npj | digital medicine

Article

Published in partnership with Seoul National University Bundang Hospital

https://doi.org/10.1038/s41746-024-01292-5

Early detection of dementia through retinal imaging and trustworthy AI

Check for updates

Jinkui Hao[1,13], William R. Kwapong[2,13], Ting Shen[3,13], Huazhu Fu[4], Yanwu Xu[5], Qinkang Lu[6], Shouyue Liu[1], Jiong Zhang[1], Yonghuai Liu[7], Yifan Zhao[8], Yalin Zheng[9], Alejandro F. Frangi[10,11], Shuting Zhang[2] ✉, Hong Qi[12] ✉ & Yitian Zhao[1,6,9] ✉

Alzheimer's disease (AD) is a global healthcare challenge lacking a simple and affordable detection method. We propose a novel deep learning framework, Eye-AD, to detect Early-onset Alzheimer's Disease (EOAD) and Mild Cognitive Impairment (MCI) using OCTA images of retinal microvasculature and choriocapillaris. Eye-AD employs a multilevel graph representation to analyze intra- and inter-instance relationships in retinal layers. Using 5751 OCTA images from 1671 participants in a multi-center study, our model demonstrated superior performance in EOAD (internal data: AUC = 0.9355, external data: AUC = 0.9007) and MCI detection (internal data: AUC = 0.8630, external data: AUC = 0.8037). Furthermore, we explored the associations between retinal structural biomarkers in OCTA images and EOAD/MCI, and the results align well with the conclusions drawn from our deep learning interpretability analysis. Our findings provide further evidence that retinal OCTA imaging, coupled with artificial intelligence, will serve as a rapid, noninvasive, and affordable dementia detection.

Dementia affects approximately 50 million individuals worldwide, and the number of patients is increasing due to the aging population[1]. The development of new methods for studying dementia's early pathophysiology is imperative. Current tests, such as magnetic resonance imaging (MRI), and biochemical quantification of proteins in cerebrospinal fluid are expensive, either time-consuming or invasive, and cannot be used on a large population.

The retina has been postulated to be a window to the brain and hence may provide an opportunity for the study of neurodegeneration and microvascular changes in the early pathophysiology of dementia[2,3]. Histopathological reports from post-mortem cases have confirmed retinal microvascular changes in patients with Alzheimer's disease (AD)[4]. Several clinical studies have supported this theory, confirming changes in retinal vasculature, vessels of different calibers, and structural changes around the optic nerve head and retinal structure in AD patients[5–8]. These changes can be detected with ophthalmic imaging modalities such as color fundus photography (CFP), optical coherence tomography (OCT), and OCT angiography (OCTA). Currently, accumulating evidence suggests that retinal imaging tools may provide useful biomarkers for the study and management of the early pathobiology of dementia.

CFP imaging has advantages in accessibility and cost. However, its analysis detects retinal microvasculature changes mainly confined to arterioles and venules (60–300 μm in diameter) due to limited image resolutions[9] and finds it challenging to detect more subtle vascular changes at the very early stages of the disease, such as mild cognitive impairment[8,9]. OCTA is a novel imaging modality that noninvasively and quickly images the retinal microvasculature including capillaries with high resolution (5–6μ m in diameter). Moreover, this technique provides in-depth information, allowing the visualization and representation of the retinal

[1]Laboratory of Advanced Theranostic Materials and Technology, Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, Ningbo, China. [2]Department of Neurology, West China Hospital, Sichuan University, Chengdu, China. [3]Department of Ophthalmology, the Second Affiliated Hospital of Zhejiang University, Hangzhou, China. [4]Institute of High-Performance Computing, Agency for Science, Technology and Research, Singapore, Singapore. [5]School of Future Technology, South China University of Technology, Guangzhou, China. [6]Department of Ophthalmology, the Affiliated People's Hospital of Ningbo University, Ningbo, China. [7]Department of Computer Science, Edge Hill University, Ormskirk, UK. [8]School of Aerospace, Transport and Manufacturing, Cranfield University, Bedfordshire, UK. [9]Department of Eye and Vision Sciences, University of Liverpool, Liverpool, UK. [10]Division of Informatics, Imaging and Data Sciences, School of Health Sciences, The University of Manchester, Manchester, UK. [11]Department of Computer Science, School of Engineering, The University of Manchester, Manchester, United Kingdom. [12]Department of Ophthalmology, Peking University Third Hospital, Beijing, China. [13]These authors contributed equally: Jinkui Hao, William R. Kwapong, Ting Shen. ✉e-mail: shutingzhang@scu.edu.cn; doctorqihong@163.com; yitian.zhao@nimte.ac.cn

microvascular network and structure of the foveal avascular zone, in different retinal layers and the choroid.

