

<https://doi.org/10.1038/s41746-025-01446-z>

Fusing multispectral information for retinal layer segmentation

Xiang He^{1,2,4}, Fuwang Wu^{1,4}, Kaixuan Hu¹, Lizhen Cui^{2,3}, Weiye Song¹ ✉ & Yi Wan¹ ✉

Extensive research on retinal layer segmentation (RLS) using deep learning (DL) is mostly approaching a performance plateau, primarily due to reliance on structural information alone. To address the present situation, we conduct the first study on the impact of multi-spectral information (MSI) on RLS. Our experimental results show that incorporating MSI significantly improves segmentation accuracy for retinal layer optical coherence tomography (OCT) images. Furthermore, we investigate the primary factors influencing MSI, including the number of multi-spectral images, spectral bandwidth, and the different spectral combinations, to assess their impacts on the accuracy of RLS. Building upon this foundation, we have incorporated MSI into RLS methods, yielding exceptional performance in segmentation outcomes, and these findings have been validated in OCT images across both the near-infrared and visible-light spectral ranges. Fusing MSI provides a novel approach to improving RLS accuracy, further demonstrating the importance of open-source MSI information in OCT devices.

The eye is often referred to as the window to the soul, with over 80% of external information perceived by humans through vision, making ocular health critically important to human life. With the continuous advancement of social science and technology, excessive eye use due to various factors has led to a rising prevalence of eye diseases¹. This situation is further exacerbated by the intensification of population aging. Among these eye diseases, a large number of patients exhibit changes in retinal structure at an early stage, such as changes in retinal layer thickness^{2–4}. Accurate detection of retinal layer changes in the early stages of high-blindness-risk fundus diseases could potentially prevent or even reverse the progression of these diseases.

OCT is a non-invasive imaging technique used for the cross-sectional visualization of retinal layers. Since its invention in 1990, it has become the gold standard for diagnosing various ophthalmic diseases⁵. Changes in retinal layer thickness often accompany many early-stage ophthalmic diseases, and OCT provides a critical means for accurately quantifying these changes⁶. Domestically and internationally researchers have conducted extensive studies on RLS based on OCT images. In recent years, DL-based RLS research has gained increasing attention^{7–9}. Compared to machine learning, DL can adaptively extract features across different dimensions, significantly reducing development difficulty^{10,11}. Since the proposal of the ReLayNet¹² model in 2017, DL-based RLS research has primarily focused on optimizing model structures and improving training strategies. However, in recent years, the overall improvement in RLS accuracy has been diminishing, and the performance of DL methods based on model structure optimization and training strategy improvement has reached its limits. We

analyze that the research based on OCT images relies solely on structural information, while the spectral information is lost, which brings disadvantages at the data source. The principle of measuring blood oxygen saturation reveals that different wavelengths of light have varying sensitivities to blood oxygen saturation^{13,14}, which is somewhat similar to the principle of multi-spectral fundus cameras¹⁵. The imaging principle of OCT in retinal layers utilizes the differences in cellular structures and elemental compositions of different retinal layers, resulting in varying backscattering rates of the same wavelength of light when it strikes different layers. Consequently, when different wavelengths of light are directed at the same retinal tissue, they also yield diverse backscattering rates. Based on this foundation, we posit that by employing OCT devices with multiple wavelength range sources for scanning retinal layers, the acquired retinal images will provide structural information and capture additional MSI. The effective utilization of this spectral information could potentially transcend the limitations imposed by relying solely on retinal structural information for segmentation, propelling DL-based research on retinal layer segmentation to a new level. However, current OCT equipment does not provide MSI, primarily for two reasons: from the perspective of OCT image-based researchers, they are unaware that the raw data collected by OCT equipment also contains rich spectral information, so they can only conduct research based on structural information in the images; From the perspective of OCT equipment manufacturers, they are not aware of the importance of multi-spectral information for medical imaging research, therefore, OCT equipment does not have open-source multi-spectral information.

¹School of Mechanical Engineering, Shandong University, Jinan, China. ²Joint SDU-NTU Centre for Artificial Intelligence Research (C-FAIR), Shandong University, Jinan, China. ³School of Software, Shandong University, Jinan, China. ⁴These authors contributed equally: Xiang He, Fuwang Wu.

✉ e-mail: songweiye@sdu.edu.cn; wanyi@sdu.edu.cn

To address the aforementioned issues, in this context, we aim to investigate whether MSI can enhance RLS. Assuming that MSI has been proven to improve the accuracy of RLS, we will further explore the impact of primary factors influencing MSI on RLS accuracy. Ultimately, the MSI will be applied to general RLS approaches, making it a simpler and more universal technique for enhancing RLS accuracy. This initiative not only calls for the OCT equipment manufacturing industry to open-source MSI but also serves as a reminder for researchers who base their work on OCT images to utilize more raw materials for processing.

In this work, our contributions are summary as follows:

1. We propose a framework to explain how to extract MSI information and how to apply it to general RLS methods, As demonstrated in Fig. 1.
2. We discover the enhancing effect of MSI on RLS accuracy. Specifically, first, compared to traditional single spectral images, using MSI images can improve the accuracy of RLS; second, we discover three main factors that may have an impact on MSI, including the spectral bandwidth, the number of multi-spectral images, and the multi-spectral combination. And further explore the mechanisms of these three factors in different spectral ranges.
3. We find that applying MSI to general RLS methods has similar improved accuracy performance in different spectral ranges, demonstrating the universality of MSI.
4. Our research indicates that OCT manufacturers have neglected a crucial factor in imaging: MSI. This will prompt the entire OCT manufacturing industry to reconsider and make corresponding product enhancements without altering existing equipment.

Results

Does the MSI can improve the accuracy of RLS ?

To ascertain whether MSI can improve the accuracy of RLS, we initially test four commonly used RLS methods on a multi-spectral OCT image dataset with equal-bandwidth spectrum and different central wavelengths of NO.1, NO.2, and NO.3, and conduct quantitative analysis. In Fig. 2a–d, under the combination of multiple-spectral, the accuracy of multiple-spectral is higher than that of single-spectral across all three segmentation metrics. In Fig. 2a, within multiple-spectral combinations, the segmentation accuracy of the Enet¹⁶ method is higher than that of the single-spectral approach. In the (No.1&2) VS (No.1) group, adding No.2 to the No.1 spectral band increased the mIoU index by 0.62 percentage points. Adding images in the No.1 and No.2 spectral bands to the No.3 spectral band also improved the segmentation accuracy of Enet to varying degrees, with the mIoU index increasing by 0.43 and 0.47 percentage points, respectively. In Fig. 2b, c, the groups with the highest accuracy improvement is (No.1&2) VS (No.1), which has improved the mIoU index by 0.91 and 1.40 percentage points. In Fig. 2d, after adding images in the 2nd and 3rd spectral band to the No.1, the segmentation accuracy of AttU_net¹⁷ increased by 0.74 and 0.77 percentage points respectively on the mIoU index, with similar trends of improvement observed in the Dice and mPA indices as well.

In addition to quantitative indicators, we also conduct qualitative analysis. In Fig. 3a–d, we can see that there is a certain degree of intra-class error in the prediction maps of each method on most single-spectral images, such as the red dashed area in the prediction map of ReLayNet¹² on No.1 in Fig. 3b, the red dashed area in the prediction map of AttU_net on No.2 and No.3 in Fig. 3c, and the red dashed area in the prediction map of TransUnet¹⁸ on No.1 in Fig. 3d. In addition, the addition of MSI has also reduced some issues of inaccurate inter-class boundaries, such as the dashed area in the predicted maps of Enet on No.1 and No.3 in Fig. 3a, the red dashed area in the predicted maps of ReLayNet on No.3 in Fig. 3b, and the red dashed area in the predicted maps of TransUnet on No.3 in Fig. 3d. In addition, we also use error distribution plots of prediction and annotation to further analyze the improvement of multi-spectral over single-spectrum. In Fig. 3e–h, the deeper the degree, the larger the error area between prediction and annotation, and the pure color background area represents complete consistency between prediction and annotation. We can clearly see from the

local zoom-in regions in Fig. 3e–h that the area of error in the multi-spectral group is generally significantly reduced.

Our study demonstrates that within spectral images of equal bandwidth but different central wavelengths, any combination of images with distinct central wavelengths enhances the segmentation accuracy compared to any single spectral image within that combination. This aligns with our previous hypothesis that light of different wavelengths incident on the same retinal layer tissue also produces varying backscattering rates, which can be further utilized to obtain MSI. This MSI effectively improves the traditional RLS accuracy based on single-spectral images.

Exploring the primary factors affecting MSI in visible-light multi-spectral (VMS) dataset

After verifying the effect of MSI on RLS, we conduct further investigations into the factors that may affect MSI, including the number of spectrum, spectral bandwidth (range of the spectrum), and spectral combinations. Figure 4a, the changes in RLS accuracy are demonstrated when sequentially adding spectral images of different central wavelengths to the raw image, by the Deformer¹⁹ method. When the No.7 spectral image is added, the segmentation accuracy improved from 85.43 to 85.82, with an increase of 0.39 on the mIoU index. On this basis, further addition of spectrum No.5 and No.6 improved the segmentation accuracy by 0.55 compared to single-spectrum. Continuing to add No.3 and No.4 spectral images sequentially, and then No.1 and No.2 spectral images, the segmentation accuracy increased by 0.54 and 0.55 respectively compared to single-spectrum.

Figure 4b shows the impact of MSI on RLS after adding spectral images with different bandwidths. Among them, there are a total of 7 spectrum with different center wavelengths in the image that is 0.05 times the raw spectrum. Three segmented spectrum with different center wavelengths can be selected and added to the raw spectrum, as shown in Fig. 4c. There are a total of 35 combinations, and taking the average of them can obtain an MSI with a spectral bandwidth of 0.05 times, which can improve the segmentation accuracy of RLS to 85.67. Similarly, adding MSI with a spectral bandwidth of 0.1 times, 0.15 times, 0.2 times, and 0.25 times can respectively improve the segmentation accuracy of RLS to 85.87, 85.98, 85.87, and 85.82. Among them, MSI with a spectral bandwidth of 0.15 times has the greatest improvement effect in this experiment.

Based on the premise that the MSI with a 0.15 times bandwidth provides the most significant enhancement in RLS segmentation accuracy, Fig. 4c illustrates all combinations of the MSI with a 0.15 times bandwidth image. The average improvement in accuracy across all combinations is 0.55. Specifically, when the MSI of combinations No.1&4&7, No.1&3&6, and No.3&4&5 are added, the accuracy improves by 0.72, 0.71, and 0.70 respectively. The combination with the least enhancement is No.1&5&7, which results in an accuracy improvement of 0.11.

