red light gradually tends to the constant DC component, indicating that $HbO_2$and Hbhave no dynamic changes. In addition, the intensity of the reflected infrared light is much greater than that of red light, indicating that there is no HbO_2in the blood of the organ. The blood circulation is suspended, and the organ has been completely and irreversibly inactivated.

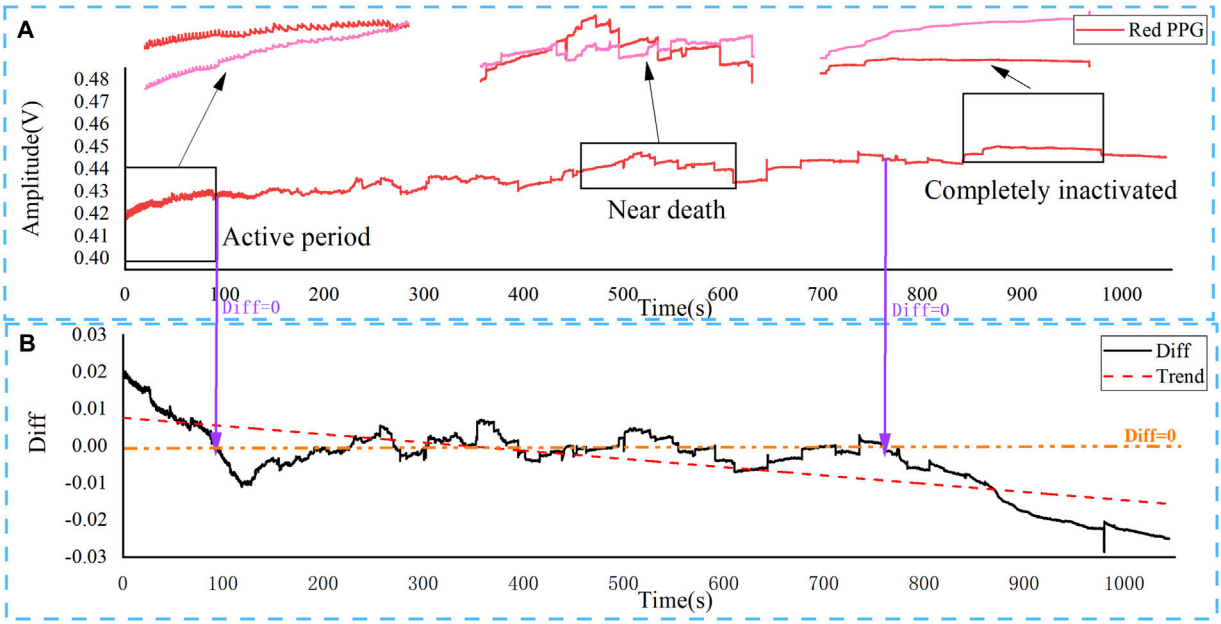


FIGURE 2
Changes of blood supply status in the brain of rats after overdose of anesthetic injection by PPG. The results show three typical periods, corresponding to different blood supply status. **(A)** Real-time change of red light (660 nm) and infrared light PPG (940 nm); **(B)** The difference (Diff) between red light and infrared light PPG.

Therefore, the advantage of this article is that necessary measures can be taken to assist clinical treatment according to the three stages.

$$Diff < 0; Energy \rightarrow 0$$

3 Results and discussion

3.1 The difference between the intensity of red and infrared light

The real-time change of PPG of the reflected red light and infrared light is illustrated in Figure 2A, respectively. For the initial period of the experiment, arterial blood still circulated in the rat brain. For the active period, the initial reflected intensity of the red-light (660 nm) was much stronger than that of the infrared light (940 nm), and the difference became smaller with time over a period of 100 s. Both the red light and infrared light PPG waveforms oscillated regularly as the rat's brain is still active as normal. The intensities of the reflected red light and infrared light from the rat brain were reversed after a 100 s and both of them varied with time irregularly. The rat was near dead at this stage. After 750 s, the reflected infrared light intensity became much stronger than that of the red light, and the difference became larger with the progress of time thereafter. The rat was completely dead after 17 min. Two features can also be noticed: 1) The regular oscillation of the PPG signal observed at the initial stage became weaker after 100 s and disappeared completely after 150 s, indicating the brain activity became weaker; 2) The PPG base line increased with time. This is because of the gradual blood coverage of the brain with time, which leds to the increased DC components of the PPG signal. Figure 2B shows the change of difference between the intensity of reflected red light and infrared light obtained by the PPG probe. Diff is equal to Red PPG minus IR PPG in Figure 2A. It can be seen that the decrease in brain activity results in an overall decrease in the difference between the red light and infrared light reflected, which reveals the changes of brain blood supply status.

3.2 Variation of SPO2 and pulserate

Figure 3A illustrates PPG of the red-light reflection as a function of time for different periods in details (The data of Figure 3 were extracted from the first 90 s data in Figure 2A for calculation, Figure 2A only shows the 30 s data). As it can be seen from Figure 3A, the pulsation intensity of the PPG decreased with the increase of time, and the according pulserate decreased

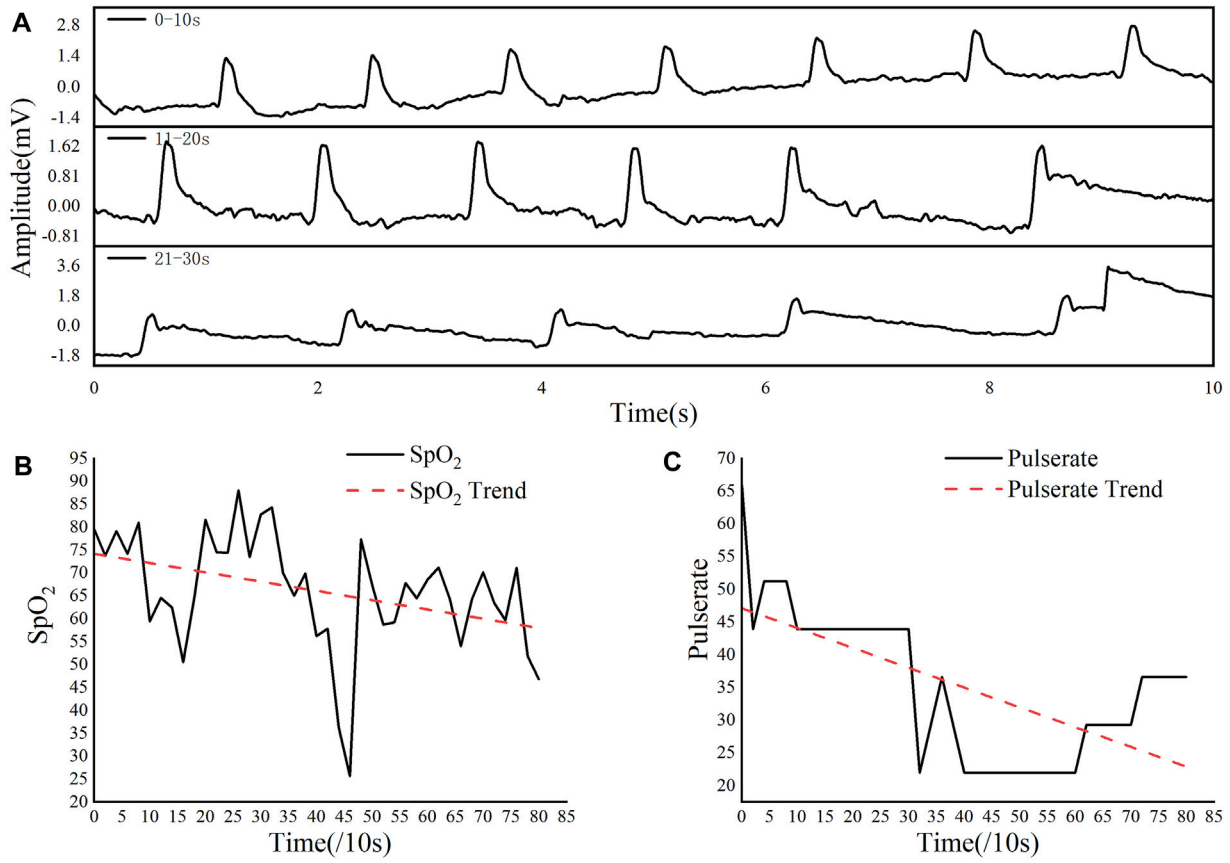


FIGURE 3
Red light PPG of the rat brain in the first 90 s. **(A)** The intensity of the red light reflected every 10 s. **(B)** The change of blood oxygen saturation gradually decrease. **(C)** Pulsation gradually decrease. Variation and analysis of Time-frequency energy spectrum characteristics.

gradually. According to Eq. 2, the change of blood oxygen saturation in the rat brain can be calculated for the 90 s period before the experiment with the result shown in Figure 3B.

As it is clear that during this period, the blood oxygen saturation fluctuated greatly, but overall it showed a downward trend with the increase of time. The pulserate of the PPG of the rat was calculated by using Eq. 3 for the 90 s before the experiment. As shown in Figure 3C, the pulsation showed a gradual decline trend during the inactivation process.

Further time-frequency analysis was conducted on PPG signals to obtain more information of the brain activity, which can be intuitive to grasp the changes in the light energy absorbed by the brain of the rats after injection of pentobarbital sodium.

A 0.01–50 Hz band-pass filtering and wavelet processing were first performed on the red light PPG signal, as to remove DC component (Such as power supply) and high-frequency interference (Such as changes in capillary density and venous blood volume, temperature changes, etc.). The result is shown in Figure 4A. It can be seen from the time domain that the PPG wave signal gradually becomes sparse during the euthanasia of rats injected with pentobarbital sodium. It indicates that the rat's heart gradually stops beating during this process, resulting in insufficient blood supply to the cerebral arteries, unchangeable light absorption, and PPG waveform that gradually stops beating.

Then we converted to the frequency domain to analyse the energy changes of the PPG signal during the euthanasia of the rat. Use Eq. 5 to perform short-time Fourier transform on the recorded PPG waveform to observe the instantaneous energy change of the PPG wave as shown in Figure 4B, and the energy gradually decreases when the fundamental frequency and harmonic components are small. Figure 4C shows the sum of the energy of all frequency components of the PPG signal of the short-time Fourier transform, that is, the sum of all frequency energy at each moment expressed in Eq. 6. It can be seen that as the blood supply of the rat brain gradually decreases, the energy of the PPG gradually decreases. The energy change obtained according to the time-frequency change more intuitively reflects the overall

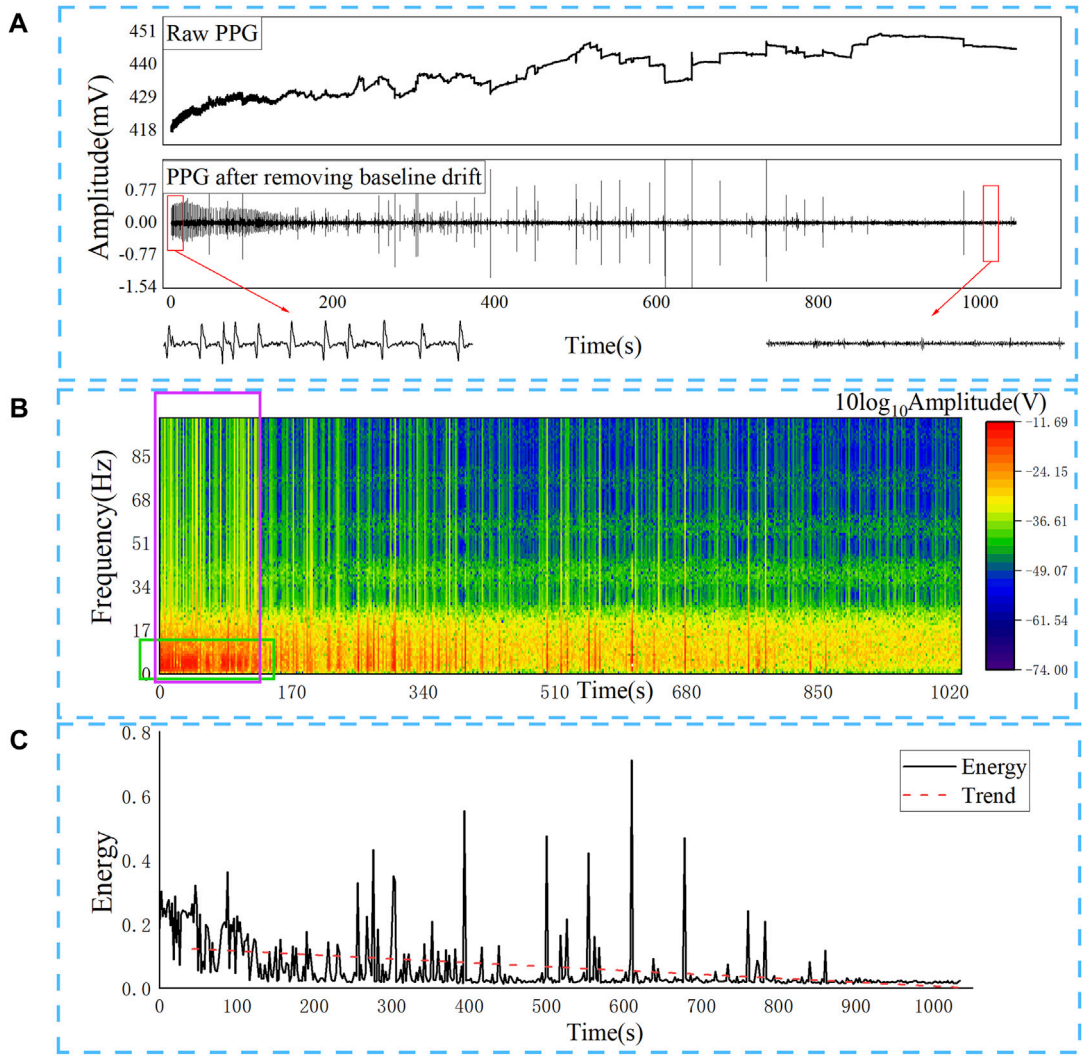


FIGURE 4
After removing the baseline drift of the original red light PPG, it can be seen that the pulsation gradually disappears, and low-frequency components in the frequency domain are gradually decreasing, and the total energy is continuously decreasing. **(A)** Red light PPG after band-pass filtering and wavelet processing; **(B)** Time-frequency change of red light PPG; **(C)** Energy change of red light PPG.

decreasing trend of PPG energy in the process of brain inactivation in rats.

4 Conclusion

A portable and flexible PPG detection probe is developed in this paper. And a new evaluation methodology of blood supply status is proposed based on this probe. Three typical indicators are proposed to comprehensively evaluate the blood supply status, which are the blood oxygen saturation and its pulsation, differential characteristics of PPG signals of different lights, and time-frequency energy spectral characteristics of PPG signal. Among them, the time-frequency energy feature is the most important discriminant indicator, followed by the red light and the Infrared light. Differential features, the indicators with lower reference are blood oxygen saturation and pulsation frequency. The probe and its evaluation methodology are verified using the brain of rats as a model. In the future work, through a large number of experimental or clinical samples, machine learning can be used to build a model that can identify the three active states, which will further improve the efficiency of discrimination. In general, this technology provides a powerful technical means for *in-situ* detection of micro-organ's blood supply status in clinical operations.

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 the Guidelines for the Care and Use of Laboratory Animals of Zhejiang University (No. ZJU20200129).

Author contributions

All authors have read and agreed to the published version of the manuscript. JX, JP, SD, GF, JL, YW, and RY designed and performed the research. FZ and LZ performed experiments. JX and JP wrote the manuscript.

Funding

This work was funded by National Key R&D Program of China (No. 2018YFA0701400) and Zhejiang Province Key R&D programs (Nos 2021C05004, 2021C03108, and 2021C03062).

Acknowledgments

All experimental protocols and live animals were approved by the IACUC of Zhejiang University (No. ZJU20200129).

Conflict of interest

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

Publisher's note

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

References

1. Xu H, Li J, Leung BHK, Poon CCY, Ong BS, Zhang Y, et al. A high-sensitivity near-infrared phototransistor based on an organic bulk heterojunction. *Nanoscale* (2013) 5(23):11850–5. doi:10.1039/C3NR03989G

2. Xue Z, Li X, Li Y, Jiang M, Ren G, Liu H, et al. A 980 Nm laser-activated upconverted persistent probe for nir-to-nir rechargeable *in vivo* bioimaging. *Nanoscale* (2017) 9(21):7276–83. doi:10.1039/C6NR09716B

3. Yang Y, Aw J, Xing B. Nanostructures for nir light-controlled therapies. *Nanoscale* (2017) 9(11):3698–718. doi:10.1039/C6NR09177F

4. Paras C, Keller M, White L, Phay J, Mahadevan-Jansen A. Near-infrared autofluorescence for the detection of parathyroid glands. *J Biomed Opt* (2011) 16(6): 067012. doi:10.1117/1.3583571

5. Solórzano CC, Thomas G, Baregamian N, Mahadevan-Jansen A. Detecting the near infrared autofluorescence of the human parathyroid. *Ann Surg* (2020) 272(6): 973–85. doi:10.1097/SLA.0000000000003700

6. Beecher HK. After the definition of irreversible coma. *N Engl J Med* (1969) 281(3):1070–1. doi:10.1056/NEJM196911062811909

7. Flowers WM, jr, Patel BR. Persistence of cerebral blood flow after brain death. *South Med J* (2000) 93(4):364–70. doi:10.1097/00007611-200004000-00003

8. Pistoia F, Johnson DW, Darby JM, Horton JA, Applegate LJ, Yonas H. The role of xenon ct measurements of cerebral blood flow in the clinical determination of brain death. *AJNR Am J Neuroradiol* (1991) 12(1):97–103.

9. Vicenzini E, Pro S, Randi F, Pulitano P, Spadetta G, Rocco M, et al. Transcranial Doppler for brain death after decompressive craniectomy: Persistence of cerebral blood flow with flat eeg. *Intensive Care Med* (2010) 36(12):2163–4. doi:10.1007/s00134-010-2008-0

10. Welschehold S, Boor S, Reuland K, Thömke F, Kerz T, Reuland A, et al. Technical aids in the diagnosis of brain death. *Dtsch Arztebl Int* (2012) 109(39): 624–30. doi:10.3238/arztebl.2012.0624

11. Kramer AH. Ancillary testing in brain death. *Semin Neurol* (2015) 35(2): 125–38. doi:10.1055/s-0035-1547541

12. Pelaez EA, Villegas ER. Led power reduction trade-offs for ambulatory pulse oximetry. *Annu Int Conf IEEE Eng Med Biol Soc* (2007) 2007:2296–9. doi:10.1109/IEMBS.2007.4352784

13. Reisner A, Shaltis PA, McCombie D, Asada HH, Warner S, Warner MA. Utility of the photoplethysmogram in circulatory monitoring. *Anesthesiology* (2008) 108(5):950–8. doi:10.1097/ALN.0b013e31816c89e1

14. Ruiz-Rodríguez JC, Ruiz-Sanmartín A, Ribas V, Caballero J, García-Roche A, Riera J, et al. Innovative continuous non-invasive cuffless blood pressure monitoring based on photoplethysmography technology. *Intensive Care Med* (2013) 39(9):1618–25. doi:10.1007/s00134-013-2964-2

15. Kamshilin AA, Zaytsev VV, Lodygin AV, Kashchenko VA. Imaging photoplethysmography as an easy-to-use tool for monitoring changes in tissue blood perfusion during abdominal surgery. *Sci Rep* (2022) 12(1):1143. doi:10.1038/s41598-022-05080-7

16. Mamontov OV, Shcherbinin AV, Romashko RV, Kamshilin AA. Intraoperative imaging of cortical blood flow by camera-based photoplethysmography at green light. *Appl Sci* (2020) 10(18):6192. doi:10.3390/app10186192

17. Unakafov AM, Möller S, Kagan I, Gail A, Treue S, Wolf F. Using imaging photoplethysmography for heart rate estimation in non-human primates. *PLoS One* (2018) 13(1):e0202581. doi:10.1371/journal.pone.0202581

18. Glazier JB, Hughes JM, Maloney JE, West JB. Measurements of capillary dimensions and blood volume in rapidly frozen Lungs. *J Appl Physiol* (1969) 26(1): 65–76. Epub 1969/01/01. doi:10.1152/jappl.1969.26.1.65

19. Kamshilin AA, Nippolainen E, Sidorov IS, Vasilev PV, Erofeev NP, Podolian NP, et al. A new Look at the essence of the imaging photoplethysmography. *Sci Rep* (2015) 5(1):10494. doi:10.1038/srep10494

20. Lipowsky HH. Microvascular rheology and hemodynamics. *Microcirculation*, 12(1994):5–15. doi:10.1080/10739680590894966

21. Driscoll P. Gray's anatomy, 39th edition. *Emerg Med J* (2006) 23(6):492. doi:10.1136/emj.2005.027847

22. Underwood W, Anthony R. Avma Guidelines for the euthanasia of animals: 2020 edition. *Retrieved on March* (2020) 2020(30):1–20.

23. Mamontov OV, Sokolov AY, Volynsky MA, Osipchuk AV, Zaytsev VV, Romashko RV, et al. Animal model of assessing cerebrovascular functional reserve by imaging photoplethysmography. *Sci Rep* (2020) 10(1):19008–10. doi:10.1038/s41598-020-75824-w

24. Beer B. Bestimmung der Absorption des rothen Lichts in farbigen Flüssigkeiten. *Ann Phys Chem* (1852) 162:78–88. doi:10.1002/andp.18521620505

25. Kim H, Kim J-Y, Im C-H. Fast and robust real-time estimation of respiratory rate from photoplethysmography. *Sensors* (2016) 16(9):1494. doi:10.3390/s16091494

26. Azmal GM, Al-Jumaily A, Al-Jaafreh A. Continuous measurement of oxygen saturation level using photoplethysmography signal. *IEEE Int Conf Biomed Pharm Eng* (2006) 504–7.

27. Kanders K, Grabovskis A, Marcinkevics Z, Aivars JI. Assessment of conduit artery vasomotion using photoplethysmography. *SPIE Proc Biophotonics Riga* (2013) 9032(1):121–9. doi:10.1117/12.2044705

28. Han C-H, Kim E, Im C-H. Development of a brain-computer interface toggle switch with low false-positive rate using respiration-modulated photoplethysmography. *Sensors* (2020) 20(2):348. doi:10.3390/s20020348

29. Sassaroli A, Pierro M, Bergethon PR, Fantini S. Low-frequency spontaneous oscillations of cerebral hemodynamics investigated with near-infrared spectroscopy: A review. *IEEE J Select Top Quan Electron.* (2012) 18(4):1478–92. doi:10.1109/jstqe.2012.2183581

30. Pilt K, Ferenets R, Meigas K, Lindberg L-G, Temitski K, Viigimaa M. New photoplethysmographic signal analysis algorithm for arterial stiffness estimation. *Scientific World J* (2013) 2013:1–9. doi:10.1155/2013/169035

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 27 September 2022
DOI 10.3389/fphy.2022.978392

Check for updates

OPEN ACCESS

EDITED BY
Liwei Liu,
Shenzhen University, China

REVIEWED BY
Dazhe Zhao,
Northeastern University, China
Rui Hu,
Shenzhen University, China

*CORRESPONDENCE
Juan Ye,
yejuan@zju.edu.cn

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

SPECIALTY SECTION
This article was submitted to Optics and Photonics, a section of the journal Frontiers in Physics

RECEIVED 26 June 2022
ACCEPTED 06 September 2022
PUBLISHED 27 September 2022

CITATION
Zhou J, Jin K, Gu R, Yan Y, Zhang Y, Sun Y and Ye J (2022), Color fundus photograph registration based on feature and intensity for longitudinal evaluation of diabetic retinopathy progression.
Front. Phys. 10:978392.
doi: 10.3389/fphy.2022.978392

COPYRIGHT
© 2022 Zhou, Jin, Gu, Yan, Zhang, Sun and Ye. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Color fundus photograph registration based on feature and intensity for longitudinal evaluation of diabetic retinopathy progression

Jingxin Zhou[1†], Kai Jin[1†], Renshu Gu[2], Yan Yan[1], Yueyu Zhang[2], Yiming Sun[1] and Juan Ye[1*]

[1]Center of Ophthalmology, The Second Affiliated Hospital of Zhejiang University School of Medicine, Hangzhou, Zhejiang, China, [2]Department of Computer Science, Hangzhou Dianzi University, Hangzhou, Zhejiang, China

Longitudinal disease progression evaluation between follow-up examinations relies on precise registration of medical images. Compared to other medical imaging methods, color fundus photograph, a common retinal examination, is easily affected by eye movements while shooting, for which we think it is necessary to develop a reliable longitudinal registration method for this modality. Thus, the purpose of this study was to propose a robust registration method for longitudinal color fundus photographs and establish a longitudinal retinal registration dataset. In the proposed algorithm, radiation-variation insensitive feature transform (RIFT) feature points were calculated and aligned, followed by further refinement using a normalized total gradient (NTG). Experiments and ablation analyses were conducted on both public and private datasets, using the mean registration error and registration success plot as the main evaluation metrics. The results showed that our proposed method was comparable to other state-of-the-art registration algorithms and was particularly accurate for longitudinal images with disease progression. We believe the proposed method will be beneficial for the longitudinal evaluation of fundus images.

KEYWORDS

registration, color fundus photograph, retinal imaging, diabetic retinopathy, disease progression

Introduction

Diabetic retinopathy (DR) is one of the major diseases that can cause blindness. It is estimated that about 600 million people will have diabetes by 2040 [1], a third of whom will be affected by DR [2]. Regular follow-up and accurate analysis of longitudinal examinations play an important part in the management of DR [3]. However, the quantitative analysis of longitudinal images is still challenging, due to the tremendous discrepancies between the images caused by vastly different photographing conditions,

involuntary eye movements, and pathological changes [4], which can disturb the observation and influence the evaluation of retinal image biomarkers [5]. Registration, which means the process of establishing pixel-to-pixel correspondence between two images, grants us the chance to eliminate these discrepancies before longitudinal assessment [6]. Therefore, a preliminary registration of two retinal images is required to reduce these effects and generate a reliable disease progression conclusion.

As a necessary work of retinal image analysis, the registration of retinal fundus images is a classic topic, in which tremendous efforts have been put into this area during past decades. From a methodological point of view, retinal image registration methods can be classified into three groups: feature-based, intensity-based, and hybrid methods. In feature-based registration methods, invariant features of the retinal images are extracted and utilized for seeking the best geometric transformation between two images. Retinal vessel bifurcations [7–11], optic disc, and fovea [12, 13] are previously commonly used features. However, some of these features rely on the segmentation of retinal structures and are sensitive to image quality. Therefore, easily-obtained and stable key-point detection is the premise for robust registration through feature-based methods, for example, Harris corner [14], scale invariant feature transform (SIFT) [15], and Speeded-Up Robust Features (SURF) [16] are classic feature points that have been extensively studied. Hernandez-Matas et al. [17, 18] introduced a feature-based registration framework exploiting the spherical eye model and pose estimation. In intensity-based methods, the intensity information is calculated and used to measure the similarity of the images and the registration performance, such as cross-correlation [19], mutual information [20], and phase correlation [21]. Hybrid registration methods combine feature-based and intensity-based methods together to seek better performance [4, 22]. Compared to single feature-based or intensity-based methods, hybrid methods have great potential for more accurate and practical image alignment, but it is less investigated. Although the registration of color fundus photographs has been intensively studied, the steps of seeking higher and more robust performance have never stopped.

Although there has been intensive research work in registration, further research is still needed. First, novel registration methods developed on other modalities should be applied to retinal images to seek better performance. Second, instead of paying attention to the improvement and development of registration methods, researchers should focus more on the clinical applicability of the proposed methods, which is extremely important for longitudinal follow-up examinations. Third, the development and evaluation of registration methods rely on the publication of open-access datasets. To the best of our knowledge, the Fundus Image Registration (FIRE) dataset is the only registration dataset that has been made publicly available [23]. We thought it would be useful to develop a registration dataset made up of longitudinal images with clarified medical diagnoses. Therefore, developing registration methods in clinical settings and establishing registration datasets would be greatly beneficial for interdisciplinary cooperation and clinical transformation of computation methods.

In this study, a robust registration method for longitudinal color fundus photographs based on both feature and intensity is proposed. An ablation study showed the necessity of combining the two main parts. A comprehensive comparison between the proposed algorithm and other state-of-the-art methods was conducted to investigate its features. The dataset will be available for registration research. We believe this work will be beneficial to follow up retinal image analysis and disease progression assessment.

Materials and methods

The proposed registration framework is a combination of feature-based and intensity-based methods. The flow of this work is shown in Figure 1.

Retinal image datasets

For the evaluation of the proposed registration method, we use two datasets, FIRE and FI-LORE, consisting of color fundus image pairs different from each other in terms of actual photographing and patient conditions. These datasets are described in detail hereinafter.

The Fundus Image Registration (FIRE) dataset [23] comprises 134 image pairs, which are further classified into three categories according to their characteristics. Category S contains 71 image pairs with more than 75% overlap area and super-resolution but no anatomical changes, while category P contains 49 image pairs with less than 75% overlap area and no anatomical changes. Category A contains 14 image pairs with high overlap and large anatomical changes due to retinopathy, which can be used to mimic practical longitudinal examinations. All the images have a resolution of 2,912 × 2,912 pixels. FIRE provides ground truths for the calculation of registration errors.

The Fundus Image for Longitudinal Registration (FI-LORE) dataset consists of 83 color fundus image pairs from 78 eyes of 54 diabetic retinopathy patients who underwent longitudinal examinations at the Second Affiliated Hospital of Zhejiang University, School of Medicine, from May 2020 to July 2020. Photograph conditions, involuntary movements of the eye, and disease progression and treatments, such as laser scars, all contribute to the differences of each image in a pair. Additionally, some of them are of low image quality because of complications, such as cataracts. FI-LORE can fully reflect practical conditions of the follow-up in clinics and test the robustness of the proposed method. All the images have a

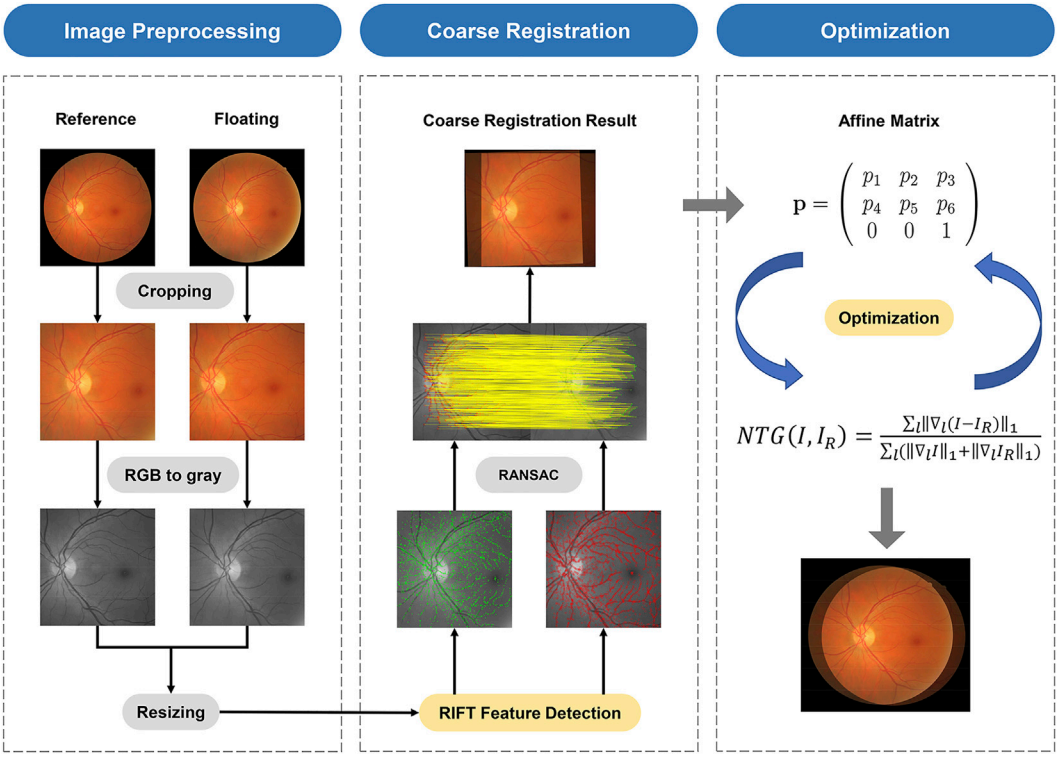


FIGURE 1
Flowchart of the proposed algorithm.

resolution of 1,500 × 1,500 pixels. Finally, to compute the registration error, we follow the annotation rule of the FIRE dataset [23], carefully choosing 10 corresponding points and repeatedly correcting the exact location of the points to guarantee the reliability of the ground truths. The FI-LORE dataset will be made publicly available.

Proposed registration framework

To normalize the images in different datasets taken at different examinations, preprocessing is the first step in the algorithm. First, the mask provided by the FIRE dataset was utilized to delete the blank margin of the original images. Second, the cropped images are further resized to 1,500 × 1,500 pixels to reach unity of the whole dataset. Furthermore, to simplify the calculation process, the RGB images are transformed into grayscale images.

Radiation-variation insensitive feature transform (RIFT) is a feature-based registration method with great robustness to non-linear radiation distortion (NRD) [24]. NRD is a rather common phenomenon that can be caused by involuntary movements of the eye. Therefore, we think it can be used for retinal image alignment tasks. The detail of the RIFT calculation can be found in the original article [24]. The alignment process is realized by the RANdom SAmple Consensus (RANSAC) algorithm [25]. Then, an affine transformation matrix $\boldsymbol{p}$ is generated on the resized image pairs (500 × 500 pixels). NTG, the normalized total gradient, was proposed by Chen et al [26], working as a registration measure. The employment of a NTG is based on the observation that the gradient difference is sparsest when the two images are perfectly aligned. The NTG is thought to outperform other intensity-based measures, such as mutual information, residual complexity, correlation ratio, and normalized cross-correlation. However, there is no study to assess the validity of the NTG in retinal images. The detail of NTG calculation is given by Chen et al. [26].

Registration evaluation

To quantitatively assess the performance of the registration result, we adopt a widely accepted registration error calculation method [23], which requires the ground truths of image pairs. Given the sets of reference points, $Y_I = \{y_I^1, y_I^2, y_I^3 \cdots y_I^{10}\} \subset \mathbb{R}^2$and $Y_R = \{y_R^1, y_R^2, y_R^3 \cdots y_R^{10}\} \subset \mathbb{R}^2$, where I and R represent the registered image and reference,

TABLE 1 Results of the ablation study in the FIRE dataset.

Category	Method	Mean registration error (pixels)		AUC of the success plot
		Mean	SD	
S (n = 71)	RIFT	2.436	0.230	0.903
	NTG	2.087	0.241	0.917
	RIFT + NTG	2.335	0.239	**0.907**
P (n = 49)	RIFT	86.702	44.233	0.510
	NTG	740.736	50.607	0.280
	RIFT + NTG	86.638	44.235	**0.512**
A (n = 14)	RIFT	7.903	1.280	0.796
	NTG	12.704	3.157	0.765
	RIFT + NTG	7.706	1.294	**0.810**
FIRE (n = 134)	RIFT	33.327	16.115	0.713
	NTG	272.598	35.595	0.556
	RIFT + NTG	33.229	16.116	**0.717**

Bold values represent the highest AUC in the specific category.

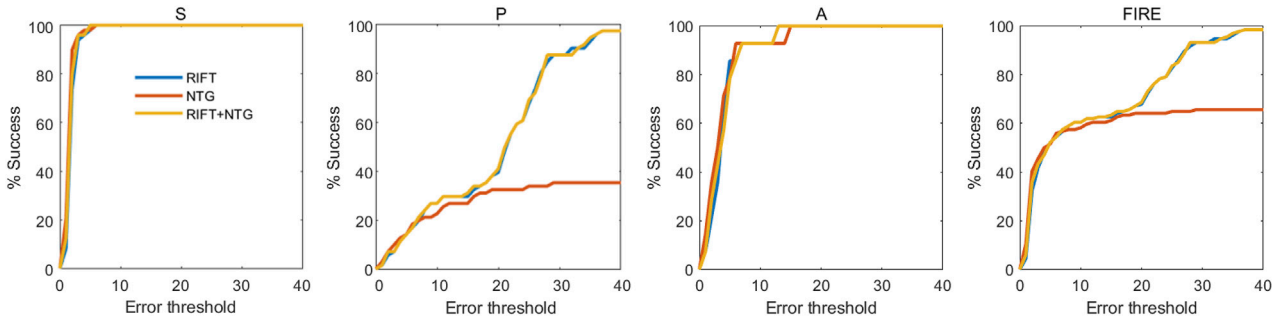


FIGURE 2
Registration success plot of the ablation study. The x-axis marks, in pixels, the registration error threshold under which registration is considered to be successful. The y-axis marks the percentage of successfully registered image pairs for a given threshold.

respectively, the mean registration error (MRE) can be calculated as

$$\mathrm{MRE}(Y_I, Y_R, p) = \frac{1}{10}\sum_{i=1}^{10}\left\|y_R^i - p(y_I^i)\right\|_2. \quad (1)$$

The $\|\|_2$ here represents the Euclidean norm. Hence, the closer *MRE* is to 0, the better the registration performance will be. To assess the registration results of a total dataset, we here utilize the success plot [18], where the x-axis marks the registration error threshold under which registration is considered to be successful and the y-axis marks the percentage of successfully registered image pairs of the given threshold. The area under the curve (AUC) is counted to quantitatively assess the registration method.

Results

Results of the ablation study

To better understand the contributions of each part of the registration framework and validate the effectiveness of the

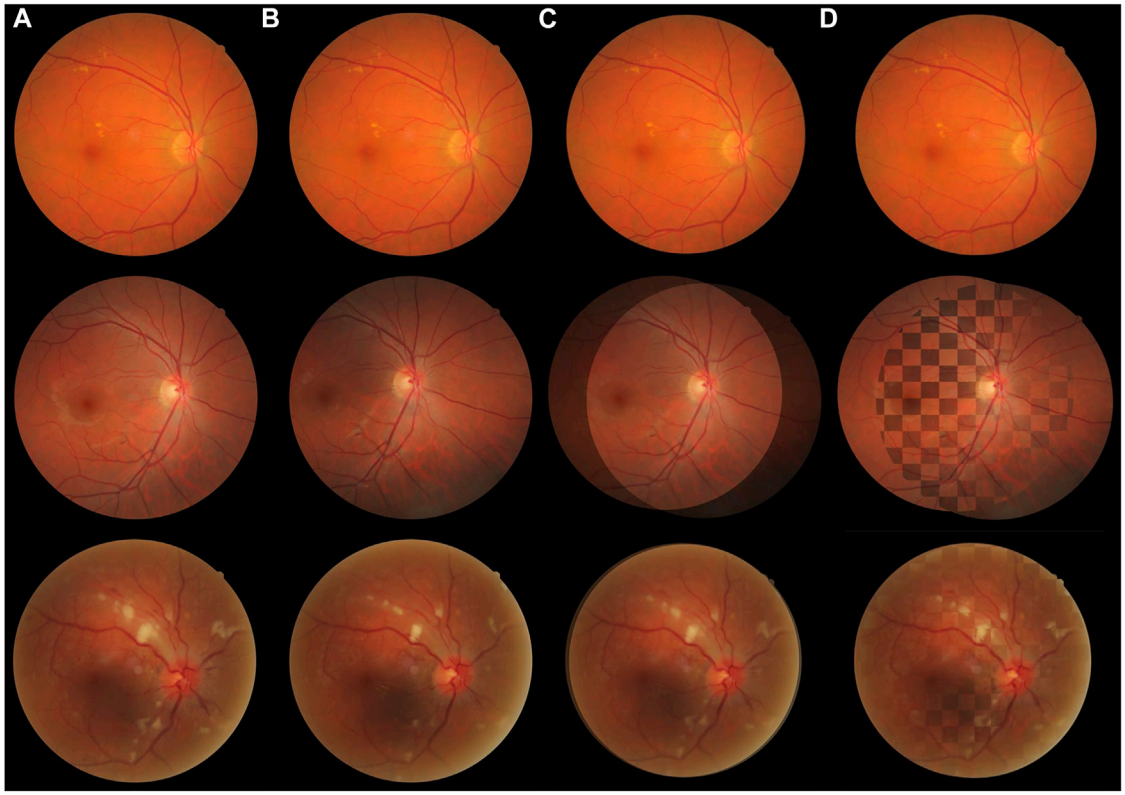


FIGURE 3
Registration results of the proposed algorithm in the FIRE dataset. The overlap decreases from the top row to bottom. **(A)** and **(B)** are image pairs without registration. **(C)** Image pairs shown in an overlaying form after registration. **(D)** Checkerboard comparisons of the proposed method.