Recent advances in artificial intelligence (AI), especially deep learning, have enabled new approaches for AD detection utilizing ocular imaging. A pioneer study using a machine learning model[10] for the analysis of CFP demonstrated significant differences in retinal vascular networks between AD patients and controls. Tian et al.[11] developed a multi-stage machine learning pipeline using vessel maps extracted from CFP images and reported an AD classification result with an AUC of 0.87. Kim et al.[12] designed an end-to-end, attention-based convolutional neural network (CNN) model for CFP-based AD classification. Most recently, Cheung et al.[7] trained a deep learning algorithm on CFP datasets from multi-centers for AD detection, obtaining an AUC of 0.93. Xie et al.[8] used a two-stage CNN network to extract the vessels in OCTA images and proposed 12 parameters to study the association of retinal structure changes with AD. While these methods show promise, most of them rely on CFP images, there is a lack of end-to-end deep learning detection models using large-scale OCTA data. Additionally, the multi-dimensional aspect of OCTA data necessitates algorithms that can effectively handle and integrate multi-layered information. Precise feature extraction and mining from OCTA images require careful attention to the interrelationships and complementarity between different projection layers. Algorithms developed for CFP data are not well-suited for this task due to their design focus on the characteristics of CFP images. More importantly, most existing end-to-end deep learning approaches lack interpretability. This makes it challenging to provide clinicians and researchers with informative disease insights.

On the other hand, there has been very little work that is focused on the early detection of dementia. For example, early-onset AD (EOAD) - diagnosed in those under 65 years old - offers a promising avenue for identifying diagnostic indicators potentially applicable to sporadic AD. Mild cognitive impairment (MCI) can be an early stage of memory or cognitive ability loss and is an intermediate stage between cognitive normalcy and AD dementia, with a high likelihood of regression to AD dementia. Cheung et al.[7] designed an AI model to distinguish AD patients from controls using CFP and achieved promising detection results with an AUC of 0.93. However, it is not clear whether their approach can be used to detect EOAD and MCI. Several studies observed that EOAD and MCI patients exhibit significant microvascular morphological changes in the deep vascular complex, but CFP lacks sufficient resolution to capture the subtle microvascular changes associated with early pathology. In consequence, A tool that can detect dementia such as EOAD or MCI will lead to enormous benefits to patients when treatments can be more effective against dementia.

In this work, we propose a novel interpretable graph-based deep learning model named Eye-AD, to identify EOAD and MCI individuals through the unique characteristics of OCTA images. This tailored model explores intra- and inter-instance relationships across diverse retinal and choroidal layers, leveraging the unique capabilities of OCTA data. We combine this innovative AI approach with conventional biomarker analysis methods. This enables a comprehensive, interpretable evaluation of the model's decision-making and reinforces its credibility. Such a combination offers further insight into the impact of dementia on retinal and choroidal vasculature. The workflow of our study is illustrated in Fig. 1. Experimental results confirm that Eye-AD has outperformed other approaches in identifying EOAD and MCI from healthy controls. The results from visualization and interpretability analysis show that Eye-AD has demonstrated decision-making patterns consistent with prior clinical findings. Our Eye-AD may therefore provide a useful tool for quick identification and rapid screening of dementia-related disease in large populations.

Results

Demographic characteristics

The OCTA images used in this study are from a multi-center case-control study that involved four centers for EOAD and MCI detection. The inclusion and exclusion criteria for participants are shown in Fig. 2. Details of the study data are described in the "Methods" section.

The demographic details of the cohort participants are presented in Table 1. Our final data for EOAD detection includes 1192 eyes from the 870 participants. There is a marginally significant age difference between EOAD and controls ($p = 0.049$), with EOAD patients being slightly older on average. The EOAD patients have a higher prevalence of hypertension and diabetes ($p < 0.001$), as well as fewer years of education ($p = 0.021$), compared to controls. The data for MCI detection include 725 eyes from 551 participants. The MCI patients were significantly older than the controls ($p = 0.002$). Additionally, MCI patients have a higher prevalence of hypertension ($p < 0.001$).