Exploring the primary factors affecting MSI in near-infrared multi-spectral (NIMS) dataset

To verify the universality of the above experimental results, we conduct the same experiment on spectral images in the near-infrared band range. Figure 5a illustrates the changes in RLS accuracy when sequentially adding spectral images with different central wavelengths to the raw image using the Deformer method. When the No.7 spectral image is added, the segmentation accuracy improved from 84.18 to 84.77, representing an increase of 0.59 in the mIoU metric. Building on this, further addition of the No.5 and No.6 spectral images enhanced the segmentation accuracy by 0.89 compared to a single spectrum. Subsequently, continuing to add the No.3 and No.4 spectral images, followed by the No.1 and No.2 spectral images, the segmentation accuracy improved by 1.02 and 0.94 respectively, relative to a single-spectrum. Figure 5b shows the effect of MSI on RLS after adding spectral images with different bandwidths. In the 0.025 times spectral bandwidth image, the average segmentation accuracy of all spectral combinations shows that MSI can improve the segmentation accuracy of RLS to 84.90. Similarly, adding MSI with a spectral bandwidth of 0.05 times, 0.1 times, and 0.2 times can respectively improve the segmentation accuracy of

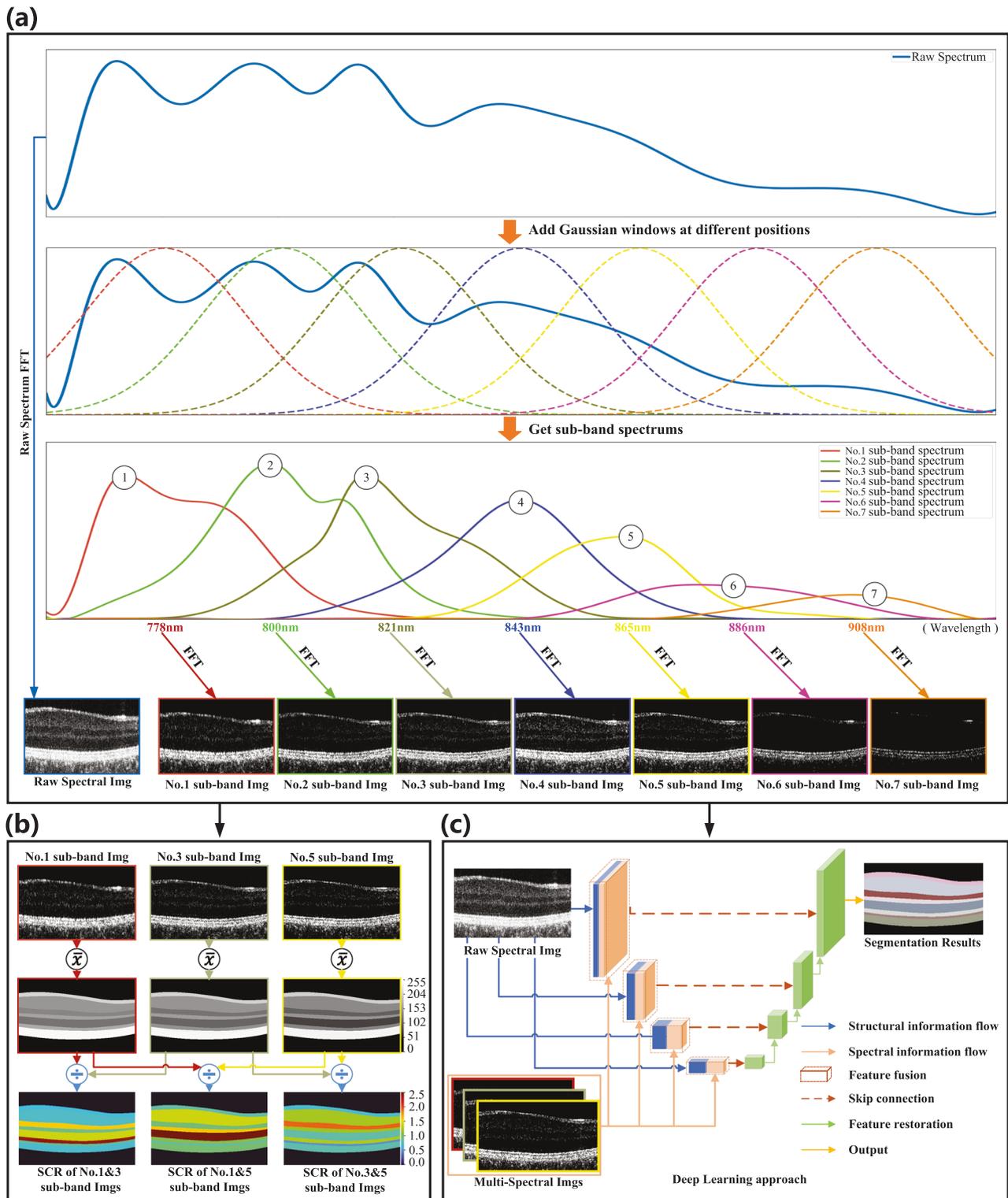


Fig. 1 | The principle demonstration and practical application of MSI. a shows how to obtain multi-spectral images from the raw spectrum, b shows the acquisition of MSI in three different sub-band spectral images, c shows the application of MSI to the general RLS method framework.

RLS to 84.95, 84.87, and 84.68. Among them, MSI with a spectral bandwidth of 0.05 times has the greatest improvement effect in this experiment. Figure 5c shows all combinations of MSI with a spectral bandwidth of 0.05 times, with an average improvement accuracy of 0.77 for all combinations. When adding MSI with No.3&5&7, No.1&2&4, and No.2&4&7, the accuracy is improved by 1.14, 1, and 1, respectively. The combination with the least improvement is No.2&3&6, with an improvement of 0.39 in accuracy.

Application of MSI to general RLS methods

After exploring the primary factors that MSI affects the accuracy improvement of RLS, we attempt to apply it to general RLS methods. To prove its universality, we still conduct experiments on both NIVS dataset and VMS dataset simultaneously. We select the most classic ReLayNet¹² method in the RLS field, as well as the latest EMV-Net²⁰ and LightReSeg²¹ methods for the experiment, and we also add Segformer²² and AttUnet¹⁷,

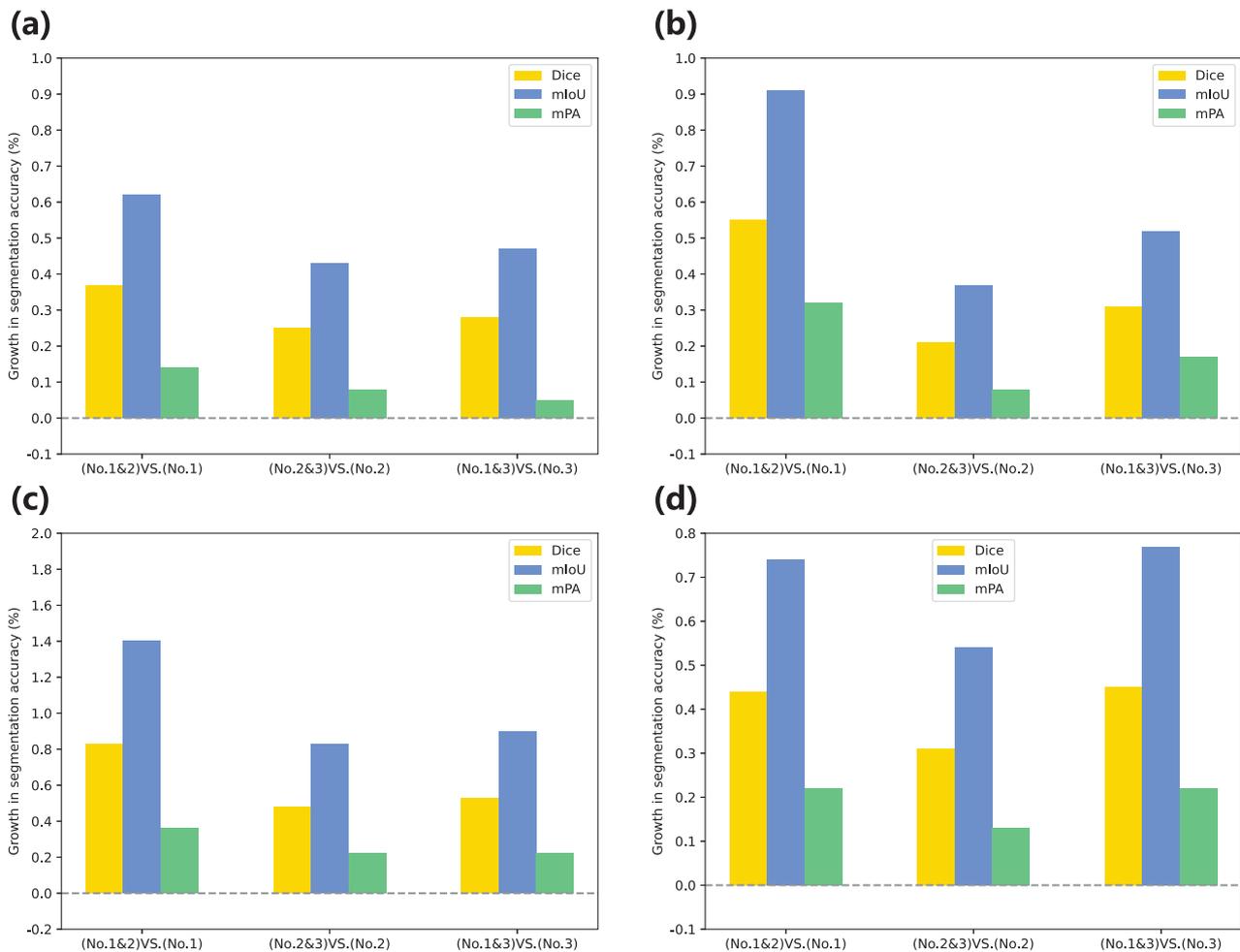


Fig. 2 | Quantitative description of the improvement in RLS accuracy of multi-spectral images compared to single spectral images. a–d show the accuracy comparison of four methods, Enet, ReLayNet, TransUNet, and AttU_net, under different spectral combinations. No.1, No.2, and No.3 in the horizontal axis represent the retinal layer images of the No.1 wavelength range (center wavelength

542 nm), the No.2 wavelength range (center wavelength 584 nm), and the No.3 wavelength range (center wavelength 627 nm), respectively, with a spectral bandwidth of 63 nm. No.1 & 2, No.2 & 3, and No.1 & 3 represent the composite of the first and second spectral images, the composite of the second and third spectral images, and the composite of the first and third spectral images, respectively.

two segmentation methods with excellent performance. As shown in Table 1, we add a combination of spectral images with a bandwidth of 0.05 times No.3&5&7 to the NIMS dataset and apply them to the selected three methods using spectral extraction modules. Among them, mIoU improved by 0.54, 1.00, 0.91, 0.60, and 1.15 on ReLayNet, EMV-Net, Segformer, LightReSeg and AttUnet, respectively. Similarly, a combination of No.1&4&7 with a bandwidth of 0.15 times spectral images, are added to the VMS dataset and applied to the selected three methods by a spectral extraction module. The mIoU metric is enhanced by 0.45, 0.61, 0.68, 0.73, and 0.71 on ReLayNet, EMV-Net, Segformer, LightReSeg and AttUnet, respectively. In addition, the addition of MSI has also improved the segmentation accuracy of each layer of the retina.