TABLE 2 Comparisons to state-of-the-art image registration methods.

Method	Year	S	P	A	FIRE
GDB-ICP [27]	2007	0.814	0.303	0.303	0.576
Harris-PIIFD [28]	2010	0.900	0.090	0.443	0.553
ED-DB-ICP [29]	2010	0.604	0.441	0.497	0.533
RIR-BS [30]	2011	0.772	0.049	0.124	0.440
SIFT + WGTM [31]	2012	0.837	0.544	0.407	0.685
SURF + WGTM [31]	2012	0.835	0.061	0.069	0.472
ATS-RGM [32]	2015	0.369	0.000	0.147	0.211
EyeSLAM [33]	2018	0.308	0.224	0.269	0.273
GFEMR [34]	2019	0.812	0.607	0.474	0.702
VOTUS [35]	2019	0.934	**0.672**	0.681	**0.812**
REMPE [18]	2020	**0.958**	0.542	0.660	0.773
Deep learning method [36]	2021	0.908	0.293	0.660	0.657
RIFT + NTG (proposed)	2021	0.907	0.512	**0.810**	0.717

Bold values represent the highest AUC in the specific category.

combination of RIFT and NTG, we conducted an ablation study to see the registration performance of these two procedures themselves. Table 1 shows the registration results of the ablation study in the FIRE dataset. Figure 2 is the success plot of the ablation study. From the results, we can see that the NTG performs better in image pairs with high overlap, but once the overlap is small, the NTG method performance is relatively poor. However, RIFT is on the opposite side of the NTG. The combination of the algorithms outperforms each one of them. Therefore, the combination of RIFT and NTG grants algorithm robustness to image pairs of different overlap areas.

Comparison to other registration methods

To further assess the accuracy of the proposed method in color fundus image registration, we compare our results to other state-of-the-art image registration methods which are already utilized in the FIRE dataset, including GDB-ICP [27], Harris-PIIFD [28], ED-DB-ICP [29], RIR-BS [30], SIFT + WGTM [31], SURF + WGTM [31], ATS-RGM [32], EyeSLAM [33], GFEMR [34], VOTUS [35], REMPE [18], and a deep-learning based registration method proposed by Rivas-Villar et al [36]. Figure 3 is the qualitative illustration of the proposed method

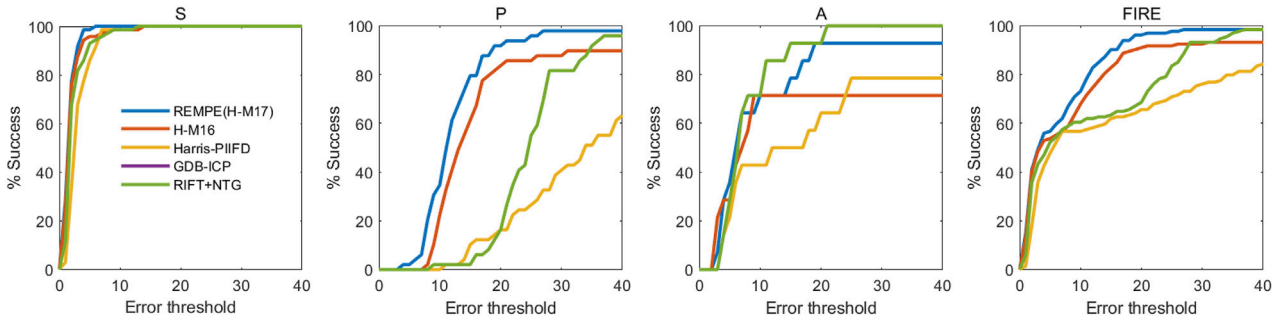


FIGURE 4
Registration success plot of the comparisons between the proposed and other registration methods. The *x*-axis marks, in pixels, the registration error threshold under which registration is considered to be successful. The *y*-axis marks the percentage of successfully registered image pairs for a given threshold.

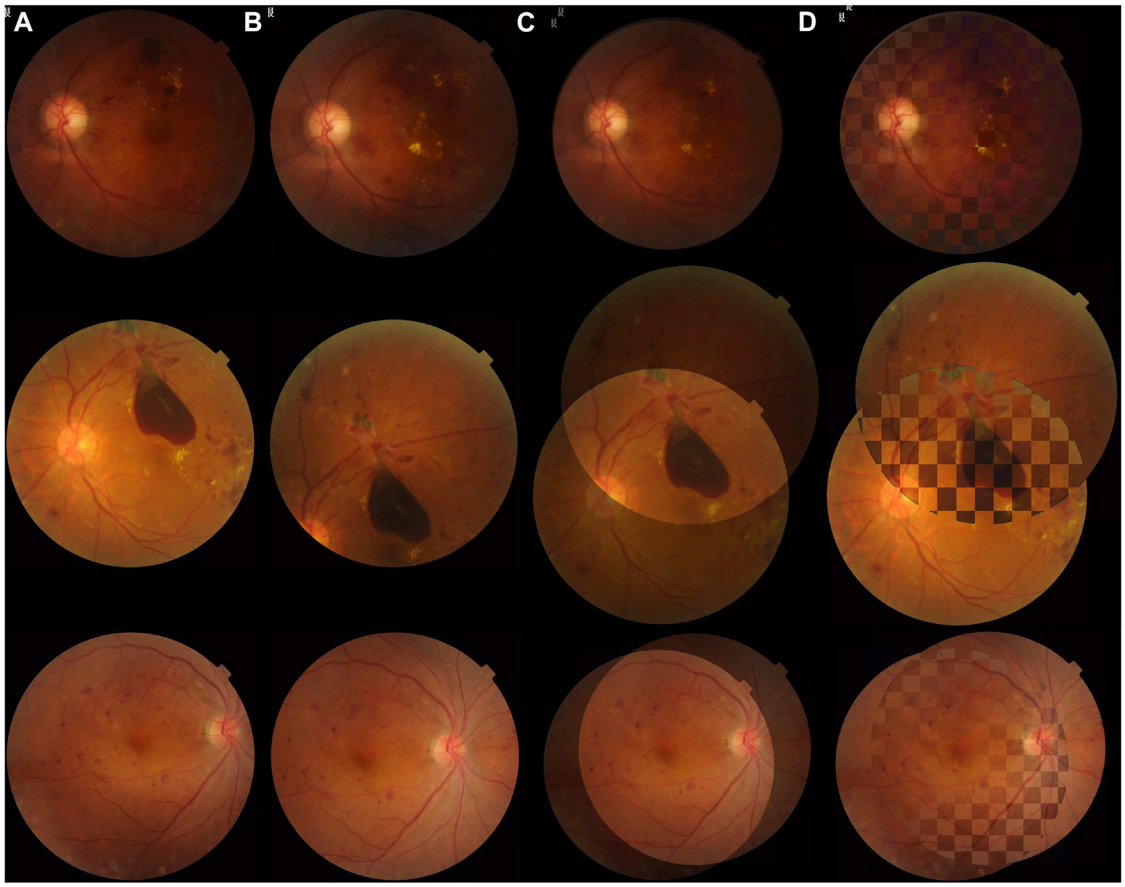


FIGURE 5
Registration results of the proposed algorithm in FI-LORE. The pairs listed, respectively, represent poor illumination quality, pathological change, and disease progression, which are common conditions in longitudinal examinations. **(A)** and **(B)** are image pairs without registration. **(C)** Image pairs shown in an overlaying form after registration. **(D)** Checkerboard comparisons of the proposed method.

registration results. Table 2 lists the methods used for comparison and the AUC of the success plot. Figure 4 contains the success plot of the proposed method and some other methods whose results are publicly available online. From the aforementioned results, one can conclude that the proposed method is competitive to the leading registration methods in

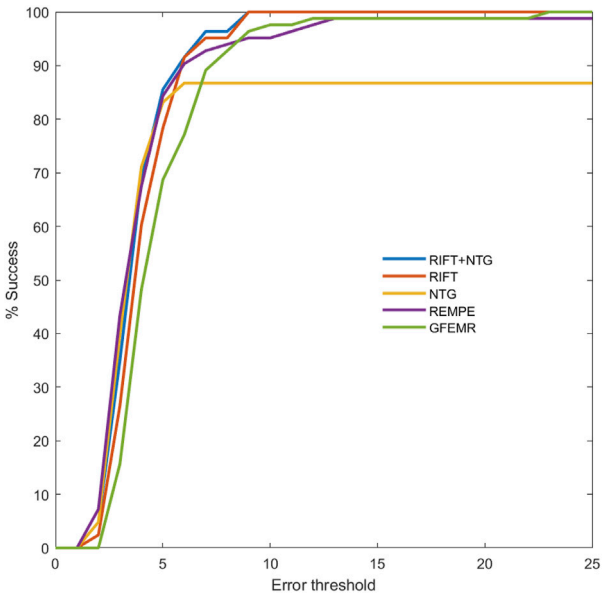


FIGURE 6
Registration success plot of the registration results in FI-LORE. The *x*-axis marks, in pixels, the registration error threshold under which registration is considered to be successful. The *y*-axis marks the percentage of successfully registered image pairs for a given threshold. The AUCs of RIFT, NTG, RITF + NTG (proposed), REMPE, and GFEMR are 0.841, 0.755, 0.850, 0.840, and 0.808, respectively.

category S. The AUC of category P clearly underperforms some algorithms. But the proposed method outperforms all the methods in category A, which stands for the longitudinal study. Therefore, we think the proposed method can remain robust under the anatomical changes and disease progression and is well worth further study.

Results in the FI-LORE dataset

Retinal images in real clinical conditions may be of poor quality due to complications such as cataracts, loss of focus, and various light conditions, posing great problems to the practical use of registration. As described earlier, FI-LORE is a collection from real ophthalmologic practice, which can be used to validate the utility of the proposed method. Figure 5 shows some image pairs before and after registration. We can see that despite these image pairs having dissimilar illumination conditions, pathological changes, and disease progression, our method can robustly align the longitudinal images to a satisfying extent. Figure 6 is the success plot of the proposed method and other state-of-the-art color fundus image registration methods in FI-LORE. The AUCs of RIFT, NTG, RITF + NTG (proposed), REMPE, and GFEMR are 0.841, 0.755, 0.850, 0.840, and 0.808, respectively. Through the quantitative analysis of the registration results in FI-LORE, we have validated the superior performance of the registration method in real clinical conditions.

Discussion

Longitudinal assessment of DR retinal images is of great importance, and longitudinal registration is an important and fundamental task which has often been neglected in clinical situations, especially for follow-up examinations. Precise alignment of different examination images is the premise for accurate detection and analysis of pathological changes, which has already been adopted in some automated retinal image analyzing devices [37]. In this study, we proposed a hybrid registration method with comprehensive experiments showing its excellent performance in longitudinal images and established a color fundus photograph dataset with pixel-wise annotation ground truth. In Table 1, the NTG shows the best performance in category S, and RIFT registered better in category P. We can conclude that the intensity-based registration method, NTG, is more precise in image pairs with large overlap while RIFT is the opposite. Thus, the combination of RIFT and NTG is reasonable and has been proved to be the best on the whole FIRE. Results from the intensive comparison experiments showed that our method is comparable to state-of-the-art image registration methods, such as GFEMR [34], VOTUS [35], and REMPE [18]. For longitudinal images in category A, the proposed method outperformed other state-of-the-art methods, for which we think it is suitable for the clinical evaluation of disease progression in follow-up examinations. This conclusion is further validated using a private dataset, FI-LORE, with more longitudinal images. Taking all of these into consideration, we believe the proposed method is good at registering longitudinal retinal images and will be beneficial in clinical use. It should be noted that in category P, which is made up of images with partial overlap, the MRE and AUC are far less than those in the other two classes. The main source of error came from several misregistered image pairs which show an MRE of nearly a thousand pixels. The same trends can also be observed in some state-of-the-art methods. The private dataset also contains images with less overlap, and the proposed method can also register them, as shown in Figure 6. Further research is needed to validate its performance and investigate the reason why these algorithms did not perform well in category P. During image pre-processing, we thought resizing might affect the final results. In the current study, the size of 500 × 500 pixels is recommended. We conducted experiments on 250 × 250 and 750 × 750 pixels, results of which can be found in Supplementary Material. Also, mutual information was evaluated to confirm the performance of the NTG, and relevant results are given in Supplementary Material.

Most of the development of registration methods focuses on either feature-based or intensity-based registration. As far

as we know, only several studies adopted hybrid methods. In 2016, Saha et al. proposed a hybrid method using Speeded-Up Robust Features (SURF) and Binary Robust Independent Elementary Features (BRIEF) [4], which are both well-known registration methods and their combination generated better performance. In our study, two registration methods, RIFT and NTG, were adopted, and further investigation revealed their own characteristics in image registration. To the best of our knowledge, this is the first study that used RIFT and NTG in ophthalmic imaging and investigated their applicability to different image overlaps. We think that the two-step hybrid registration method is promising in retinal imaging.

Deep learning has been showing great potential in medical image processing, including segmentation and registration. Several methods have been proposed to attempt to utilize deep learning in retinal image registration [36, 38–45]. However, to the best of our knowledge, there are two inherent problems limiting the development of deep learning-based registration. On the one hand, unlike other image processing issues (segmentation, enhancement, etc.), registration contains two steps theoretically, feature recognition and feature alignment. In retinal image registration, these usually mean retinal feature extraction (feature points, vessel network, etc.) and retinal feature alignment. Therefore, an inevitable question comes up, that is, when and where to adopt deep learning in the registration workflow. Different researchers provided various solutions. Some researchers adopted deep learning in the feature detection process and further aligned the feature points using conventional image alignment methods, such as RANSAC[40, 41], while some work constructed an outlier-rejection network to compute the image transformation matrix [45, 46]. There is no consensus on how deep learning should be added to the registration pipeline [46]. Moreover, in most deep learning-based registration algorithms, accurate registration relies on accurate segmentation, which is still an ongoing research topic in medical image processing. On the other hand, training and validation of deep learning networks rely on massive labeled data. In the specific topic of image registration, ground truth annotation is labor-intensive and time-consuming. For some deep learning methods using vessel segmentation, the networks also need large annotated vessel segmentation datasets. From these two perspectives, we tend to believe that although deep learning has shed light on medical image processing and analysis, it is still in the exploration stage for image registration. In the current study, we compare a state-of-the-art registration method with the proposed method, and the results showed that for longitudinal retinal image registration, our proposed method still stood out. Deep learning-assisted retinal image registration should be paid more attention to find out whether it is actually superior to conventional algorithms.

The development of retinal image registration methods is limited due to the lack of registration datasets. As far as we know, FIRE is the only dataset that focuses on retinal image registration and proposes pixel-level ground truth which can be used for the development and evaluation of registration methods. However, the longitudinal category contains only 14 image pairs, significantly small when compared to other categories. Taking this situation and the clinical use of registration methods into account, we collected and annotated 83 image pairs, especially for longitudinal image registration tasks. These image pairs are different in photograph conditions, involuntary movements of the eye, and disease progression and treatments. We believe that the adoption of this dataset can greatly benefit the study of retinal image registration.

Because of the specialty of registration in retinal image analysis, some methods have been claimed to be put into clinical use. To the best of our knowledge, there is one registration software, the DualAlign i2k software package (Clifton Park, NY), that has been made commercial. The software was developed based on the GDB-ICP algorithm, which has been compared in our work [27]. With growing interest in image registration, more novel and efficient methods have been proposed to ensure better and swifter registration performance. These methods show promise for medical image registration tasks. However, due to the lack of interdisciplinary cooperation of medical and computer science researchers, the study of these novel methods for medical use is limited. In this study, we focus on two methods and validate their performance. We believe more research is needed to provide more possibilities for more precise and swifter image analysis in real clinical use.

There are some limitations to our current study. First, the proposed algorithm performed relatively poorly in the category which stands for images with small overlay. However, there are also some similar image pairs in the FI-LORE dataset, but the proposed methods did not perform like that, which is confusing. More image pairs are needed to further test this method. Second, in the current study, we still focus on unimodality registration tasks. The performance of this method in multi-modal tasks needs more examination. Third, some state-of-the-art methods should be compared with the local dataset FI-LORE, but due to the lack of reliable source codes and our inability to completely repeat the methods, we failed to put them into further comparison. Finally, there are some artificial intelligence algorithms that have been developed for retinal image registration tasks [22, 47, 48]. We have compared one deep learning algorithm, but further studies are needed to investigate deep learning in the context of retinal image alignment.

Conclusion

RIFT can better align images with small overlap, while the NTG is more precise with large overlap image pairs. Thus, the

combination of RIFT and NTG was reasonable and outperformed single RIFT or NTG. The proposed method was comparable to other state-of-the-art registration algorithms and was especially accurate for longitudinal images with disease progression. We believe that the proposed method will be beneficial for the longitudinal evaluation of fundus images.

Data availability statement

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

Author contributions

JZ and KJ contributed to the idea, performed the experiments, analyzed the results, and wrote the manuscript. RG helped with the experiments and gave meaningful advice in the algorithm part. YY and YZ helped perform the analysis. YS helped with the preparatory work. JY contributed to the conception of the study and supervised the whole process of the experiment and writing.

Funding

The study was supported in part by the National Key Research and Development Program of China (2019YFC0118400), the Key Research and Development Program of Zhejiang Province (2019C03020), the Clinical Medical Research Center for Eye Diseases of Zhejiang Province (2021E50007), and the Natural Science Foundation of Zhejiang Province (grant number LQ21H120002).

Acknowledgments

The authors would like to thank C. Hernandez-Matas et al. for providing with the data from the FIRE database (see https://projects.ics.forth.gr/cvrl/fire/).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2022.978392/full#supplementary-material

References

1. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040. *Diabetes Res Clin Pract* (2017) 128:40–50. doi:10.1016/j.diabres.2017.03.024

2. Yau JW, Rogers SL, Kawasaki R, Lamoureux EL, Kowalski JW, Bek T, et al. Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care* (2012) 35(3):556–64. eng. Epub 2012/02/04. doi:10.2337/dc11-1909

3. Wong TY, Sun J, Kawasaki R, Ruamviboonsuk P, Gupta N, Lansingh V, et al. Guidelines on diabetic eye care. *Ophthalmology* (2018) 125(10):1608–22. EnglishCited in: Pubmed; PMID WOS:000445012100028. doi:10.1016/j.ophtha.2018.04.007

4. Saha SK, Xiao D, Frost S, Kanagasingam Y. A two-step approach for longitudinal registration of retinal images. *J Med Syst* (2016) 40(12):277. Epub 2016/10/28Cited in: Pubmed; PMID 27787783. doi:10.1007/s10916-016-0640-0

5. Ting DSW, Peng L, Varadarajan AV, Keane PA, Burlina PM, Chiang MF, et al. Deep learning in ophthalmology: The technical and clinical considerations. *Prog Retin Eye Res* (2019) 72:100759. EnglishCited in: Pubmed; PMID WOS:000488311400001. doi:10.1016/j.preteyeres.2019.04.003

6. 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

7. Zana F, Klein JC. A multimodal registration algorithm of eye fundus images using vessels detection and Hough transform. *IEEE Trans Med Imaging* (1999) 18(5):419–28. doi:10.1109/42.774169

8. Can AS, Charles V. A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. *IEEE Transactions on Pattern Analysis & Machine Intelligence* 24 (2002).doi:10.1109/34.990136

9. Laliberte F, Gagnon L, Sheng Y. Registration and fusion of retinal images--an evaluation study. *IEEE Trans Med Imaging* (2003) 22(5):661–73. Epub 2003/07/09. doi:10.1109/TMI.2003.812263Cited in: Pubmed; PMID 12846435

10. Matsopoulos GK, Asvestas PA, Mouravliansky NA, Delibasis KK. Multimodal registration of retinal images using self organizing maps. *IEEE Trans Med Imaging* (2004) 23:1557–63. doi:10.1109/tmi.2004.836547

11. Fang B, Tang YY. Elastic registration for retinal images based on reconstructed vascular trees. *IEEE Trans Biomed Eng* (2006) 53(6):1183–7. Epub 2006/06/10. doi:10.1109/TBME.2005.863927

12. Hart WE, Goldbaum MH. Registering retinal images using automatically selected control point pairs, *Proceedings of the Image processing, icip-94. IEEE International Conference* (1994). doi:10.1109/icip.1994.413740

13. Nunes JC, Bouaoune Y, Delechelle E, Bunel P. A multiscale elastic registration scheme for retinal angiograms. *Comput Vis Image Underst* (2004) 95(2):129–49. doi:10.1016/j.cviu.2004.03.007

14. Harris C, Stephens M. A combined corner and edge detector. *Proceedings of the Alvey vision conference* (1988). doi:10.5244/c.2.23

15. Lowe DG. Distinctive image features from scale-invariant keypoints. *Int J Comp Vis* (2004) 60(2):91–110. doi:10.1023/B:VISI.0000029664.99615.94

16. Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-up robust features (SURF). *Computer Vis Image Understanding* (2008) 110(3):346–59. doi:10.1016/j.cviu.2007.09.014

17. Hernandez-Matas C, Zabulis X, Argyros AA. Retinal image registration through simultaneous camera pose and eye shape estimation. *Annu Int Conf IEEE Eng Med Biol Soc* (2016) 2016:3247–51. eng. Epub 2017/03/09. doi:10.1109/embc.2016.7591421Cited in: Pubmed; PMID 28269000

18. Hernandez-Matas C, Zabulis X, Argyros A. Rempe: Registration of retinal images through eye modelling and pose estimation. *IEEE J Biomed Health Inform* (2020) 24:3362–73. eng. Epub 2020/04/06. doi:10.1109/jbhi.2020.2984483Cited in: Pubmed; PMID 32248134

19. Cideciyan AV. Registration of ocular fundus images: An algorithm using cross-correlation of triple invariant image descriptors. *IEEE Eng Med Biol Mag* (2002) 14(1):52–8. doi:10.1109/51.340749

20. Maes F, Collignon A, Vandermeulen D, Marchal G, Suetens P. Multimodality image registration by maximization of mutual information. *IEEE Trans Med Imaging* (1997) 16(2):187–98. doi:10.1109/42.563664

21. Suthaharan S, Rossi EA, Snyder V, Chhablani J, Lejoyeux R, Sahel JA, et al. Laplacian feature detection and feature alignment for multimodal ophthalmic image registration using phase correlation and Hessian affine feature space. *Signal Processing* (2020) 177:107733. Epub 2020/09/19Cited in: Pubmed; PMID 32943806. doi:10.1016/j.sigpro.2020.107733

22. ÁlvaroHervella SJR, Novo J, Ortega M. Multimodal registration of retinal images using domain-specific landmarks and vessel enhancement. *Proced Comp Sci* (2018) 126:97–104. doi:10.1016/j.procs.2018.07.213

23. Hernandez-Matas C, Zabulis X, Triantafyllou A, Anyfanti P, Douma S, Argyros AA. Fire: Fundus image registration dataset. *Model Artif Intelligence Ophthalmol* (2017) 1(4):16–28. doi:10.35119/maio.v1i4.42

24. Li J, Hu Q, Ai M. Rift: Multi-Modal image matching based on radiation-variation insensitive feature transform. *IEEE Trans Image Process* (2019) 29:3296–310. Epub 2019/12/24. doi:10.1109/TIP.2019.2959244Cited in: Pubmed; PMID 31869789

25. Hossein-Nejad Z, Nasri M, . A-RANSAC: Adaptive random sample consensus method in multimodal retinal image registration. *Biomed Signal Process Control* (2018) 45:325–38. doi:10.1016/j.bspc.2018.06.002

26. Shu-Jie C, Hui-Liang S, Chunguang L, Xin JH. Normalized total gradient: A new measure for multispectral image registration. *IEEE Trans Image Process* (2018) 27(3):1297–310. Epub 2018/07/11. doi:10.1109/TIP.2017.2776753Cited in: Pubmed; PMID 29990251

27. Yang G, Stewart CV, Sofka M, Tsai CL. Registration of challenging image pairs: Initialization, estimation, and decision. *IEEE Trans Pattern Anal Mach Intell* (2007) 29(11):1973–89. eng. Epub 2007/09/13. doi:10.1109/tpami.2007.1116in: Pubmed; PMID 17848778

28. Chen J, Tian J, Lee N, Zheng J, Smith RT, Laine AF. A partial intensity invariant feature descriptor for multimodal retinal image registration. *IEEE Trans Biomed Eng* (2010) 57(7):1707–18. eng. Epub 2010/02/18. doi:10.1109/TBME.2010.2042169Cited in: Pubmed; PMID 20176538

29. Tsai CL, Li CY, Yang G, Lin KS. The edge-driven dual-bootstrap iterative closest point algorithm for registration of multimodal fluorescein angiogram sequence. *IEEE Trans Med Imaging* (2010) 29(3):636–49. Epub 2009/08/28. doi:10.1109/TMI.2009.2030324Cited in: Pubmed; PMID 19709965

30. Chen L, Xiang Y, Chen Y, Zhang X. Retinal image registration using bifurcation structures. *2011 18th IEEE International Conference on Image Processing* (2011). p. 2169–72. doi:10.1109/icip.2011.6116041

31. Izadi M, Saeedi P. Robust weighted graph transformation matching for rigid and nonrigid image registration. *IEEE Trans Image Process* (2012) 21(10):4369–82. doi:10.1109/TIP.2012.2208980

32. Serradell E, Pinheiro MA, Sznitman R, Kybic J, Moreno-Noguer F, Fua P. Non-rigid graph registration using active testing search. *IEEE Trans Pattern Anal Mach Intell* (2015) 37(3):625–38. Epub 2015/09/10. doi:10.1109/TPAMI.2014.2343235Cited in: Pubmed; PMID 26353266

33. Braun D, Yang S, Martel JN, Riviere CN, Becker BC. EyeSLAM: Real-time simultaneous localization and mapping of retinal vessels during intraocular microsurgery. *Int J Med Robot* (2018) 14(1). Epub 2017/07/19. doi:10.1002/rcs.1848

34. Wang J, Chen J, Xu H, Zhang S, Mei X, Huang J, et al. Gaussian field estimator with manifold regularization for retinal image registration. *Signal Processing* (2019) 157:225–35. doi:10.1016/j.sigpro.2018.12.004

35. Motta D, Casaca W, Paiva A. Vessel optimal transport for automated alignment of retinal fundus images. *IEEE Trans Image Process* (2019) 28(12):6154–68. Epub 2019/07/10. doi:10.1109/TIP.2019.2925287

36. Rivas-Villar D, Hervella AS, Rouco J, Novo J. Color fundus image registration using a learning-based domain-specific landmark detection methodology. *Comput Biol Med* (2021) 140:105101. Epub 2021/12/08. doi:10.1016/j.compbiomed.2021.105101

37. Grzybowski A, Brona P, Lim G, Ruamviboonsuk P, Tan GSW, Abramoff M, et al. Artificial intelligence for diabetic retinopathy screening: A review. *Eye (London, England)* (2020) 34(3):451–60. Epub 2019/09/07Cited in: Pubmed; PMID 31488886. doi:10.1038/s41433-019-0566-0

38. Mahapatra D, Antony B, Sedai S, Garnavi R. Deformable medical image registration using generative adversarial networks, *Proceedings of the IEEE 15th international symposium on biomedical imaging (ISBI 2018)*. p. 1449–53. doi:10.1109/isbi.2018.8363845

39. Lee J, Liu P, Cheng J, Fu H. A deep step pattern representation for multimodal retinal image registration. *IEEE/CVF International Conference on Computer Vision (ICCV)* (2019). doi:10.1109/iccv.2019.00518

40. De Silva T, Chew EY, Hotaling N, Cukras CA. Deep-learning based multi-modal retinal image registration for the longitudinal analysis of patients with age-related macular degeneration. *Biomed Opt Express* (2021) 12(1):619–36. eng. Epub 2021/02/02. doi:10.1364/boe.408573. Cited in: Pubmed; PMID 33520392

41. Ding L, Kuriyan AE, Ramchandran RS, Wykoff CC, Sharma G. Weakly-supervised vessel detection in ultra-widefield fundus photography via iterative multi-modal registration and learning. *IEEE Trans Med Imaging* (2020) 40:2748–58. eng. Epub 2020. doi:10.1109/tmi.2020.3027665./09/30Cited in: Pubmed; PMID 32991281

42. Luo G, Chen X, Shi F, Peng Y, Xiang D, Chen Q, et al. Multimodal affine registration for ICGA and MCSL fundus images of high myopia. *Biomed Opt Express* (2020) 11(8):4443–57. eng. Epub 2020/09/15. doi:10.1364/boe.393178. Cited in: Pubmed; PMID 32923055

43. Tian Y, Hu Y, Ma Y, Hao H, Mou L, Yang J, et al. Multi-scale U-net with edge guidance for multimodal retinal image deformable registration. *Annu Int Conf IEEE Eng Med Biol Soc* (2020) 2020:1360–3. eng. Epub 2020/10/07. doi:10.1109/embc44109.2020.9175613. Cited in: Pubmed; PMID 33018241

44. Zhang J, An C, Dai J, Amador M, Bartsch D, Borooah S, et al. Joint vessel segmentation and deformable registration on multi-modal retinal images based on style transfer *Proceedings of the IEEE international conference on image processing* New Jersey, United States: IEEE (2019). p. 839–43. doi:10.1109/icip.2019.8802932

45. Wang Y, Zhang J, An C, Cavichini M, Jhingan M, Amador-Patarroyo MJ, et al. A segmentation based robust deep learning framework for multimodal retinal image registration *Proceedings of the Icassp 2020 - 2020 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (2020). p. 1369–73. doi:10.1109/icassp40776.2020.9054077

46. Zhang J, Wang Y, Dai J, Cavichini M, Bartsch DUG, Freeman WR, et al. Two-step registration on multi-modal retinal images via deep neural networks. *IEEE Trans Image Process* (2022) 31:823–38. doi:10.1109/TIP.2021.3135708

47. Wang Y, Zhang J, Cavichini M, Bartsch DG, Freeman WR, Nguyen TQ, et al. Robust content-adaptive global registration for multimodal retinal images using weakly supervised deep-learning framework. *IEEE Trans Image Process* (2021) 30:3167–78. eng. Epub 2021. doi:10.1109/tip.2021.3058570./02/19Cited in: Pubmed; PMID 33600314

48. Cavichini M, An C, Bartsch DG, Jhingan M, Amador-Patarroyo MJ, Long CP, et al. Artificial intelligence for automated overlay of fundus camera and scanning laser ophthalmoscope images. *Transl Vis Sci Technol* (2020) 9(2):56. Epub 2020/11/12. doi:10.1167/tvst.9.2.56Cited in: Pubmed; PMID 33173612

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 11 October 2022
DOI 10.3389/fphy.2022.1037821

Check for updates

OPEN ACCESS

EDITED BY
Minbiao Ji,
Fudan University, China

REVIEWED BY
Wei Zheng,
Shenzhen Institutes of Advanced Technology (CAS), China
Jiong Ma,
Fudan University, China

*CORRESPONDENCE
Ke Si,
kesi@zju.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Optics and Photonics,
a section of the journal
Frontiers in Physics

RECEIVED 06 September 2022
ACCEPTED 26 September 2022
PUBLISHED 11 October 2022

CITATION
Xu X, Zheng Y, Xiao X, Wang Z, Lv J, Wang Y, Zhang J, Duan S and Si K (2022), FOCMS: An ultrafast optical clearing method with quantified analysis. *Front. Phys.* 10:1037821. doi: 10.3389/fphy.2022.1037821

COPYRIGHT
© 2022 Xu, Zheng, Xiao, Wang, Lv, Wang, Zhang, Duan and Si. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

FOCMS: An ultrafast optical clearing method with quantified analysis

Xiaobin Xu[1,2,3†], Yameng Zheng[2,3†], Xiao Xiao[2,4†], Zizheng Wang[4], Jie Lv[5], Yongjie Wang[6], Jianmin Zhang[6], Shumin Duan[2,3] and Ke Si[1,2,3,4,7*]

[1]Department of Psychiatry of the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2]MOE Frontier Science Center for Brain Science & Brain-Machine Integration, NHC and CAMS Key Laboratory of Medical Neurobiology, School of Brain Science and Brain Medicine, Zhejiang University, Hangzhou, China, [3]Liangzhu Laboratory, Zhejiang University Medical Center, Hangzhou, China, [4]College of Optical Science and Engineering, Zhejiang University, Hangzhou, China, [5]School of Software Technology, Zhejiang University, Ningbo, China, [6]Department of Neurosurgery of the Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China, [7]Intelligent Optics and Photonics Research Center, Jiaxing Research Institute, Zhejiang University, Jiaxing, China

Optical clearing technology offers a prospective solution to improve the imaging depth and quality of optical microscopy, but there is still a lack of quantitative standards to accurately evaluate transparency effects so the composition and concentration of most reagents are not optimal. Here, we propose a transparency quantitative analysis method (TQAM) based on the tissue area recognition technique to achieve the high-throughput reagent concentration gradient screening. After optimizations of reagent composition, concentration, operation time and other parameters of the optical clearing, we develop a new ultrafast optical clearing method with quantified analysis (FOCMS) with excellent transparency effect, simple operation, improved imaging depth and quality, minor morphological change and outstanding fluorescence retention. Applied the FOCMS to an application of human brain tissue, significant differences are observed between glioma and normal human brain tissue, while these differences are difficult to be found without the assistance of FOCMS. Therefore, FOCMS shows great application potential in clinical diagnosis and treatment, pathological analysis and so on.

KEYWORDS

ultrafast optical clearing, transparency quantitative analysis, human brain, glioma, deep tissue imaging

Introduction

High resolution optical imaging of deep biological tissue shows great significance in the fields of tissue structure and medical diagnosis [1]. However, the strong scattering and high turbidity in biological tissue greatly hinder the penetration capability and imaging quality of the optical imaging system. Various optical imaging systems such as laser scanning confocal imaging (LSCM) [2], multiphoton imaging [3], stimulated Raman

scattering imaging [4], and photoacoustic imaging [5] have been developed to improve the penetration depth, but they are usually suffered from low imaging speed and high cost [6]. Moreover, these optical imaging systems are not enough to satisfy the requirements for biological studies such as whole brain or organ imaging [7, 8]. Optical clearing is a promising method utilizing refractive index matching and tissue decolorization, which greatly reduces the scattering of biological tissue and significantly improves the penetration depth and resolution [9–12].

There are three major clearing methods including organic solvent-based tissue clearing, aqueous-based tissue clearing, and hydrogel embedding tissue clearing. The DISCO [13, 14] technology based on organic solvents achieves the transparency of large-volume biological tissues, and transparency of intact human organs has been obtained with the development of the SHANEL method [7]. However, this organic solvent-based tissue clearing method causes tissue deformation, especially for large tissues and requires a complex operation. The aqueous-based tissue clearing technologies such as SeeDB [15], Scale [16], and CUBIC [17], have successfully achieved versatile whole-organ staining and imaging [17, 18]. They have the advantages of simple and safe operation, and good fluorescence preservation, but suffer from long clearing time and poor clearing efficiency. The advancement of the hydrogel embedding tissue clearing method well improves the optical clearing efficiency and maintains the integrity of the tissue morphology. However, high concentrations of detergents and harsh treatments may result in the loss of native biomolecules and damage to tissue architecture [12, 19, 20]. Optical clearing method is often combined with light sheet microscopy to achieve rapid three-dimensional imaging of organs and even the whole body [21]. The previous clearing methods are designed to be transparent to large biological tissues, however, they have defects of either severe fluorescence quenching, long clearing time, morphology distortion or complicated operations, which seriously limit the applications in combing with commonly used optical imaging systems such as LSCM and two-photon microscopy. We recently proposed a fast optical clearing method (FOCM) with simple protocols and common reagents, achieving efficient transparency, easy operation and less fluorescence toxicity [22]. Current optical clearing approaches use human eyes to evaluate the transparency effect in most scenarios, however, there is still a lack of accurate quantitative evaluation standards for the transparency effect. Thus, the composition and concentration of reagents are not optimal.

To solve this issue, quantitative measurements of optical clearing have recently been developed using spectrophotometer [8] and microplate reader [23], but these are not widely used due to the high cost of measurement devices. Here, we propose the TQAM based on the brain slice area recognition technique, which achieves high-throughput gradient screening of reagent concentrations. This concentration gradient screening benefits the optimization of the reagent composition, concentration, operation time and other parameters of the optical clearing, and the mutual comparison of various optical clearing reagents. Through the concentration gradient screening, we propose a novel ultrafast optical clearing method, FOCMS, only consisting of DMSO and urea. This paper focuses on optimizing the optical clearing performances on time, imaging depth and quality, transparency effect and fluorescence retention, and expands the application scenarios of FOCMS optical clearing technology. Taking human glioma as an example, we find that there are distinct differences between glioma and normal human brain tissue, however, these differences are difficult to be observed if the tissue is not cleared by FOCMS. Therefore, the FOCMS optical clearing technology is expected to greatly promote its application in clinical diagnosis and treatment, pathological analysis and other fields with the combination of optical imaging systems.

Material and method

Experimental procedure and method

For traditional immunofluorescence imaging, the mice were first subjected to operations including anesthesia, perfusion, brain removal, and sectioning, followed by immunofluorescence staining of brain slices, and finally mounting and imaging, shown in Figure 1A. To improve the quality of immunofluorescence imaging, FOCMS is a method that contains optical clearing after immunofluorescence staining, and this optical clearing process only takes 5 min to complete which is less than the imaging preparation time.

Immunofluorescence staining. After decolorization, brain slices were incubated with primary antibodies in PBST (0.01 M PBS with 0.1% Triton X-100) at 37°C for 48–96 h with shaking, followed by washing at 37°C for 1 h in PBST three times. Then incubated with secondary antibodies under the same conditions.

Optical Clearing Method. PBST should be sopped up by KimWipes after washing and labeling samples. The sample only needs to incubate for 3–10 min using 600 μl FOCMS reagent at room temperature in dishes. After FOCMS clearing Figure 1B, the sample should be mounted (Figure 1C) immediately to isolate air.