Performance of EOAD detection

We first evaluate the Eye-AD for EOAD detection over the internal dataset ROAD-I, and the quantitative results are shown in Table 2. Five state-of-the-art CNN-based methods and three GNN-based methods were selected for comparison. These methods all require multiple inputs for different scenarios, including early fusion[13], middle fusion[14], late fusion[15], MCC[16], MUCO[17], GCN[18], GAT[19], and UG-GAT[20]. We also selected two ViT-based methods for comparison, including ViT[21] and SwinV2-T[22].

As shown in Table 2, the performance of our Eye-AD on ROAD-I exceeds those of all the compared methods in terms of five evaluation metrics: accuracy of 0.8885, precision of 0.8862, F1-score of 0.8867, Kappa of 0.7018 and AUC of 0.9355. Notably, our method achieves a statistically significant improvement on AUC over all the other methods, when evaluated by the pairwise Delong's test. We can also see that CNN-based and GNN-based approaches perform better than ViT-based methods for our task, given the limited dataset size. ViT models have a very large number of parameters and require substantial training data to reach peak performance. With our current dataset scale, ViT was prone to overfitting and achieved inferior results compared to CNNs. The MCC method exhibits poor performance compared with other multi-input methods, which may be due to the fact that the method is specifically designed for multi-weight magnetic resonance imaging (MRI). The structures of the different MRI inputs are similar, while the structures of the multiple OCTA *en face* inputs vary widely, which makes MCC perform poorly in this task. In addition, the GNN-based methods achieve better performance than those employing a CNN. This is likely to be because GNNs are better at modeling and mining relationships between regions and instances, and exploiting the correlation and complementarity between different *en face* images for classification.

The generalization performance of the Eye-AD model is validated on the external dataset ROAD-II. The quantitative results are shown in Table 2. Compared to the internal dataset ROAD-I, while the overall performance of all the compared methods on ROAD-II is lower, as expected, the proposed Eye-AD achieves again the best performance, outperforming all competitors in terms of accuracy (0.8176), precision (0.8429), F1-score (0.8291), kappa (0.5865) and AUC (0.9007). This indicates that the proposed method can effectively extract discriminative features for reliable EOAD detection and that the model is stable over the datasets captured by different hospitals.

Performance of MCI detection

We also train and evaluate the Eye-AD for the MCI detection: the results are shown in Table 2. The Eye-AD achieves the highest accuracy (0.8487), precision (0.8506), F1-score (0.8410), Kappa (0.6229), and AUC (0.8630) scores when compared with the other methods. The images in ROAD and ROMCI were acquired using different OCTA systems. This allows us to evaluate the generalization performance of the model under various tasks and machine settings. In addition, ROMCI utilizes a smaller amount of data compared with ROAD, which makes the model training more susceptible to overfitting problems.

The overall performances of all the methods on ROMCI-II are worse than those on ROMCI-I, while the proposed Eye-AD achieving the best performance in terms of accuracy (0.8444), precision (0.8392), F1-score (0.8339), kappa (0.5489) and AUC (0.8037). The GNN-based methods achieve better performance than the CNN-based ones. This implies that consideration of the relationships between different regions and instances is