In addition, we further perform the statistical significance test by using the Wilcoxon rank sum test. For example, when comparing the differences on the NIMS dataset, we observe a P value of 0.0176 ($p < 0.05$) on the ReLayNet, indicating a statistically significant difference after adding the application of MSI. Similar statistically significant differences are observed, with P values of 0.01562 ($p < 0.05$), 0.00781 ($p < 0.05$), 0.01562 ($p < 0.05$), and 0.03906 ($p < 0.05$) respectively when comparing on the EMV-Net, AttUnet, SegFormer-B0 and LightReSeg. In the context of the VMS dataset, the P values obtained for each method are as follows: ReLayNet exhibits a P value of 0.01775 ($p < 0.05$), while all other methods demonstrate a uniformly lower P value of 0.00781 ($p < 0.05$). These results statistically signify that

following the application of MSI, every method indicates a significant enhancement in performance.

We observe that MSI has an enhancing effect on both the classic ReLayNet method and the latest LightReSeg method. Moreover, this conclusion is validated in datasets across both visible light and near-infrared spectral ranges, demonstrating the universality of the MSI approach. Table 1 shows that the ReLayNet method based on the raw spectrum generally under-performed LightReSeg in terms of the mIoU metric. Although accuracy improved after incorporating MSI, it did not reach the precision of LightReSeg using just a single spectrum. This indicates that MSI is a versatile and universally applicable method for improving accuracy, but it cannot compensate for the structural deficiencies in general RLS methods. Furthermore, the Table 1 reveals that relying solely on MSI falls short of surpassing segmentation outcomes based on Raw images (structural information) in this experiment. This underscores the potential for further enhancement in extracting and utilizing spectral information, which currently serves merely as a supplement to structural data.

We further conduct qualitative analysis on the NIMS dataset, as shown in Fig. 6. In Fig. 6a–c, we can clearly see that on most single-spectral images, the fusion of MSI significantly reduces intra-class and inter-class errors. The prediction graphs of ReLayNet before and after adding MSI in the first row of Fig. 6a, and the prediction graphs of EMV-Net before and after adding MSI in the first and second rows of Fig. 6b all significantly reduce intra-class

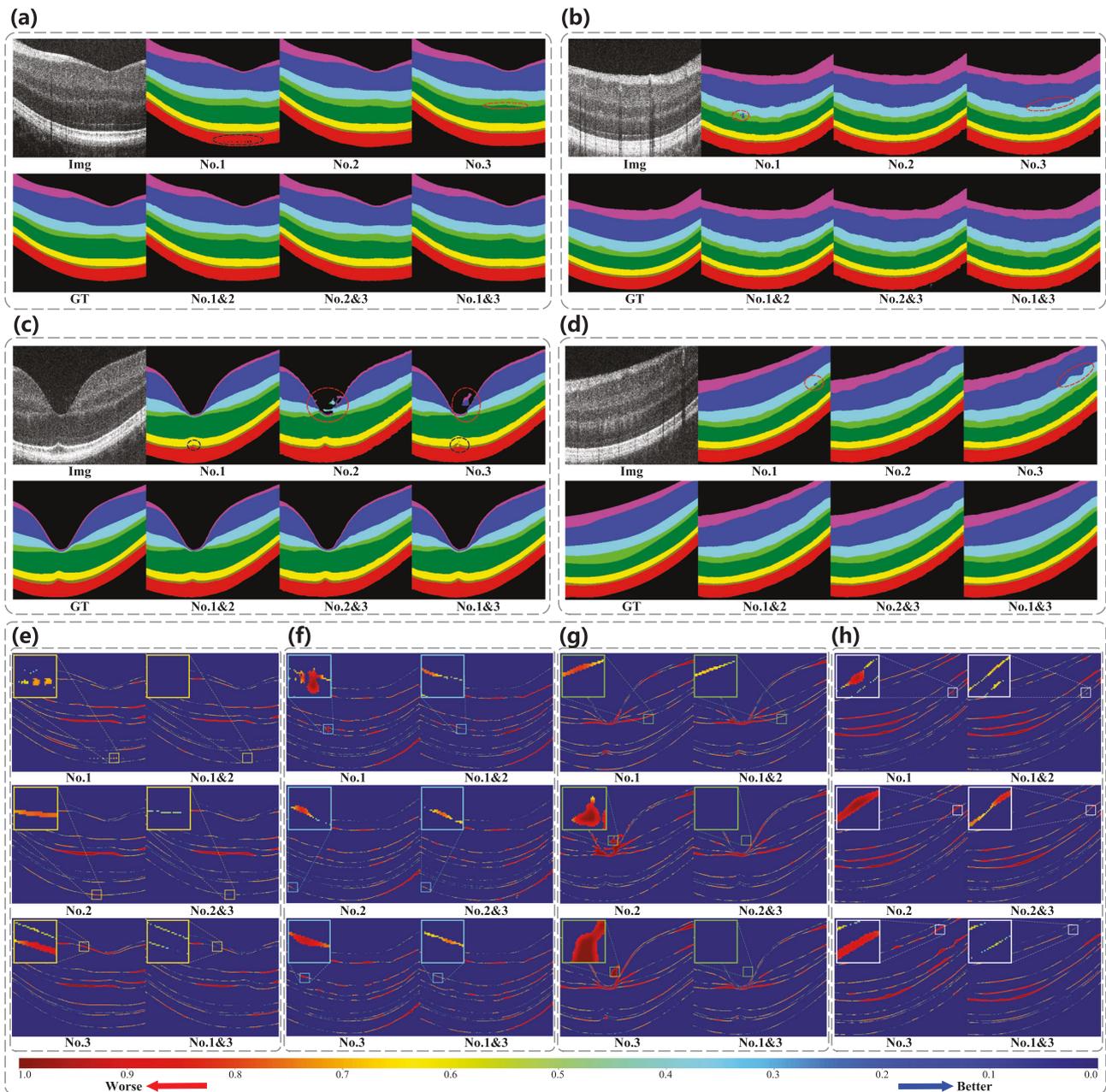


Fig. 3 | Qualitative description of multi-spectral methods for RLS. a–d shows a comparison of predictions based on four methods, Enet, ReLayNet, TransUnet, and AttU_net, under different spectral combinations. The red dashed boxes in the figures provide annotations for obvious intra-class errors or inter-class ambiguities.

e–h shows the absolute error distribution between the predicted and GT of four methods, Enet, ReLayNet, TransUnet, and AttU_net, and some regions are annotated using local zoom-in images.

errors. As shown in the prediction graphs of ReLayNet before and after adding MSI in the second and third rows of Fig. 6a, the prediction graphs of EMV-Net before and after adding MSI in the third row of Fig. 6b, and the prediction graphs of LightReSeg before and after adding MSI in Fig. 6c, all significantly reduce inter-class boundary errors. It is clear from the locally magnified areas in Fig. 6d–f that the error area of the prediction map with MSI added is usually significantly reduced.

We also conduct qualitative analysis on the VMS dataset, as shown in Fig. 7. In Fig. 7a–c, we can clearly see that on most single spectral images, the fusion of MSI significantly reduces intra-class and inter-class errors. The prediction graphs of ReLayNet before and after adding MSI in Fig. 7a, EMV-Net before and after adding MSI in the second row of Fig. 7b, and LightReSeg before and after adding MSI in the second row of Fig. 7c all

significantly reduce intra-class errors. As shown in the first and third rows of Fig. 7b, the predicted EMV-Net before and after adding MSI, and the predicted LightReSeg before and after adding MSI in the first and third rows of Fig. 7c, all significantly reduce inter-class boundary errors. It is clear from the locally magnified areas in Fig. 7d–f that the error area of the prediction map with MSI added is usually significantly reduced.

In addition, the introduction of this MSI encoder specifically designed for extracting MSI image features does indeed increase the number of model parameters. We conduct ablation experiments on the encoder using the LightReSeg method on the NIMS dataset. We find that if the LightReSeg method is not changed and only raw images are input, the mIoU index is 82.31; If the MSI encoder is added to the LightReSeg model and the raw image is simultaneously input into the MSI encoder, it will result in a