Laser Scanning Confocal Microscopy (LSCM). An inverted confocal microscopy FV1000 (Olympus) was used to perform fluorescence imaging of the brain section. A He-Ne laser (543.5 nm) and laser diode (473 and 635 nm) were used as the light source. To acquire the fluorescent images, we used different objective lenses including a ×20 air objective lens (Olympus, UPLSAPO ×20, N.A.0.75, WD 0.6 mm), ×40 air objective lens (Olympus, UPLSAPO ×40, N.A.0.95, WD 0.18 mm) and ×60 oil objective lens (Olympus,

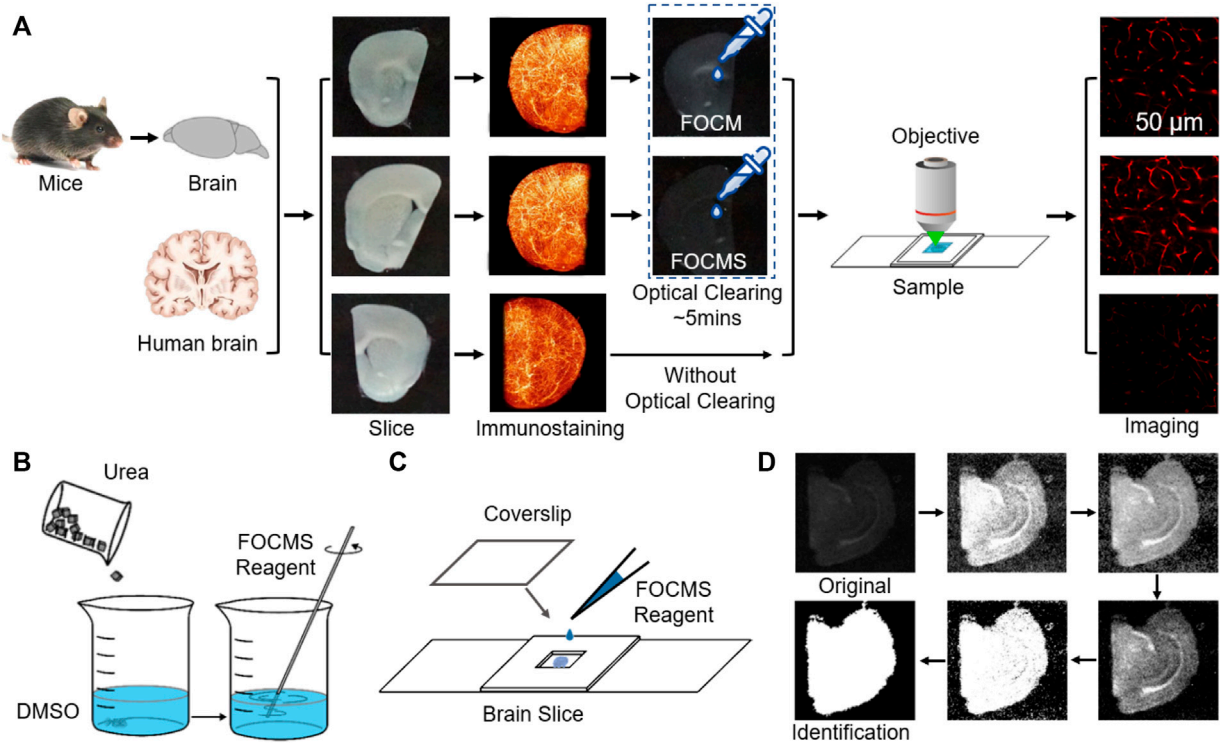


FIGURE 1
FOCMS used for immunostaining imaging and tissue recognition. **(A)** Optical imaging process including brain preparation, slicing, immunostaining, optical clearing and imaging. **(B)** Reagents preparation of FOCM. **(C)** Biological sample clearing and mounting procedure before optical imaging. **(D)** Brain slice range recognition process. The six images are results of normalized, enhancement, pre-recognition, differentiated processing and denoising, binarization of customized threshold and removing isolated noise, respectively. White: brain area, black: background. The sum of white pixels is the area of the brain slice.

UPLSAPO ×60, N.A.1.35, WD 0.15 mm). For maximum intensity projection, the imaging interval on z axis was set as the half of z-axis resolution of different objectives (1.09 μm per slice for ×20 objective, and 0.48 μm per slice for ×60 objective).

Three-Dimensional Reconstruction. The reconstruction method is to find an extremum after preprocessing of normalization, denoising, removing isolated noise, and then use this extremum as the center to judge whether the 26 adjacent points satisfy the requirements of the threshold condition. If the point satisfies the threshold condition, we judged it as cell tissue. The coordinates of cell tissues are considered as the new center, and the fluorescence intensity at the center is used as the new baseline of the threshold. Then it is judged whether the intensity of the surrounding points of the new center satisfies the new threshold condition. Repeat this until no extremum meets the threshold limit and then rejudge whether the reconstructed structure is a cell. Set the reconstructed structure as zero and loop the above process with the next extreme point until there is no extreme point that meets the preset condition.

Transparency Quantitative Analysis Method (TQAM). Transparency was measured by placing the sample in a Petri dish with a black background. The optical clearing process was imaged using a camera (Sony ilce-6000) with the same parameters. The white light source was used to provide uniform brightness and the distance between the sample and the camera was set as 4 cm.

TQAM was achieved using MATLAB R2020a. The detailed procedures are: first, normalize and automatically identify the brightness of the captured images with the assistance of black borders and white grids, and find the black area where the brain slice is located for cropping. Second, enhance the image contrast. Third, binarization and isolated noise are performed to obtain the first pre-identified brain slice area. The threshold of the black area is generally set at a large value to remove noise, while the threshold of the white area sets a small value to prevent the ventricle from being mistaken for an area with brain tissue. However, the coordinates of the brain slice region pre-identified for the first time are often inaccurate because there is no significant difference in average light intensity between the brain slice and outside of the brain slices. To address this,

differentiate the original image to reduce the noise intensity of the background using coordinates of a pre-identified brain slice. Finally, binarize, and remove the isolated noise to obtain the region where the brain slice is located, as shown in Figure 1D.

The means of brightness of the brain area before and after optical clearing are I_{t0} and I_t, respectively. And the average values of the area outside the brain slice before and after optical clearing are I_{b0} and I_b, so the transparency is defined as:

$$\mathbf{T} = \mathbf{1} - \frac{|\boldsymbol{I_t} - \boldsymbol{I_b}|}{|\boldsymbol{I_{t0}} - \boldsymbol{I_{b0}}|}. \quad (1)$$

When T = 0 without optical cleaning, the closer T is to 1, the better the optical cleaning effect. It is worth mentioning that the method is also suitable for the transparency evaluation of other biological tissues and optical clearing methods with proper parameter adjustment. For colored tissues, the brain slice needs to be placed on a Petri dish with a white background, and then the optical clearing process is photographed using a camera. Finally, the tissue area identification and transmittance calculation are performed based on the images.

Sample preparation

Mice and Mouse Brain Samples. Adult mice C57BL/6 and Thy-GFP-M mice were used in the optical clearing and imaging process. Mice were rapidly anesthetized with chloral hydrate (5% wt/vol, 0.1 ml/10 g, intraperitoneal (i.p.)), then transcardially perfused with ice-cold 0.01M PBS (Solarbio) and paraformaldehyde (4% in PBS wt/vol, Sinopharm Chemical Reagent Co., Ltd.). Brain tissues were incubated in the same paraformaldehyde solution at 4°C for 24–48 h for uniform fixation. After fixation, 300-μm-thick brain slices were sectioned using a vibrating slicer (VT 1200S, Leica).

Human Brain Sample. Human brain tissues were fixed in paraformaldehyde at 4°C for 24 h and then 300 μm thick brain slices were sectioned using a vibrating slicer. Before the immunofluorescence staining and optical clearing, the brain slices were incubated in CHAPS and N-Methyldiethanolamine solutions at 4°C for 24 h to remove the influence of the blood.

Antibodies Selection. Primary antibodies: chicken polyclonal anti-GFAP antibody (Abeam, ab4674, dilution 1: 250) for labeling astrocytes; goat polyclonal anti-CD31 antibody (R&D system, AF3628, dilution 1:100) for labeling human vascular endothelial cells; mouse polyclonal anti-CD31 antibody (R&D system, BBA7, dilution 1:100) for labeling vascular endothelial cells.

Secondary antibodies: donkey anti-chicken IgY H&L (FITC) (abcam, ab63507); goat anti-chicken IgY H&L (Alexa Fluor® 647) (abcam, ab150175); donkey anti-goat IgG H&L (Alexa Fluor® 555) (abcam, ab150134); donkey anti-mouse IgG H&L (Alexa Fluor® 647) (abcam, ab150107).

Optical Clearing Reagents Preparing. FOCMS reagent was prepared only using 30% (wt/vol) urea dissolved in DMSO, shown in Figure 1B. FOCM reagent was prepared as 30% wt/vol urea, 20% wt/vol Dsorbitol, and 5% wt/vol glycerol dissolved in DMSO. When preparing the reagent, urea and Dsorbitol were dissolved in DMSO and stirred at room temperature (25°C) overnight. After complete dissolution, glycerol was added and stirred further. The reagents can be stored at room temperature for several months and shaken gently before use.

Results

Optimization of FOCMS using concentration gradient screening

The TQAM method accurately calculates the optical clearing capacity and the area change of the brain tissue, so this method is used for high throughput gradient screening of the optical clearing reagents. Through concentration gradient screening of three solvents (urea, sorbitol and glycerol) used in FOCM for optical clearing, it is found that the optical clearing ability becomes better for the 300 μm brain slice with the decrease of sorbitol concentration (Figure 2A), when the concentration of urea and glycerol are 30% and 5% wt/vol, respectively. At the sorbitol concertation of 20% wt/vol, the transparency of brain slices is only 0.56 within 5 min, while it reaches 0.79 when the sorbitol concertation drops to 0% wt/vol, shown in Figure 2B. The transparency using FOCM achieves 0.76 after 30 min optical clearing, while that of using FOCM reagent with 0% wt/vol sorbitol is up to 0.87 under the same condition. Besides, the FOCM reagent causes severe shrinkage of the brain slice which is less than 60% of the original area and cannot recover to the original size, as shown in Figure 2C. In comparison, the FOCM reagent with 0% wt/vol sorbitol shrinks the brain slice morphology and then swells back to its original size, indicating better tissue morphology preserving ability than that of FOCM. Therefore, the sorbitol in the FOCM reagent reduces the optical clearing effect and causes morphology deterioration. Besides, we found that the increase of glycerol concentration has less influence on optical clearing ability, and deteriorates the brain slice morphology at the urea concentration of 30% wt/vol and the sorbitol concentrations of both 0% and 20% wt/vol.

Compared to FOCM, FOCMS (without sorbitol and glycerol) has the advantages of simpler operation, faster transparency time, better optical clearing effect and less tissue deformation (Figure 2D), especially in areas with higher fiber content and tighter tissues. Using the same concentration gradient screening technology, we analyzed the urea concentration in FOCMS and found that the transparency of the brain slice reaches 0.65 when the urea concentration is 0% (only DMSO solvent). The transparency ability improves with the increase of urea

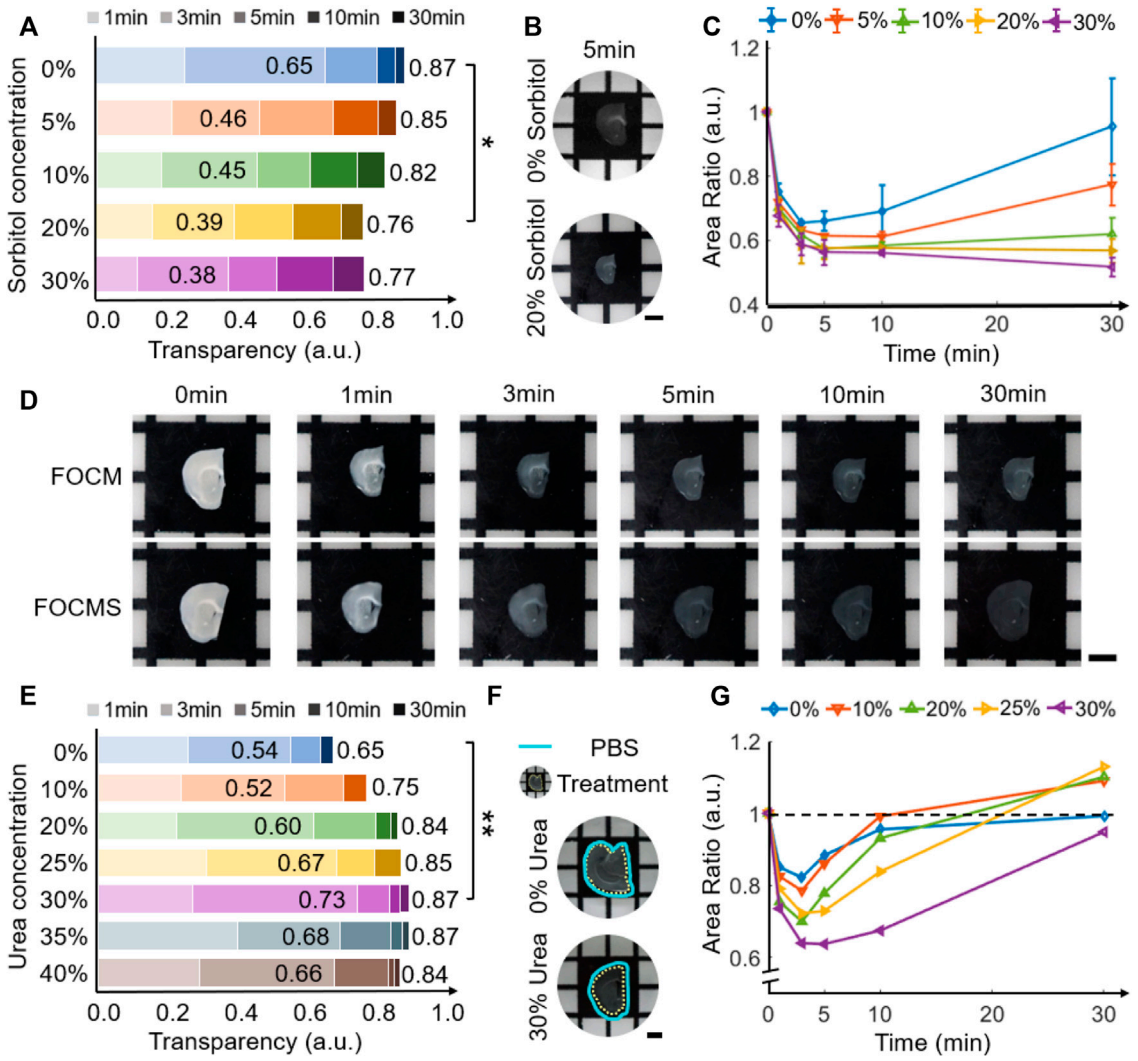


FIGURE 2
FOCM and FOCMS concentration gradient screening. **(A)** Variation of transparency with different sorbitol concentrations in FOCM reagent. **(B)** Comparison of transparency of brain slices using FOCM reagent with sorbitol concentration of 0% and 20% wt/vol for 5 min. **(C)** Changes of brain tissue area with sorbitol concentration in FOCM reagent. **(D)** Comparison of optical clearing effects between FOCM and FOCMS at 0 min, 1 min, 3 min, 5 min, 10 min and 30 min **(E)** Variation of transparency with different urea concentrations in FOCMS reagent. **(F)** Morphology caused by FOCMS reagents. The cyan solid line is the slice boundary before optical clearing, and the yellow dotted line is after optical clearing. **(G)** Changes of brain tissue area with urea concentration in FOCMS reagent. Scaler bar: 2 mm, statistical significance (**$p < 0.01$, *$p < 0.05$).

concertation from 0% to 30% wt/vol, shown in Figure 2E. At the urea concentration of 30% wt/vol, the transparency of brain slice using FOCMS is 0.73 at 3 min, then increase to more than 0.8 at 5 min and finally reaches up to 0.87, while that of using FOCM is up to 0.76 at the same condition. When the urea concentration continues to rise, the optical clearing effect deteriorates. Thus, FOCMS has the best optical clearing ability when the urea concentration is 30% wt/vol. During the optical clearing using FOCMS, brain slices shrink rapidly within 30 min first, and reach the most at 3 min, but gradually recover. Moreover, the morphology of brain slices has severe changes with the rise of the urea concentration increase. When the urea concentration is 30% wt/vol, the area shrinks the most at 3 min, reaching about 64% of the original area, and then gradually recovers, shown in Figure 2F,G. When the urea concentration exceeds 30% wt/vol, the changes in brain area are similar to that of 30% wt/vol. The area difference between and after optical clearing is less than 15% and there are minor changes after 30 min of the optical clearing. Therefore, optical clearing has an optimal time of about 5 min.

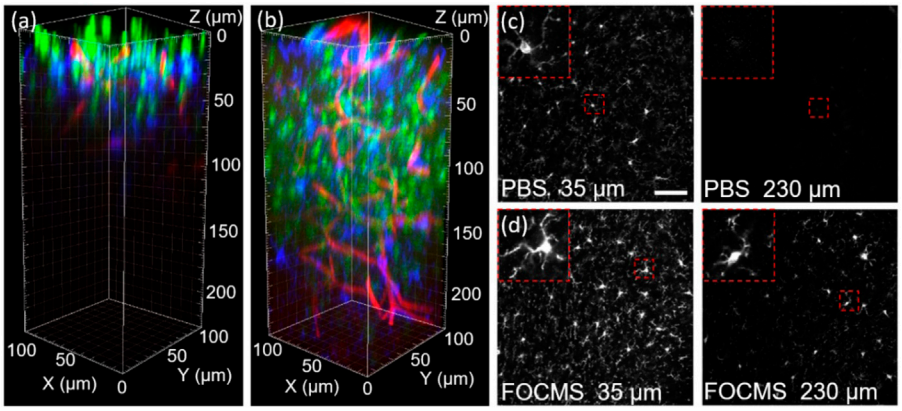


FIGURE 3
Fluorescence imaging of brain tissue **(A)** before and **(B)** after optical clearing. Green: nucleus, blue: glia cell, red: blood vessel. Comparison of glia cell imaging at the depth of 35 μm and 230 μm **(C)** before and **(D)** after clearing, respectively. Scalebar: 50 μm.

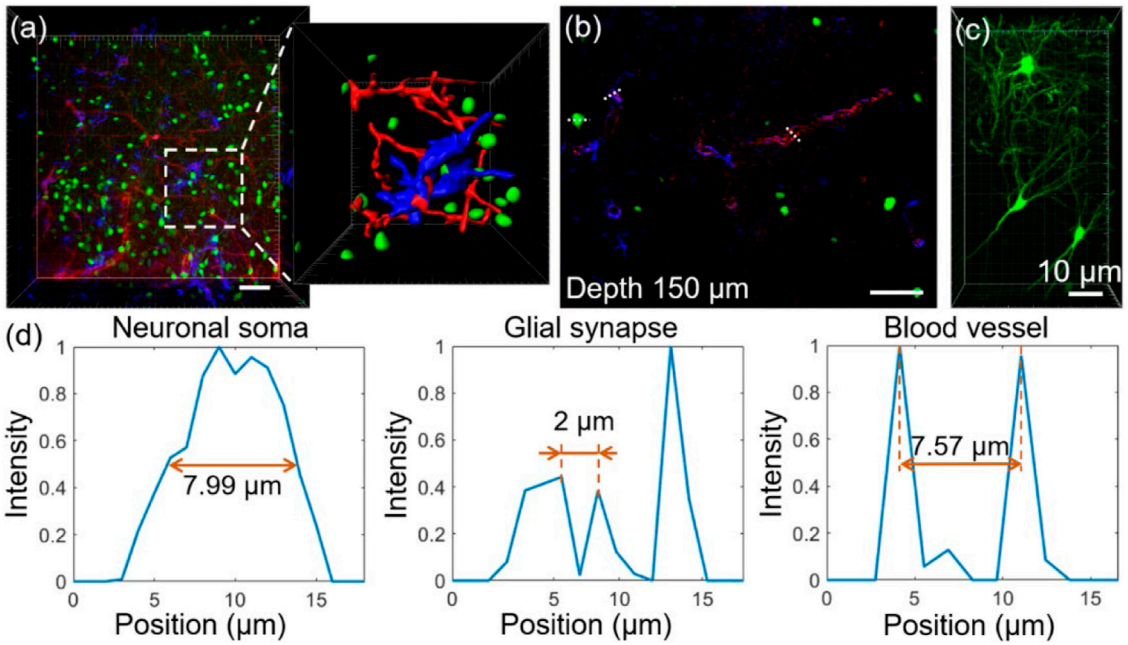


FIGURE 4
Fluorescence imaging quality after FOCMS. **(A)** Imaging of brain tissue with detailed structure exhibition after FOCMS clearing. Green: neurons, blue: glia cell, red: blood vessel. **(B)** Selected single imaging at the depth of 150 μm in **(A)**. **(C)** Imaging of neurons with high NA objective lens. Scale bar: 10 μm **(D)** Normalized fluorescence corresponding to the neuronal Soma, glial synapse, and blood vessel at the depth of 150 μm in **(B)** indicated by the dashed line, respectively. Scale bar: 50 μm.

Imaging depth and quality improvement using FOCMS

FOCMS achieves rapid optical clearing for biological tissues with simple operation, which greatly benefits the improvement of the imaging depth and resolution of fluorescence microscopy. Figure 3 compares the imaging of brain tissues before and after optical clearing using LSCM. Without optical clearing, the range of clear imaging is only about 50 μm. The SNR ratio is greatly reduced after the imaging depth of 50 μm. The SNR ratio drops

with the increase of imaging depth. At the depth of 100 μm, no more effective information is obtained when the SNR ratio is close to 0 (Figure 3A). However, the imaging depth improves to 230 μm using FOCMS, and the SNR maintains almost unchanged, shown in Figure 3B. Figures 3C,D are the imaging of glia cells at the depth of 35 and 230 μm before and after FOCMS clearing, and the upper left corner is an enlarged view of a single glia cell.

Comparing the results of Figures 3C,D, it can be seen that the synaptic intensity of glia cells after optical clearing is significantly stronger than that of before clearing, the number of glia cells is greater and the morphology preserves better, indicating that optical clearing using FOCMS significantly improves the imaging quality even in the imaging depth of LSCM (50 μm). When the depth reaches 230 μm, there is almost no effective information remaining before optical clearing, while after light clearing, some of the glial cells are still resolved, and the synapses are clearly visible.

We further evaluated the results of deep tissue imaging using the FOCMS clearing technique. Figure 4A shows the LSCM imaging after FOCMS clearing, in which neurons are in green, glial cells in blue, and blood vessels in red. The sub-image on the right is the enlarged part of the white dashed box in Figure 4A to show the detailed three-dimensional morphological structures of neurons, glial cells and blood vessels. It is clearly seen that the synapses of glial cells and the vascular walls of blood vessels in Figure 4B show the selected single imaging at the depth of 150 μm in (a). LSCM imaging with high NA clearly displays the dendrites and axons of neurons (Thy1-GFP-M mice) with an imaging resolution higher than 1 μm, as shown in Figure 4C. Figure 4D shows the normalized fluorescence intensity of neuronal Soma, glial synapse and blood vessel in (b), where the full width at half maximum (FWHM) of neuronal Soma is 7.99 μm, and the minimum distance between the three glial synapses is 2 μm and the diameter of the blood vessel is 7.57 μm, respectively. It can be found that when the imaging depth is 150 μm, the structure of cells and tissues still has a high signal-to-noise ratio. We also found that the structures of cells and brain tissues still have a high signal-to-noise ratio at the imaging depth of 150 μm. Therefore, the FOCMS optical clearing technology on brain tissues shows significant improvements in the imaging depth of the microscopic imaging system. Besides, the FOCMS does not quench the exogenous fluorescence like immunofluorescence and chemical dyes and the endogenous fluorescence such as GFP, indicating its excellent fluorescence retention ability.

Applications in human brain tissue with FOCMS

FOCMS significantly improves the depth and quality of imaging with minimal time and effort which shows great potential in the applications of clinical diagnosis and treatment, pathological analysis, *etc.* Here, we demonstrated an application using FOCMS in human brain tissue imaging. As shown in Figures 5A,B describe the three-dimensional LSCM imaging of normal human cerebral blood vessels and brain glioma (GBM) without optical clearing, respectively. The top left image is the original imaging result, the top right image is the detailed structure of blood vessel morphology, and the bottom two images are the selected single imaging at the depth of 50 μm (left) and 200 μm (right) in (a) (b), respectively. Unfortunately, using these imaging results without optical clearing, the morphological differences are not sufficient to distinguish between normal human brain tissue and glioma, and pathological analysis cannot be performed.

After using FOCMS, three-dimensional imaging of normal human brain tissue and glioma are shown in Figures 5C,D, respectively. Among them, the top left image is the original result of imaging, the top right is the detailed structure of blood vessel morphology, and the bottom two images are the selected single imaging at the depth of 50 μm (left) and 200 μm (right), respectively. Owning to the FOCMS optical clearing, the morphological differences between normal human brain tissue and glioma are obvious. Compared to normal blood vessels, blood vessels of glioma are lump-shaped, with tortuous, disordered structure, and enlarged diameter. Therefore, FOCMS technology demonstrates the ability that it can be used to solve the problem of unclear boundary identification of glioma.

Unlike mouse brain tissues, human brain tissues have higher fiber content, greater cell density, tighter tissues, and many impurities, which severely reduces the optical clearing effect. Compared to FOCM, we found that FOCMS technology has a better optical clearing effect on human brain tissue and smaller morphological changes, shown in Supplementary. Comparing the results in the white dotted circles in Figures 5A,C, it is found that there is no significant difference in the morphology of human cerebral blood vessels before and after optical clearing. Figure 5E,G shows the three-dimensional morphology of normal human cerebral blood vessels before (yellow) and after (red) FOCMS optical clearing, respectively, and found that they are well merged. Besides, there is no significant difference in the vascular morphology of gliomas before and after FOCMS optical clearing as well, indicating that the FOCMS technology will not cause obvious deviation in the imaging results of brain tissues.

Since there is no ideal imaging for comparison in practical applications, we use a three-dimensional reconstruction algorithm to separate the signal from the noise for the calculation of the signal-to-noise ratio (SNR). SNR is defined as:

$$\mathbf{SNR} = \mathbf{10} \times \mathbf{log}_{\mathbf{10}}\frac{\boldsymbol{I_s}}{\boldsymbol{I_n}}, \tag{2}$$

10.3389/fphy.2022.1037821

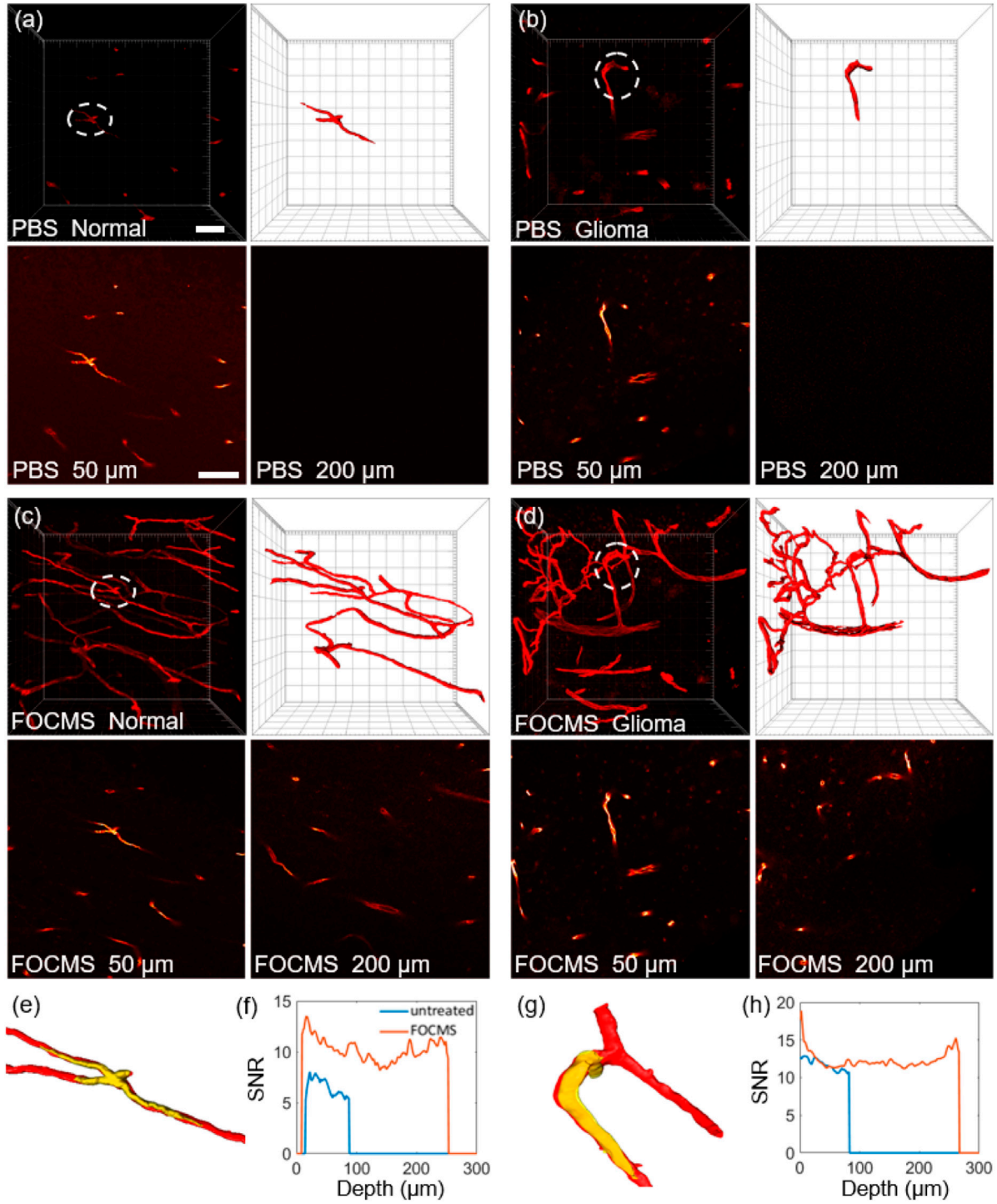


FIGURE 5
Application of FOCMS in human brain tissue imaging. Comparison of three-dimensional imaging of **(A)** normal human brain tissue and **(B)** glioma without optical clearing. Top left: the original imaging, top right: detailed structure of blood vessel morphology, bottom: selected single imaging at the depth of 50 μm (left) and 200 μm (right) in **(A)** and **(B)**, respectively. After using FOCMS, three-dimensional imaging of **(C)** normal human brain tissue and **(D)** glioma, respectively. Top left: the original imaging, top right: detailed structure of blood vessel morphology, bottom: selected single imaging at the depth of 50 μm (left) and 200 μm (right) in **(C)** and **(D)**, respectively. Comparison of **(E)** morphology changes and **(F)** SNR of normal brain tissue before and after FOCMS. **(G)** Morphology changes and **(H)** SNR for glioma with and without optical clearing. Scale bar: 50 μm.

where I_s the mean of signal and I_n is the mean of noise. The final SNR of multiple immunofluorescence staining is the mean of each channel.

Figures 5F,H represent the SNR of 3D imaging at different imaging depths with normal human brain tissue and glioma, respectively. The blue line represents before optical clearing and the orange line represents after optical clearing. For both normal human brain tissue and glioma, the results show that the SNR after optical clearing is higher than that of before optical clearing at the imaging depth of 50 μm (the range of LSCM) and 200 μm. The mean SNR of normal human brain tissue before optical clearing is 6.54, while it reaches 10.23 after FOCMS clearing. For glioma tissue, the mean SNRs are 11.72 and 12.33 before and after optical clearing, respectively. The SNR of glioma tissue is higher than that of normal human brain tissue, which may be due to the changes in the composition of brain tissue during carcinogenesis. With the increase of imaging depth, the SNR without optical clearing gradually decreases. However, the SNR after FOCMS clearing decreases first and then increases, mainly due to the staining effect of the middle position of the tissue during the staining process being worse than that of the edge position. In summary, the FOCMS method effective improve the imaging depth and quality of human brain tissues without changing the tissue morphology, demonstrating the applications in clinical diagnosis and treatment and pathological analysis.

Conclusion

In this paper, we propose a new technology, FOCMS, using a new reagent consisting of only DMSO and urea for optical clearing with better performance. The results show that the transparency reaches 0.73 which is an 87.2% improvement compared to that of FOCM when the urea concentration is 30% wt/vol at the optical clearing time of 3 min. It is achieved owing to the TQAM we proposed based on the brain slice area recognition technique. This method accurately calculates the transparency of optical clearing reagents and the morphological changes in brain slices. Furthermore, high-throughput gradient screening is performed on the reagent composition, concentration, processing time and other parameters of the optical clearing technology using this method. Using the TQAM, we analyze the influences on the three solutes (urea, sorbitol and glycerol) of FOCM *via* concentration gradient screening. It is found that the increased sorbitol concentration reduces the optical clearing effect and prevents morphological shrinkage from returning to the pre-clearing area. Moreover, glycerol has less effect on optical clearing and made the area change more severe. Compared to FOCM, FOCMS without sorbitol and glycerol reagents has the advantages of simpler operation, faster transparency, better optical clearing effect, and less morphological change.

The FOCMS technology achieves ultra-fast and effective optical clearing with little fluorescence quenching only by simply incubating the brain tissue sample for about 5 min. After FOCMS clearing, the results show that the glial cells show stronger synaptic signals, better morphology, and significantly improved imaging quality at the imaging depth of 50 μm. Besides, more than 200 μm of imaging depth is achieved, and the SNR keeps almost unchangeable, enabling clear imaging of neurons and glial synapses, blood vessel walls and other tissues with a resolution of more than 1 μm. Moreover, FOCMS technology does not quench most of the exogenous and endogenous fluorescence and has outstanding fluorescence retention ability. The FOCMS technology has been successfully applied to the human brain tissue and it is shown that FOCMS is superior to FOCM in improvements of optical clearing effect, SNR (56.4% stronger) and preservation of morphology. We also demonstrate an application of solving the problem of unclear identification of brain glioma boundaries, which shows great potential in clinical diagnosis and treatment, pathological analysis, etc.

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

All animal studies were approved by Animal Advisory Committee at Zhejiang University. All of the human specimens were obtained from the Department of Neurosurgery, Second Affiliated Hospital of Zhejiang University School of Medicine. And these studies on human specimens were approved by the Human Research Ethics Committee of the Second Affiliated Hospital of Zhejiang University School of Medicine.

Author contributions

KS, SD, and JZ conceived the concept and designed the study. XXu and JL performed the image processing. YZ, XXi, ZW, and YW performed the experiments. XXu and YZ analyzed the data. All the authors contributed to the preparation of the manuscript.

Funding

National Natural Science Foundation of China (61735016). Key R&D Program of Zhejiang Province (2021C03001). Natural Science Foundation of Zhejiang Province of China (LR22F050007). CAMS Innovation Fund for Medical Sciences (2019-I2M-5-057).

Fundamental Research Funds for the Central Universities. China Postdoctoral Science Foundation (2021M692816).

Conflict of interest

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

Publisher's note

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

References

1. Brenna C., Simioni C., Varano G., Conti I., Costanzi E., Melloni M., Neri L. M. Optical tissue clearing associated with 3D imaging: Application in preclinical and clinical studies. *Histochem Cel Biol* (2022) 157, 497PMC9114043–511. doi:10.1007/s00418-022-02081-5

2. Hu Y., Tang W., Cheng P., Zhou Q., Tian X., Wei X., et al. Monitoring circulating tumor cells *in vivo* by a confocal microscopy system. *Cytometry A* (2019) 95(6):657–63. doi:10.1002/cyto.a.23702

3. Wang S., Li Y., Zhao Y., Lin F., Qu J., Liu L. Investigating tunneling nanotubes in ovarian cancer based on two-photon excitation FLIM-FRET. *Biomed Opt Express* (2021) 12(4):1962–73. doi:10.1364/BOE.418778

4. Liu Z., Su W., Ao J., Wang M., Jiang Q., He J., et al. Instant diagnosis of gastroscopic biopsy via deep-learned single-shot femtosecond stimulated Raman histology. *Nat Commun* (2022) 13, 4050PMC9279377. doi:10.1038/s41467-022-31339-8

5. Vu T., Razansky D., Yao J. Listening to tissues with new light: Recent technological advances in photoacoustic imaging. *J Opt* (2019) 21. doi:10.1088/2040-8986/ab3b1a

6. Xu X, Chen J, Zhang B, Huang L, Zheng Y, Si K, et al. Enlarged field of view based on Schwartz modulation for light sheet fluorescence microscopy in deep tissue. *Opt Lett* (2020) 45(17):4851–4. doi:10.1364/OL.398985

7. Zhao S, Todorov MI, Cai R, -Maskari RA, Steinke H, Kemter E, et al. Cellular and molecular probing of intact human organs. *Cell* (2020) 180, 796PMC7557154–812. doi:10.1016/j.cell.2020.01.030

8. Zhu J, Yu T, Li Y, Xu J, Qi Y, Yao Y, et al. Macs: Rapid aqueous clearing system for 3D mapping of intact organs. *Adv Sci (Weinh)* (2020) 7, 81903185. doi:10.1002/advs.201903185

9. Pende M, Vadiwala K, Schmidbaur H, Stockinger AW, Murawala P, Saghafi S, et al. A versatile depigmentation, clearing, and labeling method for exploring nervous system diversity. *Sci Adv* (2020) 6, eaba0365. doi:10.1126/sciadv.aba0365

10. Xu F, Shen Y, Ding L, Yang CY, Tan H, Wang H, et al. High-throughput mapping of a whole rhesus monkey brain at micrometer resolution. *Nat Biotechnol* (2021) 39(12):1521–8. doi:10.1038/s41587-021-00986-5

11. Geng J, Zhang X, Prabhu S, Shahoei SH, Nelson ER, Swanson KS, et al. 3D microscopy and deep learning reveal the heterogeneity of crown-like structure microenvironments in intact adipose tissue. *Sci Adv* (2021) 7(8):eabe2480. doi:10.1126/sciadv.abe2480

12. Ku T, Guan W, Evans NB, Sohn CH, Albanese A, Kim JG, et al. Elasticizing tissues for reversible shape transformation and accelerated molecular labeling. *Nat Methods* (2020) 17, 609PMC8056749–613. doi:10.1038/s41592-020-0823-y

13. Nudell V, Wang Y, Pang Z, Lal NK, Huang M, Shaabani N, et al. HYBRiD: Hydrogel-reinforced DISCO for clearing mammalian bodies, *Nat Methods* (2022) 19, 479PMC9337799–485. doi:10.1038/s41592-022-01427-0

14. Hong SM, Jung D, Kiemen A, Gaida MM, Yoshizawa T, Braxton AM, et al. Three-dimensional visualization of cleared human pancreas cancer reveals that sustained epithelial-to-mesenchymal transition is not required for venous invasion. *Mod Pathol* (2020) 33(4):639–47. doi:10.1038/s41379-019-0409-3

15. Davis FM, Lloyd-Lewis B, Harris OB, Kozar S, Winton DJ, Muresan L, Watson CJ. Single-cell lineage tracing in the mammary gland reveals stochastic clonal dispersion of stem/progenitor cell progeny. *Nat Commun* (2016) 7, 13053. doi:10.1038/ncomms13053

16. Kolesová H, Čapek M, Radochová B, Janáček J, Sedmera D. Comparison of different tissue clearing methods and 3D imaging techniques for visualization of GFP-expressing mouse embryos and embryonic hearts. *Histochem Cel Biol* (2016) 146(2):141–52. doi:10.1007/s00418-016-1441-8

17. Susaki EA, Shimizu C, Kuno A, Tainaka K, Li X, Nishi K, et al. Versatile whole-organ/body staining and imaging based on electrolyte-gel properties of biological tissues, *Nat Commun* (2020) 11. doi:10.1038/s41467-020-15906-5

18. Tainaka K, Murakami TC, Susaki EA, Shimizu C, Saito R, Takahashi K, et al. Chemical landscape for tissue clearing based on hydrophilic reagents. *Cell Rep* (2018) 24(8):2196–210.e9. doi:10.1016/j.celrep.2018.07.056

19. Murray E, Cho JH, Goodwin D, Ku T, Swaney J, Kim SY, et al. Simple, scalable proteomic imaging for high-dimensional profiling of intact systems. *Cell* (2015) 163, 1500PMC5275966–14. doi:10.1016/j.cell.2015.11.025

20. Park YG, Sohn CH, Chen R, McCue M, Yun DH, Drummond GT, et al. Protection of tissue physicochemical properties using polyfunctional crosslinkers. *Nat Biotechnol* (2018) 37:73–83. doi:10.1038/nbt.4281

21. Avilov SV. Navigating across multi-dimensional space of tissue clearing parameters. *Methods Appl Fluoresc* (2021) 9(2):022001. doi:10.1088/2050-6120/abe6fb

22. Zhu X, Huang L, Zheng Y, Song Y, Xu Q, Wang J, et al. Ultrafast optical clearing method for three-dimensional imaging with cellular resolution. *Proc Natl Acad Sci U S A* (2019) 116(23):11480–9. doi:10.1073/pnas.1819583116

23. Liu L, Xia X, Xiang F, Gao Y, Li X, Li H, Zheng W. F-CUBIC: A rapid optical clearing method optimized by quantitative evaluation. *Biomed Opt Express* (2021) 13, 237–51. doi:10.1364/BOE.442976

frontiers | Frontiers in Physics

TYPE Brief Research Report
PUBLISHED 17 October 2022
DOI 10.3389/fphy.2022.1041577

Check for updates

OPEN ACCESS

EDITED BY
Liwei Liu,
Shenzhen University, China

REVIEWED BY
Juanjuan Zheng,
Xidian University, China
Lei Gong,
University of Science and Technology of China, China

*CORRESPONDENCE
Tianyu Zhao,
zhaotianyu@xjtu.edu.cn
Ming Lei,
ming.lei@mail.xjtu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Optics and Photonics,
a section of the journal
Frontiers in Physics

RECEIVED 11 September 2022
ACCEPTED 03 October 2022
PUBLISHED 17 October 2022

CITATION
Wang M, Zhao T, Wang Z, Feng K, Ren J, Liang Y, Wang S and Lei M (2022), Three-dimensional natural color imaging based on focus level correlation algorithm using structured illumination microscopy.
Front. Phys. 10:1041577.
doi: 10.3389/fphy.2022.1041577

COPYRIGHT
© 2022 Wang, Zhao, Wang, Feng, Ren, Liang, Wang and Lei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Three-dimensional natural color imaging based on focus level correlation algorithm using structured illumination microscopy

Mengrui Wang[1,2†], Tianyu Zhao[1†*], Zhaojun Wang[1], Kun Feng[1], Jingrong Ren[1], Yansheng Liang[1], Shaowei Wang[1] and Ming Lei[1*]

[1]MOE Key Laboratory for Nonequilibrium Synthesis and Modulation of Condensed Matter, Shaanxi Province Key Laboratory of Quantum Information and Quantum Optoelectronic Devices, School of Physics, Xi'an Jiaotong University, Xi'an, China, [2]Department of Engineering Mechanics, SVL and MMML, Xi'an Jiaotong University, Xi'an, China

Taking advantages of high-resolution, natural color restoration, and high imaging speed, optical sectioning structured illumination microscopy (OS-SIM) plays an important role in geology, biology, and material science. However, when encountering chromatic aberration or dealing with samples with semitransparent surface, the HSV (Hue, Saturation, and Value) decoding algorithm suffers intensity deviation and fading color. In this paper, we propose a focus level correlation algorithm for 3D color image reconstruction in OS-SIM. Simulations and experiments demonstrate that the algorithm can restore color of sample authentically, and improve the image processing speed by about 45%. This new algorithm successfully improves the results and the speed of optical sectioning reconstruction, expanding the application of OS-SIM.