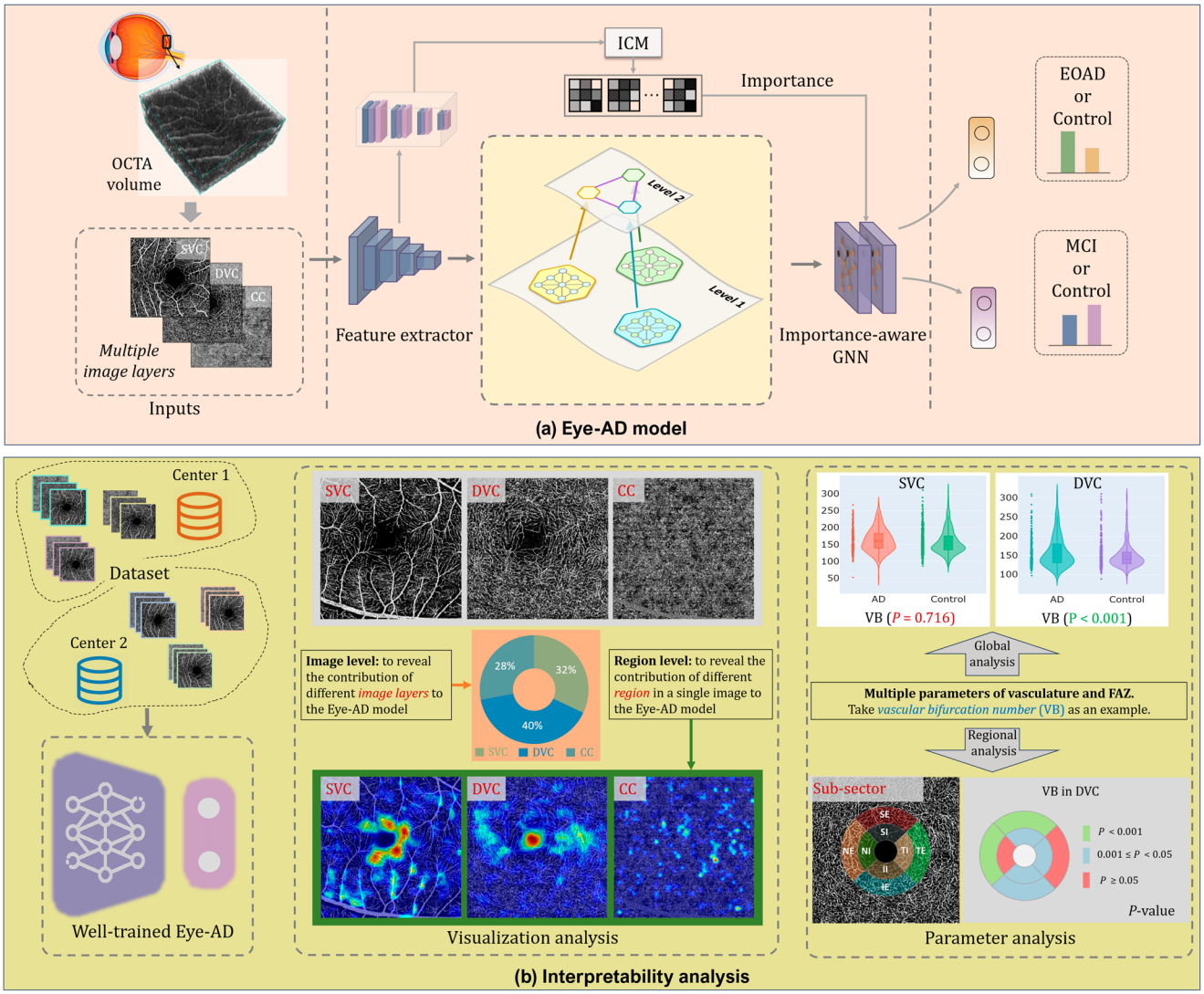


Fig. 1 | A schematic illustration of the Eye-AD model and the model's interpretability analysis. a Three *en face* images of a retina are fed into a feature extractor to obtain initial features for constructing a multilevel graph. Then an importance-aware graph neural network (GNN) module is used to obtain the prediction results. **b** We conducted visualization and parameter analysis for the trained Eye-AD, and compared the significant patterns (image- and pixel-level) learned by Eye-AD, with the conventional statistical analysis of handcrafted features (global and regional). Multiple global vascular and foveal avascular zone (FAZ) parameters were analyzed to compare the importance of different *en face*. A sub-sector parametric analysis for the vasculature surrounding the FAZ was also performed to determine if there were significant differences between the EOAD/MCI and control groups in these regions.

crucial for MCI detection. The results validate the effectiveness and generalization capability of our model for the task of MCI detection over the datasets captured by different hospitals.

Interpretability analysis

To understand the decision-making process of the Eye-AD model and identify discernible patterns in its ability to detect AD, we perform extensive visualization and statistical analyses. We compare Eye-AD's decision patterns with those obtained through conventional parameter analysis[8]. Through these comparisons, we aim to evaluate the consistency of the Eye-AD's decision-making process with the findings of previous studies and to determine whether it offers any significant advantages over conventional parametric statistical approaches. Specifically, We first train the Eye-AD on the ROAD dataset and obtain the importance distribution at the image level and pixel level for each OCTA case. We also compute the average importance distribution of all the samples, allowing us to identify general trends and patterns in the Eye-AD's decision-making process. The same preprocessing steps are also applied to the ROMCI dataset. Results are presented in Fig. 3.