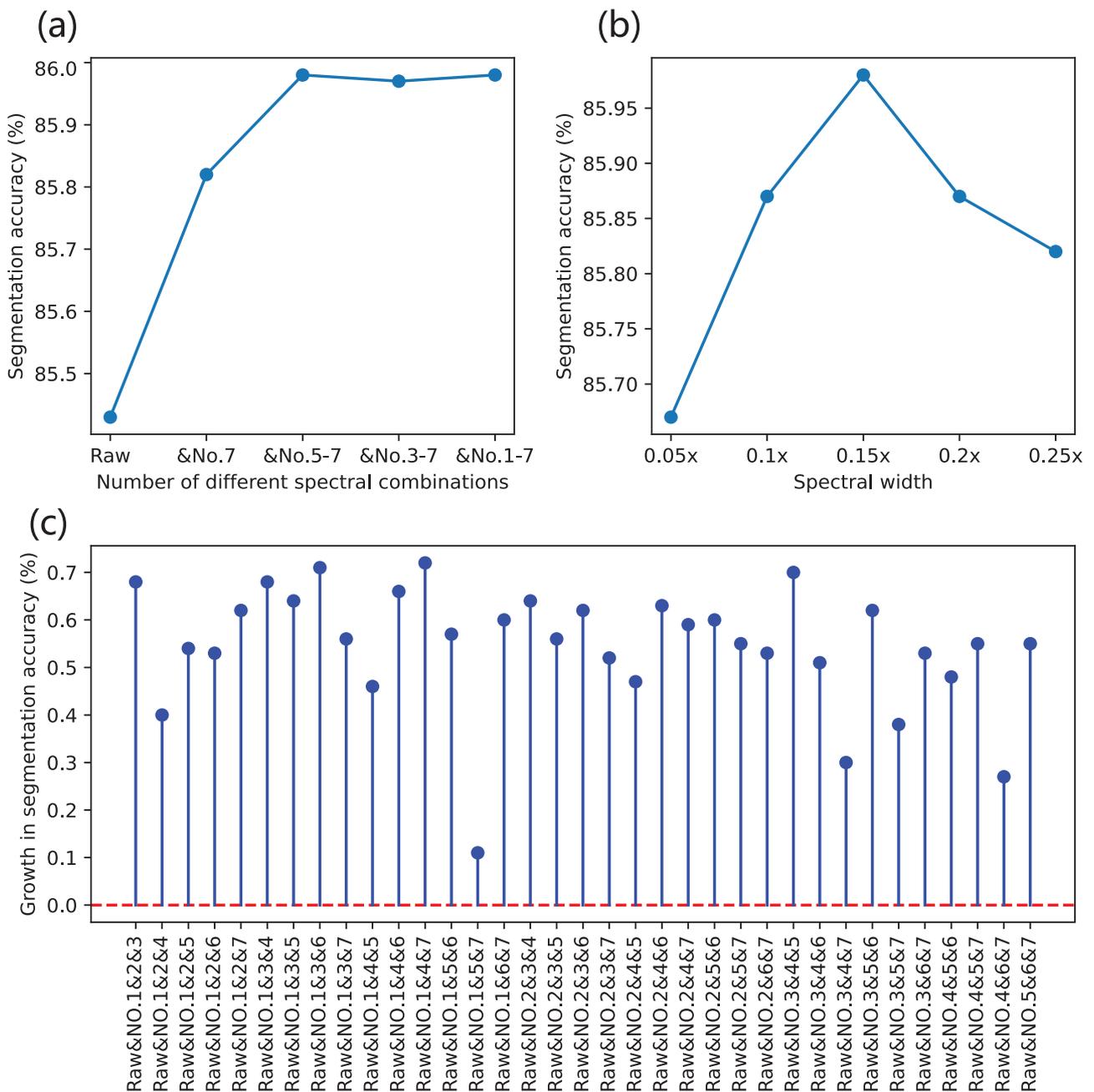


Fig. 4 | The primary factors of MSI on improving RLS accuracy in the VMS dataset. **a** explores the sequential addition of different numbers of visible-light spectral images to the raw-spectral image. Raw on the horizontal axis represents the raw spectral image (center wavelength 585 nm, spectral bandwidth of 127 nm), while &No.7, &No.5-7, &No.3-7, &No.1-7 represent the addition of the 7th, 5th to 7th, 3rd to 7th, and 1st to 7th spectral images to the raw spectral image, respectively, with a spectral bandwidth of 19 nm. No.1 (center wavelength 532 nm), No.2 (center wavelength 542 nm), No.3 (center wavelength 560 nm), No.4 (center wavelength 574 nm), No.5 (center wavelength 596nm), No.6 (center wavelength 610nm), No.7 (center wavelength 635 nm). **b** explores the sequential addition of visible-light spectral images with different bandwidths to the raw-image, with the abscissa representing multiples of the spectral bandwidth of the raw-image. **c** explores the sequential addition of different combinations of visible-light spectral images on the raw-image.

decrease in the accuracy of retinal layer segmentation, with a mIoU of 81.67; If the MSI image is input into the MSI encoder, the segmentation accuracy is significantly improved, with a mIoU index of 82.91. Therefore, the MSI encoder used in this method has specificity for MSI images, and only performs better in feature extraction of MSI images.

Discussion

Since the emergence of the ReLayNet¹² method in 2017, research on RLS based on DL has shown a rapid growth trend. However, most studies have focused primarily on improvements to model architecture or optimizations

of training strategies, leading to a fixed mindset in DL-based RLS that struggles to break through from other perspectives to further enhance the limitations of segmentation accuracy. Moreover, DL-based RLS methods often only demonstrate good segmentation performance in one type of data, and their universality is often unsatisfactory across a broader range of data. To address these issues, we propose an RLS method study based on MSI, verifying whether MSI has the capability to enhance RLS, exploring the main influencing factors of MSI, including bandwidth, quantity, and combinations, and applying it across general RLS methods and multiple spectral ranges of data.

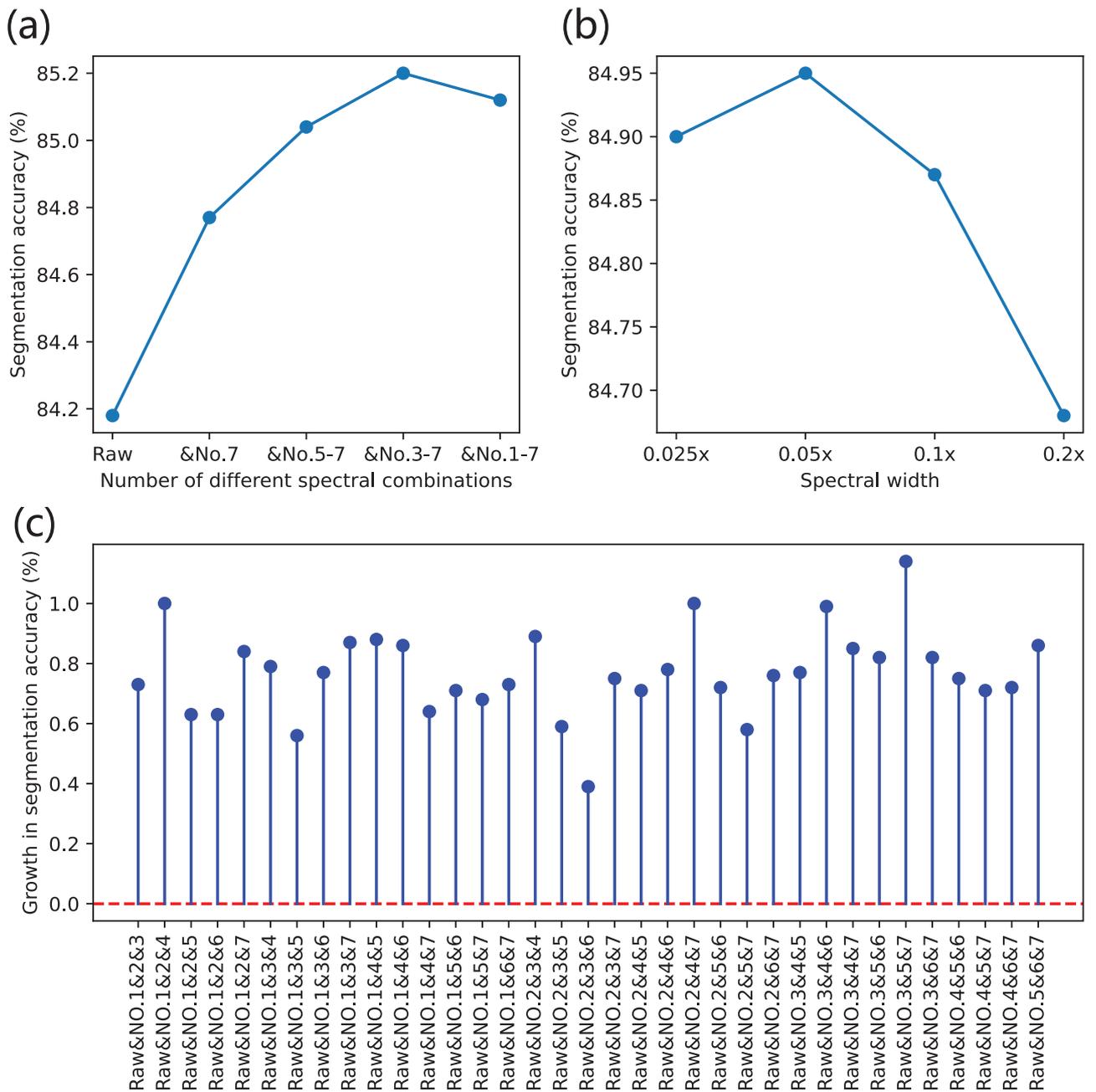


Fig. 5 | The primary factors of MSI on the improvement of RLS accuracy in the NIMS dataset. **a** explores the sequential addition of different numbers of near-infrared spectral images to the raw-spectral image. Raw on the horizontal axis represents the raw spectral image (center wavelength 844 nm, spectral bandwidth of 173 nm), while &No.7, &No.5-7, &No.3-7, &No.1-7 represent the addition of the 7th, 5th to 7th, 3rd to 7th, and 1st to 7th spectral images to the raw spectral image, respectively, with a spectral bandwidth of 8.6 nm. No.1 (center wavelength 778 nm), No.2 (center wavelength 800 nm), No.3 (center wavelength 821 nm), No.4 (center wavelength 843 nm), No.5 (center wavelength 865 nm), No.6 (center wavelength 886 nm), and No.7 (center wavelength 908 nm). **b** explores the sequential addition of near-infrared spectral images with different bandwidths to the raw-image, with the abscissa representing multiples of the spectral bandwidth of the raw-image. **c** explores the sequential addition of different combinations of near-infrared spectral images on the raw-image.

In exploring the primary factors influencing the effectiveness of MSI in the NIMS dataset and VMS dataset, we identified several commonalities: Firstly, as the number of spectral images increases, the trend of MSI's enhancement on RLS accuracy initially grows and then stabilizes. It is easy to understand from the perspective of information theory because the amount of input information has increased. Secondly, within a certain range of spectral bandwidths, there exists an optimal bandwidth range where MSI has the maximum effect on RLS. Lastly, among different MSI combinations, their impact on RLS accuracy varies. Besides these commonalities, we also observed differences between the two datasets. For instance, in Fig. 4a, the

trend of MSI's enhancement on RLS accuracy eventually stabilizes with the increase in the number of spectral images, whereas in Fig. 5a, it ultimately declines. We attribute this phenomenon to two main reasons: Firstly, as the number of spectral images increases, the weight of key spectral images in the DL model is diluted, leading to a reduced ability to extract effective MSI features. Secondly, with the addition of spectral images, the differences between multi-spectral images are diminished, which prevents the DL model from effectively utilizing MSI despite the increased input of spectral images. Regarding MSI combinations, the trends of enhancement also differ between the two spectral ranges. For example, in VMS dataset, the

Table 1 | Quantitative analysis of general RLS methods using MSI on NIMS dataset and VMS dataset, respectively