KEYWORDS

structured illumination microscopy, optical sectioning, natural color, 3D image, focus level correlation algorithm

Introduction

Due to the advantages of noncontact and minimally invasive observation, optical microscopy plays an essential role in morphology analyzing, which is one of the most important research areas in material science and geological fields [1, 2]. However, the conventional wide-field microscopy cannot obtain optical section directly, as the image obtained by the camera is the superposition of the in-focus information and the defocused background. For this reason, many three-dimensional (3D) imaging technologies have been developed, such as confocal laser scanning microscopy (CLSM) and two photon microscopy [3–5].

CLSM is one of the most widely used technique obtaining high-resolution 3D scans of micrometer-scale specimens [3]. It uses a highly focused laser to scan the sample point by point. The reflective or fluorescence signal is filtered by the detection

pinhole and collected by the photomultiplier tube. Since only the signal excited at the laser focus can pass through the detection pinhole, CLSM can reduce the out-of-focus information and obtain a background-free image [3]. However, CLSM is too time-consuming to implement scans for centimeter-scale creatures, and the focused laser is detrimental for the samples. Furthermore, CLSM ignores the color information of the specimens. Color is an essential feature in geological research, which reflects the optical properties of crystals and is the direct basis for the classification and identification of minerals [6, 7]. In this case, it is necessary to develop an imaging technology which is able to accurately present the natural color of samples with complex microstructures.

Structured illumination microscopy (SIM) is another way obtaining optical section, and has found widely applications for 3D imaging and measurement in biological, medical and geological fields due to its advantages of high spatial resolution, fast imaging speed and natural color imaging [8–11]. The structured illumination was first applied on microscopy by Neil et al. to eliminate the out-of-focus background encountered in wide-field microscopy [12]. OS-SIM uses sinusoidal fringes with different phases to illuminate the sample, which appear only near the focal plane. The contrast of the structured illumination rapidly attenuates in the out-of-focus area and become uniform wide-field illumination. With decoding algorithm, the in-focus information can be separated from the background, and a 3D image of the sample can be reconstructed from optical sections. In our previous work, a DMD (digital micromirror device) based LED-illumination SIM has been reported that is suitable for 3D imaging [13–15]. By using Hilbert transform reconstruction, the image acquisition time can be reduced by 1/3 [16]. An HSV (Hue, Saturation, and Value) decoding algorithm realized full-color optical section [17]. With this well-developed OS-SIM system, we successfully obtained numbers of full-color 3D images of typical insect samples, such as shining leaf chafer, tiger beetles and so on [18]. However, due to the dependence on modulation depth of the HSV decoding algorithm, the sectioning images often suffer from intensity deviation and residual fringes under chromatic aberration. In addition, for samples with semitransparent surface, the modulation depth of the fringes changes considerably between surfaces of different reflective properties, leading to the variation of the optical section intensity and fading color of the reconstructed 3D image.

In this paper, we propose a focus level correlation (FLC) algorithm to reconstruct 3D natural color image in OS-SIM. For samples with varying colors and semitransparent surface like minerals, the results of reconstruction are improved with authentic colors and about 45% faster speed. This new algorithm is expected to expand the application of OS-SIM.

Materials and methods

Principle of optical sectioning with structured illumination microscopy

In the conventional wide-field microscope, the targeted in-focus information is inevitably merged with an unwanted out-of-focus background because the image is limited by the depth-of-field. SIM provides an approach to extract the in-focus information from the raw images by illuminating the sample with a nonuniform structured light. Only the in-focus part is modulated by the structured illumination, and the modulation depth attenuates rapidly as the defocus distance increases. Thus, the in-focus information can be decoded by algorithms. A typical kind of structured illumination is sinusoidal fringe with light intensity distribution like

$$\boldsymbol{S}(\boldsymbol{x}) = \boldsymbol{1} + \boldsymbol{m} \cdot \boldsymbol{\sin}\ (\boldsymbol{2\pi\nu x} + \boldsymbol{\varphi}) \tag{1}$$

where m denotes the modulation depth, ν is the spatial frequency, and φ is the spatial phase. Under the illumination of the sinusoidal fringe, the image of the sample is

$$\boldsymbol{I} = \boldsymbol{I}_{\mathbf{out}} + \boldsymbol{I}_{\mathbf{in}} \cdot \left[\boldsymbol{1} + \boldsymbol{m} \cdot \boldsymbol{\sin}\left(\boldsymbol{2\pi\nu x} + \boldsymbol{\varphi}\right)\right] \tag{2}$$

where I_{out} and I_{in} respectively represent the out-of-focus and the in-focus part of the sample. The optical sectioning image I_{OS} can be obtained by the RMS operation of three raw images with an adjacent phase-shift of 2π/3:

$$\boldsymbol{I}_{\mathbf{OS}} = \frac{\sqrt{\boldsymbol{2}}}{\boldsymbol{3}}\sqrt{\left(\boldsymbol{I}_{\mathbf{0^\circ}} - \boldsymbol{I}_{\mathbf{120^\circ}}\right)^{\boldsymbol{2}} + \left(\boldsymbol{I}_{\mathbf{120^\circ}} - \boldsymbol{I}_{\mathbf{240^\circ}}\right)^{\boldsymbol{2}} + \left(\boldsymbol{I}_{\mathbf{240^\circ}} - \boldsymbol{I}_{\mathbf{0^\circ}}\right)^{\boldsymbol{2}}} \tag{3}$$

In addition, the wide-field image I_{WF} can also be acquired simultaneously with

$$\boldsymbol{I}_{\mathbf{WF}} = \boldsymbol{I}_{\mathbf{out}} + \boldsymbol{I}_{\mathbf{in}} = \frac{\boldsymbol{1}}{\boldsymbol{3}}\left(\boldsymbol{I}_{\mathbf{0^\circ}} + \boldsymbol{I}_{\mathbf{120^\circ}} + \boldsymbol{I}_{\mathbf{240^\circ}}\right) \tag{4}$$

3D color image reconstruction based on focus level correlation algorithm

The most usual way to reconstruct color image is transforming the raw images into RGB (Red, Green, and Blue) space. Results can be received after calculating the RMS of each channel with Eq. 3 and recombining three channels. However, because lights with different wavelength focus on different planes and the modulation depth is related to the focus level, the modulation depths of R, G, and B are not identical in practice. The sectioning images suffer from color distortion caused by the chromatic aberration. In theory, the intensity of the optical sectioning image is proportional to the modulation depth according to Eqs 2, 3. Therefore, the ratio of R, G, and B intensities will change compared with the raw image, leading to the color distortion.

10.3389/fphy.2022.1041577

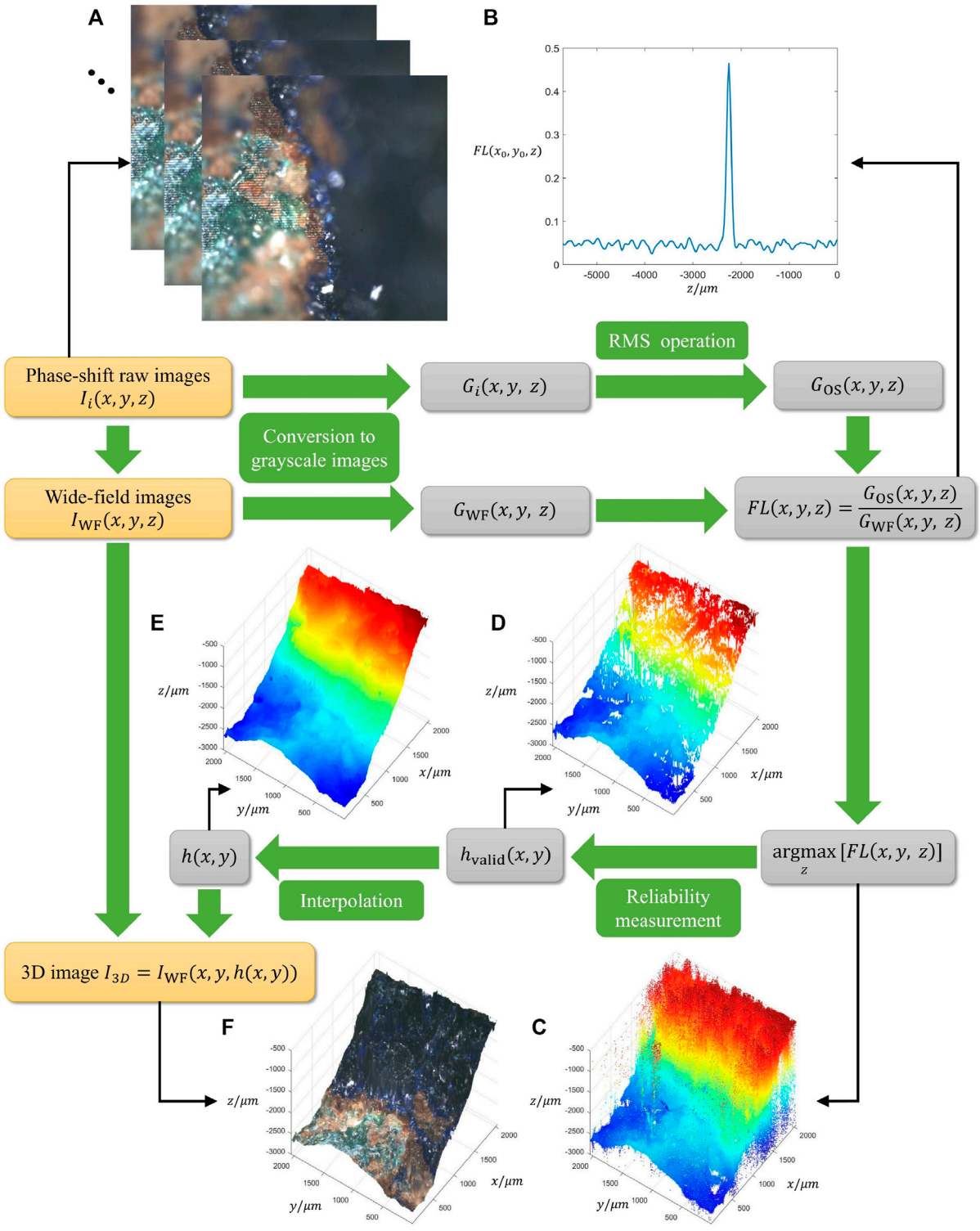


FIGURE 1
Flowchart diagram of the 3D color image reconstruction algorithm based on focus level correlation. **(A)** Phase-shift raw images. **(B)** Modulation depth curve of a pixel (x_0, y_0). **(C)** Height map of maximum focus level. **(D)** Valid height map. **(E)** Height map with interpolation. **(F)** 3D image.

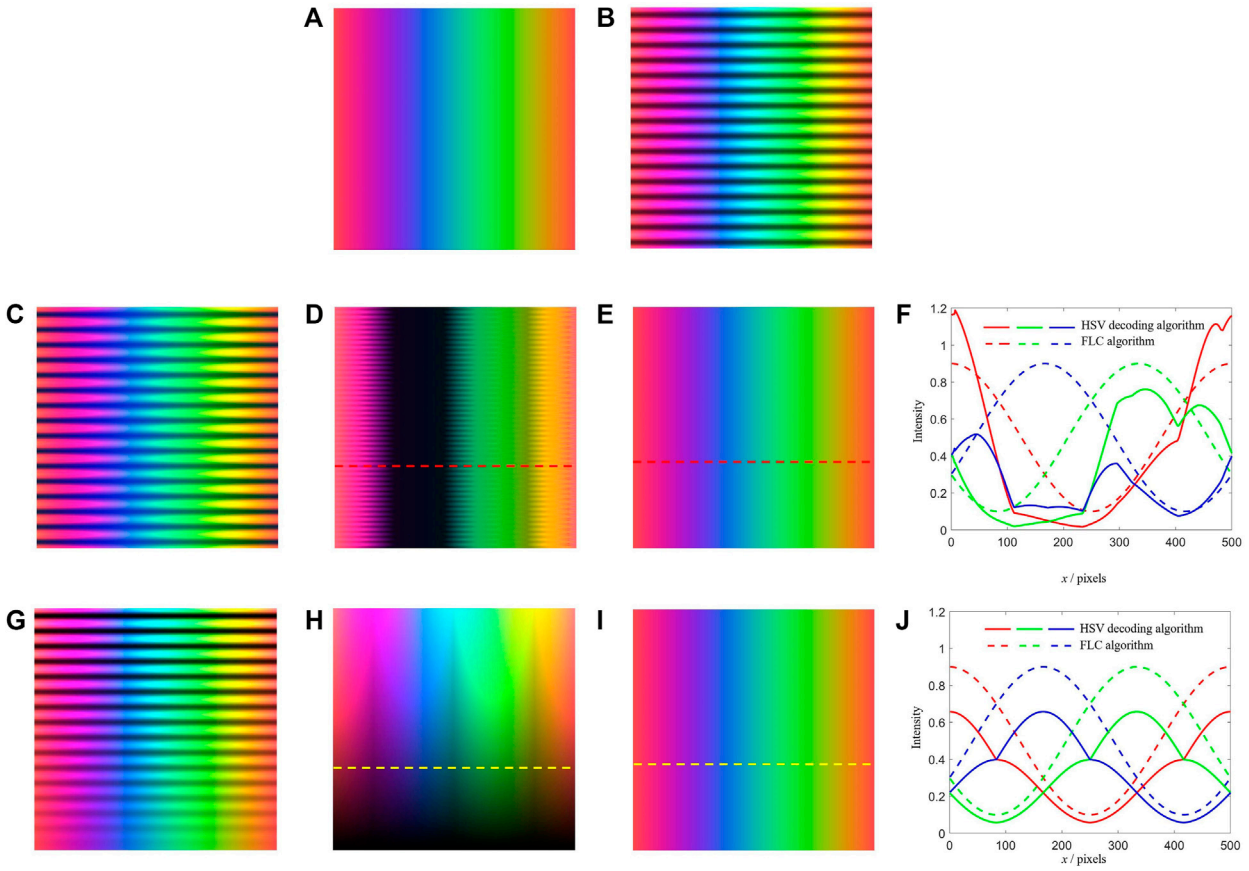


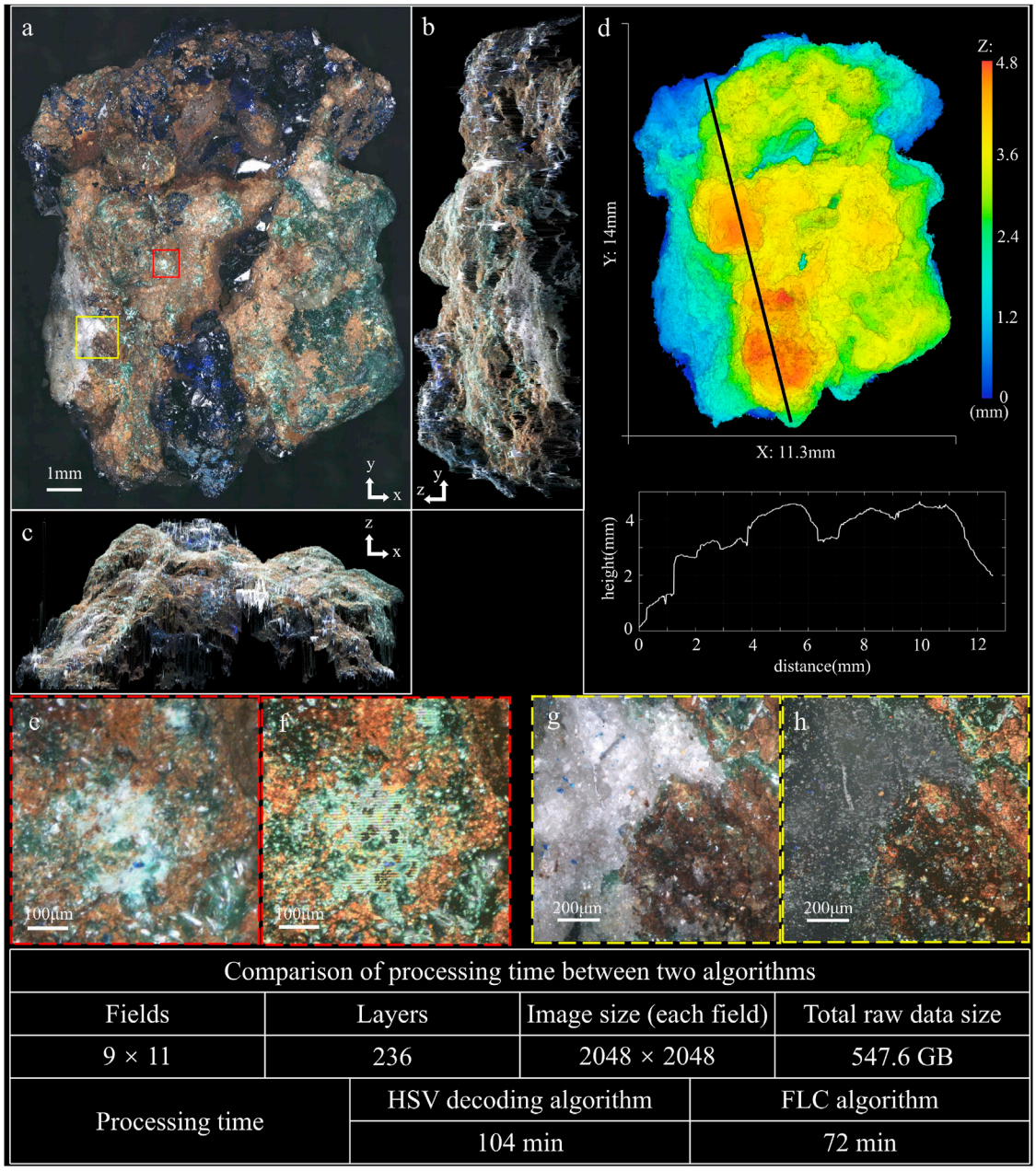
FIGURE 2
Simulation of the HSV decoding algorithm and the FLC algorithm in this paper. **(A)** Raw image. **(B)** Image added with normal sinusoidal fringe. **(C)** Image added with sinusoidal fringe with chromatic aberration. **(D)** Image restored by the HSV decoding algorithm under chromatic aberration. **(E)** Image restored by the FLC algorithm under chromatic aberration. **(F)** Intensity curve of R, G, and B along the red dashed line in **(D,E)**. **(G)** Image added with sinusoidal fringe with modulation depth variation. **(H)** Image restored by the HSV decoding algorithm under modulation depth variation. **(I)** Image restored by the FLC algorithm under modulation depth variation. **(J)** Intensity curve of R, G, and B along the yellow dashed line in **(H,I)**.

To restore the original color, a new 3D color image reconstruction algorithm named focus level correlation (FLC) is employed, whose flowchart is shown in Figure 1. With phase-shift raw images $I_i(x, y, z)$ $(i = 1, 2, 3)$ captured (Figure 1A), the wide-field image $I_{WF}(x, y, z)$ can be directly obtained by Eq. 4. Both of the raw images and wide-field images are converted to grayscale images $G_i(x, y, z)$ and $G_{WF}(x, y, z)$. The optical sections $G_{OS}(x, y, z)$ can be obtained by RMS decoding operation, which indicate the averaged sectioning images of the three colors RGB. To calculate the focus position for each (x, y) and noticing that the intensity of optical section is maximum at the focus position, we define the focus level $FL(x, y, z)$ as the proportion of $G_{OS}(x, y, z)$ and $G_{WF}(x, y, z)$:

$$\boldsymbol{FL}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) = \frac{\boldsymbol{G}_{\mathbf{OS}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})}{\boldsymbol{G}_{\mathbf{WF}}(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z})} \tag{5}$$

Figure 1B illustrates the focus level curve of a pixel (x_0, y_0), showing a distinct peak at the focal plane. The maximum of focus level along z axis and the corresponding argument are acquired for each (x, y). For pixels which have low SNR (signal-to-noise ratio) of focus level, they tend to result in wrong focus positions (Figure 1C). Therefore, a reliability measurement is implemented that the maximum of focus level along z axis has to exceed a threshold to be considered valid (Figure 1D). After interpolation on the invalid (x, y) by the valid height map $h_{valid}(x, y)$, the whole height map $h(x, y)$ can be obtained (Figure 1E). By restoring the color of wide-field images on corresponding position of height map, the 3D color image I_{3D} can be finally reconstructed (Figure 1F).

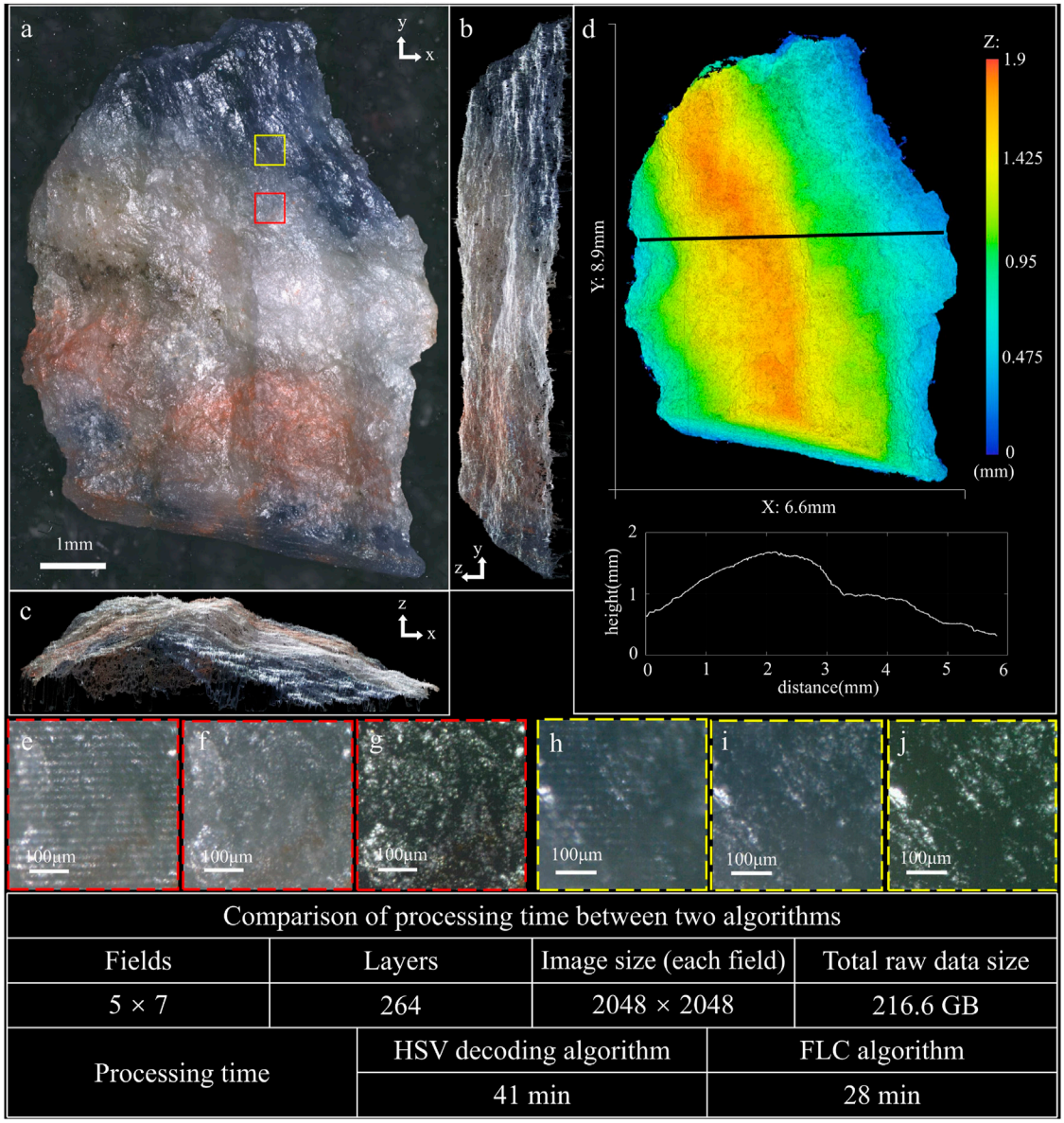
To demonstrate the color aberration in OS-SIM, we simulate the HSV decoding algorithm [17] and the FLC algorithm. The raw image (Figure 2A) is a rectangle with color gradients, generated by varying R, G, and B values to cover different colors. The image added with normal sinusoidal fringe of which modulation depth is 0.6 is shown in Figure 2B. For simulation of the chromatic



Comparison of processing time between two algorithms			
Fields	Layers	Image size (each field)	Total raw data size
9 × 11	236	2048 × 2048	547.6 GB
Processing time	HSV decoding algorithm	FLC algorithm	
	104 min	72 min	

FIGURE 3
3D imaging result of azurite (see Supplementary Video S1). **(A)** MIP image along *z* axis. **(B)** MIP image along *x* axis. **(C)** MIP image along *y* axis. **(D)** Height map of the sample, with the profile along the black line-scan. **(E)** Detail image of the malachite in **(A)**. **(F)** Image restored by the HSV decoding algorithm corresponding to the region of **(E)**. **(G)** Detail image of the white crystal in **(A)**. **(H)** Image restored by the HSV decoding algorithm corresponding to the region of **(G)**.

aberration, the image is added with sinusoidal fringe of different R, G, and B modulation depth which is 0.9, 0.5, and 0.1 respectively (Figure 2C). As a result, the image decoded by the HSV decoding algorithm (Figure 2D) significantly deviates from the raw image in intensity and appears residual fringes. In contrast, the image decoded by the FLC algorithm (Figure 2E) restores the raw image with 100% accuracy. Figure 2F shows the intensity curve of R, G, and B along the red dashed line in Figures 2D,E for comparison. To simulate the modulation depth variation caused by varying



Comparison of processing time between two algorithms			
Fields	Layers	Image size (each field)	Total raw data size
5 × 7	264	2048 × 2048	216.6 GB
Processing time	HSV decoding algorithm	FLC algorithm	
	41 min	28 min	

FIGURE 4
3D imaging result of blue-veins stone (see Supplementary Video S2). **(A)** MIP image along *z* axis. **(B)** MIP image along *x* axis. **(C)** MIP image along *y* axis. **(D)** Height map of the sample, with the profile along the black line-scan. **(E)** Raw image of the white area with focused modulated fringe. **(F)** Image restored by the FLC algorithm corresponding to the region of **(E)**. **(G)** Image restored by the HSV decoding algorithm corresponding to the region of **(E)**. **(H)** Raw image of the blue area with focused modulated fringe. **(I)** Image restored by the FLC algorithm corresponding to the region of **(H)**. **(J)** Image restored by the HSV decoding algorithm corresponding to the region of **(H)**.

reflective properties, the raw image is added with sinusoidal fringe with modulation depth changing from 1 to 0 (Figure 2G). As a result, the image decoded by the HSV decoding algorithm (Figure 2H) loses intensity at low modulation depth areas, while the image decoded by the FLC algorithm (Figure 2I) still restores the raw image authentically. Figure 2J shows the intensity curve of R, G, and B along the yellow dashed line in Figures 2H,I for comparison. In addition, the HSV decoding algorithm needs to compute the RMS operation in three channels H, S, and V [17], while the FLC algorithm firstly converts the color images to grayscale images and then compute the RMS operation in the single grayscale channel, thus the processing time of optical sectioning is saved by 67% theoretically.

Natural color optical sectioning structured illumination microscopy system

The natural color OS-SIM system is same as our previous work [11, 19]. The LED (SOLIS-3C, Thorlabs Inc.) light enters the total internal reflected (TIR) prism and is then reflected onto the DMD (V7000, ViALUX GmbH, Germany) to be modulated into a binary grating. After that, the light is collimated by an achromatic collimating lens (f = 200 mm), passes through a 50/50 beam-splitter, and is focused by the objective lens (×4, NA 0.2, Thorlabs Inc., United States) to illuminate the sample. Volume data of the sample can be obtained by axially moving the translation stage (3-M-122.2DD1, 25 mm travel range, Physik Instrumente GmbH and Co., KG, Germany), while moving in landscape orientation enables the extension of the field of view (FOV). A color sCMOS camera (pco.edge 5.5 CLHS, 100 fps at 2,560 × 2,160 pixels, PCO AG, Germany) is used to capture the 2D images. DMD patterns generation, stage movement and image record are controlled by custom software programmed in C++ to implement hardware synchronization.

Results

To demonstrate the better color restoration capability of the FLC algorithm, we capture images of two mineral samples with varying natural colors. The first mineral sample is a piece of azurite with rich colors, which consists of dark blue and semitransparent crystal, accompaning with green malachite in the brown oxidized zone of copper lodes. Figure 3 shows the imaging result of the azurite, and Supplementary Video S1 presents its reconstructed 3D image. The max intensity projection (MIP) images along z, x, and y axis are shown in Figures 3A–C, respectively. Figure 3D shows the height map of the sample with the profile along the black line-scan. Figure 3E shows a detail on the malachite, and Figure 3F shows the corresponding image restored by the HSV decoding algorithm. It can be seen that the HSV decoding algorithm arises residual fringes while the FLC algorithm is not affected. Figure 3G shows a detail on the white crystal, and Figure 3H shows the corresponding image restored by the HSV decoding algorithm. The HSV decoding algorithm fails to restore the color on the white crystal that the brightness loses seriously. The whole 3D image is stitched from 99 data sets (9 rows and 11 columns) with 236 layers at 20 μm axial intervals and 2,048 × 2,048 pixels for each field, which means the total raw data is 547.6 GB, and the whole 3D volume is 11.3 mm × 14 mm × 4.7 mm. The total image processing time of the HSV decoding algorithm and the FLC algorithm are 104 min and 72 min, which means the processing speed is improved by 44% (Windows 10, 32GB RAM, Intel Xeon W-2123 at 3.6 GHz, MATLAB R2017a).

Another mineral sample is blue-veins stone, with mixed colors of blue, white and red. Its blue surface is semitransparent, on which the modulated fringe is difficult to be projected. Figure 4 shows the imaging result of the blue-veins stone, and Supplementary Video S2 presents its reconstructed 3D image. The MIP images along z, x, and y axis are shown in Figures 4A–C, respectively. Figure 4D shows the height map of the sample with the profile along the black line-scan. Figures 4E,H, respectively show the raw images of the white area and blue area with focused modulated fringe, that the modulation depth of fringe in Figure 4H is quite lower than that in Figure 4E. As a result, the HSV decoding algorithm restores the image of the white area fairly well (Figure 4G) compared to the image restored by the FLC algorithm (Figure 4F). However, the color of blue area restored by the HSV decoding algorithm (Figure 4J) fades a lot due to the low modulation depth of fringe, while the FLC algorithm still restores the raw blue color (Figure 4I). The whole 3D image is stitched from 35 data sets (5 rows and seven columns) with 264 layers at 7 μm axial intervals and 2,048 × 2,048 pixels for each field, which means the total raw data is 216.6 GB, and the whole 3D volume is 6.6 mm × 8.9 mm × 1.8 mm. The total image processing time of the HSV decoding algorithm and the FLC algorithm are 41 min and 28 min, which means the processing speed is improved by 46%.

Discussion and conclusion

The RMS operation in Eq. 3 determines that the intensity of optical sections is proportional to the modulation depth. The HSV decoding algorithm is aimed to solve the color distortion caused by chromatic aberration, but it still does not get rid of the dependence on modulation depth and performs poorly when the surface is weakly modulated. The FLC algorithm abandons direct computation of the color optical sections, but firstly figures out the height map through the focus level and then restores the color from wide-field images which are not affected by the modulation depth. Therefore, it thoroughly solves the color restoration problem.

In addition, the FLC algorithm also improves the axial resolution of the reconstructed 3D image. The optical section of traditional OS-SIM has a depth corresponding to the full width at half maximum (FWHM) of spatial frequency intensity distribution [14], leading to unclear profile of its 3D image. However, the FLC algorithm condenses each pixel in the 3D image to its focus position, so the 3D image can precisely display the surface of the sample. Moreover, the 3D image reconstructed by traditional OS-SIM often contains wrong focused pixels (Figure 1C) which appear to be noises and decline the 3D image quality, while the reliability measurement in FLC algorithm eliminates and replaces them, ensuring the accuracy and fineness of the 3D image.

In summary, we present a focus level correlation algorithm for 3D color image reconstruction in OS-SIM, proved to restore color authentically and without affected by chromatic aberration

or modulation depth variation, and improve the reconstruction speed by about 45%. Capability of high precision height map measurement and natural color restoration allows OS-SIM to apply in 3D imaging on various kinds of minerals. This technique may find potential applications in geology, material science, and biology, especially analyzing translucent and colorful samples.

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

MW conceived the idea and finished the manuscript. TZ and ML gave guidance and supervised the project. All the authors contributed to the discussion on this manuscript.

Funding

This work was supported by the Natural Science Foundation of China (NSFC) (62135003, 62205267, 62005208, and 62205265), Innovation Capability Support Program of Shaanxi (Program No. 2021TD-57), and Natural Science Basic Research Program of Shaanxi (2022JZ-34, 2020JQ-072, 2022JQ-069, 2022JM-321).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2022.1041577/full#supplementary-material

References

1. Stoops G, Delvigne J. Morphology of mineral weathering and neoformation. II neoformations. *Dev Soil Sci* (1990) 19:483–92. doi:10.1016/S0166-2481(08)70363-5

2. Kanouo NS, Zaw K, Yongue RF, Sutherland FL, Meffre S, Njonfang E, et al. Detrital mineral morphology and geochemistry: Methods to characterize and constrain the origin of the Nsanaragati blue sapphires, south-western region of Cameroon. *J Afr Earth Sci* (2012) 70:18–23. doi:10.1016/j.jafrearsci.2012.05.002

3. Michels J. Confocal laser scanning microscopy: Using cuticular autofluorescence for high resolution morphological imaging in small crustaceans. *J Microsc* (2007) 227:1–7. doi:10.1111/j.1365-2818.2007.01787.x

4. Conchello JA, Lichtman JW. Optical sectioning microscopy. *Nat Methods* (2005) 2:920–31. doi:10.1038/nmeth815

5. Helmchen F, Denk W. Deep tissue two-photon microscopy. *Nat Methods* (2005) 2:932–40. doi:10.1038/nmeth818

6. Aligholi S, Lashkaripour RG, Khajavi R, Razmara M. Automatic mineral identification using color tracking. *Pattern Recognit DAGM* (2017) 65:164–74. doi:10.1016/j.patcog.2016.12.012

7. Baykan NA, Yılmaz N. Mineral identification using color spaces and artificial neural networks. *Comput Geosci* (2010) 36:91–7. doi:10.1016/j.cageo.2009.04.009

8. Winter PW, Chandris P, Fischer RS, Wu Y, Waterman CM, Shroff H. Incoherent structured illumination improves optical sectioning and contrast in multiphoton super-resolution microscopy. *Opt Express* (2015) 23:5327–34. doi:10.1364/OE.23.005327

9. Karadaglić D, Wilson T. Image formation in structured illumination wide-field fluorescence microscopy. *Micron* (2008) 39:808–18. doi:10.1016/j.micron.2008.01.017

10. Krzewina LG, Kim MK. Single-exposure optical sectioning by color structured illumination microscopy. *Opt Lett* (2006) 31:477–9. doi:10.1364/ol.31.000477

11. Wang ZJ, Feng K, Yang F, Liang YS, Yun X, Tihelka E, et al. Breathing colour into fossils: A tomographic system for reconstructing the soft tissue microstructure of amber inclusions. *Opt Lasers Eng* (2022) 148:106775. doi:10.1016/j.optlaseng.2021.106775

12. Neil MA, Juskaitis R, Wilson T. Method of obtaining optical sectioning by using structured light in a conventional microscope. *Opt Lett* (1997) 22:1905–7. doi:10.1364/ol.22.001905

13. Dan D, Lei M, Yao BL, Wang W, Winterhalder M, Zumbusch A, et al. DMD-based LED-illumination Super-resolution and optical sectioning microscopy. *Sci Rep* (2013) 3:1116. doi:10.1038/srep01116

14. Dan D, Yao BL, Lei M. Structured illumination microscopy for super-resolution and optical sectioning. *Chin Sci Bull* (2014) 59:1291–307. doi:10.1007/s11434-014-0181-1

15. Ren FF, Wang ZJ, Qian J, Liang YS, Dang SP, Cai YN, et al. Multi-view object topography measurement with optical sectioning structured illumination microscopy. *Appl Opt* (2019) 58:6288–94. doi:10.1364/AO.58.006288

16. Zhou X, Lei M, Dan D, Yao BL, Qian J, Yan SH, et al. Double-Exposure optical sectioning structured illumination microscopy based on Hilbert transform reconstruction. *PLoS One* (2015) 10:e0120892. doi:10.1371/journal.pone.0120892

17. Qian J, Lei M, Dan D, Yao BL, Zhou X, Yang YL, et al. Full-color structured illumination optical sectioning microscopy. *Sci Rep* (2015) 5:14513. doi:10.1038/srep14513

18. Qian J, Dang SP, Wang ZJ, Zhou X, Dan D, Yao BL, et al. Large-scale 3D imaging of insects with natural color. *Opt Express* (2019) 27:4845–57. doi:10.1364/OE.27.004845

19. Zhao TY, Feng K, Liu PY, Ren JR, Wang ZJ, Liang YS, et al. Reconstructing the color 3D tomography of lunar samples. *Spectrosc* (2022) 43:6–12. doi:10.46770/AS.2022.009

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 19 October 2022
DOI 10.3389/fphy.2022.1033422

Check for updates

OPEN ACCESS

EDITED BY
Xunbin Wei,
Peking University, China

REVIEWED BY
Jinchao Feng,
Beijing University of Technology, China
Xiaohua Qian,
Shanghai Jiao Tong University, China

*CORRESPONDENCE
Ting Li,
liting@bme.cams.cn
Aiming Yang,
yangaiming@medmail.com.cn

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to Optics and Photonics,
a section of the journal
Frontiers in Physics

RECEIVED 31 August 2022
ACCEPTED 03 October 2022
PUBLISHED 19 October 2022

CITATION
Wang L, Yang Y, Li J, Tian W, He K, Xu T, Fang Z, Yang A and Li T (2022), Automatic classification of gastric lesions in gastroscopic images using a lightweight deep learning model with attention mechanism and cost-sensitive learning.
Front. Phys. 10:1033422.
doi: 10.3389/fphy.2022.1033422

COPYRIGHT
© 2022 Wang, Yang, Li, Tian, He, Xu, Fang, Yang and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Automatic classification of gastric lesions in gastroscopic images using a lightweight deep learning model with attention mechanism and cost-sensitive learning

Lingxiao Wang[1†], Yingyun Yang[2†], Jingyang Li[3], Wei Tian[4], Kun He[2], Tianming Xu[2], Zhaohui Fang[5], Aiming Yang[2*] and Ting Li[1*]

[1]Institute of Biomedical Engineering, Chinese Academy of Medical Sciences and Peking Union Medical College, Tianjin, China, [2]Department of Gastroenterology, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China, [3]School of Life Sciences, Tiangong University, Tianjin, China, [4]Liang Xiang Hospital Affiliated to Capital Medical University, Beijing, China, [5]Department of Gastroenterology, Ningcheng Central Hospital, Chifeng, China

Magnification endoscopy with narrow-band imaging (ME-NBI) technology is widely used in the early diagnosis of precancerous lesions and gastric cancer, which is critical to reducing the incidence of gastric cancer and improving the survival rate of gastric cancer patients. The diagnosis based on ME-NBI image is mostly in manual way in clinics. In order to improve its objectivity and efficiency, here we proposed a lightweight attention mechanism deep learning model to automatically classify ME-NBI images for artificial intelligence (AI) diagnosis of early gastric cancer, low-grade intraepithelial neoplasia, and non-neoplasm. We collected 4,098 images from 990 patients for model training and validation and evaluated the performance of our model by comparisons with that of other models, such as the benchmark model. An additional 587 images from 296 patients were collected as an independent test set to further evaluate our method's performance. The validation set showed that the overall accuracy, recall, precision, F1 score, and the area under the curve of our method were higher than those of other methods. The independent test set showed that our method achieved state-of-the-art classification for low-grade intraepithelial neoplasia (accuracy = 93.9%, sensitivity = 92.6%). Our method displayed the advantages of lightweight and high effectiveness in classifying effectiveness, which is the potential for AI diagnosis of early gastric cancer, low-grade intraepithelial neoplasia, and non-neoplasm.