In addition, we also extract eight different parameters characterizing both retinal microvasculature and foveal avascular zone (FAZ)[8], including vascular length density (VLD), vascular area density (VAD), vascular bifurcation number (VB), vascular fractal dimension (VFD), FAZ Area (FA), FAZ Circularity (FC), FAZ roundness (FR) and FAZ Solidity (FS). We then investigate the difference between EOAD/MCI and control groups. The results are shown in Fig. 4.

Regarding image-level importance, it can be observed that the DVC generally contributes more than either the SVC or the CC, obtaining an average importance score of 40% and 49% for EOAD and MCI cases, respectively, as shown in Fig. 3. The global parameter statistical analysis shows a similar pattern, as shown in Fig. 4. For the ROAD dataset, three of the eight parameters in SVC showed significant differences between the EOAD and control groups, while five parameters in the DVC showed significant differences. Parameter statistical analysis between the MCI and control groups further supports these conclusions, as illustrated in Fig. 4. The DVC showed more parameters with significant differences (two parameters) than the SVC (one parameter). This study demonstrates strong and consistent associations between the interpretability of Eye-AD and

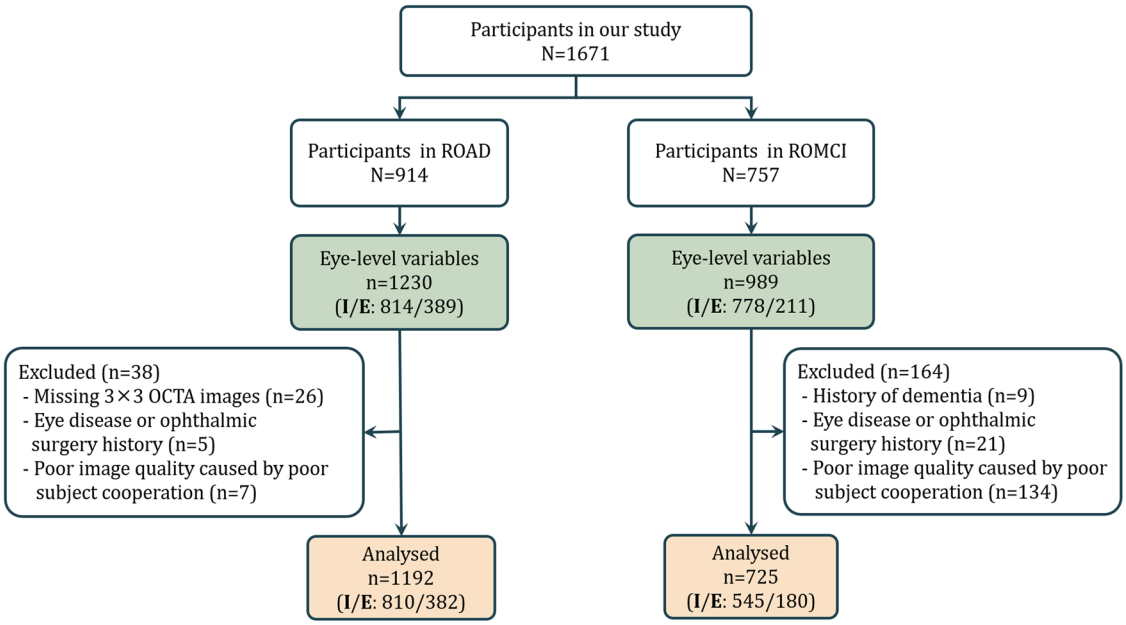


Fig. 2 | Patient flow diagram for the inclusion and exclusion of participants for the analysis. The study data consists of data for Retinal OCTA-based EOAD detection (ROAD-I, ROAD-II), and Retinal OCTA-based MCI detection (ROMCI-I and ROMCI-II). **I** and **E** denote the internal and external datasets, respectively.