NIMS Dataset													
Method	Settings	IoU		GCL	INL	OPL	ONL	ELM	OS	RPE	mIoU	Dice	mPA
		NFL	NFI										
ReLayNet	Raw	81.57 ± 2.13	90.64 ± 0.72	80.90 ± 2.48	68.46 ± 3.03	89.77 ± 2.17	84.35 ± 2.15	73.12 ± 5.47	88.66 ± 3.01	82.19 ± 2.35	90.20 ± 1.43	97.24 ± 0.39	
	Raw+MSI	82.82 ± 2.30	90.64 ± 1.62	81.76 ± 2.56	69.65 ± 2.99	90.07 ± 2.29	84.65 ± 2.28	73.52 ± 5.12	88.70 ± 2.55	82.73 ± 2.33	90.53 ± 1.41	97.29 ± 0.46	
	Growth	1.25	0.00	0.86	1.19	0.30	0.29	0.40	0.04	0.54	0.32	0.05	
EMV-Net	Raw	79.64 ± 3.71	90.29 ± 1.54	82.46 ± 1.13	68.33 ± 4.46	89.32 ± 3.00	84.09 ± 2.82	73.30 ± 4.86	88.00 ± 3.11	81.93 ± 2.81	90.04 ± 1.72	97.18 ± 0.47	
	Raw+MSI	82.11 ± 3.32	90.92 ± 1.17	82.43 ± 1.87	70.35 ± 3.28	89.94 ± 2.63	84.46 ± 2.98	74.78 ± 4.99	88.47 ± 3.50	82.93 ± 2.57	90.65 ± 1.55	97.33 ± 0.47	
	Growth	2.46	0.63	-0.04	2.02	0.62	0.37	1.48	0.47	1.00	0.60	0.15	
SegFormer_B0	Raw	81.02 ± 1.97	90.54 ± 0.52	81.47 ± 1.38	68.48 ± 4.32	89.39 ± 3.01	83.73 ± 2.48	72.15 ± 5.59	88.10 ± 3.01	81.86 ± 2.41	90.00 ± 1.47	97.18 ± 0.42	
	Raw+MSI	81.68 ± 2.91	91.00 ± 0.87	82.42 ± 1.96	71.27 ± 1.58	90.26 ± 2.29	83.55 ± 3.15	73.32 ± 5.31	88.64 ± 2.65	82.77 ± 2.38	90.55 ± 1.44	97.34 ± 0.42	
	Growth	0.67	0.46	0.96	2.80	0.87	-0.17	1.17	0.54	0.91	0.55	0.16	
LightReSeg	Raw	82.51 ± 2.00	90.18 ± 2.02	81.68 ± 2.79	69.55 ± 4.48	89.22 ± 3.99	84.66 ± 2.57	72.70 ± 6.31	87.95 ± 3.84	82.31 ± 3.26	90.26 ± 1.99	97.21 ± 0.61	
	Raw+MSI	82.45 ± 1.86	91.01 ± 1.12	82.24 ± 2.13	70.43 ± 3.63	89.77 ± 3.02	84.39 ± 3.25	74.51 ± 5.88	88.49 ± 3.70	82.91 ± 2.84	90.63 ± 1.71	97.32 ± 0.50	
	Growth	-0.06	0.82	0.56	0.88	0.56	-0.27	1.80	0.53	0.60	0.37	0.11	
AtrUnet	Raw	82.19 ± 1.57	90.59 ± 1.37	82.24 ± 2.22	69.43 ± 3.97	89.70 ± 2.78	84.19 ± 2.49	72.42 ± 4.15	87.62 ± 2.90	82.30 ± 2.28	90.27 ± 1.39	97.24 ± 0.42	
	Raw+MSI	83.37 ± 2.02	91.27 ± 1.00	83.42 ± 0.89	71.46 ± 2.23	90.51 ± 2.18	84.70 ± 2.36	74.51 ± 5.87	88.34 ± 3.89	83.45 ± 2.29	90.96 ± 1.37	97.42 ± 0.43	
	Growth	1.18	0.68	1.18	2.02	0.82	0.51	2.09	0.72	1.15	0.69	0.19	
VMS Dataset													
Method	Settings	IoU		GCL	INL	OPL	ONL	ELM	OS	RPE	mIoU	Dice	mPA
		NFL	NFI										
ReLayNet	Raw	82.66 ± 4.27	91.48 ± 1.33	84.69 ± 1.19	74.08 ± 2.35	91.41 ± 1.18	89.27 ± 1.23	73.53 ± 1.61	91.44 ± 0.55	84.82 ± 0.93	91.78 ± 0.54	96.34 ± 0.47	
	Raw+MSI	83.00 ± 3.86	91.98 ± 1.17	85.23 ± 1.02	74.88 ± 2.24	91.91 ± 1.03	89.42 ± 1.25	74.32 ± 1.18	91.44 ± 0.60	85.27 ± 0.90	92.05 ± 0.52	96.46 ± 0.46	
	Growth	0.35	0.51	0.54	0.79	0.51	0.15	0.79	0.00	0.45	0.26	0.13	
EMV-Net	Raw	83.34 ± 4.04	92.25 ± 1.21	85.80 ± 0.88	75.17 ± 1.29	92.03 ± 0.97	89.25 ± 0.82	73.93 ± 2.63	91.45 ± 0.56	85.40 ± 0.53	92.12 ± 0.31	96.53 ± 0.37	
	Raw+MSI	83.96 ± 3.41	92.53 ± 1.16	86.13 ± 1.00	76.62 ± 1.58	92.31 ± 0.81	89.65 ± 1.18	75.15 ± 1.08	92.48 ± 1.00	86.01 ± 0.78	92.47 ± 0.45	96.64 ± 0.43	
	Growth	0.62	0.28	0.33	1.45	0.28	0.40	1.23	1.03	0.61	0.35	0.11	
SegFormer_B0	Raw	82.69 ± 3.41	91.92 ± 1.05	85.23 ± 0.98	74.83 ± 2.43	91.67 ± 1.03	88.78 ± 1.11	72.48 ± 2.42	91.21 ± 0.65	84.85 ± 0.49	91.8 ± 0.28	96.38 ± 0.34	
	Raw+MSI	83.35 ± 3.38	92.28 ± 1.04	85.74 ± 0.92	76.01 ± 1.63	92.19 ± 0.94	89.10 ± 1.02	74.00 ± 1.09	91.58 ± 0.53	85.53 ± 0.71	92.2 ± 0.42	96.58 ± 0.40	
	Growth	0.66	0.37	0.51	1.18	0.52	0.33	1.52	0.37	0.68	0.40	0.20	
LightReSeg	Raw	82.85 ± 4.12	92.04 ± 1.30	85.38 ± 1.12	75.06 ± 2.18	91.97 ± 1.00	89.29 ± 1.32	73.77 ± 1.99	91.01 ± 0.70	85.17 ± 0.85	91.99 ± 0.50	96.44 ± 0.45	
	Raw+MSI	83.71 ± 3.72	92.39 ± 1.17	86.18 ± 1.08	76.38 ± 1.83	92.21 ± 0.83	89.61 ± 1.10	75.31 ± 0.81	91.40 ± 0.57	85.90 ± 0.80	92.41 ± 0.46	96.60 ± 0.39	
	Growth	0.85	0.35	0.80	1.32	0.24	0.32	1.54	0.39	0.73	0.42	0.17	
AtrUnet	Raw	83.77 ± 3.63	92.41 ± 0.98	85.80 ± 0.74	75.74 ± 2.05	92.17 ± 0.92	89.71 ± 1.16	74.06 ± 2.20	91.27 ± 0.35	85.62 ± 0.46	92.25 ± 0.27	96.57 ± 0.37	
	Raw+MSI	84.42 ± 3.40	92.64 ± 0.86	86.35 ± 0.61	80.59 ± 8.14	92.46 ± 1.00	90.24 ± 1.05	75.83 ± 1.70	92.08 ± 0.44	86.33 ± 0.59	92.66 ± 0.34	96.74 ± 0.37	
	Growth	0.65	0.23	0.55	4.84	0.29	0.53	1.77	0.81	0.71	0.41	0.17	

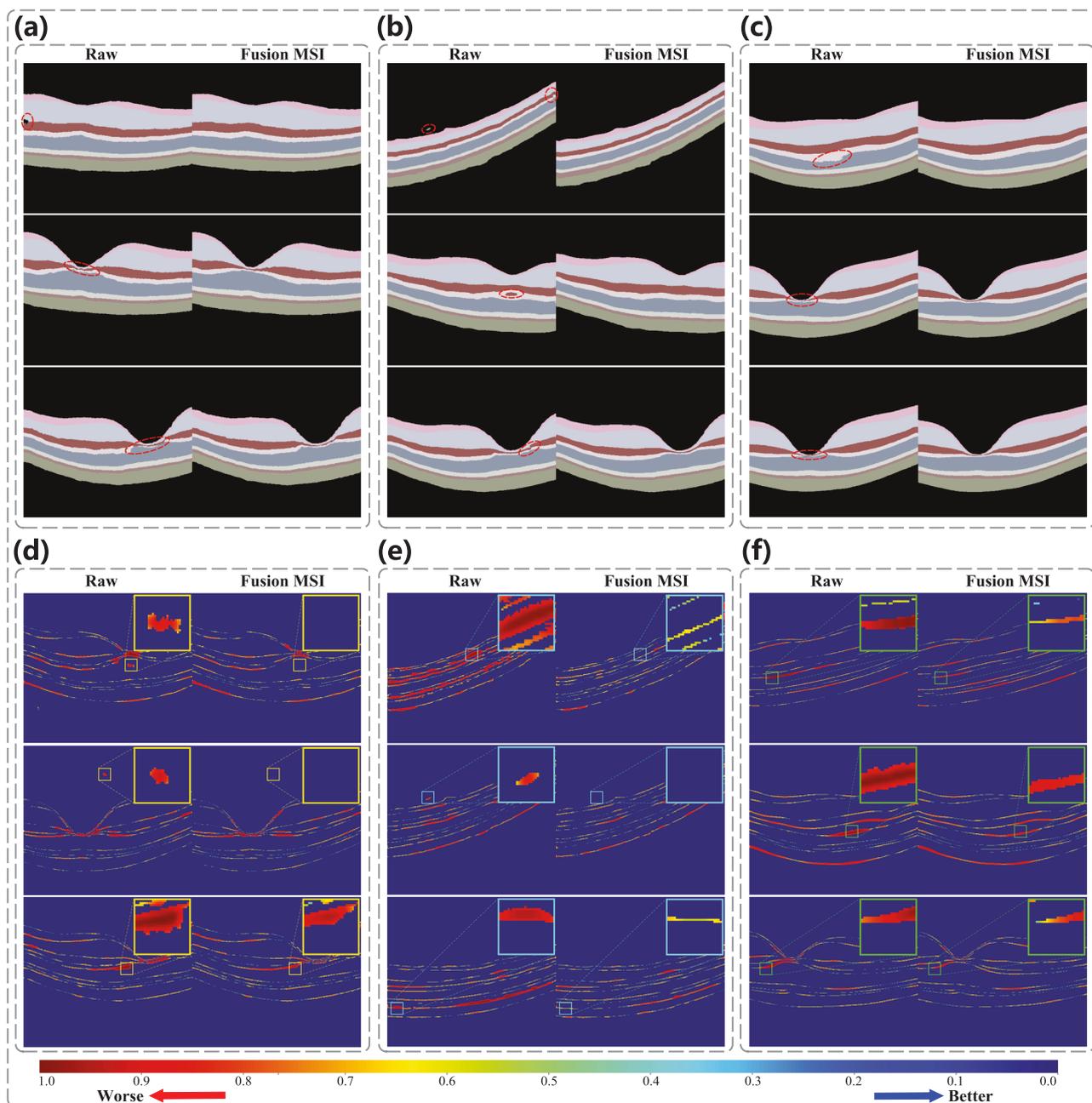


Fig. 6 | Apply MSI to general RLS methods on NIMS dataset. a–c show a comparison of predictions before and after adding MSI on three methods, ReLayNet, EMV_Net, LightReSeg. The red dashed boxes in the figures provide annotations for

obvious intra-class errors or inter-class ambiguities. d–f shows the absolute error distribution between the predicted and true values of three methods, ReLayNet, EMV_Net, LightReSeg, with some regions annotated using local zoom-in images.

combinations with the maximum and minimum enhancement are No.1&4&7 and No.1&6&7, respectively, while in NIMS dataset, they are No.3&5&7 and No.2&3&6. The magnitude of RLS accuracy improvement also varies significantly between the two datasets. We analyze that this phenomenon is primarily due to the different backscattering rates and penetration capabilities of light at different wavelengths in the same tissue or cells.