KEYWORDS

gastric cancer and precancerous lesions, deep learning, magnification endoscopy with narrow-band imaging, artificial Intelligence, automatic classification

1 Introduction

Gastric cancer (GC) is the third leading cause of cancer death worldwide, with one million new cases and 783,000 additional deaths reported globally in 2021 [1]. GC has a series of evolution processes, gradually developing from chronic gastritis (CGT), intestinal metaplasia (IM), and low-grade intraepithelial neoplasia (LGIN) [2]. With the progression of lesions, the risk of cancer significantly increases. For example, patients with IM and LGIN are about 10 and 25 times more likely to develop GC than ordinary people[2, 3]. The 5-year overall survival rate of patients with pathological early gastric cancer (EGC) is higher than 90% [4], while in Asian countries such as South Korea and Japan, the survival rate for patients with advanced gastric cancer is less than 20% [5, 6]. Therefore, timely detection and accurate classification of precancerous lesions and EGC are crucial for providing proper treatment, reducing the incidence of gastric cancer, and improving the survival rate of EGC patients.

Endoscopy is the primary tool for examining and diagnosing various gastric injuries [7]. Standard endoscopes include white-light imaging (WLI) endoscopy, narrow-band imaging (NBI) endoscopy, and magnification endoscopy with narrow-band imaging (ME-NBI). Compared with WLI and NBI, ME-NBI can observe the morphology of gastric mucosal surface microstructures and microvessels in detail, which has been widely used in diagnosing clinical gastric lesions, especially EGC [8, 9]. However, there are many problems in clinical gastroscopy. First, the diagnosis of endoscopic images is often subjective and depends on the professional knowledge and experience of endoscopists; when endoscopists are inexperienced, misdiagnosis or missed diagnosis will be caused [10, 11]. Secondly, endoscopy will produce many images, so the artificial classification of endoscopic images is a labor-intensive process; when the number of patients is large, the lack of endoscopists and fatigue will further aggravate the missed diagnosis or misdiagnosis of patients[12]. Computer-aided diagnosis (CAD) provides an objective and automatic classification method; it can help doctors make more effective decisions in a shorter diagnosis cycle. Studies have shown that CAD can improve the efficiency and accuracy of diagnosing gastrointestinal lesions [13–15].

Early CAD systems were usually based on manual feature extraction algorithms[16–19]. Kanesaka et al.[14] designed 8 gray level co-occurrence matrix (GLCM) features and developed a CAD system based on these manual features and the coefficient of variation of feature vectors for GC detection. Van D S F et al. [16] developed a CAD system for early esophageal cancer detection in high-definition endoscopic images by calculating local color and texture features based on the original and Gabor-filtered images. However, these methods rely on human-designed algorithms for feature extraction and fail to realize automatic identification of gastric lesions. Moreover, the manual features are insufficient to fit the diversified features of actual lesions[20] and cannot be generalized to practical diagnostic applications.

Deep learning (DL) can automatically capture the subtle features in images and has better accuracy and flexibility than manual feature extraction methods. It has made significant progress in the application of computer vision[21–24] and has been widely used in the field of medical images to solve automatic classification [25–28], segmentation [29–31], localization [32–34], and other tasks. Several scholars have recently demonstrated DL's applicability in automatic endoscopic image analysis[35–38]. Horiuchi et al.[39] proposed a classification method based on GoogleNet, which could automatically identify EGC and gastritis in ME-NBI images with an accuracy of 85.3%. Yan T et al. [40] developed a diagnostic system for IM detection based on the EfficientNet B4 network, and the accuracy for patients reached 88.8%. In terms of detecting LGIN, Cho et al.[41] developed a classification system for gastric lesions based on 5017 WLI images and the DL model, and the accuracy for LGIN was 78.5%, lower than that of endoscopists. Lui et al.[42] developed a classification system based on 3000 NBI images that could automatically classify LGIN, high-grade intraepithelial neoplasia (HGIN), and GC and achieved better classification performance than primary endoscopists. Liu et al.[43] automatically classified gastritis, LGIN, and EGC based on the transfer learning method and ME-NBI images and achieved an average accuracy of 96%.

However, there are still some problems in the automatic classification of gastric lesions based on deep learning. First, existing studies use transfer learning methods, but there is no study on using an end-to-end training model to classify images in small sample datasets. The transfer learning method can solve the problem that training on small sample datasets is challenging to fit, but this method assumes that the source domain datasets and the target domain datasets are correlated, which may reduce the accuracy when the images of the two datasets are quite different [44]. In addition, the traditional DL model has relatively high structural complexity and large volume; even when combined with transfer learning, it still has problems of slow convergence or overfitting on small sample datasets. Second, no research group has classified non-neoplasm (gastritis, IM), LGIN, and EGC based on ME-NBI images and deep learning. Third, the classification performance of precancerous lesions in existing studies needs to be improved.

The attention mechanism can improve the classification performance of deep learning models [45], which has been confirmed in medical image analysis tasks [46, 47]. Inspired by F. Wang et al. [48], this study proposed a deep learning model of attention mechanism. Based on this model, an automatic classification framework for non-neoplasm, low-grade intraepithelial neoplasia, and early gastric cancer based on ME-NBI images was developed. The main contributions of this paper are as follows:

1) A deep learning model with the attention mechanism and cost-sensitive learning was proposed, which can strengthen the discrimination ability of subtle feature differences of gastric lesions and solve the problem of class imbalance in the dataset. This is the first time an attention mechanism model has been introduced into the endoscopic analysis of gastric lesions.
2) The lightweight model was implemented, which allows the model to be trained from scratch on the small sample gastroscopy dataset and does not require pre-trained weights, and significantly reduces the training and deployment time of the model.
3) An automatic classification method was developed based on the proposed model, and the classification of non-neoplasm (including gastritis and IM), LGIN, and EGC based on ME-NBI images were performed for the first time, and state-of-the-art performances were obtained in the classification of LGIN. This method can be used as an additional diagnostic tool in diagnosing gastric lesions under clinical endoscopy.

2 Materials and methods

2.1 Materials

This study was conducted in accordance with the Declaration of Helsinki and approved by the Institutional Review Board of Peking Union Medical College Hospital, Beijing, China. This was a retrospective study, and the data were analyzed anonymously, so informed consent from patients was not required. Endoscopic images were captured using the GIF-H260 Z endoscope with an EVIS LUCERA CV-290 endoscopic video imaging system (Olympus OpticalCorp, Tokyo, Japan), and saved as graphic files of type JPEG (Joint Photographic Experts Group) with two resolutions: 1920 × 1080 pixels and 1440 × 1080 pixels.

All ME-NBI images were collected retrospectively from patients admitted to the Department of Gastroenterology, Peking Union Medical College Hospital, from February 2014 to February 2020. Images of poor quality due to under-inflation, defocus, mucus, blur, and lack of pathological diagnosis were excluded from the study. Finally, 4098 ME-NBI images from 990 patients were collected, including 336 EGC images from 101 patients, 1182 LGIN images from 324 patients, and 2580 non-neoplasm images from 565 patients.

All patients were confirmed by biopsy pathology, with pathological diagnosis as the gold standard. The final pathological results were determined by endoscopic submucosal dissection or surgical resection for suspected cancerous lesions during endoscopy. Two Peking Union Medical College Hospital pathologists made the pathological diagnosis based on tissue sections. Patients with EGC, LGIN, and non-neoplasm confirmed by histology were eligible for this study. Three endoscopists from the Department of Gastroenterology, Peking Union Medical College Hospital, evaluated and classified the gastroscopic images: first, two endoscopists (with more than 7 years of gastroscopy experience) carefully reviewed all ME-NBI images according to the pathology report, those images that did not match the anatomical location in the pathology report were discarded, and the final retained images were captured at almost the exact location as the biopsy or surgical location where the doctor suspected the abnormality, and have corresponding tissue samples; later, according to pathological diagnosis, two endoscopists divided the remaining images into EGC, LGIN, and non-neoplasm (including gastritis and IM); when two endoscopists have different opinions on image classification, the third endoscopist (with more than 10 years of gastroscopy experience) will review, verify, and determine the image category.

In this study, a 5-fold cross-validation was performed on the training set. The training set was divided into five groups using a patient-based random sampling method. The images of a single patient with a type of lesion were only assigned to one group. When a patient had different lesions, the images of each type of lesion might appear in different groups, and the number of patients with the same type of lesion in different groups was the same. Then, five iterations of training and validation were performed, with one different group for validation in each iteration and the remaining four groups for training. The number of images acquired in this study is minor. In order to improve the robustness of the system, the data augmentation strategy was implemented for the training group. We augmented the training group by rotation (±15°), flip (vertical and horizontal), and other transformations that did not affect the image features, while the validation group was not augmented.

Another new dataset was collected and used as the independent test set. All ME-NBI images with pathologically confirmed were collected from consecutive patients who underwent gastroscopy at Peking Union Medical College Hospital from March 2020 to December 2020. With the same exclusion criteria as above, a total of 587 ME-NBI images from 296 patients were finally collected. The overall median age of the test set was 56, with a range of 24–89, and the sex ratio between males and females was 177/119. Table 1 shows the image category composition of the datasets used in this study and the population characteristics of the patients in the test set.

2.2 Methods

2.2.1 Data preprocessing and cost-sensitive learning

The original gastroscopy image contains a black border and text information that does not contribute to disease identification and may contain patient information. Therefore, in the preprocessing process, the black border of the original image

TABLE 1 Demographics of the dataset used in this study.

	The training set		The test set			
	No. Of images	No. Of patients	No. Of images	No. Of patients	Median age (range)	Sex (M/F)
Overall	4,098	990	587	296	59 (24–89)	177/119
EGC	336	101	50	36	68 (36–87)	23/13
LGIN	1,182	324	169	96	57 (32–89)	57/39
non-neoplasm	2,580	565	368	164	52 (24–77)	97/67

is firstly removed by cropping, then the text information is removed to protect the patient's privacy. After that, the image size was uniformly adjusted to 224 × 224 pixels.

There is a severe class imbalance problem in our dataset; that is, the number of samples in the minority class is significantly less than that in the majority class, which will make the model obtain more prior information from the majority class samples, resulting in the classification results biased toward the majority class, and thus affecting the classification accuracy. Cost-sensitive learning uses a unique penalty term or weight value for each class to characterize the importance of different classes. Usually, it uses a smaller weight value for the majority class and a larger weight value for the minority class to adjust the loss value of each class to minimize the cost of misclassification and alleviate the bias of class imbalance. In order to solve the problem of class imbalance in our dataset, this study introduced cost-sensitive learning into the model and redesigned the loss function. We first set each category's penalty or weight value; the method for obtaining the weight value is shown in Equation 1:

$$W_j = \frac{N_{total}}{C \cdot N_j}. \tag{1}$$

In the above equation, j represents the category, W_j represents the weight value of the category, N_{total} represents the total number of samples, C represents the number of categories and N_j represents the number of samples of category j. Then, we introduced the weight value into the loss function and got the weighted loss function, as shown in Equation 2:

$$Loss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} W_j \cdot Y_{ij} \cdot \log\left(P_{ij}\right). \tag{2}$$

In the above equation, N represents the number of batch samples,i represents the sample, Y_{ij} represents the actual label of samples, and P_{ij} represents the predicted probability value. In the calculation process, Y_{ij} is presented as one-hot and contains only 0 and one elements. In order to avoid unnecessary calculation, only Y_{ij} with the value of one and the predicted value of the corresponding position are kept. The modified weighted loss function is shown in Equation 3:

$$Loss = -\frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{C} W_j \cdot 1_{[ij']} \cdot \log\left(P_{ij'}\right). \tag{3}$$

In Equation 3, j' represents the actual category label of the ith sample.

2.2.2 Lesion classification framework

The Attention Module [48] adopts the (bottom-up, top-down) encoding and decoding structure, which can refine the attention to the image in feedforward learning, highlight the subtle feature differences between different lesions, enhance feature selection, and promote gradient update in feedback learning. The separable convolution layer [49] maps cross-channel correlation and spatial correlation separately, improving the model's performance without increasing the number of model parameters and computational complexity. The Inception module [50] can effectively reduce parameters, extend network depth through factorizing convolution, and combine multi-layer feature fusion technology to achieve better performance with lower computational cost. Dilated convolution [51, 52] expands the convolution range by inserting spaces between the elements of the small-size convolution kernel; it can obtain the same "receptive field" and feature learning ability as the large-size convolution kernel under the condition of occupying a few parameters. Inspired by these techniques, this study designs a new lightweight convolution neural network model and develops a framework for automatically classifying EGC, LGIN, and non-neoplasm gastroscopy images based on this model. The details are as follows:

Firstly, by referring to the structural characteristics of the attention module, this study designs two types of attention modules, the AttenSeparableConv block and the AttenInception block. The two attention modules have the same branching structure. AttenSeparableConv block uses stacked separable convolution layers as convolution units of

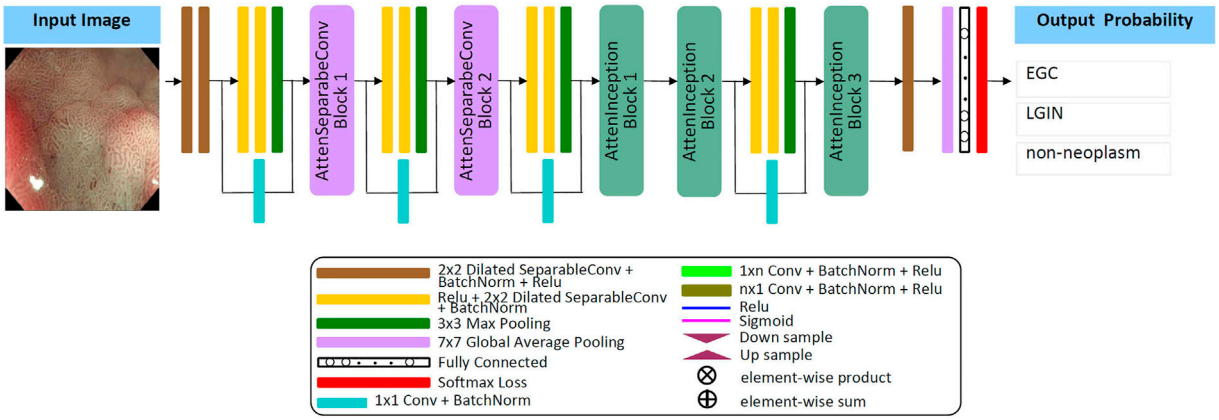


FIGURE 1
The framework of the classification method of gastric lesions.

trunk branches, while AttenInception block uses stacked Inception units as convolution units of trunk branches. Secondly, the unit composed of two separable convolution layers and the Maxpooling layer is used as the dimension reduction unit to reduce the dimension of the feature map, and the trade-off between the bottleneck of the feature map and pooling can be effectively balanced. In order to reduce the number of model parameters and computational complexity, we replace the convolution kernels of all separable convolution layers in the model with the dilated convolution kernels of size 2×2 and dilated rate 2 and name the new convolution layer as Dilated SeparableConv layer. The final model structure is shown in Figure 1. The model contains four dimension reduction units, two AttenSeparableConv blocks, and three AttenInception blocks. The dimension reduction units are located at the starting position and the low, middle, and high-level feature extraction positions to reduce the dimension of the feature map timely. Affected by structural differences, at the same position of the model, the number of parameters occupied by the separable convolution layer far exceeds that occupied by the factorization convolutional layer. For example, the number of parameters occupied by AttenInception block 1 is 1.28 million. If an AttenSeparableConv block 3, in which the structure is the same as AttenSeparableConv block1-2, is set at this location, the number of parameters occupied by this block is 2.38 million. And the parameters difference between the two blocks increases with the depth of the position. Therefore, we only use AttenSeparableConv blocks in the shallow layer of the model and use AttenInception blocks in the deeper layer of the model to extend the network depth, to effectively balance the number of parameters and performance. We set the AttenSeparableConv block after the first and second dimension reduction units for low-level feature extraction. The AttenInception block is set after the third and fourth dimension reduction units for middle and high-level feature extraction. Another three separate Dilated SeparableConv layers are used for head and tail feature learning, respectively. The above components make up the feature extractor. The global average pooling layer, the fully connected layer, and the softmax loss function compose the classifier. After the ME-NBI image is input to the network, the feature extractor identifies the feature, and the classifier outputs the probabilities that the image belongs to three gastric lesions.

The schematic diagrams of the AttenSeparableConv block and AttenInception block are shown in Figure 2. The trunk of the AttenSeparableConv block is composed of two stacked Unit1 units, each of which contains three Dilated SeparableConv layers. We added a RELU activation function and a batch normalization layer before and after each Dilated SeparableConv layer. Relevant studies have shown that a short-time connection can improve the classification performance of attention-oriented structures [48], so we added a short-time connection to Unit1 to improve the classification performance of AttenSeparableConv blocks. The trunk of the AttenInception block is composed of stacked Unit2 units, and different AttenInception blocks contain different amounts of Unit2 units in the trunk. The trunk of AttenInception block1-2 at the middle-level feature extraction position contains three Unit2 units, and the trunk of AttenInception block3 at the high-level feature extraction position contains two Unit2 units. Unit2 adopts the Inception block. In the original literature, the Inception block has convolution kernels of various sizes ($1\times n$, $n\times 1$, $n = 3,5,7$) [50]. In this study, a relatively large convolution kernel ($n = 5$ or 7) is used in the Unit2 of

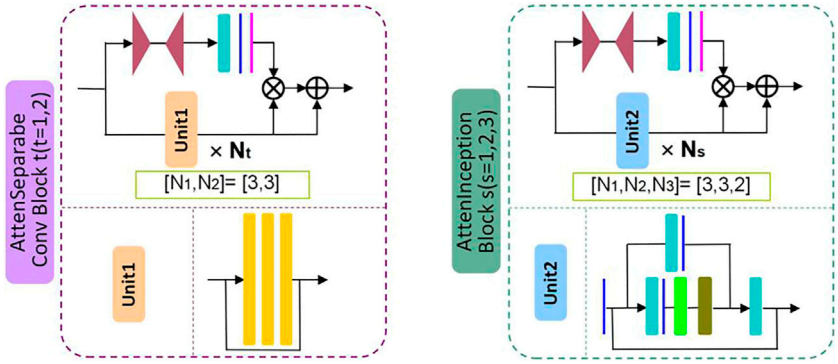


FIGURE 2
The proposed AttenSeparableConv Block and AttenInception Block.

AttenInception block1-2, while a small convolution kernel (n = 3) is used in the Unit2 of AttenInception block3. The attention branches of the two types of attention modules both use the Maxpooling layer as the bottom-up structure and the linear interpolation layer as the top-down structure to retain the characteristics of the original attention branch structure, and the Sigmoid function is retained for the mixed attention constraint. Unlike the original attention branch structure, this study only retained one convolution layer and added the RELU function to this convolution layer to enhance the learnable feature change space. The weight of the trunk feature map is realized by multiplying and then adding the output of the attention branch and the output of the trunk feature map.

2.2.3 Training details

Adam was used as the network optimizer for all CNN models with a learning rate of 0.0001, training epochs of 150, and a batch size of 8. All experiments were performed on an AMD Ryzen 7-1700X eight-core processor CPU and a GeForce GTX 1080 Ti GPU(graphics processing units).

2.2.4 Evaluation metrics

Accuracy (ACC), recall (RE), precision (PRE), F1 score (F1), and the area under the curve (AUC) were used to evaluate the classification performance of the model, and each metric was calculated for an independent category. ACC and F1 evaluation comprehensive classification ability, RE represents the susceptibility to disease, PRE representative disease recognition accuracy, and receiver operating characteristic (ROC) curve can measure classifier robustness. The AUC value is automatically calculated according to the ROC curve, which can intuitively reflect the comprehensive classification ability of the model, and its range is between 0 and 1. The larger the value, the better the performance of the classification model. The calculation of each metric is shown in equations (4)–(7):

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}, \tag{4}$$

$$RE = \frac{TP}{TP + FN}, \tag{5}$$

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

$$F1 = 2 \times \frac{PRE \times RE}{PRE + RE}. \tag{7}$$

TP, TN, FP, and FN stand for True Positive, True Negative, False Positive, and False Negative, respectively. To evaluate the overall performance of the model, the overall accuracy (OA), recall (OR), precision (OP), F1 Score (OF1), and AUC (O-AUC) were obtained by adding and calculating the average metric of each category. For the 5-fold cross-validation experiments, the average results of multiple cross-validation experiments were evaluated using the evaluation metrics with 95% confidence intervals (CI).

3 Results

We first evaluated the model's performance on the validation set. In order to demonstrate the effectiveness of the proposed method for three types of gastric lesions classification, we performed ablation studies, cost-sensitive learning tests, and comparison tests with other advanced methods on the validation set. We used the 5-fold cross-validation method; the whole training set was divided into five groups and performed five experiments. In each experiment, a different group was used as the validation set, and the remaining four groups were used for training. We evaluated the model performance on the validation set. The final result was the average of five experiments. OA, OR, OP, OF1, and O-AUC were used as evaluation metrics to evaluate the classification

TABLE 2 Statistical comparison of ablation studies.

Methods	OA,% (95% CI)	OR,% (95% CI)	OP,% (95% CI)	OF1,% (95% CI)	O-AUC,% (95% CI)	P (M)	Time (ms)
Inception-ResNetV2 [50]	91.6 (88.4–94.8)	78.5 (69.9–87.1)	87.6 (82.0–93.2)	81.2 (72.8–89.6)	96.0 (94.0–98.0)	55.9	0.96
Xception [49]	92.9 (90.0–95.8)	82.2 (72.9–91.5)	88.2 (83.8–92.6)	84.4 (77.0–91.8)	97.4 (96.1–98.7)	22	0.60
Our model 1 (N1, N2)	90.8 (85.5–96.1)	81.9 (72.7–91.1)	86.7 (78.6–94.8)	82.6 (71.7–93.5)	96.1 (93.0–99.2)	10.4	0.54
Our model 2 (N1, Y2)	92.5 (88.8–96.2)	83.5 (82.5–84.5)	87.3 (80.2–94.4)	84.6 (76.5–92.7)	96.5 (93.8–99.2)	10.8	0.54
MainNet (Y1, Y2)	**93.7 (90.4–97.0)**	**84.9 (74.8–95.0)**	**88.7 (82.2–95.2)**	**85.4 (82.1–88.7)**	**97.5 (95.6–99.4)**	11.4	**0.54**

N1, no use of attention branch in all blocks of the model; N2, The small size factorization convolution kernel of 1* 5,5 *1 was used in Unit2 of AttenInception Block1 to three; Y1, attention branch was used in all blocks of the model; Y2, a large size factorization convolution kernel of 1* 7,7 *1 was used in Unit2 of AttenInception Block1 to three; P, parameters; M, million.

performance. The number of parameters of each model was counted to indicate the computational complexity of different models. In addition, the classification performance of the proposed method was further evaluated on an independent test set.

3.1 Ablation studies

We first performed ablation studies to demonstrate the proposed attention mechanism model's effectiveness. Xception [49]and Inception-ResNetV2 [50] with ImageNet pre-trained weights were used as benchmark models and compared with our three models: Our model 1 and Our model 2, which were obtained by removing the attention branch, and MainNet (the model in Figure 1), which was obtained by adding attention branch based on Our model 2. The difference between Our model 1 and Our model 2 is the factorization convolution kernel in Unit2 of AttenInception block1 to 2 (Figure 1). In Our model 1, the convolution kernel size was 1×5,5×1, while in Our model 2, the convolution kernel size was 1×7,7×1. The classification performance of our three models (Our model 1, Our model 2, and MainNet) was compared with that of the benchmark model to prove the performance of the designed models, and the proposed models can be compared with each other to show the role of attention mechanism further. In order to ensure the fairness of the comparison, the training conditions of all models were the same (learning rate = 1E-4, batch size = 8, epochs = 150), and the input image size was consistent with the original network, that was, the input image size of Inception-ResNetV2 and Xception model was 299×299, while the input image size of our three models was 224×224.

Table 2 summarizes the overall classification performance of five methods for three kinds of gastric lesions. The values in bold in the Table 2 represent the optimal values. Compared with other methods, MainNet obtained the highest OA, OR, OP, OF1, and O-AUC, which were 93.7%, 84.9%, 88.7%, 85.4%, and 97.5%, respectively. MainNet had 11.4M parameters, nearly half the number of parameters for the Xception and less than one-fifth of the number for Inception-ResNetV2. These results showed that our attention mechanism model achieves lightweight while maintaining high classification performance. In addition, Our model 1 and Our mode 2 achieved comparable classification performance with Inception-ResNetV2, proving the superiority of our backbone model. In terms of speed, the prediction time of our three models was 0.54 ms for each image, which was less than the prediction time of the benchmark model, which verified that the model trained from scratch proposed in this study was more conducive to the classification of endoscopic images of gastric lesions than the transfer learning model pre-trained on the natural image dataset. In addition, the overall classification performance of MainNet was better than that of Our model 1 and Our model 2; Our model 2's classification performance was better than that of Our model 1. These comparison results showed that using the attention mechanism can effectively improve the model's classification performance, and using a large convolution kernel in the middle-level feature extractor helps improve the model's classification accuracy.

3.2 Cost-sensitive learning test

In order to prove the effectiveness of cost-sensitive learning in solving the problem of class imbalance in dataset, the performance of MainNet with and without cost-sensitive learning is compared. The results are shown in Table 3. As can be seen from Table 3, the classification performance of MainNet with cost-sensitive learning was significantly better than that of MainNet without cost-sensitive learning, and the difference in OR was particularly prominent, reaching 3.3%. This indicated that combining cost-sensitive learning with our attention mechanism model can effectively improve the model's sensitivity to minority samples in the imbalanced dataset and the overall classification performance.

TABLE 3 Statistical comparison of MainNet with different class imbalance processing methods.

Methods	OA,% (95%CI)	OR,% (95%CI)	OP,% (95%CI)	OF1,% (95%CI)	O-AUC,% (95%CI)
MainNet (without CSL)	92.5 (89.1–95.9)	81.6 (73.2–90.0)	86.8 (80.7–92.9)	83.5 (75.7–91.3)	96.2 (92.7–99.7)
MainNet (with CSL)	**93.7 (90.4–97.0)**	**84.9 (74.8–95.0)**	**88.7 (82.2–95.2)**	**85.4 (82.1–88.7)**	**97.5 (95.6–99.4)**

CSL, cost-sensitive learning.

TABLE 4 Statistical comparison with other related advanced methods.

Methods	OA,% (95% CI)	OR,% (95% CI)	OP,% (95% CI)	OF1,% (95% CI)	O-AUC,% (95% CI)	P (M)	Time (ms)
[53]	90.6 (88.4–92.8)	77.7 (71.5–83.9)	83.2 (77.8–88.6)	79.4 (75.1–83.7)	95.4 (93.9–96.9)	8.1	0.45
[40]	91.7 (89.6–93.8)	78.5 (70.4–86.6)	83.8 (78.8–88.8)	80.1 (74.5–85.7)	96.1 (94.4–97.8)	19.5	0.60
[54]	91.4 (88.4–94.4)	79.1 (71.2–87.0)	87.3 (83.1–91.5)	81.8 (74.7–88.9)	96.2 (94.4–98.0)	25.6	0.79
MainNet	**93.7 (90.4–97.0)**	**84.9 (74.8–95.0)**	**88.7 (82.2–95.2)**	**85.4** (82.1–88.7)	**97.5 (95.6–99.4)**	11.4	0.54

P, parameters; M, million.

3.3 Comparison with other advanced methods

To further verify the proposed method's effectiveness, we compared our method's classification performance with other advanced methods on our dataset. We selected some characteristic classification methods for gastric lesions, which either have lighter models[53] or better classification performance [40, 54]. These advanced methods include: the classification method of intestinal metaplasia proposed by Yan et al. [40], the classification method of chronic atrophic gastritis proposed by Zhang et al. [53], and the classification method of atrophic gastritis and intestinal metaplasia proposed by Zheng et al.[54]. The models in these advanced methods are pre-trained in ImageNet. While for our model, we trained it from scratch on the gastric lesion dataset. Other training conditions of all methods are the same to ensure the fairness of comparison. The statistical results of the evaluation metrics are shown in Table 4. As can be seen from Table 4, the OA, OR, OP, OF1, and O-AUC obtained based on MainNet are significantly higher than those obtained by other methods. In terms of the number of parameters, our model's parameters are higher than that of Zhang et al.'s method (11.4M vs 8.1M), which is much lower than that of Yan T et al.'s 19.5M and Zheng et al.'s 25.6M. In terms of computational efficiency, the average test time of our method is 0.54 ms per image, while the average test time of Zhang et al., Yan T et al., and Zheng et al. are 0.45 ms, 0.60 ms, and 0.79 ms per image, respectively. Zhang et al.'s method show lower parameters and higher computational efficiency. However, compared with their method, the proposed method achieves 3.1%, 7.2%, and 5.5% performance increments in OA, OR, and OP, respectively. The results show that compared with other advanced methods, our method can achieve a better balance among classification performance, number of parameters and computational efficiency.

Supplementary Table S1 summarizes the performance of several methods in per-category classification. It can be seen that in the identification of EGC, the ACC, RE, F1, and AUC obtained based on MainNet were higher than those obtained by other advanced methods, and only the PRE was lower than that obtained by Zheng et al. [54]. In identifying LGIN, our method obtained the highest ACC, PRE, F1, and AUC, and only the RE was lower than that of Zheng et al. [54]. In identifying non-neoplasm, the proposed method performed better than all other methods and achieved the highest values of ACC, RE, PRE, F1, and AUC.

3.4 Model evaluation on the independent test set

We further evaluated MainNet's classification performance on the independent test set. A total of 587 ME-NBI images from 296 patients were included in the test set, including 50 EGC images from 36 patients, 169 LGIN images from 96 patients, and 368 non-neoplasm images from 164 patients. The classification performance of the MainNet on the test set was compared with the results of other studies, and our best-performing model was compared with human experts.

3.4.1 Comparison with other research results

First, MainNet was trained on the training set using the 5-fold cross-validation method, and then the average results on the independent test set were counted and compared with the

TABLE 5 Per-category average classification performance according to the 5-fold cross-validation.

Methods	Classes	ACC,% (95%CI)	SE,% (95%CI)	SP,% (95%CI)	AUC,% (95%CI)
[41]	EGC	80.8 (77.9–83.4)	52.4 (45.0–59.8)	89.2 (86.5–91.5)	70.8 (67.5–73.9)
	LGIN	87.1 (84.6–89.3)	22.2 (14.5–31.7)	96.1 (94.4–97.4)	59.1 (55.7–62.6)
	non-neoplasm	83.1 (80.4–85.6)	92.2 (89.0–94.8)	75.8 (71.6–79.7)	84.0 (81.3–86.5)
MainNet	EGC	96.8 (95.1–98.5)	93.2 (89.9–96.5)	97.1 (95.0–99.2)	98.8 (98.0–99.6)
	LGIN	93.9 (93.4–94.4)	92.6 (88.4–96.8)	95.4 (91.5–99.3)	97.6 (97.1–98.1)
	non-neoplasm	92.9 (91.7–94.1)	82.9 (74.2–91.6)	97.3 (95.5–99.1)	98.5 (98.4–98.6)

TABLE 6 Statistical comparison of two endoscopists and the best model's performance in three-category classification.

Methods	OA (%)	OS (%)	O-SP (%)	Kappa
Endoscopist 1	94.7	89.4	94.5	0.842
Endoscopist 2	94.2	90.2	95.2	0.834
The best model	**95.0**	**93.2**	**96.1**	**0.859**

average results of Cho et al. [41] which was obtained in their independent test based on 5-fold cross-validation. We adopted the same evaluation metric as the literature[41]. The results are shown in Table 5. It can be observed that the accuracy (ACC), sensitivity (SE), specificity (SP) and AUC of the proposed method for each category were higher than those of Cho, et al., and only the SP of LGIN (95.4% vs. 96.1%) and SE of non-neoplasm (82.9% vs. 92.2) were lower than those of Cho, et al. The comparison results showed that the proposed method improved classification performance in classifying EGC, LGIN, and non-neoplasm gastric lesions images. Based on deep learning, Liu et al.[43] divided gastric ME-NBI images into EGC, LGIN, and chronic gastritis (CGT). Their recall (which means the same as our SE) for EGC and LGIN were 99% and 92%, respectively. Compared with the SE of our method for EGC and LGIN(93.2%, 92.6%, respectively), it can be seen that the classification performance of our method for LGIN was better than that of Liu et al., but the SE for EGC was lower than that of Liu et al.

3.4.2 Comparison with endoscopists

Our best model's classification results were further compared with those of two endoscopists with more than 8 years of endoscopy experience on the independent test set. The best model refers to a MainNet model that achieves the best classification performance on the test set by 5-fold cross-validation. Two endoscopists classified the images without knowing the image category labels. Overall accuracy (OA), sensitivity (OS), specificity (O-SP), and Kappa is used to measure the agreement among raters, which can quantify the degree of agreement between the classifier and the gold standard.

Table 6 summarizes the overall classification performance of our best model and two endoscopists on EGC, LGIN, and non-neoplasm. It can be seen that the Kappa, OA, OS, and O-SP of the best model were 0.859, 95.0%, 93.2%, and 96.1%, respectively, which were higher than those of Endoscopist 1 and Endoscopist 2 (the optimal values were bolded). Figure 3 compares accuracy, sensitivity, and specificity for each category between our best model and the two endoscopists. It can be seen that the best model's accuracy for EGC and LGIN was better than that of the two endoscopists, while the accuracy for non-neoplasm was slightly lower than that of Endoscopist 1 and the same as that of Endoscopist 2. Besides, our best model showed higher sensitivity to EGC and LGIN, but slightly lower sensitivity to non-neoplasm.

We calculated the correlation between the predicted and actual labels for each image by the best model and the two endoscopists and reflected the results in the confusion matrix shown in Figure 4. As can be seen from Figure 4, in the identification of EGC images, the best model had the lowest number of misclassified images; two out of 50 images were misclassified as non-neoplasm, while there were four misclassified images for Endoscopist 1 and Endoscopist 2. In the identification of LGIN, the number of misclassified images of the best model was 16, much lower than the 38 images of Endoscopist 1 and 27 images of Endoscopist 2. In identifying non-neoplasm, the number of images correctly identified by the best model was 342, which was lower than that of Endoscopist 1 and Endoscopist 2 (363 and 348, respectively).

3.4.3 Model binary classification

To further evaluate the binary classification performance of MainNet on gastric lesions, three groups of tests were performed, including classification tests for cancer or non-cancer (EGC vs Others), LGIN or non-LGIN(LGIN vs Others), non-neoplasm or neoplasm (non-neoplasm vs Others). We plotted the ROC curve of the best binary classification model in each group of tests and obtained the AUC, as shown in Figure 5. In the discrimination of cancer or non-cancer, the AUC reached 0.981; In the discrimination of LGIN or non-LGIN, the AUC reached 0.984; In the discrimination of non-neoplasm or neoplasm, the AUC reached 0.988. The binary classification results show that the performance of MainNet was well decomposed.

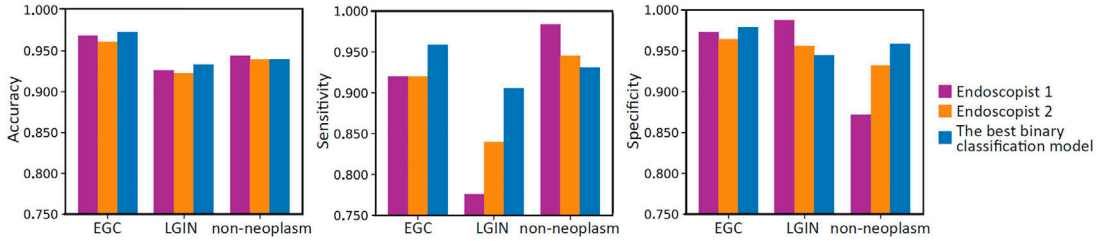


FIGURE 3
Bar charts compare the per-category accuracy, sensitivity, and specificity between two endoscopists and the best model.