Table 1 | Demographic data for all the subjects involved in this study

	Data for EOAD detection			Data for MCI detection		
	Control	EOAD	*P*-value	Control	MCI	*P*-value
Subject-level variables	631	239	-	455	96	-
Age, years (SD)	58.79 (7.34)	66.42 (8.58)	0.049	63 (15.48)	72.16 (9.73)	0.002
Number of females (%)	411 (65.1%)	155 (64.8%)	0.446	268 (58.9%)	59 (61.4%)	0.643
Hypertension (%)	84 (13.3%)	93 (38.9%)	<0.001	210 (46.1%)	27 (28.1)	<0.001
Diabetes (%)	35 (5.5%)	52 (21.7%)	<0.001	79 (17.3%)	11 (11.4%)	0.807
Education level, year (SD)	10.61 (4.88)	8.51 (4.61)	0.021	9.63 (4.04)	6.17 (3.16)	0.057
MMSE (SD)	28.75 (1.55)	18.81 (6.80)	<0.001	27.44 (3.13)	21.09 (3.90)	0.044
MoCA (SD)	24.79 (3.75)	13.52 (8.23)	<0.001	23.77 (4.61)	16.17 (2.74)	0.007

The *p*-value of age, education level, Mini-Mental State Examination (MMSE), and Montreal Cognitive Assessment (MoCA) were calculated using the *t*-test, and the *p*-values of gender, hypertension, and diabetes were calculated using the chi-square test.

feature analysis of retinal microvasculature and FAZ, indicating that the DVC plays a more important role in EOAD and MCI detection. Additionally, the most significant difference in the interpretable analyses of EOAD and MCI is that EOAD exhibits greater between-group differences than MCI in both the SVC and DVC layers. This implies that EOAD causes more substantial alterations to the fundus structure than MCI. The primary reason for this could be that although EOAD develops in individuals younger than 65 years of age, its symptoms can be as severe as other forms of AD, leading to a greater impact on the fundus structure. On the other hand, MCI represents an early stage of dementia with milder symptoms, resulting in less pronounced changes in the fundus. This suggests that the effect of dementia on the fundus structure becomes more evident as symptoms worsen. Furthermore, there are also similar patterns in the interpretable results of EOAD and MCI. For both conditions, the between-group differences in the DVC layer are larger than those in the SVC layer. This consistency indicates that certain changes in the deeper layers of the retina might be more sensitive indicators of early dementia stages, regardless of the specific type.

Figure 3 shows the attention maps of each *en face* image of both EOAD/MCI and control cases. It reveals that in the SVC, the areas of activation for control cases are mainly around the FAZ and large vessels; while in the EOAD cases, regions of small vessels between the larger ones are activated. In the DVC, the activation area is mainly concentrated in the FAZ and its surrounding vessels. Compared to control cases, the activation area in EOAD cases is larger. The degree of activation of the CC differs greatly between AD and control cases, with more activation areas in AD cases. The average activation maps of all the samples indicate that the FAZ and its surrounding vessels have the highest probability of being activated, suggesting that EOAD is most likely to cause changes in capillaries surrounding the FAZ. To confirm this, we perform a sub-sector parametric analysis for the microvasculature surrounding the FAZ, as presented in the bottom row in Fig. 4. We find that for the parameters of microvasculature in both the SVC and the DVC, several sub-sectors demonstrate significant differences between the EOAD and control groups, indicating that the microvasculature surrounding the FAZ undergoes detectable alterations in EOAD. The results of the ROMCI analysis are similar to those of ROAD, except that fewer parameters and regions are significantly different in ROMCI than in ROAD, as shown in Fig. 4.

Discussion

We proposed a novel deep learning-based model namely Eye-AD, by leveraging OCTA images for the detection of both EOAD and MCI. Our