For researchers in RLS, the advent of MSI offers a direct and significant benefit. Historically, these researchers have invested considerable effort in developing methodologies to conduct RLS research based on structural information in images, progressing from simple thresholding techniques to sophisticated mathematical modeling approaches^{23–26}. They have also transitioned from machine learning algorithms that required manual feature extraction to deep learning

models capable of autonomous feature extraction^{27–29}. Despite these advancements, they have consistently sought to incorporate novel methods while preserving the insights of their predecessors, all in the pursuit of enhancing the accuracy of RLS. However, until now, they have not considered the possibility of unreasonable “compression” in dataset data or the potential neglect of certain information. The introduction of MSI has introduced a qualitative shift in the field of RLS segmentation. This innovative approach is poised to equip researchers with new tools for achieving higher precision in RLS. Consequently, this will significantly improve the diagnostic accuracy of various early-stage diseases that rely on RLS segmentation outcomes. As a result, certain retinal disorders, which have high blindness rates if left untreated in their initial stages, can now be effectively managed and suppressed right from their onset.

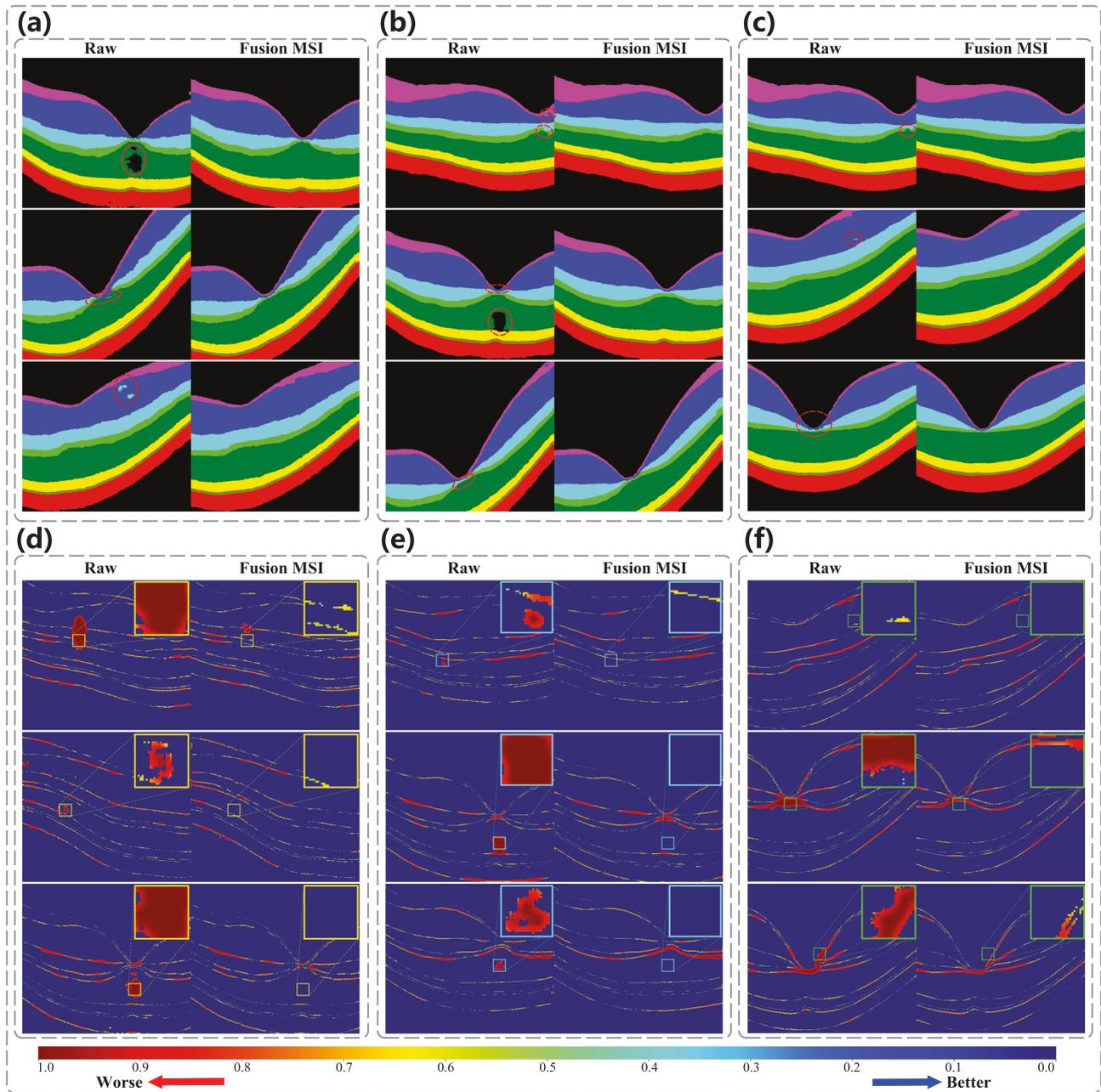


Fig. 7 | Apply MSI to general RLS methods on VMS dataset. a–c shows a comparison of predictions before and after adding MSI on three methods, ReLayNet, EMV_Net, LightReSeg. The red dashed boxes in the figures provide annotations for

obvious intra-class errors or inter-class ambiguities. d–f shows the absolute error distribution between the predicted and true values of three methods, ReLayNet, EMV_Net, LightReSeg, with some regions annotated using local zoom-in images.

For researchers based on OCT images, this is also a stunning discovery. This is because it is highly probable that new dimensions of variables will be introduced into their research, including retinal layer segmentation, fundus vascular segmentation³⁰, lesion detection³¹ and classification³² based on OCT images, and so on. In most cases, the increased volume of information is bound to contain certain factors beneficial to the task, and achieving high accuracy and reliability is the ultimate goal for medical image researchers. Therefore, researchers based on OCT images should also consider incorporating MSI into their next research steps. This is like 3D OCT scan data having one more depth information than 2D photos taken by fundus cameras.

For manufacturers of OCT equipment, we believe that MSI will usher in a new technological revolution in the fabrication of OCT. Firstly, the

neglect of MSI signifies an insufficient exploration by the industry on the raw data collected by OCT devices, with only a focus on conventional structural information (referring to 3D structural data). However, the amount of information supplied by OCT devices is far more extensive than this, as demonstrated by MSI; Secondly, the capacity to provide MSI will undoubtedly emerge as a highlight of OCT devices, and offering additional functions also reflects the technical prowess of the company; Finally, the continuous quest for innovative and valuable technologies will become an enduring and valuable custom for businesses. We call upon the entire OCT device manufacturing sector to incorporate this technology into their machinery, which involves almost zero cost.

In reality, this study also has several aspects that require further investigation. For instance, among different combinations of spectral images, some combinations significantly improve the accuracy of RLS, such as

No.1&4&7 in Fig. 4c and No.3&5&7 in Fig. 5c, while others only marginally enhance it, such as No.1&5&7 in Fig. 4c and No.2&3&6 in Fig. 5c. However, the internal patterns are not easy to discover. We believe that addressing this issue could start from several aspects. Firstly, the light of different central wavelengths incident on the retinal layers causes varying degrees of backscattering due to the physiological structure and elemental composition differences of each layer. Yet, it is challenging to precisely quantify how these differences contribute to the enhancement of RLS accuracy. Certainly, we can begin with fundamental optical incidence and reflection experiments, ranging from basic chemical elements to various retinal layer cells, ultimately conducting experiments on animals and even humans. This would thoroughly investigate the differences in backscattering rates of different spectrum on retinal layers and identify the main factors causing these differences. Secondly, there is a need to further explore the internal mechanisms of DL models, to comprehensively understand how DL models extract MSI from different spectral combinations and in which combinations DL models excel at obtaining MSI. However, current experimental conditions do not allow for the smooth resolution of the aforementioned uncertainties.

In addition, the MSI in this study is limited by current technological means and is derived from a continuous fixed spectrum with narrow bandwidth, such as the spectral range of VMS dataset being 521 nm–648 nm and that of NIMS dataset being 757 nm–930 nm. If images are selected over a wider spectral bandwidth range, it may result in a more significant enhancement of RLS. If the aforementioned optimal MSI combinations cannot be explored by analyzing spectral patterns, it raises the possibility of testing all potential combinations using an exhaustive approach. Taking the combination of three different central wavelengths as an example, assuming that spectral images are taken at intervals of 1nm center wavelengths on the VMS dataset, there would be at least approximately 2 million combinations. Therefore, computational power and time cost are the main constraints for the exhaustive method. Besides, the quality of single spectral images is affected by two main factors: one is the center wavelength, and the other is the spectral intensity of the original OCT light source in different wavelength ranges. Therefore, it is challenging to evaluate the contribution of different single spectral images to the segmentation accuracy of the original image.

When applying MSI to general retinal layer methods, we design a specialized MSI encoder specifically designed to extract spectral information from MSI images, which indeed increases the computational complexity of the model. The reason why we design the MSI encoder is that: firstly, the sub-band spectral width of MSI images $\Delta\lambda$ is much lower than that of raw images. According to the OCT axial resolution formula $\delta_z = \frac{2\ln 2}{\pi} \frac{\lambda_0^2}{\Delta\lambda}$, the optical axial resolution of MSI images δ_z is lower than that of raw images, this leads to the inability to improve segmentation performance by directly using concatenating as an input during the research process; Secondly, we hope to conveniently apply the spectral information to general retinal layer segmentation methods, so we design the MSI encoder specifically for MSI images. Our immediate objective is to advance in this direction by designing a MSI encoder that strikes an optimal balance between being light-weight and ensuring high segmentation accuracy.