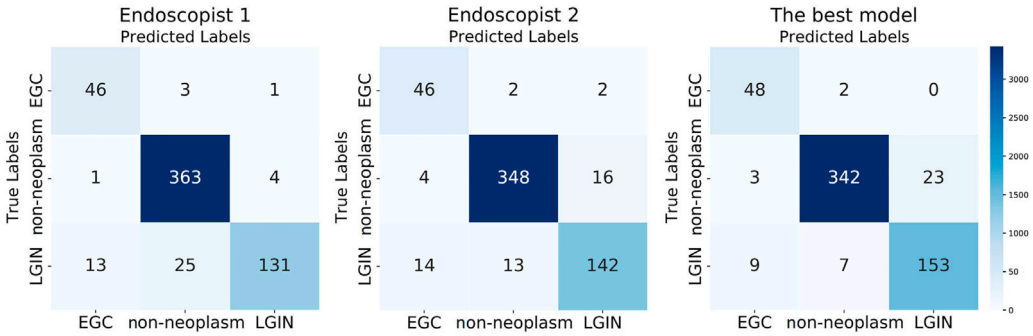


FIGURE 4
Three confusion matrixes for two endoscopists and the best model's predictions, respectively.

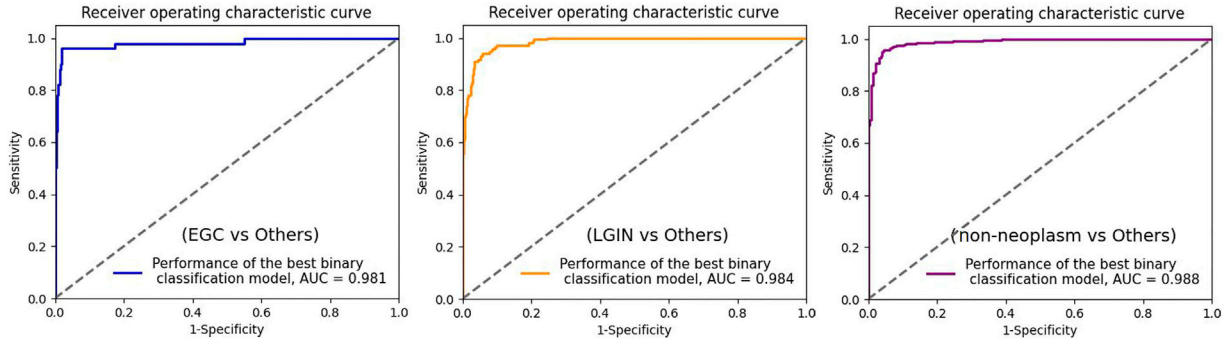


FIGURE 5
ROC curve for the three best binary classification models. ROC, receiver operating characteristic; AUC, area under the curve; EGC, early gastric cancer; LGIN, low-grade intraepithelial neoplasia.

4 Discussion and conclusion

Although ME-NBI can provide imaging of gastric mucosal surface microstructure and microvascular morphology and is widely used in clinical diagnosis of early gastric cancer and precancerous lesions, qualitative assessment of ME-NBI images requires much training, and even experienced endoscopists may misdiagnose or miss a diagnosis. In addition, in the case of limited medical resources, it is unrealistic to manually examine a large number of ME-NBI images, which may cause missed diagnosis or misdiagnosis. Therefore, it is significant to realize the accurate and

automatic diagnosis of gastric lesions in clinical practice. In this study, we developed an automated diagnosis system based on a deep learning model with the attention mechanism, which could automatically classify ME-NBI images in a small sample dataset into EGC, LGIN, and non-neoplasm. The classification performance of the proposed method is verified on the validation set and independent test set, respectively.

In this paper, we first conduct ablation studies on the validation set to verify the role of the attention module. The attention module in MainNet adopted the structure of the attention branch combined with the trunk. The encoding and decoding structure of the attention branch can refine the attention to the image and use the branch's output to weight the trunk's output to enhance feature selection. Then the attention module was embedded into the model in a stacked manner to refine attention to subtle lesion features. We used efficient convolution layers such as the separable and factorization convolutional layers as the basis and introduced a dilated convolution kernel to reduce the parameters further. By adopting this attention-guided feature learning mode, combined with the parameter reduction strategy, MainNet achieves better classification performance than the backbone and the benchmark models when the parameters are only 11.4M. Our proposed three models are relatively lightweight (10.4M, 10.8M, and 11.4M parameters, respectively) but achieve similar or even better classification performance than Inception-ResNetV2. This phenomenon is related to model complexity. Inception-ResNetV2 demonstrates excellent classification performance on large natural image datasets by increasing network complexity and parameters; however, in practical applications, these heavyweight characteristics limit the implementation of Inception-ResNetV2 on small sample datasets (such as most medical image datasets). On the contrary, Our models are relatively simple and effectively balance computational complexity and task requirements, ultimately achieving classification performance comparable to Inception-ResNetV2.

Our dataset has a severe class imbalance problem, and the deep learning model will tend to over-classify the majority category in the imbalanced dataset. When the appropriate solution is not taken, the accuracy of the majority category may be high, while the overall accuracy is low. As seen in Table 3, the classification performance of MainNet using cost-sensitive learning was higher than that of MainNet without cost-sensitive learning. This indicates that the introduction of cost-sensitive learning can effectively solve the problem of class imbalance in data sets and improve the classification accuracy of gastric lesions.

In the comparison experiment with other advanced methods, the overall classification performance of the MainNet was better than other methods, OA was improved by at least 2.0%, OR was improved by at least 5.8%, and most of the evaluation metrics of per-category classification were also significantly improved. It is worth noting that MainNet outperforms other advanced methods even without using cost-sensitive learning. This indicates that the proposed method could effectively improve the classification performance. It is especially noted that for these advanced methods, we adopted the same training mode as in the original literature; we fine-tuned these pre-trained models on the gastric lesion dataset. While for our model, we trained it from scratch on the gastric lesion dataset. Experimental results show that our end-to-end training model outperforms these pre-trained models on the gastric lesions dataset.

The classification performance of MainNet on an independent test set was further compared with the results of related studies. As seen in Table 5, compared with Cho et al. [41], our method's ACC for EGC, LGIN, and non-neoplasm increased by 16.8%, 6.8%, and 9.8%, respectively. We believe that the model's superiority and the practical solution to the class imbalance problem are one of the reasons that make our ACC better than those of Cho et al. Still, the most important reason is the difference in image modality. ME-NBI images were used in this study, while Cho et al. conducted their analysis based on WLI images. However, gastric lesions usually show changes in mucosal surface microstructure. It is difficult for conventional WLI to capture subtle disease features, especially for LGIN. Liu et al. [43] performed EGC, LGIN, and chronic gastritis (CGT) classification based on deep learning and ME-NBI images. Liu et al. obtained 92% and 99% sensitivity for LGIN and EGC, respectively, while the sensitivity of our method for LGIN and EGC was 92.6% and 93.2%, respectively. In general, the sensitivity of our method for LGIN exceeds that of Liu et al. and achieves state-of-the-art classification performance. However, the sensitivity of our method to EGC is lower than that of Liu et al., which may be related to the small sample size of EGC in our dataset. In this study, there are only 369 EGC images in the training set and 50 EGC images in the test set. In such conditions, the feature variation space that the model in training can learn is limited, and the generalization effect and accuracy of the test set are affected. This problem can be solved by including more EGC samples.

Besides, the classification performance of the best model was compared with that of human experts in the independent test set. The results in Table 6 showed that the overall classification performance of the proposed method was better than that of the two endoscopists. In addition, it can be observed from Figure 3 that the diagnostic accuracy of EGC, non-neoplasm, and LGIN decreased successively in both the proposed method and the two endoscopists. For endoscopists, EGC is a severe gastric disease with significant imaging features associated with significant mortality, so they will emphasize identifying such lesions. For our method, although the number of EGC samples in our dataset is small, the attention mechanism model can strengthen the learning of subtle feature differences, and cost-sensitive learning can solve the classification bias that tends to the majority classes. Those make our method achieve better classification performance than endoscopists. On the contrary,

LGIN is a lesion characterized by cellular atypia of mucosal structure, different degrees of atypia and different feature patterns also make it difficult to identify accurately, so the accuracy of LGIN by both us and endoscopists is relatively low. However, benefiting from the attention-guided feature learning model, the accuracy of our method for LGIN exceeds that of endoscopists.

Although our method performs better than other methods in gastric ME-NBI image lesion classification, this study still has some limitations. First of all, the data of this study were collected in the same hospital, and the endoscopic equipment in different centers and the characteristics of different populations may have an impact on the method's performance; in the subsequent study, we will collect endoscopic images from different centers and analyze them. Secondly, there are few EGC and non-neoplasm samples in the dataset of this study, which may affect the accuracy; more samples will be included in subsequent studies to improve the system's accuracy further. In addition, we only analyzed gastric lesions in this study; after more cases were included, esophagitis and early esophageal cancer will be included in the system to increase the clinical application value of the system. Besides, grading the severity of the lesions is crucial to the prognosis and formulation of treatment plans; so far, only the classification of lesions has been completed in this study, and the severity of lesions will be further graded in subsequent studies. Finally, the proposed method is a supervised learning method, which still requires doctors to carry out a large amount of data annotation work in the early stage, causing a specific workload for doctors; in subsequent studies, methods based on self-supervised deep learning models can be used to solve this problem.

In this study, we designed an attention-guided deep learning model and introduced cost-sensitive learning into the model. Based on this model, we developed an automatic classification method for gastric lesions, which achieved good diagnostic performance on a limited number of ME-NBI images and outperformed other advanced methods. In addition, we achieved the most advanced classification performance for LGIN and non-neoplasm. Through an in-depth literature review, we found that this was the first time to automatically classify non-neoplasm, LGIN, and EGC based on the deep learning model and gastric ME-NBI images. And the first time to introduce the attention mechanism model into the automatic classification of gastric lesions. The number of parameters in our model was only 11.4 million, which allowed the model to be trained end-to-end on small sample medical datasets and can shorten the prediction time per image. The prediction time of each image of our method was only 0.54 ms, which met the demand for real-time diagnosis. In conclusion, our approach can provide objective and accurate guidance information for endoscopists in real time and has an excellent clinical application prospect.

Data availability statement

The datasets presented in this article are not readily available due to patient confidentiality. Requests to access the datasets should be directed to the corresponding authors.

Author contributions

TL and AY designed the study. LW, YY, and JL performed the study and carried out the implementation. WT, KH, TX, and ZF analyzed the data. LW and TL wrote the manuscript. TL and AY reviewed the manuscript and were responsible for the master plan.

Funding

This study was supported by the National Natural Science Foundation of China (No.81971660), Tianjin Outstanding Youth Fund Project (No. 20JCJQIC00230), Capital's Funds for Health Improvement and Research (2022-4-4019), National High-level Hospital Clinician Research Funding (APL221003101004060, 2022-PUMCH-A-017), Chinese Academy of Medical Science health innovation project (2021-I2M-042, 2021-I2M-058), Program of Chinese Institute for Brain Research in Beijing(2020-NKX-XM-14), Sichuan Science and Technology Program (No. 2021YFH0004), and Basic Research Program for Beijing-Tianjin-Hebei Coordination (19JCZDJC65500(Z)).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2022.1033422/full#supplementary-material

References

1. Bray F., Ferlay J., Soerjomataram I., Siegel R. L., Torre L. A., Jemal A., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*. (2018) 68(6): 394-424. doi:10.3322/caac.21492

2. Correa P., Piazuelo M. B. The gastric precancerous cascade. *Journal of digestive diseases* (2012) 13:2–9. doi:10.1111/j.1751-2980.2011.00550.x

3. Li D., Bautista M. C., Jiang S. F., Daryani P., Brackett M., Armstrong M. A., Hung Y., Postlethwaite D., Ladabaum U., Risks and Predictors of gastric Adenocarcinoma in patients with gastric intestinal metaplasia and Dysplasia: A population-based study. *Official journal of the american college of gastroenterology* (2016) 111(8):1104–13. doi:10.1038/ajg.2016.188

4. Suzuki H., Oda I., Abe S., Sekiguchi M., Mori G., Nonaka S., Yoshinaga S., Saito Y., High rate of 5-year survival among patients with early gastric cancer undergoing curative endoscopic submucosal dissection. *Gastric Cancer* (2016) 19(1):198–205. doi:10.1007/s10120-015-0469-0

5. Katai H., Ishikawa T., Akazawa K., Isobe Y., Miyashiro I., Oda I., Tsujitani S., Ono H., Tanabe S., Fukagawa T., Five-year survival analysis of surgically resected gastric cancer cases in Japan: A retrospective analysis of more than 100, 000 patients from the nationwide registry of the Japanese gastric cancer association (2001-2007). *Gastric Cancer* (2018) 21(1):144–154. doi:10.1007/s10120-017-0716-7

6. Chun H. J., Keum B., Kim J. H., Seol S. Y., Current status of endoscopic submucosal dissection for the management of early gastric cancer: A Korean perspective. *World journal of gastroenterology* (2011) 17(21):2592–2596. doi:10.3748/wjg.v17.i21.2592

7. Take I., Shi Q., Zhong Y. S. Progress with each passing day: Role of endoscopy in early gastric cancer. *Translational Gastrointest Cancer* (2015) 4(6):423–428. doi:10.3978/j.issn.2224-4778.2015.09.04

8. Buxbaum J. L., Hormozdi D., Dinis-Ribeiro M., Lane C., Dias-Silva D., Sahakian A., Jayaram P., Pimentel-Nunes P., Shue D., Pepper M., Narrow-band imaging versus white light versus mapping biopsy for gastric intestinal metaplasia: A prospective blinded trial. *Gastrointestinal endoscopy* (2017) 86(5):857–865. doi:10.1016/j.gie.2017.03.1528

9. Sumiyama K. Past and current trends in endoscopic diagnosis for early stage gastric cancer in Japan. *Gastric Cancer* (2017) 20(Suppl. 1):20–27. doi:10.1007/s10120-016-0659-4

10. Decharatanachart P., Chaiteerakij R., Tiyarattanachai T., Treeprasertsuk S. Application of artificial intelligence in chronic liver diseases: A systematic review and meta-analysis. *BMC Gastroenterol* (2021) 21(1):10. doi:10.1186/s12876-020-01585-5

11. Menon S., Trudgill N. How commonly is upper gastrointestinal cancer missed at endoscopy? A metaanalysis. *Endoscopy international open* (2014) 2(2):E46–50. doi:10.1055/s-0034-1365524

12. Gao Y., Zhang Z. D., Li S., Guo Y. T., Wu Q. Y., Liu S., Yang S., Ding L., Zhao B., Li S., Deep neural network-assisted computed tomography diagnosis of metastatic lymph nodes from gastric cancer. *Chinese Medical Journal* (2019) 132(23):2804–2811. doi:10.1097/CM9.0000000000000532

13. Ali H., Yasmin M., Sharif M., Rehmani M. H., Computer assisted gastric abnormalities detection using hybrid texture descriptors for chromoendoscopy images. *Comput Methods Programs Biomed* (2018) 157:39–47. doi:10.1016/j.cmpb.2018.01.013

14. Kanesaka T., Lee T. C., Uedo N., Lin K. P., Chen H. Z., Lee J. Y., Wang H., Chen H. Computer-aided diagnosis for identifying and delineating early gastric cancers in magnifying narrow-band imaging. *Gastrointest Endosc* (2018) 87(5): 1339–1344. doi:10.1016/j.gie.2017.11.029

15. Wang S., Zhu Y., Yu L., Chen H., Lin H., Wan X., Fan X., Heng P. Recalibrated multi-instance deep learning for whole slide gastric image classification. *Medical image analysis* (2019) 58:101549. doi:10.1016/j.media.2019.101549

16. Van D. S., Zinger S., Schoon E. J., With D. P. Supportive automatic annotation of early esophageal cancer using local Gabor and color features. *Neurocomputing* (2014) 144:92–106. doi:10.1016/j.neucom.2014.02.066

17. Liu D. Y, Gan T., Rao N. N., Xing Y. W., Zheng J., Li S., Luo C., Zhou Z., Wan Y. Identification of lesion images from gastrointestinal endoscope based on feature extraction of combinational methods with and without learning process. *Medical image analysis* (2016) 32:281–294. doi:10.1016/j.media.2016.04.007

18. Zhou R., Yang C., Meng M. Q. H., Xu G., Hu C., Li B. Capsule endoscopy images classification by random forests and ferns. In Proceedings of the 4th IEEE International Conference on information Science and technology, Shenzhen, China: IEEE (2014). p. 414–417. doi:10.1109/ICIST.2014.6920505

19. Riaz F., Silva F. B., Ribeiro M. D., Coimbra M. T. Invariant Gabor texture descriptors for classification of gastroenterology images. *IEEE Transactions Biomedical Engineering* (2012) 59(10):2893–2904. doi:10.1109/TBME.2012.2212440

20. Ali H., Sharif M., Yasmin M., Rehmani M. H., Riaz F. A survey of feature extraction and fusion of deep learning for detection of abnormalities in video endoscopy of gastrointestinal-tract. *Artificial Intelligence Review* (2020) 53: 2635–2707. doi:10.1007/s10462-019-09743-2

21. Gando G., Yamada T., Sato H., Oyama S., Kurihara M. Fine-tuning deep convolutional neural networks for distinguishing illustrations from photographs. *Expert Systems with Applications* (2016) 66:295–301. doi:10.1016/j.eswa.2016.08.057

22. Zhao Z., Jiao L., Zhao J., Gu J., Zhao J. Discriminant deep belief network for high-resolution SAR image classification. *Pattern Recognition* (2017) 61:686–701. doi:10.1016/j.patcog.2016.05.028

23. Bochkovskiy A, Wang CY, Liao H. YOLOV4: Optimal speed and accuracy of Object detection (2020). 10934, https://arxiv.org/abs/2004.10934.

24. Lin T. Y., Dollár P., Girshick R., He K., Hariharan B., Belongie S. Feature pyramid networks for object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. Hawaii: IEEE (2017). p. 2117–25. doi:10.48550/arXiv.1612.03144

25. Shankar K., Zhang Y., Liu Y., Wu L., Chen C. H. Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. *IEEE Access* (2020) 8: 118164–73. doi:10.1109/ACCESS.2020.3005152

26. Maghdid H. S., Asaad A. T., Ghafoor K. Z., Sadiq A. S., Khan M. K. Diagnosing COVID-19 pneumonia from x-ray and CT images using deep learning and transfer learning algorithms. *Multimodal Image Exploitation and Learning*, 2021:11734. doi:10.1117/12.2588672

27. Bulten W., Pinckaers H., van Boven H., Vink R., de Bel T., van Ginneken B., van der Laak J., Hulsbergen-van de Kaa C., Litjens G. Automated deep-learning system for Gleason grading of prostate cancer using biopsies: A diagnostic study. *Lancet Oncology* (2020) 21(2):233–41. doi:10.1016/S1470-2045(19)30739-9

28. Sajjad M., Khan S., Muhammad K., Wu W., Ullah A., Baik S. W. Multi-grade brain tumor classification using deep CNN with extensive data augmentation. *Journal of Computer Science* (2019) 30:174–182. doi:10.1016/j.jocs.2018.12.003

29. Huo Y., Xu Z., Xiong Y., Aboud K., Parvathaneni P., Bao S., 3D whole brain segmentation using spatially localized atlas network tiles. *Neuroimage* (2019) 194: 105–119. doi:10.1016/j.neuroimage.2019.03.041

30. Xie W., Jacobs C., Charbonnier J. P., van Ginneken B. Relational modeling for Robust and efficient Pulmonary Lobe segmentation in CT Scans. *IEEE Transaction Medical Imaging* (2020) 39(8):2664–75. doi:10.1109/TMI.2020.2995108

31. Zilly J., Buhmann J. M., Mahapatra D. Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. *Computerized Medical Imaging Graphics* (2017) 55:28–41. doi:10.1016/j.compmedimag.2016.07.012

32. Sun W, Tseng TB, Zhang J, Qian W. Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. *Computerized Medical Imaging Graphics* (2017) 57:4–9. doi:10.1016/j.compmedimag.2016.07.004

33. Huang R, Xie W, Vp-Nets ANJ. VP-Nets : Efficient automatic localization of key brain structures in 3D fetal neurosonography. *Medical Image Analysis* (2018) 47:127–39. doi:10.1016/j.media.2018.04.004

34. Alison Q., Kascenas A., Henry J., Wyeth D., Shepherd M., Beverid E., Clunie L., Sansom C. Attaining human-level performance with atlas location autocontext for anatomical landmark detection in 3D CT data. In: Proceedings of the European conference on computer vision (ECCV) Workshops (2019). Available at: https://link.springer.com/conference/eccv.

35. de Groof A. J., Struyvenberg M. R., van der Putten J., van der Sommen F., Fockens K. N, Curvers W. L., . Deep-learning system Detects neoplasia in patients with Barrett's Esophagus with higher accuracy than endoscopists in a Multistep training and validation study with benchmarking. *Gastroenterology* (2020) 158(4): 915–29.e4. doi:10.1053/j.gastro.2019.11.030

36. Ohmori M., Ishihara R., Aoyama K., Nakagawa K., Iwagami H., Matsuura N., Endoscopic detection and differentiation of esophageal lesions using a deep neural network. *Gastrointest Endosc* (2020) 91(2):301–9.e1. doi:10.1016/j.gie.2019.09.034

37. Pannala R., Krishnan K., Melson J., Parsi M. A., Schulman A. R., Sullivan S., Artificial intelligence in gastrointestinal endoscopy. *VideoGIE* (2020) 5(12): 598–613. doi:10.1016/j.vgie.2020.08.013

38. Wang C., Li Y., Yao J., Chen B., Song J., Yang X. Localizing and identifying intestinal metaplasia based on deep learning in Oesophagoscope. In Proceedings of the 8th International Symposium on Next Generation Electronics (ISNE). Zhengzhou, China: IEEE (2019). 1–4. doi:10.1109/ISNE.2019.8896546

39. Horiuchi Y., Aoyama K., Tokai Y., Hirasawa T., Yoshimizu S., Ishiyama A., Convolutional neural network for Differentiating gastric cancer from gastritis using magnified endoscopy with narrow band imaging. *Digestive Diseases and Sciences* (2020) 65(5):1355–63. doi:10.1007/s10620-019-05862-6

40. Yan T, Wong P. K., Choi I. C., Vong C. M., Yu H. H. Intelligent diagnosis of gastric intestinal metaplasia based on convolutional neural network and limited number of endoscopic images. *Computuers in Biology and Medicine* (2020) 126: 104026 doi:10.1016/j.compbiomed.2020.104026

41. Cho BJ, Bang CS, Park SW, Yang YJ, Seo SI, Lim H. Automated classification of gastric neoplasms in endoscopic images using a convolutional neural network. *Endoscopy* (2019) 51(12):1121–9. doi:10.1055/a-0981-6133

42. Lui T. K. L., Wong K. K. Y., Mak L. L. Y., To E. W. P., Tsui V. W. M., Deng Z., Guo J., Ni L., Cheung M. K. S., Leung W. K., Feedback from artificial intelligence improved the learning of junior endoscopists on histology prediction of gastric lesions. *Endosc Int Open* (2020) 8(2):E139–46. doi:10.1055/a-1036-6114

43. Liu X., Wang C., Bai J., Liao G. Fine-tuning pre-trained convolutional neural networks for gastric precancerous disease classification on magnification narrow-band imaging images. *Neurocomputing* (2020) 392:253–67. doi:10.1016/j.neucom.2018.10.100

44. Weiss K., Khoshgoftaar T. M., Wang D. A survey of transfer learning. *Journal of Big Data* (2016) 3(1):9–40. doi:10.1186/s40537-016-0043-6

45. Fu J., Zheng H., Mei T. Look Closer to See better: Recurrent attention convolutional neural network for fine-grained image recognition. In: Proceedings -30th IEEE Conference.On computer vision and pattern recognition. Hawaii: IEEE (2017). p. 4476–84. doi:10.1109/CVPR.2017.476

46. Guan H. Y., Zhong Z., Zheng Z., Zheng L., Yang Y. Diagnose like A Radiologist: Attention guided convolutional neural network for Thorax disease classification. arxiv: 1801.09927(2018).

47. Du W., Rao N., Dong C., Wang Y., Hu D., Zhu L., Zeng B., Gan T. Automatic classification of esophageal disease in gastroscopic images using an efficient channel attention deep dense convolutional neural network. *Biomedical Optic Express* (2021) 12(6):3066–81. doi:10.1364/BOE.420935

48. Wang F., Jiang M., Qian C., Yang S., Li C., Zhang H., Wang X., Tang X. Residual attention network for image classification. In: Proceedings -30th IEEE Conference.On computer vision and pattern recognition. Hawaii: IEEE (2017). p. 6450–8. doi:10.1109/CVPR.2017.683

49. Chollet F. Xception: Deep learning with depthwise separable convolutions (2016). arxiv:1610.02357v2.

50. Szegedy C, Ioffe S, Vanhoucke V. Inception-v4, inception-resnet and the impact of residual connections on learning. arxiv:1602.07261. (2016).

51. Yu F, Koltun V. Multi-scale context aggregation by dilated convolutions. arxiv: 1511.07122, (2015).

52. Zhang Y, Kang B, Hooi B, Yan S., Feng J., Deep long-tailed learning: A survey (2021). Available at: https://arxiv.org/abs/2110.04596.

53. Zhang Y., Li F., Yuan F., Zhang K., Huo L., Dong Z., . Diagnosing chronic atrophic gastritis by gastroscopy using artificial intelligence. *Digestive and Liver Disease* (2020) 52(5):566–72. doi:10.1016/j.dld.2019.12.146

54. Zheng W, Yu T, Lin N, Ye G, Zhu X, Shen Y, Tu1075 deep convolutional neural networks for recognition of atrophic gastritis and intestinal metaplasia based on endoscopy images. *Gastrointest Endosc* (2020) 91(6):AB533–4. doi:10.1016/j.gie.2020.03.3247

frontiers | Frontiers in Physics

TYPE Brief Research Report
PUBLISHED 11 November 2022
DOI 10.3389/fphy.2022.1057002

Check for updates

OPEN ACCESS

EDITED BY
Wei Gong,
Zhejiang University, China

REVIEWED BY
Leiting Pan,
Nankai University, China
Jiaye He,
National Innovation Center for Advanced Medical Devices, China

*CORRESPONDENCE
Yiyan Fei,
fyy@fudan.edu.cn
Lan Mi,
lanmi@fudan.edu.cn
Jiong Ma,
jiongma@fudan.edu.cn

SPECIALTY SECTION
This article was submitted to Optics and Photonics, a section of the journal Frontiers in Physics

RECEIVED 29 September 2022
ACCEPTED 24 October 2022
PUBLISHED 11 November 2022

CITATION
Liu Y, Chen Q, Fei Y, Mi L and Ma J (2022), Design of an ellipsoidal mirror assisted oblique plane microscopy. *Front. Phys.* 10:1057002. doi: 10.3389/fphy.2022.1057002

COPYRIGHT
© 2022 Liu, Chen, Fei, Mi and Ma. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Design of an ellipsoidal mirror assisted oblique plane microscopy

Yuzhe Liu[1], Qiushu Chen[1], Yiyan Fei[1]*, Lan Mi[1]* and Jiong Ma[1,2,3]*

[1]Department of Optical Science and Engineering, Shanghai Engineering Research Center of Ultra-Precision Optical Manufacturing, Key Laboratory of Micro and Nano Photonic Structures (Ministry of Education), Fudan University, Shanghai, China, [2]Institute of Biomedical Engineering and Technology, Academy for Engineer and Technology, Fudan University, Shanghai, China, [3]Shanghai Engineering Research Center of Industrial Microorganisms, The Multiscale Research Institute of Complex Systems (MRICS), School of Life Sciences, Fudan University, Shanghai, China

Using one objective for excitation and detection simultaneously, oblique plane microscopy (OPM) provides a mounting-friendly approach for optical sectioning. Unfortunately, the original OPM has three major defects: the mechanical constraints when placing the objectives, the phase loss and the resulting anisotropy of the point spread function (PSF). In order to alleviate the above defects, an ellipsoidal mirror assisted oblique plane microscopy (EM-OPM) was proposed. By inserting an ellipsoidal mirror into the optical path to help collect the light beam, the problem of placing the objectives was solved. The numerical calculation results showed that EM-OPM can obtain higher relative light intensity and larger effective area of exit pupil than OPM when the tilt angle of the light sheet becomes larger. The imaging simulation results showed that EM-OPM effectively solves the problem of resolution reduction in the Y direction of OPM. In addition, optimization of the higher-order terms of the ellipsoidal mirror further improved the imaging ability of EM-OPM in large field of view (FOV).

KEYWORDS

oblique plane microscopy, (point spread function), phase loss, mechanical constraint, ellipsoidal mirror, light-sheet fluorescence microscopy, anisotropy

1 Introduction

In order to achieve high-speed, high-resolution, low damage and large FOV imaging simultaneously, light sheet fluorescence microscopy (LSFM) came into being [1–3]. The biggest difference between light sheet microscopy and wide field microscopy comes from different illuminating methods. In classical configuration of LSFM, there are two orthogonally placed objective lenses with the sample at the common focal point. One objective lens, also known as the illumination objective, focus laser light to a thin sheet which is always perpendicular to the optical axis of the other objective also known as the detection objective. The key point of LSFM is that only the sample located in the focal plane of the detection objective lens is illuminated, while the upper and lower samples are

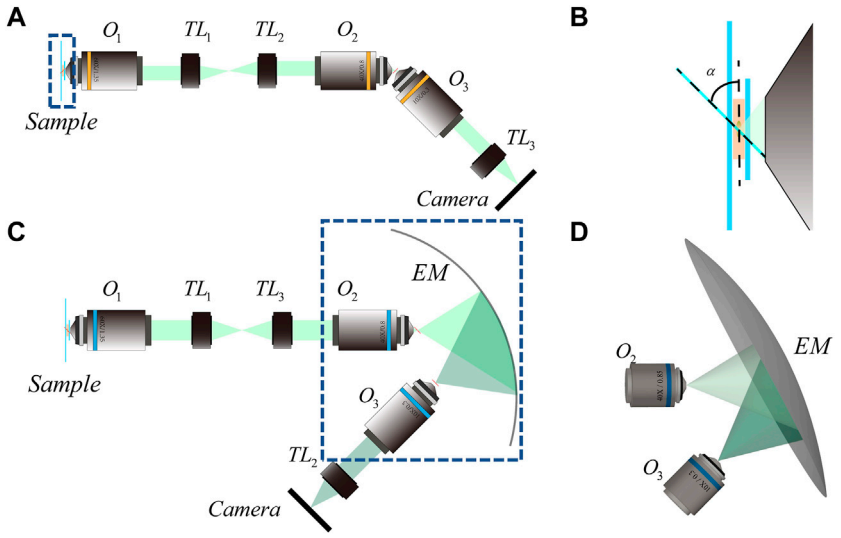


FIGURE 1
Schematic diagram of OPM system and EM-OPM system **(A)** 2D optical path diagram of OPM system **(B)** Enlarged view of the part bounded by dotted lines in Figure **(A–C)** 2D optical path diagram of EM-OPM system **(D)** The three-dimensional schematic diagram of the dotted line in Figure **(B)**.

not affected. Therefore, LSFM can be used for long-term biological studies with high axial resolution and minimal phototoxicity.

Although LSFM has the above advantages, it has some limitations. The dual objective geometry and the need for side-on illumination restricts the type of sample that can be mounted between the objectives. In addition, in a high-resolution LSFM system, the sample has to directly contact the non-sterile optical surface to ensure the high numerical aperture of the immersion objective lens. These challenges have led to innovative strategies for optical path design, including LSFM that uses the same objective lens for illumination and detection [4, 5].

In 2008, Dunsby proposed oblique plane microscopy (OPM) to enable LSFM using a single objective to illuminate the specimen and collect the resulting fluorescence at the same time [6]. As shown in Figure 1A, the light sheet emitted from the edge of O_1 illuminate the sample obliquely. The fluorescence illuminated by the light sheet is also captured by O_1 but cannot be imaged by it directly because of the severe optical aberrations. The solution in OPM is to employ a one-to-one magnification system which simultaneously follow both the sine and Herschel condition by introducing O_2 [7]. The oblique intermediate real image is then brought to lie perpendicular to the optical axis of O_3 and can be imaged in a conventional way by a camera. Due to the improvement of illumination path, OPM can be assembled based on standard inverted microscopes, and is suitable for traditional sample placement and laser based auto focusing. However, there are three main defects in the original OPM: 1, mechanical constraint between O_2 and O_3; 2, Loss of phase information on pupil plane (Short for Phase Loss); 3. Anisotropy of PSF [8, 9]. For the first defect, in addition to sacrificing the numerical aperture of O_3 for working distance, installing a micro mirror after O_2 can also eliminate the mechanical size limitation between O_2 and O_3 [10–12]. However, the fine fabrication and accurately installation of a micro mirror is technically challenging. Using a diffractive grating can achieve the same goal, but the intensity will be dispersed by the diffractive element [13]. To mitigate the impact of the second defect, Yang et al introduced a small special water container at the rear focus to change the refractive index of the light beam on both sides, so that almost all beams from O_2 can be collected by O_3 [14]. Although the phase loss of the system is effectively reduced, the addition of the water container increases the difficulty of focusing the system and intensifies the mechanical constraint between O_2 and O_3. A similar idea was also put forward later [15], while facing the same challenge. For the third defect, Kim et al introduced a polarizer to change the polarization state of the beam thus improve the anisotropy of the overall PSF [16]. However, the improvement is unstable and the polarized beam splitter (PBS) in the optical path makes the phase loss more serious. To sum up, OPM and its existing improved system cannot solve the three defects mentioned above at the same time.

In this paper, inspired by the optical property of ellipse, we proposed an ellipsoidal mirror assisted oblique plane microscopy (EM-OPM), which provides some advantages over the original OPM and other OPMs. In this configuration, the light beam emitted by O_2 is collected by O_3 after being reflected by the ellipsoidal mirror, thus avoiding the top-to-top placement of O_2

and O_3. Numerical results demonstrate the advantages of EM-OPM in terms of relative light intensity and relative effective pupil area. Imaging simulation results prove that the resolution of EM-OPM in Y direction is improved. In addition, the optimization of ellipsoidal mirror further improves the large FOV imaging capability of EM-OPM.

2 Materials and methods

2.1 EM-OPM

As mentioned above, the original OPM has three defects. In order to solve these three problems, we proposed a new optical configuration, namely EM-OPM, which is expected to solve these three defects simultaneously. Except for the introduction of an ellipsoidal mirror (EM) between O_2 and O_3, EM-OPM uses elements similar to the original OPM. The expression of the EM is:

$$z = \frac{cr^2}{1 + \sqrt{1 - (1 + k)c^2r^2}} \tag{1}$$

where c represents the curvature, k represents the conic coefficient. In our design, $c = 0.01$ and $k = 0.01$. r is the unit radial coordinate and z is the height of ellipsoidal mirror.

The way the sample is illuminated is depicted in Figure 1B. The inclined light sheet used for illuminating the sample in OPM is formed by scanning an inclined beam with an angle of α to the focal plane. The layout of EM-OPM is shown in Figure 1C. The back focus of O_2 is coincident with one of the focuses of EM when back focus of O_3 is coincident with the other. By adjusting the tilting angle of each element, the major axis of EM is parallel to the light sheet and perpendicular to the optical axis of O_3 at the same time. Although the real image of the sample is still inclined to the focal plane of O_1, it lies parallel to the focal plane of O_3, thus make the ordinary flat field imaging using O_3 possible. Our design can be understood as an extension of the remote focusing (RF) system [17], because the information of the sample is further copied from the back focal plane of O_2 to a more flexible space by an EM.

2.2 Mechanical constraint

The mechanical constraint of O_2 and O_3 is mainly due to the fact that the working distances of the two objective lenses is far less than their mechanical dimensions. To make sure that the objectives will not grind against each other, the working distance (WD_2 for O_2 and WD_3 for O_3) and the radius of the glass cover plate (r_{CG2} for O_2 and r_{CG3} for O_3) should meet at least one of the following requirements:

$$\begin{cases} WD_2 + WD_3 \cdot \tan\alpha > r_{CG2} \cdot \tan\alpha \\ WD_2 \cdot \cos\alpha + r_{CG2} \cdot \sin\alpha > r_{CG3} \end{cases} \tag{2}$$

In the system proposed by Dunsby, O_2 is a 40×/0.85 air objective with $r_{CG2} = 3.750$ mm and $WD_2 = 0.200$ mm. To meet the requirements above when $\alpha = 45$ for example, WD_3 must be greater than 2.510mm, otherwise r_{CG3} must be less than 2.793 mm. We can use d, the distance between the vertices of O_2 and O_3, to indicate the severity of the mechanical constraint. In OPM:

$$d = \sqrt{WD_2^2 + WD_3^2 + 2 \cdot WD_2 \cdot WD_3 \cdot \sin\alpha} \tag{3}$$

In EM-OPM, d is determined by the parameters of the EM, which can be calculated as:

$$d = 2 \cdot \sqrt{\frac{k}{c^2(1 - k)^2}} \tag{4}$$

2.3 Phase loss

The phase loss mainly occurs between O_2 and O_3 (Some OPM using a planar micro mirror lose extra 50% intensity due to the introduction of PBS). To make it clear, it is the reduction in NA and the tilting placement of O_3 that result in the phase loss. In order to calculate the phase loss, it is necessary to calculate the initial pupil ignoring O_3 and the effective pupil of the complete system. The calculation of the initial and effective pupil can be based on either the strict analytic geometry method [18] or the Monte Carlo algorithm. In this paper, the three-dimensional point clouds representing the pupils of OPM and EM-OPM are obtained after ray tracing using Monte Carlo algorithm. The Delaunay triangulation algorithm is used to convert the three-dimensional point cloud into a triangular mesh [19]. Then, all the meshes are traversed, and the area of each triangular mesh is calculated separately. After accumulation, the approximate effective pupil area can be obtained. Based on the theory above, we compared the effective pupil of OPM and EM-OPM.

2.4 Point spread function

The third defect of OPM is the anisotropy of PSF. Because objective lenses with large numerical aperture are used, the paraxial approximation scalar diffraction theory is no longer applicable [20]. Based on the vector diffraction theory [21], the electromagnetic field distribution at any point $p\ (x, y, z)$ in the focus area of the objective lens can be regarded as the superposition of all diffracted plane waves that can pass

through the pupil of the objective lens within the solid angle (determined by NA).

$$E(x,y,z) = -\frac{i}{\lambda}\int_0^{\beta}\int_0^{2\pi} \sin\theta\cos\theta A(\theta,\varphi)P(\theta,\varphi) \times \exp\left[ikn\left(z\cos\theta + x\sin\theta\cos\varphi + y\sin\theta\sin\varphi\right)\right]d\theta d\varphi \quad (5)$$

With $0<\theta<\beta$, where β is the maximum focus angle of the objective lens. φ represents the azimuth angle in the object plane, λ represents the wavelength. $A\ (\theta,\ \varphi)$ is the amplitude of the incident beam. $p\ (\theta,\ \varphi)$ indicates the polarization state of the EM field in the focal region.

The effective pupil can be cut into a large number of sub regions by using the Monte Carlo algorithm to trace a large number of rays. The position of each sub region can be approximately determined by one of the points $u_i\ (\theta_i,\ \varphi_i,\ r_i)$. It is approximately considered that the plane wave from each sub region to the point$p$ has a unique direction vector. Then Eq. (5) can be written by a discretely way as:

$$E(x,y,z) = \sum_{(u_i(\theta_i,\varphi_i,r_i)\in R)} \sin\theta_i\cos\theta A(\theta_i,\varphi_i)P(\theta_i,\varphi_i) \times \exp\left[ikn\left(z\cos\theta_i + x\sin\theta_i\cos\varphi_i + y\sin\theta_i\sin\varphi_i\right)\right] \quad (6)$$

In this study, the numerical calculation is programmed by a personal code script based on Matlab software. The code allows to simulate PSFs for OPM and EM-OPM under different conditions, such as different NA, different light sheet angles, etc.