Table 2 | The detection results of EOAD and MCI using different methods

Dataset	Method	Accuracy	Precision	F1-score	Kappa	AUC (*P*-value)
ROAD-I	ViT	0.6254	0.6431	0.6302	0.2376	0.6958 (<0.001)
	SwinV2-T	0.6615	0.6702	0.6645	0.2990	0.7278 (<0.001)
	Early fusion	0.8047	0.7924	0.7930	0.4361	0.7470 (<0.001)
	Middle fusion	0.8126	0.8013	0.7969	0.4399	0.7313 (<0.001)
	Late fusion	0.8179	0.8074	0.8053	0.4660	0.7761 (<0.001)
	MCC	0.7744	0.7524	0.7514	0.3109	0.6902 (<0.001)
	MUCO	0.8284	0.8256	0.8089	0.4688	0.8256 (<0.001)
	GCN	0.8545	0.8522	0.8438	0.5765	0.9135 (<0.001)
	GAT	0.8508	0.8482	0.8390	0.5627	0.9029 (<0.001)
	UG-GAT	0.8610	0.8578	0.8529	0.6036	0.9154 (<0.05)
	Eye-AD	**0.8885**	**0.8862**	**0.8867**	**0.7018**	**0.9355**
ROAD-II	ViT	0.6133	0.6337	0.6185	0.2168	0.6747 (<0.001)
	SwinV2-T	0.6024	0.6169	0.7930	0.4361	0.7470 (<0.001)
	Early fusion	0.6806	0.6744	0.6752	0.3125	0.6982 (<0.001)
	Middle fusion	0.7356	0.7318	0.7398	0.4315	0.7906 (<0.001)
	Late fusion	0.7853	0.8021	0.7991	0.5169	0.8525 (<0.001)
	MCC	0.7303	0.7373	0.7116	0.3909	0.7052 (<0.001)
	MUCO	0.6178	0.5942	0.6771	0.0940	0.8173 (<0.001)
	GCN	0.7696	0.7732	0.7598	0.4906	0.8197 (<0.001)
	GAT	0.7670	0.7921	0.7867	0.4682	0.8301 (<0.001)
	UG-GAT	0.7382	0.7628	0.7642	0.3971	0.7988 (<0.001)
	Eye-AD	**0.8167**	**0.8429**	**0.8291**	**0.5865**	**0.9007**
ROMCI-I	ViT	0.7242	0.7538	0.7370	0.1869	0.6349 (<0.001)
	SwinV2-T	0.7610	0.7320	0.7443	0.1156	0.6168 (<0.001)
	Early fusion	0.8051	0.8156	0.8098	0.3925	0.6773 (<0.001)
	Middle fusion	0.7830	0.8222	0.7970	0.3933	0.6769 (<0.001)
	Late fusion	0.8308	0.8236	0.8267	0.4186	0.7267 (<0.001)
	MCC	0.7904	0.7731	0.7803	0.2500	0.6742 (<0.001)
	MUCO	0.8051	0.7757	0.7843	0.2401	0.6593 (<0.001)
	GCN	0.8103	0.8056	0.8063	0.5438	0.8490 (<0.05)
	GAT	0.8240	0.8211	0.8166	0.5662	0.8579 (<0.05)
	UG-GAT	0.8209	0.8194	0.8114	0.5524	0.8558 (<0.05)
	Eye-AD	**0.8487**	**0.8506**	**0.8410**	**0.6229**	**0.8630**
ROMCI-II	ViT	0.7333	0.7436	0.7383	0.0769	0.6390 (<0.001)
	SwinV2-T	0.7250	0.7406	0.7324	0.0660	0.6385 (<0.001)
	Early fusion	0.5583	0.7868	0.6131	0.1263	0.5717 (<0.001)
	Middle fusion	0.5416	0.7464	0.5997	0.0517	0.5040 (<0.001)
	Late fusion	0.5833	0.7693	0.6363	0.1071	0.5080 (<0.001)
	MCC	0.7333	0.7333	0.7333	0.0400	0.5585 (<0.001)
	MUCO	0.7666	0.7935	0.7782	0.2500	0.6830 (<0.001)
	GCN	0.7851	0.7735	0.7768	0.4018	0.7574 (<0.001)
	GAT	0.7851	0.7796	0.7819	0.4246	0.7637 (<0.001)
	UG-GAT	0.7703	0.7643	0.7669	0.3850	0.7760 (<0.001)
	Eye-AD	**0.8444**	**0.8392**	**0.8339**	**0.5489**	**0.8037**

The best performance is highlighted in boldface. *P*-value is calculated using Delong's test.

method has demonstrated its potential to bridge the gap in accurate identification and rapid screening of EOAD/MCI for large-scale populations. Furthermore, it provides a trustworthy interpretable result, leveraging the rich spatial information in OCTA data for improved analysis. Previous works have reported a correlation between biomarkers found in OCTA images and AD[8,23], thus the utilization of retinal OCTA images may prove advantageous in expediting AD screening. Retinal imaging is less expensive, simpler, and faster, and has greater feasibility for smaller hospitals or community screening programs, when compared with a conventional AD diagnosis protocol, which relies on a series of complicated and expensive