In summary, we have discovered for the first time the enhancing effect of MSI on RLS accuracy. Specifically, this includes: First, compared to traditional single-spectrum imagery, our use of equal-bandwidth MSI images yields higher segmentation accuracy; Second, we have identified that bandwidth, quantity, and spectral combinations are the main factors affecting MSI; Third, the application of MSI to general RLS methods shows significant improvement, with similar patterns across different spectral ranges, proving the universality of MSI. In the future, we will build upon the current research to further explore precise segmentation of lesion areas in fundus diseases based on MSI, as well as classification studies of common ophthalmic diseases using MSI. At the same time, further promoting this work to the OCT manufacturing industry and researchers based on OCT images promotes the transformation and upgrading of the entire OCT field.

Methods

The principle of MSI

The fundamental principle of OCT involves directing a light beam onto the tissue or specimen to be imaged. The light beam is reflected by microscopic structures at various depths. By measuring the reflected light, as well as the intensity of the reflected or backscattered light, and converting the reflection information obtained from different positions into digital signals, which are then processed by a computer, it is possible to transform this data into two-dimensional or three-dimensional images. This process reveals the microscopic structures of the layers within the imaged tissue⁵.

As shown in Fig. 1a, firstly, we perform FFT transformation on the raw spectrum to obtain the raw spectral image. Then, we convolve the raw spectrum of OCT images by adding identical Gaussian windows at seven different positions, resulting in seven sub-band spectrums with center wavelengths of 778 nm, 799 nm, 820 nm, 841 nm, 863 nm, 884 nm, and 905nm, respectively, and they are all of equal-bandwidth spectrum. By using these 7 new sub-band spectrums to simultaneously image the retina, OCT images containing 7 different spectral information at the same location can be obtained after FFT transformation. At this point, we complete the first step, which is how to obtain multi-spectral images.

We randomly select three images generated from the No.1, No.3 and No.5 spectrum, as shown in Fig. 1b, and then we average the pixel values of each layer in No.1, No.3 and No.5 images to obtain denoised smooth images. It's difficult for us to distinguish the difference between these denoised images with the naked eye, but according to the formula:

$$SCR = \frac{LPV_{No.n}}{LPV_{No.m}}, \quad (1)$$

where the LPV represents the pixel values of each layer, No.n and No.m represent the nth and mth spectral images, $m \neq n$, and SCR represents the spectral ratio. Upon calculation, under the premise of 0.3 times the raw spectral bandwidth, the SCR values for No.1 and No.3 are as follows: NFL is 1.041, GCL is 1.057, INL is 1.370, OPL is 1.131, ONL is 1.298, ELM is 1.665, OS is 1.119, RPE is 1.054, respectively. And the SCR values for No.1 and No.5 are as follows: NFL is 1.100, GCL is 1.301, INL is 2.044, OPL is 1.376, ONL is 2.588, ELM is 2.077, OS is 1.097, RPE is 1.148, respectively. The SCR values for No.3 and No.5 are as follows: NFL is 1.056, GCL is 1.230, INL is 1.490, OPL is 1.216, ONL is 1.991, ELM is 1.247, OS is 0.980, RPE is 1.090, respectively. We can observe distinct differences among the various spectral images.

When different combinations are used, these differences undergo significant changes. As the number of spectral images increases, the information about these differences becomes more abundant. We refer to this differential information as a manifestation of MSI. Furthermore, we have employed a univariate analysis approach to compare the SCR performance on different spectral image combinations for each retinal layer, as presented in Table 2. For example, when comparing the differences between No.1&2 and the other combinations on the NIMS dataset, we observe the *P* values as follows: *P* value of 0.00785 ($p < 0.05$), 0.00785 ($p < 0.05$), 1.0, 0.00785 ($p < 0.05$), and 0.00785 ($p < 0.05$) when comparing on the No.1&2 with No.1&3, No.1&4, No.1&5, No.1&6, No.1&7, respectively. On the VMS dataset, we observe the *P* values as follows: *P* value of 0.00785 ($p < 0.05$), 0.00785 ($p < 0.05$), 0.00785 ($p < 0.05$) and 0.05468 when comparing on the No.1&2 with No.1&3, No.1&4, No.1&5, No.1&6, No.1&7, respectively. Our findings reveal significant differences in MSI across most of these combinations. Additionally, the performance of MSI varies across different spectral ranges. This further underscores that MSI encompasses a wealth of nonlinear information, offering an abundant resource for DL-based RLS.

The framework of MSI applied to RLS method

Traditional DL-based RLS methods adopt an encoder-decoder structure and typically include a multi-scale feature extraction mechanism. For instance, ReLayNet utilizes a U-shaped network that first performs multi-

Table 2 | The SCR performance on different spectral combinations, and it's a univariate analysis conducted by randomly selecting one sample of data from NIMS dataset and VMS dataset, respectively

SCR on NIMS Dataset										
Combination	NFL	GCL	INL	OPL	ONL	ELM	OS	RPE		
No.1&No.2	1.174E+00	1.440E+00	1.817E+00	1.557E+00	1.572E+00	2.091E+00	1.154E+00	1.082E+00		
No.1&No.3	1.041E+00	1.057E+00	1.370E+00	1.131E+00	1.298E+00	1.665E+00	1.119E+00	1.054E+00		
No.1&No.4	8.070E-01	6.750E-01	8.490E-01	7.580E-01	8.580E-01	1.146E+00	9.810E-01	9.860E-01		
No.1&No.5	1.100E+00	1.301E+00	2.044E+00	1.376E+00	2.588E+00	2.077E+00	1.097E+00	1.148E+00		
No.1&No.6	2.625E+00	1.607E+01	4.741E+01	2.047E+01	1.578E+02	3.409E+00	1.519E+00	1.839E+00		
No.1&No.7	1.379E+01	6.360E+02	1.305E+03	1.872E+03	1.244E+03	2.469E+00	3.183E+00	4.061E+00		
No.2&No.3	8.870E-01	7.340E-01	7.530E-01	7.260E-01	8.250E-01	7.960E-01	9.690E-01	9.730E-01		
No.2&No.4	6.870E-01	4.690E-01	4.670E-01	4.860E-01	5.450E-01	5.480E-01	8.500E-01	9.110E-01		
No.2&No.5	9.370E-01	9.030E-01	1.123E+00	8.830E-01	1.645E+00	9.930E-01	9.500E-01	1.061E+00		
No.2&No.6	2.236E+00	1.115E+01	2.606E+01	1.314E+01	1.003E+02	1.630E+00	1.316E+00	1.698E+00		
No.2&No.7	1.175E+01	4.414E+02	7.175E+02	1.202E+03	7.907E+02	1.180E+00	2.757E+00	3.751E+00		
No.3&No.4	7.740E-01	6.380E-01	6.190E-01	6.700E-01	6.600E-01	6.880E-01	8.770E-01	9.350E-01		
No.3&No.5	1.056E+00	1.230E+00	1.490E+00	1.216E+00	1.991E+00	1.247E+00	9.800E-01	1.090E+00		
No.3&No.6	2.520E+00	1.519E+01	3.457E+01	1.809E+01	1.214E+02	2.046E+00	1.357E+00	1.745E+00		
No.3&No.7	1.324E+01	6.012E+02	9.516E+02	1.654E+03	9.574E+02	1.482E+00	2.844E+00	3.854E+00		
No.4&No.5	1.364E+00	1.926E+00	2.406E+00	1.816E+00	3.013E+00	1.812E+00	1.118E+00	1.165E+00		
No.4&No.6	3.254E+00	2.378E-01	5.582E+01	2.701E+01	1.837E+02	2.974E+00	1.548E+00	1.865E+00		
No.4&No.7	1.709E+01	9.414E+02	1.537E+03	2.470E+03	1.449E+03	2.154E+00	3.244E+00	4.120E+00		
No.5&No.6	2.386E+00	1.234E+01	2.317E+01	1.487E+01	6.083E+01	1.641E+00	1.384E+00	1.601E+00		
No.5&No.7	1.253E+01	4.886E+02	6.379E+02	1.360E+03	4.798E+02	1.188E+00	2.901E+00	3.537E+00		
No.6&No.7	5.252E+00	3.933E+01	2.678E+01	9.088E+01	6.887E+00	7.240E-01	2.096E+00	2.209E+00		
SCR on VMS Dataset										
Combination	NFL	GCL	INL	OPL	ONL	ELM	OS	RPE		
No.1&No.2	7.250E-01	6.420E-01	5.450E-01	5.180E-01	4.660E-01	4.000E-01	4.970E-01	6.580E-01		
No.1&No.3	5.250E-01	4.120E-01	3.020E-01	2.760E-01	2.160E-01	1.610E-01	2.800E-01	4.560E-01		
No.1&No.4	4.710E-01	3.410E-01	2.370E-01	2.170E-01	1.610E-01	1.160E-01	2.390E-01	4.000E-01		
No.1&No.5	4.640E-01	3.250E-01	2.250E-01	2.050E-01	1.510E-01	1.100E-01	2.310E-01	3.700E-01		
No.1&No.6	4.880E-01	3.610E-01	2.610E-01	2.300E-01	1.770E-01	1.320E-01	2.460E-01	3.710E-01		
No.1&No.7	6.810E-01	6.280E-01	5.840E-01	4.570E-01	4.630E-01	3.590E-01	4.650E-01	4.570E-01		
No.2&No.3	7.240E-01	6.410E-01	5.540E-01	5.320E-01	4.640E-01	4.030E-01	5.640E-01	6.920E-01		
No.2&No.4	6.500E-01	5.310E-01	4.350E-01	4.180E-01	3.440E-01	2.900E-01	4.810E-01	6.070E-01		
No.2&No.5	6.400E-01	5.070E-01	4.120E-01	3.960E-01	3.240E-01	2.750E-01	4.660E-01	5.620E-01		
No.2&No.6	6.740E-01	5.630E-01	4.790E-01	4.430E-01	3.800E-01	3.300E-01	4.960E-01	5.640E-01		
No.2&No.7	9.390E-01	9.780E-01	1.070E+00	8.810E-01	9.930E-01	8.970E-01	9.370E-01	6.940E-01		
No.3&No.4	8.970E-01	8.270E-01	7.860E-01	7.870E-01	7.410E-01	7.190E-01	8.530E-01	8.780E-01		
No.3&No.5	8.830E-01	7.900E-01	7.440E-01	7.460E-01	6.980E-01	6.830E-01	8.260E-01	8.120E-01		

Table 2 (continued) | The SCR performance on different spectral combinations, and it's a univariate analysis conducted by randomly selecting one sample of data from NIMS dataset and VMS dataset, respectively