3 Results and discussion

As a single-objective LSFM methods, OPM offers a convenient approach for optical sectioning using a conventional epi-fluorescence microscope. The trade-offs it has to make in order to achieve the convenience of sample mounting sacrifice the imaging quality. Unfortunately, all of the improved configurations up to now have failed to reduce the three defects mentioned above simultaneously. By inserting an ellipsoidal mirror into the space near the focal region of O_2 and titling the following elements accordingly, we reconstructed OPM into EM-OPM. EM-OPM has obvious progress compared with OPM, especially in the alleviation of the three defects. We used numerical calculations to illustrate these advances.

We proved the alleviation of the first defect by comparing the space between O_2 and O_3 in EM-OPM and OPM. Except for some extremely high NA objectives for special purposes, only a series of small NA objectives with ×10 magnification can meet the requirements in Eq. (5). Therefore, O_3 is a 10×/0.3 air objective which has a working distance of 3.100 mm in original OPM. Even it is very easy to align and the working distance is plenty, the distance between the two vertices is only 3.241 mm (Figure 2A). In contrast, in EM-OPM, the distance between the two vertices is relaxed to 20.202 mm (Figure 2B). It significantly eased the mechanical constraint and therefore increased the flexibility of O_3 selection.

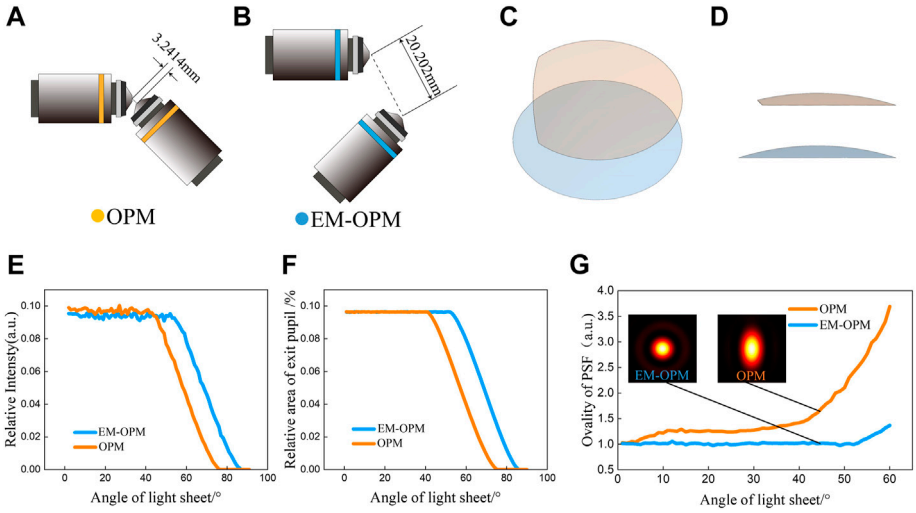


FIGURE 2
Comparison between OPM system and EM-OPM system **(A)** In OPM system, the distance between two objective lens vertices is 3.241 mm **(B)** In the EM-OPM system, the distance between the two objective lens vertices is 20.202 mm **(C)** Comparison of effective pupil of OPM system and EM-OPM system (45 top view) **(D)** Comparison of effective pupil of OPM system and EM-OPM system (front view) **(E)** The relative intensity of the two systems varies with α from 0 to 90 **(F)** The relative effective exit pupil area of the two systems varies with α from 0 to 90 **(G)** The ellipticity of the PSF of the two systems with different α and the PSF distribution of the two systems when α = 45.

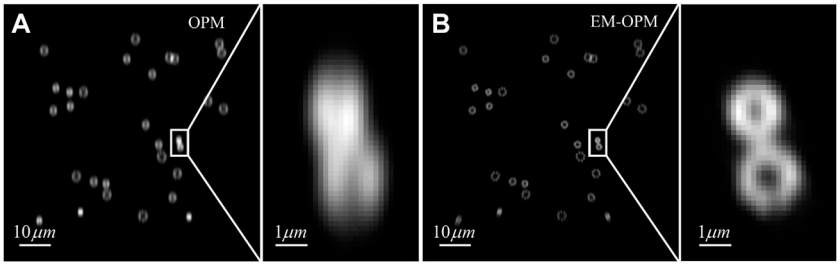


FIGURE 3
Simulated image obtained by imaging with OPM and EM-OPM **(A)** Simulation image of OPM **(B)** Simulation image of EM-OPM.

The effective pupil of OPM and EM-OPM when $\alpha = 45$ are shown in Figures 2C,D. Both remaining a small part of the spherical shell, the effective pupil of OPM is cut out while the effective pupil of the EM-OPM is still complete. Due to the reduction of NA of objective lens, the pupil of both systems becomes smaller. Effective pupil of EM-OPM faithfully reflects the reduction of NA but avoids the influence of the inclined placement of O_2 and O_3. When $\alpha < 42$, the relative light intensity and the relative effective pupil area of OPM and EM-OPM are approximately the same (Figure 2E). However, EM-OPM system has higher relative light intensity and larger effective pupil area when $\alpha \geq 42$ (Figure 2F). When α is too large, both systems are unable to collect the light from O_2. For OPM, α should be smaller than 75, while for EM-OPM, α should be smaller than 85. Calculation result shows that EM-OPM can make more effective use of fluorescence. This is exactly the goal of researchers to reduce phase loss when improving OPM.

The third defect of OPM system, namely the anisotropy of PSF, can also be understood as the loss of resolution in Y direction. Generally, the more complete the effective pupil, the more complete the high-frequency information contained in the optical system, which means that the resolution of the system in each direction is not lost. According to the results in the previous section, the PSF of EM-OPM is predicted to be more isotropic. The numerical calculation results also agree with this idea. In Figure 2G, the anisotropy of PSF is measured by a defined parameter, μ calculated as:

$$\mu = \frac{\text{FWHMy of PSF}}{\text{FWHMx of PSF}} \tag{7}$$

With the increase of α, μ of the OPM increases, while μ of EM-OPM is relatively stable. For example, when $\alpha = 45$, the PSF of EM-OPM presents a perfect quasi-Gaussian distribution, while the PSF of OPM is elongated along the y-direction. In other words, EM-OPM can recover the Y direction loss of resolution in OPM.

We then demonstrate the power of EM-OPM by presenting the imaging simulation of OPM and EM-OPM. We focus on the imaging performance of Fluorescence conjugated to the focal plane of O_3, therefore the simulation results are 2D images. Firstly, we generate the ground truth pattern used for simulation. Here to highlight the difference in resolution, we choose hollow circular structure to be the sample pattern which is aimed to model nuclear pore complexes (NPCs). This pattern consists of octagons labelled at their vertices for the convenience of calculation. The positions of the centers of octagons randomly distributed within an active area of 81.92 × 81.92 μm2 and the octagon radius is randomly distributed between 400 nm and 1200 nm. Secondly, we simulate the imaging process. In most algorithms used for single molecule localization microscopy (SMLM) simulation [22, 23], PSF is usually space-invariant. In order to accurately study the imaging results under different FOV, we have established the communication between our imaging simulation algorithm and Zemax/OpticStudio™. Because the built-in ray tracing package of Zemax can easily calculate the PSF at different FOVs. Therefore, the PSF of different areas in the FOV can be obtained and then convolved with the ground truth image. Figure 3A is the simulation result obtained using the OPM system, and Figure 3B is the one obtained using the EM-OPM system. The improvement of EM-OPM can be found by comparing the imaging results of the entire FOV or ROI area delineated by white solid lines. OPM will significantly reduce the Y resolution of the image, while EM-OPM can achieve higher resolution in both directions.

EM-OPM has significant advantages over OPM, but the ellipsoidal mirror will introduce serious coma, which limits the use of EM-OPM in large FOV. It is also illustrated by the severe degradation of imaging quality in the large FOV area in Figure 3B (note the two octagons marked with arrows). In order to improve the imaging capability in large FOV, we designed an optimization function to comprehensively evaluate the imaging ability of each vertex in the field of 60 × 60 μm2, and optimized the high-order terms in the expression of EM. The optimization results are shown in Figure 4A,B. It can be seen that the normalized Strehl ratio of the optimized EM-OPM at each

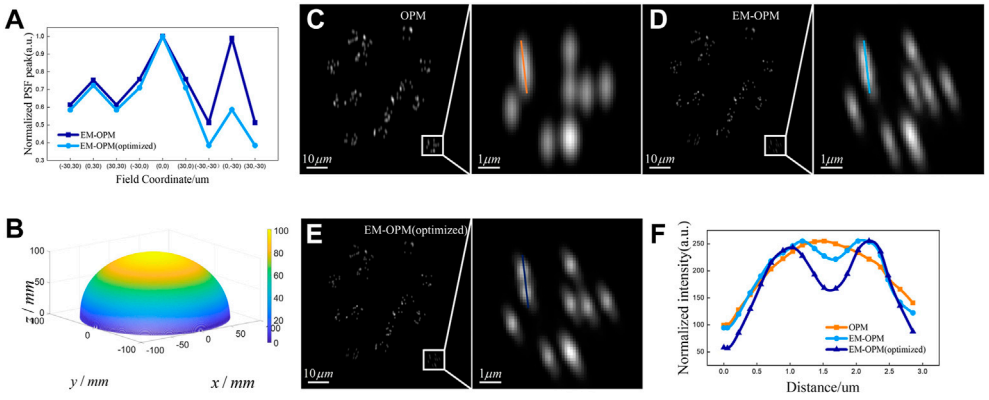


FIGURE 4
Optimization results and imaging simulation comparison of OPM, EM-OPM and EM-OPM (optimized) **(A)** Normalized Strehl ratio of EM-OPM and EM-OPM (optimized) at each FOV point **(B)** 3D profile of the optimized ellipsoidal mirror **(C)** Simulation imaging results of OPM **(D)** Simulation imaging results of EM-OPM before optimization **(E)** Simulation imaging results of optimized EM-OPM **(F)** Resolution comparison of three systems.

FOV point is higher than that of the pre-optimized EM-OPM. The optimization result indicates that the imaging contrast of the system at each position in the large FOV is improved.

The imaging simulations of OPM, EM-OPM and EM-OPM (optimized) were then carried out respectively. To compare the resolution and edge distortion of these three systems at the same time, fluorescence beads distributed randomly are chosen to be the ground truth pattern. Several fluorescent beads are set to be in a group, and the center of each group are still randomly distributed within an active area of 81.92 × 81.92 μm2. We also define that the fluorescent beads of each group are randomly distributed in a circle with a radius of 3000 nm. The simulation results of the three systems are shown in Figures 4C–E respectively. The simulation results prove again that the original OPM system has the problem of reduced resolution in the Y direction, which makes it impossible to distinguish some fluorescent spheres with close distance using OPM. In EM-OPM, the problem of PSF anisotropy has been significantly alleviated. However, the decrease of image contrast and serious side lobe indicate that the imaging capability in large FOV is still unsatisfying. Fortunately, the optimized EM-OPM make some change. A line segment is used to extract the intensity distribution in the ROI regions in Figures 4C–E and the resolution comparison of the three systems is shown in Figure 4F. It can be seen that OPM is unable to distinguish the two fluorescent beads that are close to each other, EM-OPM can barely distinguish them, and the optimized EM-OPM further improves the resolution.

4 Conclusion

EM-OPM proposed in this paper is an improvement based on OPM. An ellipsoidal mirror is inserted into the space between O_2 and O_3. The optical property of the ellipse is used to collect the beam so that the mechanical constraint between O_2 and O_3 can be solved. The change of relative intensity and the change of relative effective pupil area under different tilt angles of the light sheet are studied and compared by numerical calculation method. The results prove that: the relative light intensity and the relative effective pupil area of OPM and EM-OPM are approximately the same when $\alpha < 42$, and the relative light intensity and the relative effective pupil area of EM-OPM are larger than that of OPM when $\alpha \geq 42$. This means EMOPM can make more effective use of fluorescence and retain as much optical high-frequency information as possible. The alleviation in phase loss is also the basis for the improvement of PSF anisotropy. Compared to OPM, EM-OPM has higher resolution in Y direction for any tilt angle of the light sheet.

Due to the complexity of the actual imaging system, image simulation based on space-variant PSFs is carried out. Simulation results show that OPM loses resolution in Y direction, while EM-OPM can achieve higher resolution in both directions. The simulation results of imaging also reveal the weakness of EM-OPM in the large FOV imaging. In order to improve the imaging ability in large FOV, we carried out optimization for high order terms of the ellipsoidal mirror. The optimization improved the normalized Strehl ratio of the system at each vertex within the field of 60 × 60 μm2. The later imaging simulation proves that the optimized EM-OPM further improves the resolution and contrast in large FOV. With the development of ultra-precision machining capability of free-form surface, the optimized EM-OPM is expected to replace OPM in low and medium resolution applications.

Data availability statement

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

Author contributions

YL and JM conceived the project; YL performed the experiments; YL wrote the manuscript and performed simulation and data analysis; QC, YF, LM, and JM revised the manuscript.

Funding

This work was financially supported by the National Key R&D Program of China (2021YFF0502900), National Natural Science Foundation of China (62175034, 62175036ml, 82030106fyy, 32271510), Shanghai Natural Science Foundation (grant No. 20ZR1405100, 20ZR1403700fyy), Science and Technology Research Program of Shanghai (grant No. 19DZ2282100), Shanghai key discipline construction plan (2020–2022) (grant No. GWV-10.1-XK01), Shanghai Engineering Technology Research Center of Hair Medicine (19DZ2250500), Medical Engineering Fund of Fudan University (yg2021-022), Pioneering Project of Academy for Engineering and Technology, Fudan University (gyy2018-001, gyy2018-002), Yantai Returned Scholars' Pioneering Park.

Acknowledgments

The authors wish to thank JW for help with the use of Zemax, MJ, YC for help in revising the manuscript.

Conflict of interest

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

Publisher's note

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

References

1. Stelzer EH, Strobl F, Chang BJ, Preusser F, Preibisch S, McDole K, et al.Light sheet fluorescence microscopy. *Nat Rev Methods Primers* (2021) 1(1):73–25. doi:10.1038/s43586-021-00069-4

2. Olarte OE, Andilla J, Gualda EJ, Loza-Alvarez P. Light-sheet microscopy: A tutorial. *Adv Opt Photon* (2018) 10:111–79. doi:10.1364/AOP.10.000111

3. Weber M, Huisken J. Light sheet microscopy for real-time developmental biology. *Curr Opin Genet Dev* (2011) 21:566–72. doi:10.1016/j.gde.2011.09.009

4. Gagliano G, Nelson T, Saliba N, Vargas-Hernández S, Gustavsson AK. Light sheet illumination for 3D single-molecule super-resolution imaging of neuronal synapses. *Front Synaptic Neurosci* (2021) 13:761530. doi:10.3389/fnsyn.2021.761530

5. You R, McGorty R. Light sheet fluorescence microscopy illuminating soft matter. *Front Phys* (2021) 9:760834. doi:10.3389/fphy.2021.760834

6. Dunsby C. Optically sectioned imaging by oblique plane microscopy. *Opt Express* (2008) 16(25):20306–16. doi:10.1364/OE.16.020306

7. Botcherby EJ, Juskaitis R, Booth MJ, Wilson T. Aberration-free optical refocusing in high numerical aperture microscopy. *Opt Lett* (2007) 32:2007–9. doi:10.1364/OL.32.002007

8. Theer P, Dragneva D, Knop M. πSPIM: High NA high resolution isotropic light-sheet imaging in cell culture dishes. *Sci Rep* (2016) 6:32880. doi:10.1038/srep32880

9. Smith CW, Botcherby EJ, Wilson T. Resolution of oblique-plane images in sectioning microscopy. *Opt Express* (2011) 19(3):2662–9. doi:10.1364/OE.19.002662

10. Bouchard MB, Voleti V, Mendes CS, Lacefield C, Grueber WB, Mann RS, et al.Swept confocally-aligned planar excitation (SCAPE) microscopy for high-speed volumetric imaging of behaving organisms. *Nat Photon* (2015) 9(2):113–9. doi:10.1038/nphoton.2014.323

11. Anselmia F, Ventalon C, Begue A, Ogden D, Emiliani V. Three-dimensional imaging and photostimulation by remote-focusing and holographic light patterning. *Proc Natl Acad Sci U S A* (2011) 108(49):19504–9. doi:10.1073/pnas.1109111108

12. Li T, Ota S, Kim J, Wong ZJ, Wang Y, Yin X, et al.Axial plane optical microscopy. *Sci Rep* (2015) 4:7253. doi:10.1038/SREP07253

13. Hoffmann M, Judkewitz B. Diffractive oblique plane microscopy. *Optica* (2019) 6(9):1166–70. doi:10.1364/OPTICA.6.001166

14. Yang B, Chen X, Wang Y, Feng S, Pessino V, Stuurman N, et al.Epi-illumination SPIM for volumetric imaging with high spatial-temporal resolution. *Nat Methods* (2019) 16(6):501–4. doi:10.1038/s41592-019-0401-3

15. Sapoznik E, Chang BJ, Huh J, Ju RJ, Azarova EV, Pohlkamp TA versatile oblique plane microscope for large-scale and high-resolution imaging of subcellular dynamics. *Elife* (2020) 9:e57681. doi:10.7554/eLife.57681

16. Kim J, Li T, Wang Y, Zhang X. Vectorial point spread function and optical transfer function in oblique plane imaging. *Opt Express* (2014) 22(9):11140. doi:10.1364/OE.22.011140

17. Botcherby EJ, Juškaitis R, Booth MJ, Wilson T. An optical technique for remote focusing in microscopy. *Opt Commun* (2008) 281(4):880–7. doi:10.1016/j.optcom.2007.10.007

18. Kim J, Wojcik M, Wang Y, Moon S, Zin EA, Marnani N, et al.Oblique-plane single-molecule localization microscopy for tissues and small intact animals. *Nat Methods* (2019) 16:853–7. doi:10.1038/s41592-019-0510-z

19. Barber CB, Dobkin DP, Huhdanpaa H. The quickhull algorithm for convex hulls. *ACM Trans Math Softw* (1993) 22(4):469–83. doi:10.1145/235815.235821

20. Gu M. *Advanced optical imaging theory*. Berlin; Heidelberg: Springer-Verlag (2000).

21. Richards B, Wolf E. Electromagnetic diffraction in optical systems. II. Structure of the image field in an aplanatic system. *Proc R Soc A. Math Phys Eng Sci* (1959) 253(1274):358. doi:10.1098/rspa.1959.0200

22. Sinkó J, Kákonyi R, Rees E, Metcalf D, Knight AE, Kaminski CF, et al.TestSTORM: Simulator for optimizing sample labeling and image acquisition in localization based super-resolution microscopy. *Biomed Opt Express* (2014) 5(3):778–87. doi:10.1364/BOE.5.000778

23. Venkataramani V, Herrmannsdörfer F, Heilemann M, Kuner T. SuReSim: Simulating localization microscopy experiments from ground truth models. *Nat Methods* (2016) 13(4):319–21. doi:10.1038/nmeth.3775

frontiers | Frontiers in Physics

TYPE Original Research
PUBLISHED 02 December 2022
DOI 10.3389/fphy.2022.1083558

Check for updates

OPEN ACCESS

EDITED BY
Wei Gong,
Zhejiang University, China

REVIEWED BY
Lifang Wei,
Fujian Agriculture and Forestry University, China
Shaowei Wang,
Xi'an Jiaotong University, China

*CORRESPONDENCE
Zhengxia Wang,
zxiawang@hainanu.edu.cn

SPECIALTY SECTION
This article was submitted to Optics and Photonics, a section of the journal Frontiers in Physics

RECEIVED 29 October 2022
ACCEPTED 24 November 2022
PUBLISHED 02 December 2022

CITATION
Feng Q, Song Q, Yan M, Huang ZL and Wang Z (2022), MSDenoiser: Muti-step adaptive denoising framework for super-resolution image from single molecule localization microscopy. *Front. Phys.* 10:1083558. doi: 10.3389/fphy.2022.1083558

COPYRIGHT
© 2022 Feng, Song, Yan, Huang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

MSDenoiser: Muti-step adaptive denoising framework for super-resolution image from single molecule localization microscopy

Qianghui Feng[1], Qihang Song[2], Meng Yan[1], Zhen Li Huang[2] and Zhengxia Wang[1]*

[1]School of Computer Science and Technology, Hainan University, Haikou, China, [2]Key Laboratory of Biomedical Engineering of Hainan Province, School of Biomedical Engineering, Hainan University, Haikou, China

Recent developments in single-molecule localization microscopy (SMLM) enable researchers to study macromolecular structures at the nanometer scale. However, due to the complexity of imaging process, there are a variety of complex heterogeneous noises in SMLM data. The conventional denoising methods in SMLM can only remove a single type of noise. And, most of these denoising algorithms require manual parameter setting, which is difficult and unfriendly for biological researchers. To solve these problems, we propose a multi-step adaptive denoising framework called MSDenoiser, which incorporates multiple noise reduction algorithms and can gradually remove heterogeneous mixed noises in SMLM. In addition, this framework can adaptively learn algorithm parameters based on the localization data without manually intervention. We demonstrate the effectiveness of the proposed denoising framework on both simulated data and experimental data with different types of structures (microtubules, nuclear pore complexes and mitochondria). Experimental results show that the proposed method has better denoising effect and universality.

KEYWORDS

noise reduction, super-resolution image processing, multi-step denoising framework, adaptive parameter selection, localization data

1 Introduction

The spatial resolution of conventional optical microscopy techniques is limited to about half the wavelength of light. This is mainly due to the diffraction of light: when the light source passes through the optical imaging system, it will form a spot on the focal plane, which we call point spread function (PSF). PSF has normally a central peak of about 200–300 nm in width. Super-resolution microscopy overcomes the resolution limit, and achieves a spatial resolution on the order of 10–20 nm [1]. Among a variety of super-resolution imaging techniques, single-molecule localization microscopy (SMLM) with a

straightforward principle and outstanding spatial resolution gains intensive attention from researchers [2]. SMLM is primarily based on the fact that the spatial coordinates of single fluorescent molecules can be established with high precision, if their PSFs do not overlap with each other. Currently, SMLM is able to image subcellular structures (such as nuclear pores complexes, viruses, chromatin complexes, and cytoskeletal filaments) with unprecedented details, and provides great opportunities for biomedical researchers in resolving biological structures at the nanoscale [3]. Unlike many other super-resolution microscopy strategies, such as Stimulated Emission Depletion (STED) [4] and Structured Illumination Microscopy (SIM) [5] that generate grayscale images directly, SMLM requires a series of complicated procedures to processes the raw images and finally reconstructs a final super-resolution image, as can be seen from the description below. Taking single color SMLM as an example, we label biological structures with a specific type of fluorescent molecules, and separate the spatially overlapping fluorescent emissions from these molecules into a series of sub images using the photophysical characteristics of the fluorescent molecules. In each diffraction-limited region, we control to excite only one or two fluorescent molecule. In this case, we can keep a low overlapping probability. After accumulating thousands of image frames of randomly distributed fluorescent molecule images, we apply a proper single molecule localization algorithm to the raw images to precisely find the center locations of each molecule. We combine all of the gathered localization points to create a localization table, which contains at least x, and y positions of individual emitters and sometimes complements by information on localization uncertainty, and emitter intensity, etc. Finally, we use the localization table to render a super-resolution image, which can present super-resolution topography information of the observed structures.

A typical SMLM image is usually suffered from a large amount of mixed and complex background noises, which are originated from autofluorescence, out-of-focus fluorescence, camera noises, as well as non-specific labelled fluorescent molecules. These background noises lead to the degradation of super-resolution image quality and affect the subsequent data analysis and processing [6, 7]. For example, in a cluster analysis task, background noises may cause excessive molecule counts, and this overcounting might lead to bias in cluster analysis and wrong interpretations of the biological findings [8]. Therefore, the localization table in SMLM should be cleaned before any further quantification, and background noise removal in SMLM data has an important engineering significance.

To remove background noises and improve the quality of super-resolution image, researchers have made many attempts. Usually, they first optimize the hardware in the imaging system to obtain high quality raw images. For example, illumination *via* Total Internal Reflection Fluorescence (TIRF) is introduced to improve the signal to noise ratio (SNR) of the raw image [9]. And, the selection of image sensors with high sensitivity, such as electron-multiplying charge-coupled device (EMCCD) and scientific complementary metal-oxide semiconductor (sCMOS) cameras, can further improve image SNR [10]. Due to the complexity of the imaging process, a large amount of noises are still introduced into the localization table, and thus should be processed before reconstructing a final super-resolution image. Although many image analysis strategies have been established for conventional fluorescence microscopy images [11, 12], these strategies cannot apply directly to localization-based super-resolution images, because a conventional fluorescence image is composed of pixels or voxels, while an SMLM image is composed of a series of 2D or 3D localization coordinates. The data form of the localization table makes many trivial operations on conventional images (such as thresholding and subtraction) to become challenging. One usual solution to these challenges is to transform the localization table into a grayscale image (that is, a reconstructed super-resolution image), and then perform denoising analysis on the grayscale image. However, this will inevitably lead to a loss of the precise localization information, and affects the analysis results in downstream tasks (such as clustering, co-location analysis) [13]. According to above discussions, it would be greatly beneficial to perform denoising directly from a localization table.

Most of the current denoising methods for localization table are based on clustering algorithms. Andronov et al adopt Voronoi Tessellation for clustering of protein complexes, where the clustered localization points have smaller Voronoi cell areas than the non-clustered points. The Voronoi Tessellation method uses this feature to achieve the purpose of denoising [14]. For another example, a background noise filter is included in the ThunderSTORM plug-in [15], which adopts Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to filter out the localization point with poor localization. However, DBSCAN requires a careful parameter setting on radius and minimum number, which seriously affects the efficiency of denoising [6]. Machine learning-based clustering methods are also propose to denoise the localization data. For example, Williamson et al used neural networks to classify points from localization table as either clustered or non-clustered, based on a sequence of values derived from each point's nearest-neighbor distances, and the non-clustered points are defined as noise [16]. However, these denoising algorithms can only remove a certain kind of noises, which are not sufficient for the remove of complex heterogeneous background noises. In addition, most of these methods need to determine manually the threshold or parameters, which is difficult and unfriendly for biological researchers.

Inspired by the fusion denoising algorithms and techniques in the field of image and point cloud processing [17–20], here we propose a multi-step adaptive denoising framework for super-resolution localization data, called MSDenoiser. This framework makes full use of the different characteristics of reported denoising algorithms (including Voronoi Tessellation [21],

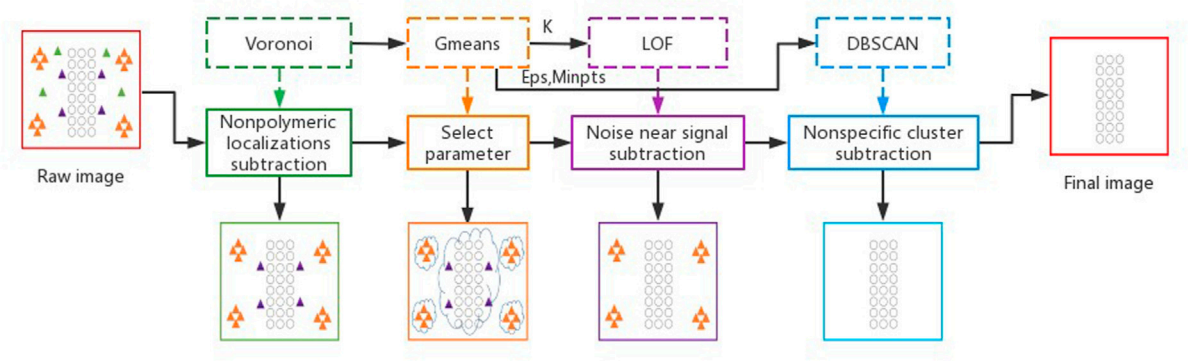


FIGURE 1
The multi-step adaptive noise reduction framework for SMLM image. The circles represent signal points, the triangles represent different type of noise points. Orange triangles represent nonspecific clusters, green triangles represent free non-polymer localization points, and purple triangles represent non-polymer localization points near the signal point.

Local Outlier Factor (LOF) [22] and DBSCAN), and gradually removes the free non-polymer localization points, non-polymer localization points near the sample signal point area, and non-specific localization points. To solve the problem of manual parameter determination in the LOF and DBSCAN algorithms, our MSDenoiser framework uses the G-means algorithm to automatically generate a set of clusters with centers, and adaptively estimates the parameters in LOF and DBSCAN algorithms without manual intervention. At the same time, to evaluate the effect of the denoising framework on experimental microtubule datasets without groundtruth, we propose an evaluation index based on skeleton information. We proved that the evaluation index is consistent with the existing denoising evaluation index. We verified the performance of the proposed MSDenoiser framework in simulated and experimental datasets (microtubule, nuclear pore complexes, and mitochondrial protein). From experimental results, we found that the MSDenoiser framework can effectively eliminate the mixed noises, achieve less detail loss and higher image SNR. Compared with the commonly used denoising algorithm in SMLM, we showed that the proposed framework has better performance and convenience in processing localization data from different types of biological structures.

2 Materials and methods

2.1 The multi-step adaptive noise reduction framework for single-molecule localization microscopy image

The proposed MSDenoiser framework includes four steps (as shown in Figure 1). Firstly, a Voronoi Tessellation-based method is used to remove free non-polymer localization points (The green box in Figure 1). Secondly, G-means algorithm is used to generate a group of clusters with centers. Features of the clusters are counted, which are passed to the LOF and DBSCAN algorithms as parameters (The orange box in Figure 1). Thirdly, LOF is used to remove non-polymer localization points near the sample signal area (The purple box in Figure 1). Finally, DBSCAN is used to eliminate non-specific localization clusters (The blue box in Figure 1).

2.1.1 Remove non-polymeric localizations using Voronoi Tessellation

Voronoi Tessellation has been applied in various fields from mathematics to natural sciences, and is usually used for clustering tasks in the field of super-resolution imaging [14, 23]. In Voronoi Tessellation, an image is divided into multiple polygonal regions centered on a set of points (seeds), with a single localization point at the center [24]. Voronoi cell represents the affected area of seed points, and the cell area provides an accurate measurement of local density of seed points. This property makes Voronoi Tessellations more suitable for describing the properties and neighborhoods of single molecules. Large Voronoi cells will be generated in low density area or randomly distributed points. Therefore, for a set of localization points with density of less than a given threshold, we can define it as noise.

2.1.2 Remove the noise near the structure point using local outlier factor

Local Outlier Factor (LOF) is an unsupervised outlier detection method [25]. LOF determines a point as an outlier by comparing the density of each point with its k neighborhood points, and considers the samples with densities much lower than those of its neighbors to be outliers. However, some boundary points may be excluded from the signal region, because their

density is lower than that of the signal region. These characteristics enable LOF to be a good algorithm to deal with edge effect of point clouds, because LOF can remove non-polymer localization points near the signal point.

2.1.3 Remove nonspecific clusters using density-based spatial clustering of applications with noise

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a typical density-based clustering method. DBSCAN divides the region with sufficient density into clusters, and can find clusters of arbitrary shapes in noisy spatial dataset. DBSCAN defines two parameters: neighborhood search radius (Eps), the minimum number of points within the search radius (MinPts) [6]. The algorithm starts from the point that has not been visited, and divides the data points into three types according to these two parameters: core points, boundary points, and noise points. DBSCAN has good noise recognition ability, and can filter the background noise. However, the setting of algorithm parameters is subjective, and it is difficult to determine appropriate parameters to ensure the quality of denoising.

2.1.4 Adaptive parameter selection using G-means algorithm

LOF and DBSCAN algorithms are used to identify high-density and low-density regions of point sets. The LOF algorithm eliminates as many non-clustered localization points as possible to minimize their interference to the DBSCAN algorithm in the next step. In this way, LOF effectively separates biologically relevant clusters from non-biologically relevant spurious clusters. But, the LOF algorithm requires an input parameter: k. The DBSCAN algorithm is able to distinguish noise points and signal points, from any shapes of clusters. However, two parameters are required for DBSCAN: Eps and MinPts.

To solve the problem of parameter selection in LOF and DBSCAN, we automatically generate a group of clusters with their centers by G-means algorithm, and count the features of the clusters to estimate the parameters of LOF and DBSCAN. Since the cluster centers generated by G-means can be affected by isolated noise and offset, which will further affect the parameters estimation, here we reprocess the data by Voronoi Tessellation to remove the isolated noise, so that the cluster center can better represent each cluster.

2.2 Simulated and experimental data

We simulated two representative kinds of biological structures (filament and ring) with different localization densities (1000 μm^{-2}~10,000 μm^{-2}) to evaluate the feasibility of the proposed denoising framework, which can cover most experimental scenes.

Step 1: Generation of groundtruth dataset with no localization error or background noise. We firstly obtained a ring structure image with a radius of 150 nm and a structure diameter of 60 nm using Python language. Secondly, we downloaded the microtubules data from an open dataset website (https://srm.epfl.ch/Datasets). We used the QC-STORM plug-in to locate and render the microtubule data, and obtained the filament structure image [26]. Finally, we merged the images of the two structures and obtained the groundtruth dataset by ThunderSTORM plug-in.

Step 2: Generation of noise dataset. The noise dataset includes background clusters and random noise. To simulate false localization events due to background fluorescence, we generated 30 background clusters with the same localization density as the foreground. To better evaluate the effectiveness of the proposed denoising framework, we generated random noise with different noise level (from 5% to 50%), and the noise level was defined as a percentage of the number of foreground localization points.

Step 3: Generating simulated localization dataset with filament and ring structures. We combined the datasets in Step 1 and 2 to obtain a localization dataset with filament and ring structures.

To verify the effectiveness of the framework in the context of complex biological structures, we performed analysis on experimental dataset of microtubules, nuclear pore complexes and mitochondrial protein. Microtubules (alpha-tubulin) were obtained by indirect immunolabeling with AlexaFluor647 (DOL 1–4) in COS7 cell staining. Nuclear pore complexes data were obtained by staining the gp210 protein of the Xenopus nuclear pore complex with Alexa647. Mitochondrial protein data were obtained by labeling mitochondrial protein TOM22 with a secondary antibody immunolabeling strategy in COS7 cells. These experimental datasets include tube-like and amorphous structures, and thus are excellent reference structures to verify the denoising effect. These data are downloaded from ShareLoc.XY (https://shareloc.xyz/#/).

2.3 Evaluation criteria

We evaluated the performance of algorithm using two commonly used metrics in the simulated dataset with groundtruth: Recall and F1-score. The calculation process is formulated as follows:

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall}$$

where Precision = TP/(TP + FP), and TP, FP, FN are true-positive (TP), false-positive (FP) and false-negative (FN), respectively. Recall represents the proportion of the total noise

that is correctly removed. The larger of the value of recall, the more noise is removed. The F1-score metric balances the results of Precision and Recall. The value of F1-score ranges from 0 to 1, with 1 representing the best result of the algorithm and 0 representing the worst output result.

Because the experimental dataset lacks groundtruth, we try to extract skeleton from microtubule data as groundtruth, and use pixel accuracy (PA) as the evaluation metric of microtubule experimental data. The generation process of skeleton groundtruth is as follows:

Step1: Calculate the microtubule width. We select a relatively straight microtubule structure as our ROI, and then process the ROI vertically. We calculate the width of the horizontal cross section on each pixel of the ROI, and then fit it using a Gaussian function. The full width at half maxima (FWHM) is used to represent the microtubule width.

Step2: Extract the skeleton. We use the method mentioned in [27] to extract the skeleton information of microtubule data.

Step3: Expand the skeleton and get the groundtruth. The skeleton is expanded based on the calculated microtubule width. The expansion process is to add pixel values to the edge of the skeleton to expand the overall pixel value, and thus achieve the skeleton expansion. In this way, we obtain the groundtruth of microtubule data.

Based on skeleton groundtruth information, pixel accuracy (PA) can be defined as follows:

$$PA = \frac{\sum_{i=0}^{n}\sum_{j=0}^{m} Y_{ij}}{\sum_{i=0}^{n}\sum_{j=0}^{m} X_{ij}}$$

where m, n denote the number of vertical and horizontal pixel of the image, respectively, X denotes the noisy Groundtruth image, and Y denotes the noisy image calculated by the algorithm. PA indicates the proportion of correctly labeled noise pixel to the total noise pixel. Therefore, PA can be used to evaluate the denoising efficiency of the algorithm.

3 Results and discussion

3.1 Validation based on skeleton evaluation criteria

We simulated filament and ring data to verify the effectiveness of the proposed evaluation metric (PA) based on skeleton information. We selected the localization data with a localization density of 4000 μm^{-2}, and added random noise with a noise level of 50% and 30 localization background clusters. We used skeleton extraction algorithm in Section 2.3 to extract the skeleton of the rendered image (see in Figure 2B). We calculated the average microtubule width of the original image to be 45 nm. To eliminate the error caused by the uneven distribution of localization density of super-resolution image, we rounded up the calculated microtubule width and took 50 nm as the basis for the skeleton expansion. The expansion results were served as the groundtruth image of the structure signal point in our experiments, as shown in Figure 2C. We used the difference operation between the signal point groundtruth image and the original image to obtain the groundtruth image of noise data.

We performed quantitative evaluation using three parameters (Recall, F1-score and PA) under different noise reduction ratio (from 10% to 90%), as shown in Figure 2D. We can see that the trends of all parameters (PA, Recall and F1-score) are generally consistent with each other. When the proportion of noise reduction is low, the correlation between the three parameters is strong and keeps rising rapidly. With the increase of the proportion of noise reduction, the rising of F1-score value is not as fast as those in Recall and PA. The reason is that, with the increase of proportion of noise reduction, some structure signal points may also be recognized as noise and thus removed, leading to the decrease of Precision and the affecting of the F1-score value. In addition, the evaluation metric based on skeleton has a disadvantage: it cannot evaluate the noise reduction in an area close to the structure signal points. This is also the reason for the difference in the value of Recall and PA, despite the strong correlation. From this analysis, we shown that the skeleton-based evaluation index PA is consistent with Recall and F1-score, and thus PA can quantitatively evaluate the denoising effect of experimental microtubule data.