Combination	NFL	GCL	INL	OPL	ONL	ELM	OS	RPE
No.3&No.6	9.300E-01	8.780E-01	8.650E-01	8.340E-01	8.170E-01	8.190E-01	8.790E-01	8.150E-01
No.3&No.7	1.297E+00	1.525E+00	1.932E+00	1.657E+00	2.137E+00	2.227E+00	1.661E+00	1.003E+00
No.4&No.5	9.850E-01	9.550E-01	9.470E-01	9.470E-01	9.410E-01	9.500E-01	9.680E-01	9.250E-01
No.4&No.6	1.037E+00	1.061E+00	1.101E+00	1.059E+00	1.102E+00	1.139E+00	1.030E+00	9.280E-01
No.4&No.7	1.446E+00	1.843E+00	2.458E+00	2.105E+00	2.882E+00	3.098E+00	1.947E+00	1.143E+00
No.5&No.6	1.054E+00	1.110E+00	1.163E+00	1.118E+00	1.171E+00	1.199E+00	1.064E+00	1.004E+00
No.5&No.7	1.468E+00	1.930E+00	2.596E+00	2.222E+00	3.061E+00	3.261E+00	2.012E+00	1.235E+00
No.6&No.7	1.394E+00	1.738E+00	2.232E+00	1.987E+00	2.614E+00	2.719E+00	1.890E+00	1.231E+00

scale feature extraction on the input image in the encoder part, followed by upsampling to merge these multi-scale features, ultimately producing the target image. Our approach involves extracting and integrating multi-scale features of the same dimensions as the original method, as illustrated in Fig. 1c. Initially, the raw spectral image undergoes multi-level feature extraction using the encoder part of the RLS method. Similarly, the MSI images are processed through a spectral MSI encoder to obtain multi-scale features of identical dimensions. Subsequently, features at the same level are merged, resulting in a feature map that encompasses both retinal structural information and MSI. This merged feature map is then fed into the RLS decoder to complete the image upsampling restoration process, then final segmented image is obtained. The simplicity of this method is highlighted by the feature fusion step, which only requires a concatenation operation. For methods with fewer multi-scale features, such as DeepLab³³ with only two levels of scale, partial feature fusion suffices. Integrating multi-spectral information, regardless of the number of scales employed, enhances the performance of RLS methods.

The MSI dataset

The spectral range of the NIMS dataset is 757 nm–930 nm, with its raw spectral images utilizing the full bandwidth. The MSI consists of seven segments, namely No.1 (central wavelength at 778 nm), No.2 (central wavelength at 800 nm), No.3 (central wavelength at 821 nm), No.4 (central wavelength at 843 nm), No.5 (central wavelength at 865 nm), No.6 (central wavelength at 886 nm), and No.7 (central wavelength at 908 nm), each with a bandwidth that can be defined according to needs. There are two methods to obtain MSI images: one is to customize OCT devices with different waveband sources, and the other is to employ spectral shaping techniques. The latter method is adopted in this study due to the higher experimental costs associated with the former. The subjects are volunteers recruited for the study, totaling 100 healthy individuals. The inclusion criteria are as follows: (1) absence of any ophthalmic diseases; (2) visual acuity greater than 0.6 in both eyes; (3) normal eye position, capable of central fixation, clear refractive media, and able to undergo ocular OCT examination; (4) 50 males and 50 females. The exclusion criterion is (1) unclear structural display on OCT imaging. The dataset consisted of 500 B-scan samples, each containing 8 types of spectral data images and a ground truth image supervised and marked by ophthalmology experts, thus the NIMS dataset contained a total of 4500 images. The dataset is divided into training and test datasets at an 8 : 2 ratio, with the training dataset containing 400 samples (3600 images) and the test dataset containing 100 samples (900 images). Each image has a dimension of 480 × 400. The spectral range of the VMS dataset is 521 nm–648 nm, and its raw spectral images utilize the full bandwidth. The MSI for this dataset is also divided into 7 segments, namely No.1 (central wavelength at 532 nm), No.2 (central wavelength at 542 nm), No.3 (central wavelength at 560 nm), No.4 (central wavelength at 574 nm), No.5 (central wavelength at 596 nm), No.6 (central wavelength at 610 nm), and No.7 (central wavelength at 635 nm). It's inclusion criteria and dataset partitioning criteria are consistent with those of the NIMS dataset.

The DL component of this study is conducted using the PyTorch framework, with training performed using the Adam optimizer featuring cross-entropy loss. It adopts a 5-fold cross-validation approach and the maximum number of training epochs is 200 on both datasets. The initial learning rate is set to 0.001 and is gradually reduced by half every 20 epochs. Data augmentation includes a horizontal flip with a probability of $P = 0.5$, and random central rotation within ± 10 degrees with a probability of $P = 0.5$. All experiments are carried out on four NVIDIA A100 GPUs.

Code availability

The underlying code for this study are available from the corresponding author upon reasonable request.

Received: 19 September 2024; Accepted: 10 January 2025;

Published online: 17 January 2025

References

- Sarla, G. S. Excessive use of electronic gadgets: health effects. *Egypt. J. Intern. Med.* **31**, 408–411 (2019).
- Leung, C. K. et al. Comparison of macular and peripapillary measurements for the detection of glaucoma: an optical coherence tomography study. *Ophthalmology* **112**, 391–400 (2005).
- Xie, H. et al. Arterial hypertension and retinal layer thickness: the beijing eye study. *Br. J. Ophthalmol.* **108**, 105–111 (2024).
- Bsteh, G. et al. Retinal layer thickness predicts disability accumulation in early relapsing multiple sclerosis. *Eur. J. Neurol.* **30**, 1025–1034 (2023).
- Huang, D. et al. Optical coherence tomography. *Science* **254**, 1178–1181 (1991).
- Wolf, S. & Wolf-Schnurrbusch, U. Spectral-domain optical coherence tomography use in macular diseases: a review. *Ophthalmologica* **224**, 333–340 (2010).
- Shen, Y. et al. Graph attention u-net for retinal layer surface detection and choroid neovascularization segmentation in oct images. *IEEE Trans. Med. Imaging* **42**, 3140–3154 (2023).
- Wang, M. et al. Mstganet: automatic drusen segmentation from retinal oct images. *IEEE Trans. Med. Imaging* **41**, 394–406 (2021).
- Gende, M. et al. Automatic segmentation of retinal layers in multiple neurodegenerative disorder scenarios. *IEEE J. Biomed. Health Inform* **27**, 5483–5494 (2023).
- Cheung, C. Y. et al. A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre. *Nat. Biomed. Eng.* **5**, 498–508 (2021).
- Li, Z. et al. Preventing corneal blindness caused by keratitis using artificial intelligence. *Nat. Commun.* **12**, 3738 (2021).
- Roy, A. G. et al. Relaynet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks. *Biomed. Opt. express* **8**, 3627–3642 (2017).
- Hickam, J. B., Frayser, R. & Ross, J. C. A study of retinal venous blood oxygen saturation in human subjects by photographic means. *Circulation* **27**, 375–385 (1963).
- Chen, S., Yi, J. & Zhang, H. F. Measuring oxygen saturation in retinal and choroidal circulations in rats using visible light optical coherence tomography angiography. *Biomed. Opt. express* **6**, 2840–2853 (2015).
- Ma, F., Yuan, M. & Kozak, I. Multispectral imaging: Review of current applications. *Surv. Ophthalmol.* **68**, 889–904 (2023).
- Paszke, A., Chaurasia, A., Kim, S. & Culurciello, E. Enet: a deep neural network architecture for real-time semantic segmentation. *arXiv preprint arXiv:1606.02147* (2016).
- Oktay, O. et al. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999* (2018).
- Chen, J. et al. TransUNet: Rethinking the U-Net architecture design for medical image segmentation through the lens of transformers. *Med. Image Analysis* **97**, 103280 (2024).
- Yin, B. et al. Dformer: Rethinking rgb-d representation learning for semantic segmentation. *arXiv preprint arXiv:2309.09668* (2023).
- He, X. et al. Exploiting multi-granularity visual features for retinal layer segmentation in human eyes. *Front. Bioeng. Biotechnol.* **11**, 1191803 (2023).
- He, X. et al. Light-weight retinal layer segmentation with global reasoning. *IEEE Transact. Instrum. Measure* **73**, 2520214 (2024).
- Xie, E. et al. Segformer: simple and efficient design for semantic segmentation with transformers. *Adv. Neural Inf. Process. Syst.* **34**, 12077–12090 (2021).
- Akram, M. U. & Khan, S. A. Multilayered thresholding-based blood vessel segmentation for screening of diabetic retinopathy. *Eng. Comput.* **29**, 165–173 (2013).
- Niu, S. et al. Automated retinal layers segmentation in sd-oct images using dual-gradient and spatial correlation smoothness constraint. *Comput. Biol. Med.* **54**, 116–128 (2014).
- Wang, C., Wang, Y. X. & Li, Y. Automatic choroidal layer segmentation using markov random field and level set method. *IEEE J. Biomed. Health Inform.* **21**, 1694–1702 (2017).
- Gan, M. et al. Robust layer segmentation of esophageal oct images based on graph search using edge-enhanced weights. *Biomed. Opt. Express* **9**, 4481–4495 (2018).
- Ruan, Y. et al. Multi-phase level set algorithm based on fully convolutional networks (fcn-mls) for retinal layer segmentation in sd-oct images with central serous chorioretinopathy (csc). *Biomed. Opt. express* **10**, 3987–4002 (2019).
- Li, J. et al. Multi-scale gcn-assisted two-stage network for joint segmentation of retinal layers and discs in peripapillary oct images. *Biomed. Opt. Express* **12**, 2204–2220 (2021).
- Huang, K. et al. Diverse data generation for retinal layer segmentation with potential structure modelling. *IEEE Transact. Med. Imaging* **43**, 3584–3595 (2024).
- Qin, Q. & Chen, Y. A review of retinal vessel segmentation for fundus image analysis. *Eng. Appl. Artif. Intell.* **128**, 107454 (2024).
- Lian, J. & Liu, T. Lesion identification in fundus images via convolutional neural network-vision transformer. *Biomed. Signal Process. Control* **88**, 105607 (2024).
- Al-Fahdawi, S. et al. Fundus-deepnet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images. *Inf. Fusion* **102**, 102059 (2024).
- Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K. & Yuille, A. L. Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. *IEEE Trans. Pattern Anal. Mach. Intell.* **40**, 834–848 (2017).

Acknowledgements

This work is supported in part by the Key Technology Research and Development Program of Shandong Province under Grant 2024CXGC010201, 2024JMRH0301 and 2024JMRH0207.

Author contributions

X.H. contributed to the study design, data collection, and analysis, and drafted the manuscript. F.W. and K.H. contributed to data collection. L.Z., W.S., and Y.W. contributed to the study design, supervision, and manuscript review. All authors reviewed and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Weiye Song or Yi Wan.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025