3.2 Comparing the denoising performance using simulated data

To test the image denoising performance of MSDenoiser, we compared it with DBSCAN, Statistical Outlier Removal Filter (SORF) [28] and Radius Outlier Removal Filter (RORF) [29] using simulated data. DBSCAN is a commonly used denoising method in the field of super-resolution microscopy, while SORF and RORF are commonly used denoising methods in the field of point cloud. We adopted grid search strategy to select the optimal parameters of DBSCAN, RORF and SORF respectively in the following experiments, as shown in Figure 3. DBSCAN algorithm requires two parameters, Eps and MinPts. We set the input value of Eps to be 20–200 nm and the input value of MinPts to be 10–100, and used a total of 100 sets of parameters to find the best parameter combination. Similarly, the radius value of the RORF algorithm ranges from 20 nm to 200 nm, and the num_points value ranges from 10 to 100. The value of std_ratio for SORF ranges from 0.2 to 2, and the value of num_neighours ranges from 10 to 100. Note that RORF and SORF select the best combination from sets of parameters. Under different noise levels and localization densities, because the optimal parameters of the compared algorithms will be different, we select the optimal parameter combination for DBSCAN, RORF and SORF. For example, we show the heat map with optimal

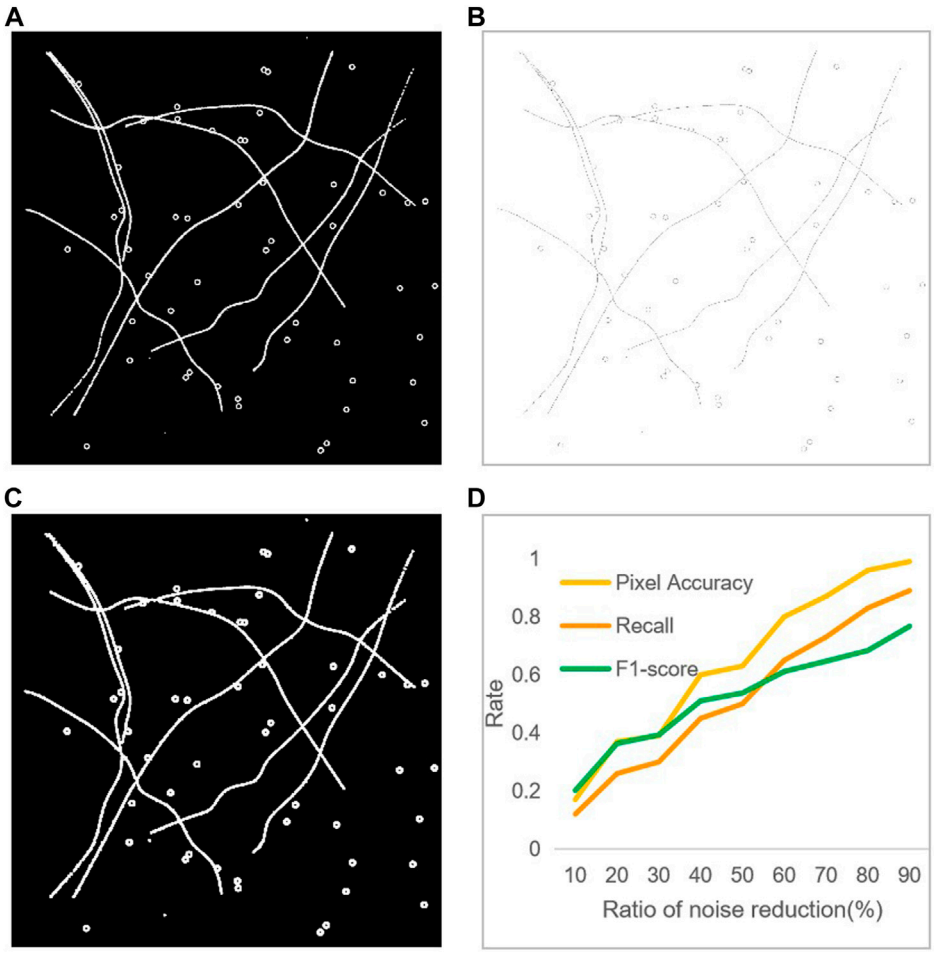


FIGURE 2
Comparison of the different evaluation metric. **(A)** Raw image. **(B)** The skeleton image. **(C)** The groundtruth image, which is the expanded skeleton image. **(D)** Three evaluation results with different noise reduction ratio.

parameter selection under different localization densities in Figure 3. In contrast, our proposed MSDenoiser framework can automatically compute parameters without manual intervention.

Firstly, we compared the denoising performance of MSDenoiser and the three reported algorithms (DBSCAN, RORF and SORF) under the same localization density and different noise levels. The localization density of simulated data is 4000 μm^{-2} and the noise level ranges from 5% to 50%. As shown in Figure 4, we find that all the four algorithms have small detail loss, but MSDenoiser achieves a balance of less detail loss and better SNR, and thus improves the quality of super-resolution images. However, although the results were similar at different noise levels, the value of Recall and F1-score changed. This is because the amount of non-specific clustering noise is unchanged. This finding also points out the bottleneck of the compared algorithms, that is, they cannot remove the non-specific clustering. At low noise level, the denoising performance of MSDenoiser is slightly worse than RORF, but it is still better than DBSCAN and SORF. The Recall value of MSDenoiser is not affected by the increase of noise, and reaches the maximum value of 0.86, when the noise level is 50%. The F1-score reaches the maximum value of 0.92. While the three reported algorithms benefit slightly from the careful selecting parameters, our proposed MSDenoiser can still achieve a good overall performance, without involving manual parameter search. Experiments with different noise levels demonstrate the robustness of our MSDenoiser method.

Then, we compared the denoising performanceX of MSDenoiser and the other three compared algorithms under the same noise level and different localization densities. The noise level of simulated data is 50%, and the localization densities ranges from 1000 μm^{-2} to 10,000 μm^{-2}. As shown in Figure 5, MSDenoiser performs slightly worse on

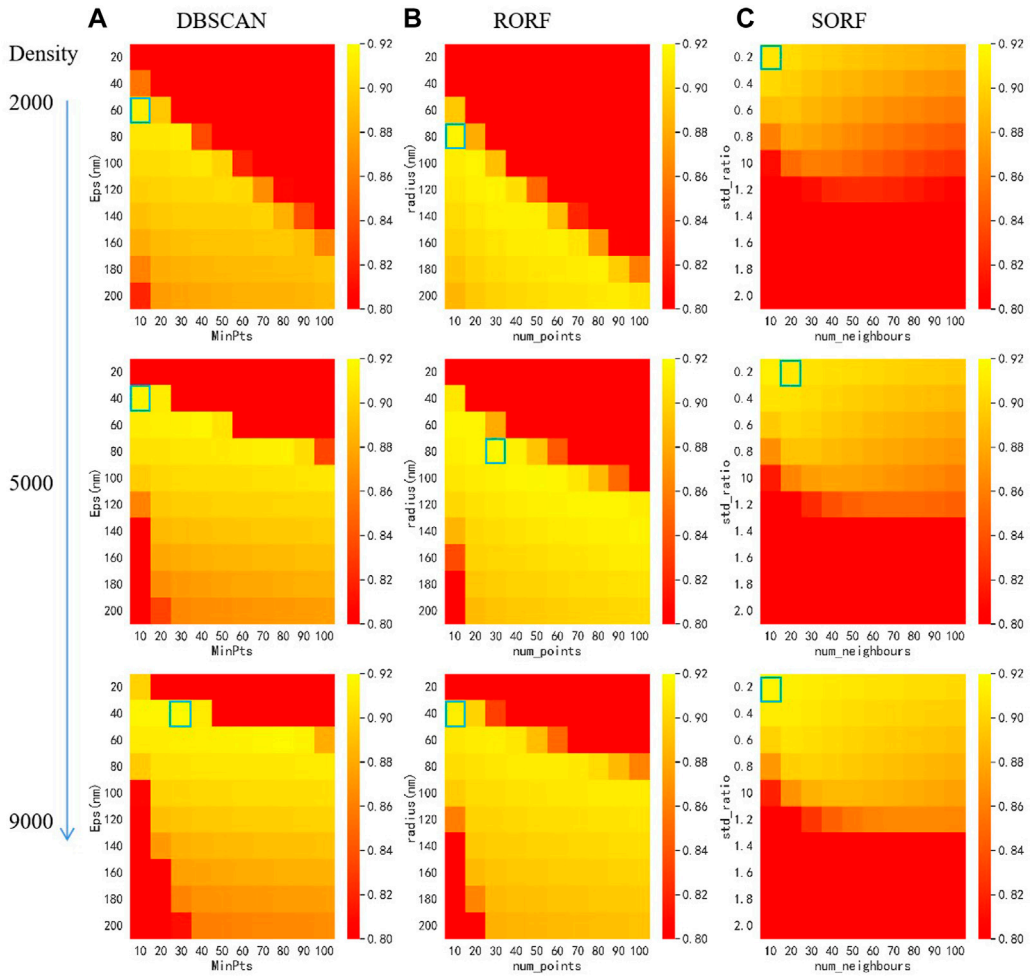


FIGURE 3
Heat map for parameter selection of different algorithms with different localization density using grid search. **(A)** Heat mat of DBSCAN. **(B)** Heat mat of RORF. **(C)** Heat map of SORF. The blue boxes indicate the optimal parameters.

simulated data with localization densities ranging from 1000 μm^{-2} to 3000 μm^{-2} (see in Figure 5G). When the localization density increases to larger than 4000 μm^{-2}, MSDenoiser achieves the best denoising performance. However, it is important to emphasize again, the MSDenoiser produces the best overall denoising results without the need of labor-intensive manual parameter setting.

Using the above discussions, we showed that MSDenoiser is capable to provide good denoising performance on simulated data with different localization densities and different noise levels. This new framework fully integrates the advantages of Voronoi Tessellation, LOF and DBSCAN, and selects parameters adaptively according to localization data, without reducing the effectiveness in removing mixed noise in SMLM data.

3.3 Comparing the denoising performance using experimental microtubule data

To verify the denoising ability of MSDenoiser in real dataset, we compared the performance among MSDenoiser, DBSCAN, RORF and SORF, using experimental microtubule data. Non-polymeric localizations and nonspecific clusters are distributed randomly within the field of view. We downloaded the experimental microtubule data from ShareLoc.XYZ, which includes 1231693 localization coordinates in the field-of-view of 300 μm^2, and the localization density is 4106 μm^{-2}. We showed in Figure 6E the raw image and its local enlarged details (rendered at 100 nm and 20 nm, respectively). We extracted the skeleton as the groundtruth of experimental microtubule data, and used them for quantitatively evaluating the denoising algorithms.

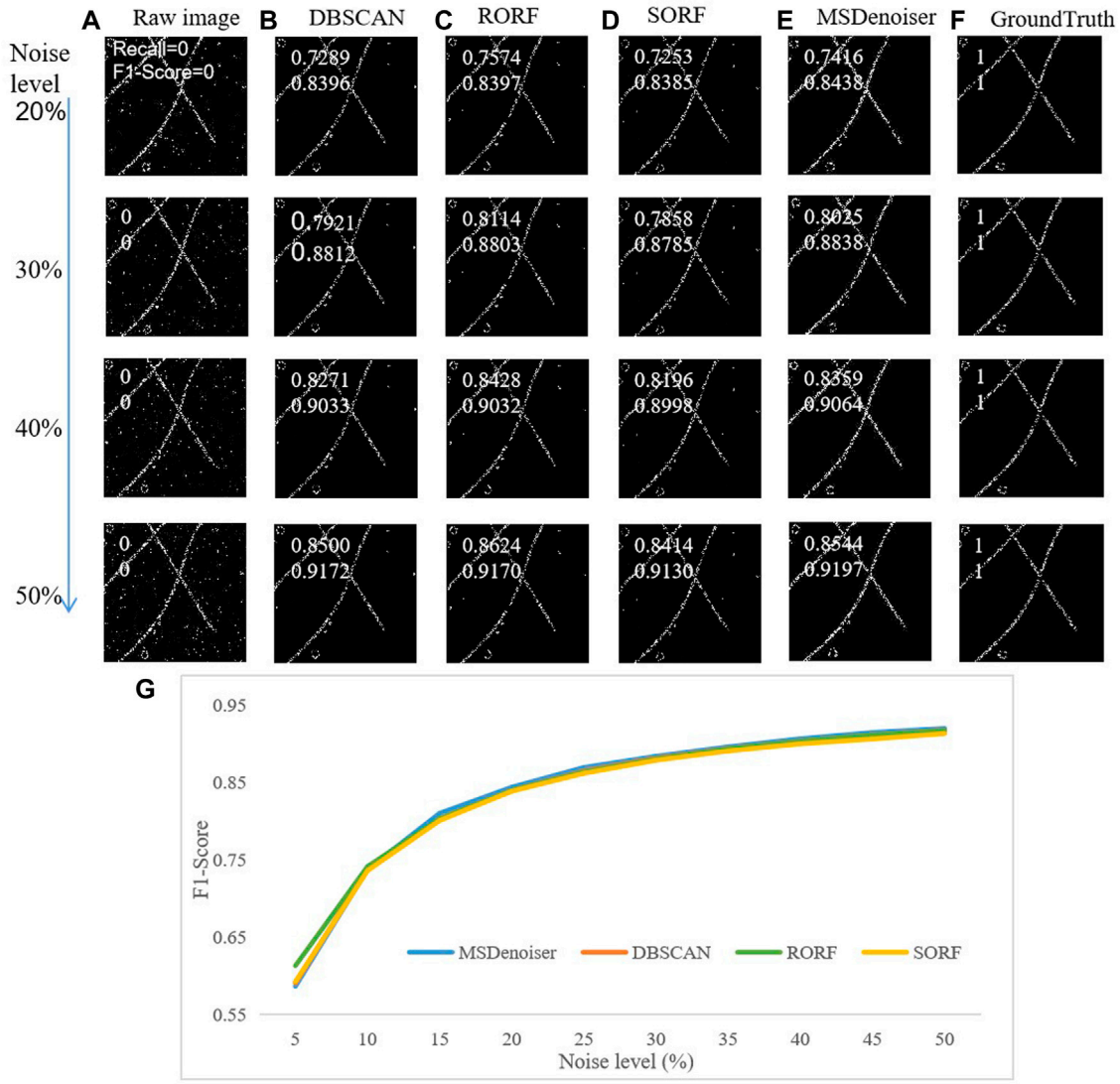


FIGURE 4
Comparison on the denoising performance in simulated microtubule data with different noise levels. **(A)** Raw rendered image. **(B)** Denoised image from DBSCAN. **(C)** Denoised image from RORF. **(D)** Denoised image from SORF. **(E)** Denoised image from MSDenoser. **(F)** Skeleton based groundtruth. **(G)** The dependence of F1-score on noise level.

Firstly, we discussed the parameter selection strategy for the four algorithms. The MSDenoiser can adaptively determine the parameters in the denoising framework. The parameters, Eps and MinPts, which are automatically calculated by MSDenoiser, can also be used by DBSCAN. In fact, for the experimental microtubule data, MSDenoiser sent the two parameters (Eps = 67 nm, MinPts = 215) to DBSCAN. The parameters in RORF and SORF were estimated by experience. The average localization density of experimental data was calculated to be 4106 μm^{-2}. The initial parameters were found according to the localization density in simulated data, and adjusted repeatedly according to the denoising effect. Therefore, the parameter are not required to be optimal at the beginning. The parameters in RORF are radius = 60 nm and num_points = 100, respectively. The parameters in SORF are std_ratio = 0.2 and num_neighbors = 80, respectively. Note that MSDenoiser framework does not require any manual intervention.

Then, we evaluated the denoising performance among these methods using PA value. We used experimental microtubule data as input data (as shown in Figure 6E, which were rendered from localization table data), and applied the four denoising methods to the localization table. We further analyzed the denoising performance from these methods. We found that MSDenoiser achieves the best denoising effect (that is, the highest PA value). Actually, as seen in Figures 6A–D, the PA value are 0.90 from DBSCAN, 0.92 from RORF, 0.90 from SORF, and 0.94 from MSDenoiser, respectively. From the local enlarged image, we

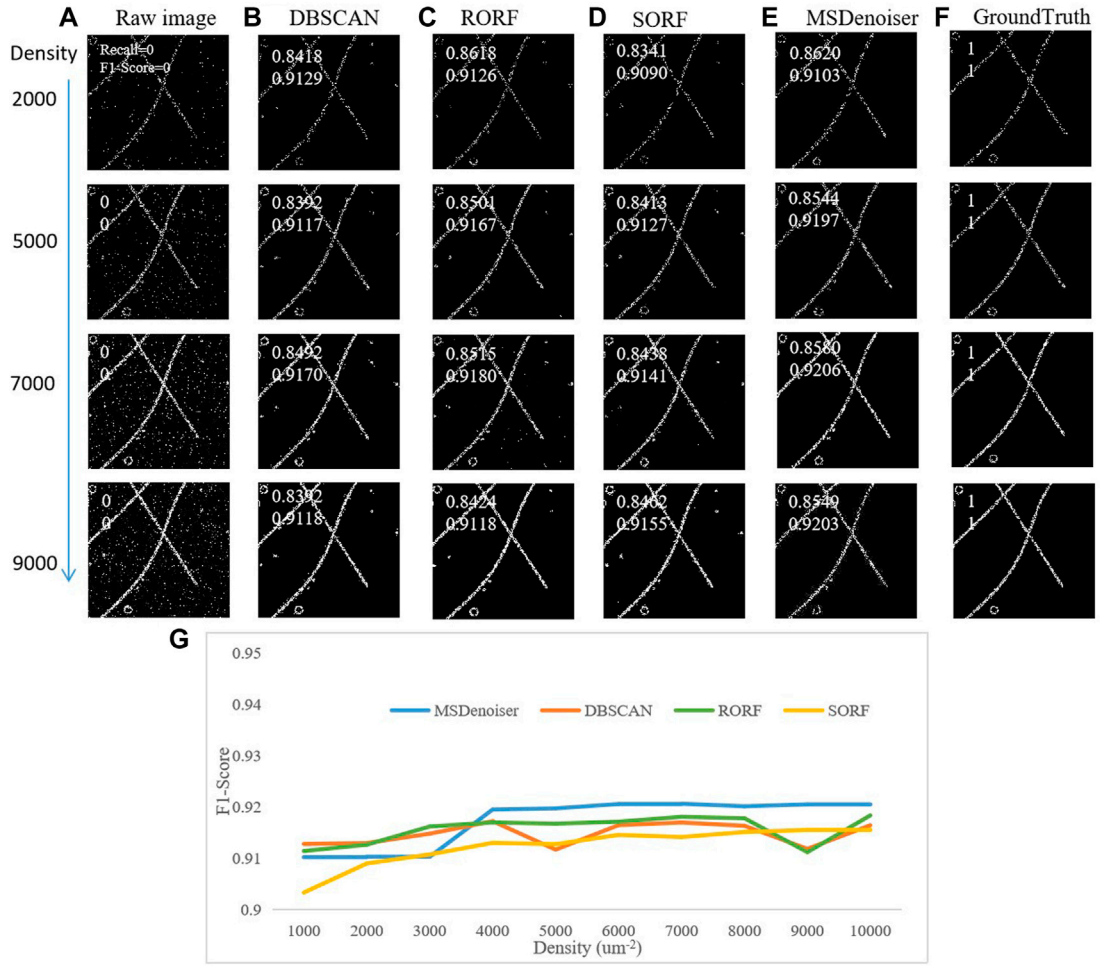


FIGURE 5
Comparison on the denoising performance in simulated patch microtubule data with different localization density by different methods. **(A)** Raw rendering image. **(B)** Denoised image from DBSCAN. **(C)** Denoised image from RORF. **(D)** Denoised image from SORF. **(E)** Denoised image from MSDenoiser. **(F)** Skeleton based groundtruth. **(G)** The dependence of F1-score on localization density.

observed that MSDenoiser removes nonspecific clusters, while DBSCAN does not remove this type of noise (see the white arrow in Figures 6A–D). Other methods removed these noises, along with the signal points of the microtubule structure, since we can see broken structures (see the yellow arrow in Figures 6A–D). MSDenoiser removed noise, but kept more signal points from microtubule structures.

3.4 Comparing the denoising performance using experimental amorphous structure data

In previous section, we demonstrated the applicability of MSDenoiser in relatively uniform and tube-like structures (microtubules). Here, we showed the denoising performance of MSDenoiser in datasets from nuclear pore complexes and mitochondrial protein (which contain amorphous features commonly seen in SMLM). We compared the denoising performance among MSDenoiser, DBSCAN, RORF, and SORF.

We downloaded experimental nuclear pore complexes and mitochondrial protein datasets from ShareLoc.XYZ. In the nuclear pore complexes dataset, a total number of 631214 fluorophores were identified from a field-of-view of 191 μm^2, and the localization density is 3305 μm^{-2}. In the mitochondrial protein dataset, a total number of 2270989 fluorophores were identified from a field-of-view of 2668 μm^2, and the localization density is 1002 μm^{-2}. Since there are no groundtruth for these experimental datasets, we only showed the results before and after denoising.

The parameter selection strategy in these two datasets is similar to that used in the experimental microtubule dataset.

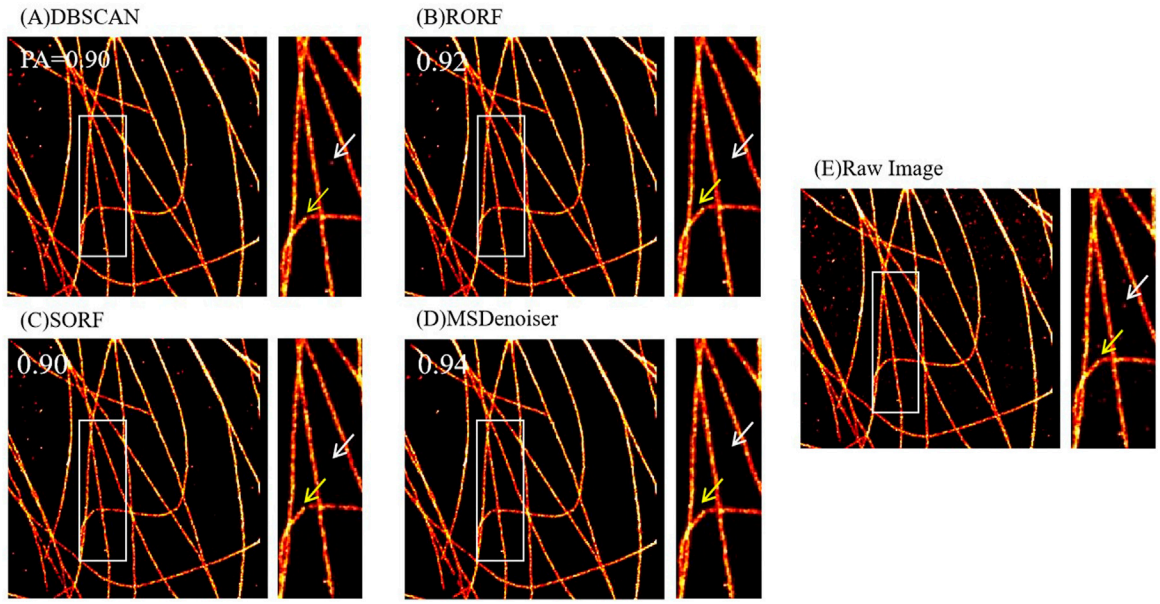


FIGURE 6
Comparison on the denoising performance of different methods using experimental microtubule data. **(A)** Denoised image from DBSCAN. **(B)** Denoised image from RORF. **(C)** Denoised image from SORF. **(D)** Denoised image from MSDenoiser. **(E)** Raw rendered image. The white arrows in **(A–C)** indicate noise, and the white arrow in **(D)** indicates the better denoising performance of MSDenoiser at the same location. The yellow arrows in **(A–C)** indicate broken microtubule structure, and the yellow arrow in **(D)** indicates the structure continuity from MSDenoiser at the same location.

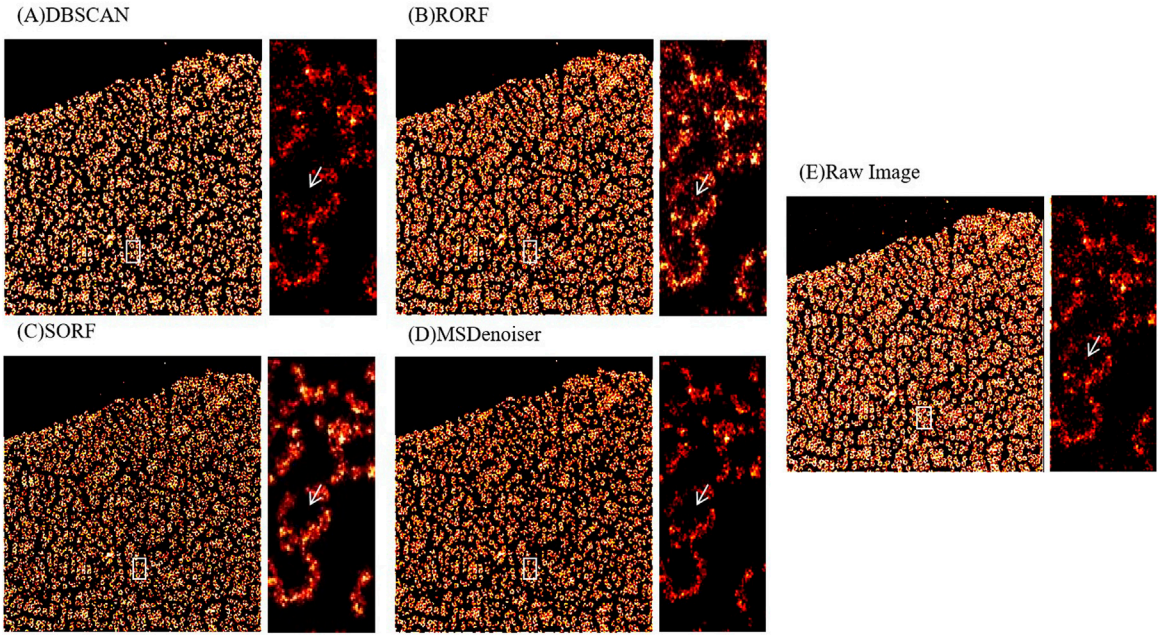


FIGURE 7
Comparison on the denoising performance of different methods using experimental nuclear pore complexes data. **(A)** Denoised image from DBSCAN. **(B)** Denoised image from RORF. **(C)** Denoised image from SORF. **(D)** Denoised image from MSDenoiser. **(E)** Raw rendered image. The white arrows in **(A–E)** indicate noise in ring structure, the white arrow in **(D)** indicates the better denoising effect of MSDenoiser at the same location.

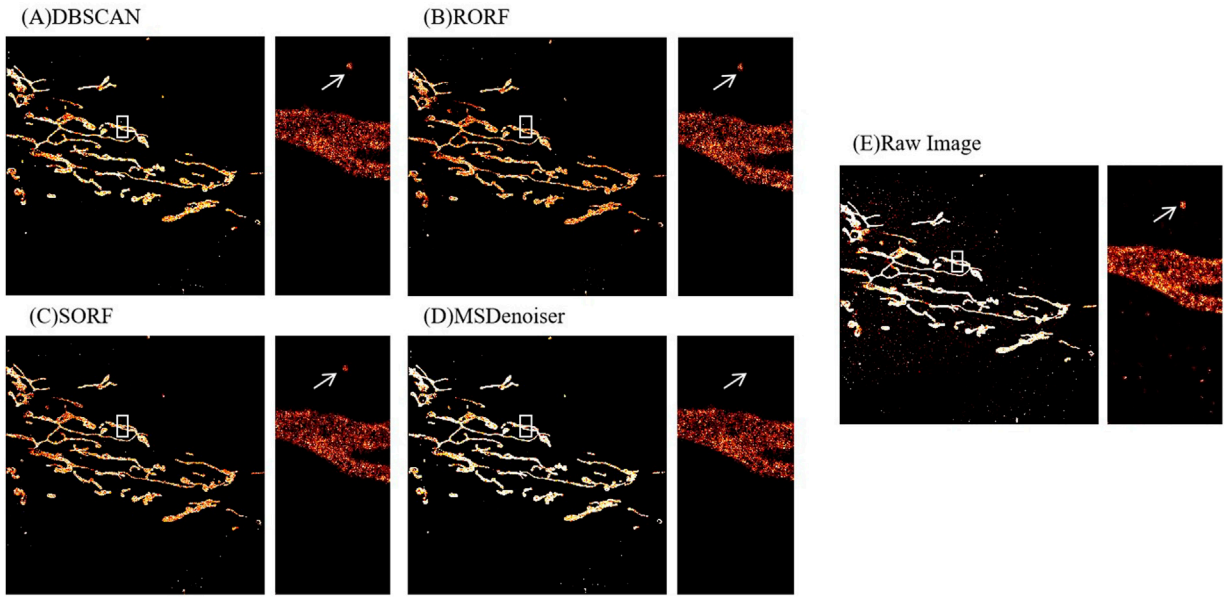


FIGURE 8
Comparison on the denoising performance of different methods using experimental mitochondrial protein data. **(A)** Denoised image from DBSCAN. **(B)** Denoised image from RORF. **(C)** Denoised image from SORF. **(D)** Denoised image from MSDenoiser. **(E)** Raw rendered image. The white arrows in **(A–E)** indicate nonspecific clustering, the white arrow in **(D)** indicates the better denoising effect of MSDenoiser at the same location.

The parameters in DBSCAN were determined from the parameters automatically calculated by MSDenoiser. For the experimental nuclear pore complexes dataset, the parameters in DBSCAN are Eps = 42 nm and MinPts = 52, respectively. For the experimental mitochondrial protein dataset, the parameters in DBSCAN are Eps = 45 nm and MinPts = 58, respectively. According to the average localization density of experimental data (3305 μm^{-2} for nuclear pore complexes, and 1002 μm^{-2} for the mitochondria), we empirically selected the parameters in the RORF and SORF algorithms. For the experimental nuclear pore complexes dataset, the parameters in RORF are radius = 70 nm and num_points = 100, respectively, and the parameters in SORF are std_ratio = 0.2 and num_neighbor = 40, respectively. For the experimental mitochondrial protein data, the parameters in RORF are radius = 30 nm and num_points = 90, respectively, and the parameters in SORF are std_ratio = 0.2 and num_neighbor = 50, respetively. The MSDenoiser framework can adaptively determine parameters without manual intervention.

We then analyzed the denoising performance of the four algorithms qualitatively using experimental nuclear pore complexes and mitochondrial protein datasets. The results are shown in Figures 7, 8. The reported three algorithms only removed part of the nonpolymer localization points, and the noise in the ring structure of the nuclear pore complexes was not effectively removed (see the white arrows in Figures 7A–E). Similarly, the reported algorithms did not remove the nonspecific clustering from the mitochondrial protein data (see the white arrows in Figures 8A–E). However, these problems can be effectively solved by our new MSDenoiser algorithm, which is able to retain maximum details of biological structures and remove most of the noises. The experimental results on nuclear pore complexes and mitochondrial protein datasets show that our proposed algorithm also have good ability in removal of the mixed noises that are common on experimental data.

From the above results, we show that our proposed MSDenoiser framework can be applied to localization data from different biological structures. Using this new framework, we can not only obtain better denoising results, but also adaptively determine parameters without manually intervention.

4 Conclusion

We developed a new denoising framework, called MSDenoiser, for improving the image quality in SMLM. In this framework, we combine the advantages of different reported denoising algorithms (Voronoi Tessellation, LOF and DBSCAN) to remove the noise of different features, and adaptively estimate the parameters required by the new framework using the G-means algorithm. In this framework, we aim to remove heterogeneous noises (free non-polymer

localization points, non-polymer localization points near the sample signal point area, and non-specific localization points). Using simulated and experimental datasets, we compared the denoising performance of MSDenoiser with DBSCAN (a commonly used denoising method in SMLM), RORF and SORF (commonly used denoising methods in point cloud). We found that MSDenoiser achieves better denoising effect than the three reported methods in the removal of mixed noises. Moreover, as compared with DBSCAN, RORF and SORF, the MSDenoiser can adaptively obtain parameters without manual intervention. We also verified the applicability of MSDenoiser in amorphous biological structures (nuclear pore complexes and mitochondrial protein), and proved that MSDenoiser has good robustness on different biological structures. Of course, there are still some limitations on our proposed MSDenoiser framework. The main limitation is the denoising speed, which needs to be improved for processing large SMLM localization table data. Because the MSDenoiser framework is composed of multiple algorithms, it takes a large amount of time on the automatic parameter selection process (G-means). When the field of view increases, the time spent in the parameter selection will increase significantly. In the future, we plan to develop a more efficient algorithm to take the place of G-means algorithm, so that we could shorten the processing time without reducing the denoising performance. Nevertheless, we believe this study points out a new strategy for solving the noise reduction challenge in SMLM.

Data availability statement

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

Author contributions

QF and ZW designed the study. QF generated simulated, analyzed the data, performed experiments, and drafted the manuscript. QS and MY generated simulated data, performed some experiments. ZH and ZW revised the manuscript. All the authors read and approved the final manuscript.

Funding

This work was partly supported by Key Research and Development Program of Hainan province (No. ZDYF2021GXJS017), National Natural Science Foundation of China (No. 82160345), the Major Science and technology plan of Hainan (ZDKJ2021016), Key science and technology plan project of Haikou (2021-016).

Conflict of interest

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

Publisher's note

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

References

1. Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS, et al. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* (2006) 313(5793):1642–5. doi:10.1126/science.1127344

2. Schermelleh L, Ferrand A, Huser T, Eggeling C, Sauer M, Biehlmaier O, et al. Super-resolution microscopy demystified. *Nat Cel Biol* (2019) 21(1):72–84. doi:10.1038/s41556-018-0251-8

3. Lelek M, Gyparaki MT, Beliu G, Schueder F, Griffié J, Manley S, et al. Single-molecule localization microscopy. *Nat Rev Methods Primers* (2021) 1(1):39. doi:10.1038/s43586-021-00038-x

4. Vicidomini G, Bianchini P, Diaspro A. STED super-resolved microscopy. *Nat Methods* (2018) 15(3):173–82. doi:10.1038/nmeth.4593

5. Wu YC, Shroff H. Faster, sharper, and deeper: Structured illumination microscopy for biological imaging. *Nat Methods* (2018) 15(12):1011–9. doi:10.1038/s41592-018-0211-z

6. Mazouchi A, Milstein JN. Fast optimized cluster algorithm for localizations (FOCAL): A spatial cluster analysis for super-resolved microscopy. *Bioinformatics* (2016) 32(5):747–54. doi:10.1093/bioinformatics/btv630

7. Whelan DR, Bell TDM. Image artifacts in single molecule localization microscopy: Why optimization of sample preparation protocols matters. *Sci Rep* (2015) 5:7924. doi:10.1038/srep07924

8. Khater IM, Nabi IR, Hamarneh G. A review of super-resolution single-molecule localization microscopy cluster analysis and quantification methods. *Patterns (New York, NY)* (2020) 1(3):100038. doi:10.1016/j.patter.2020.100038

9. Guo M, Chandris P, Giannini JP, Trexler AJ, Fischer R, Chen J, et al. Single-shot super-resolution total internal reflection fluorescence microscopy. *Nat Methods* (2018) 15(6):425–8. doi:10.1038/s41592-018-0004-4

10. Herbert S, Soares H, Zimmer C, Henriques R. Single-molecule localization super-resolution microscopy: Deeper and faster. *Microsc Microanal* (2012) 18(6):1419–29. doi:10.1017/s1431927612013347

11. Chen J, Sasaki H, Lai H, Su Y, Liu J, Wu Y, et al. Three-dimensional residual channel attention networks denoise and sharpen fluorescence microscopy image volumes. *Nat Methods* (2021) 18(6):678–87. doi:10.1038/s41592-021-01155-x

12. Chan KG, Streichan SJ, Trinh LA, Liebling M. Simultaneous temporal superresolution and denoising for cardiac fluorescence microscopy. *IEEE Trans Comput Imaging* (2016) 2(3):348–58. doi:10.1109/tci.2016.2579606

13. Schnitzbauer J, Wang Y, Zhao S, Bakalar M, Nuwal T, Chen B, et al. Correlation analysis framework for localization-based superresolution microscopy. *Proc Natl Acad Sci U S A* (2018) 115(13):3219–24. doi:10.1073/pnas.1711314115

14. Andronov L, Orlov I, Lutz Y, Vonesch J-L, Klaholz BP. ClusterViSu, a method for clustering of protein complexes by Voronoi tessellation in super-resolution microscopy. *Sci Rep* (2016) 6:24084. doi:10.1038/srep24084

15. Ovesny M, Krizek P, Borkovec J, Svindrych ZK, Hagen GM. ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics* (2014) 30(16):2389–90. doi:10.1093/bioinformatics/btu202

16. Williamson DJ, Burn GL, Simoncelli S, Griffie J, Peters R, Davis DM, et al. Machine learning for cluster analysis of localization microscopy data. *Nat Commun* (2020) 11(1):1493. doi:10.1038/s41467-020-15293-x

17. Wang J, Zhang H, Gao J, Xiao D. Dust removal from 3D point cloud data in mine plane areas based on orthogonal total least squares fitting and GA-TELM. *Comput Intelligence Neurosci* (2021) 2021:1–8. doi:10.1155/2021/9927982

18. Regaya Y, Fadli F, Amira A. Point-Denoise: Unsupervised outlier detection for 3D point clouds enhancement. *Multimed Tools Appl* (2021) 80(18):28161–77. doi:10.1007/s11042-021-10924-x

19. Yang X, Xu Y, Quan Y, Ji H. Image denoising via sequential ensemble learning. *IEEE Trans Image Process* (2020) 29:5038–49. doi:10.1109/tip.2020.2978645

20. Ma X, Ugurbil K, Wu X. Denoise magnitude diffusion magnetic resonance images via variance -stabilizing transformation and optimal singular -value manipulation. *Neuroimage* (2020) 215:116852. doi:10.1016/j.neuroimage.2020.116852

21. Shivanasab P, Abbaspour RA. An incremental algorithm for simultaneous construction of 2D Voronoi diagram and Delaunay triangulation based on a face-based data structure. *Adv Eng Softw* (2022) 169:103129. doi:10.1016/j.advengsoft.2022.103129

22. Duan L, Xu L, Liu Y, Lee J. Cluster-based outlier detection. *Ann Oper Res* (2009) 168(1):151–68. doi:10.1007/s10479-008-0371-9

23. Levet F, Hosy E, Kechkar A, Butler C, Beghin A, Choquet D, et al. SR-tesseler: A method to segment and quantify localization-based super-resolution microscopy data. *Nat Methods* (2015) 12(11):1065–71. doi:10.1038/nmeth.3579

24. Aurenhammer F. Voronoi diagrams—A survey of a fundamental geometric data structure. *ACM Comput Surv* (1991) 23(3):345–405. doi:10.1145/116873.116880

25. Breunig MM, Kriegel H-P, Ng RT, Sander JLOF. Identifying density-based local outliers. In: *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*. Dallas, Texas, USA: Association for Computing Machinery (2000). p. 93

26. Sage D, Kirshner H, Pengo T, Stuurman N, Min J, Manley S, et al. Quantitative evaluation of software packages for single-molecule localization microscopy. *Nat Methods* (2015) 12(8):717–24. doi:10.1038/nmeth.3442

27. Li M, Song Q, Xiao Y, Wu J, Kuang W, Zhang Y, et al. LuckyProfiler: An ImageJ plug-in capable of quantifying FWHM resolution easily and effectively for super-resolution images. *Biomed Opt Express* (2022) 13(8):4310–25. doi:10.1364/boe.462197

28. Minh-Hai L, Cheng C-H, Liu D-G, Thanh-Tuan N. An adaptive group of density outlier removal filter: Snow particle removal from LiDAR data. *Electronics* (2022) 11(19):2993. doi:10.3390/electronics11192993

29. Duan Y, Yang C, Li H. Low-complexity adaptive radius outlier removal filter based on PCA for lidar point cloud denoising. *Appl Opt* (2021) 60(20):E1–E7. doi:10.1364/ao.416341

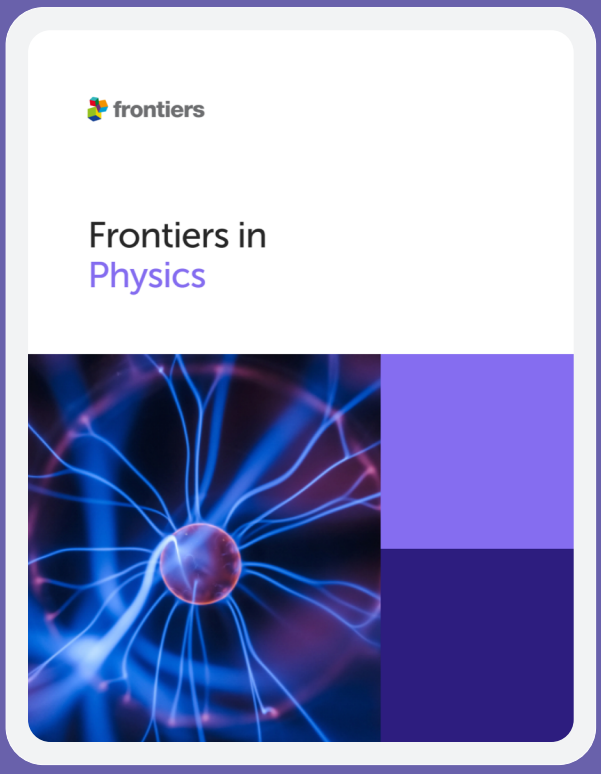
Frontiers in Physics

Investigates complex questions in physics to understand the nature of the physical world

Addresses the biggest questions in physics, from macro to micro, and from theoretical to experimental and applied physics.

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