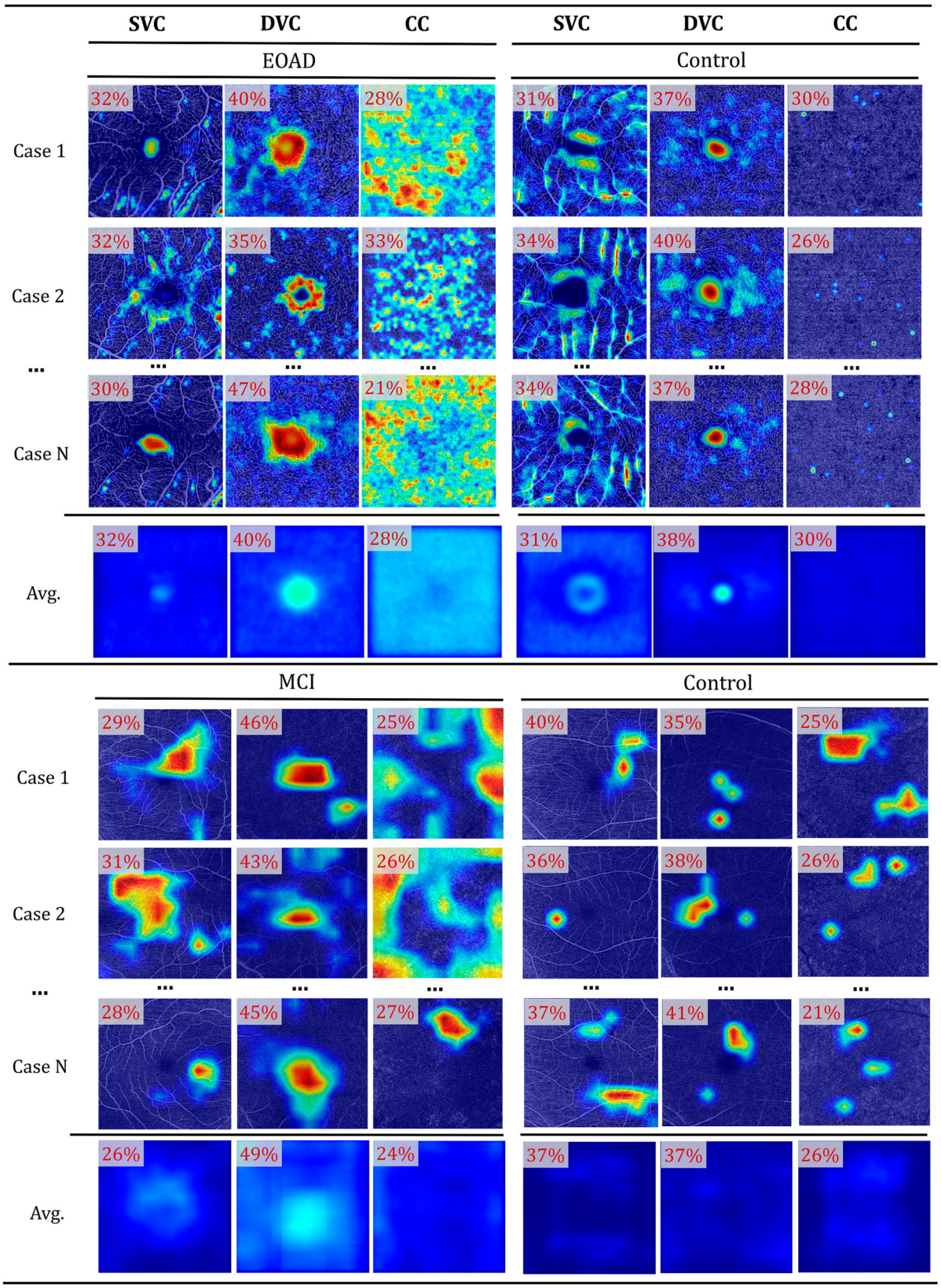


Fig. 3 | Visualization and interpretability analysis of the Eye-AD model. We present both image and region-level importance for different inputs. The image-level importance is presented as a percentage (in red) at the top-left of each image. The last row shows the average importance distribution of all cases, where we may identify general trends and patterns in the Eye-AD's decision-making process.

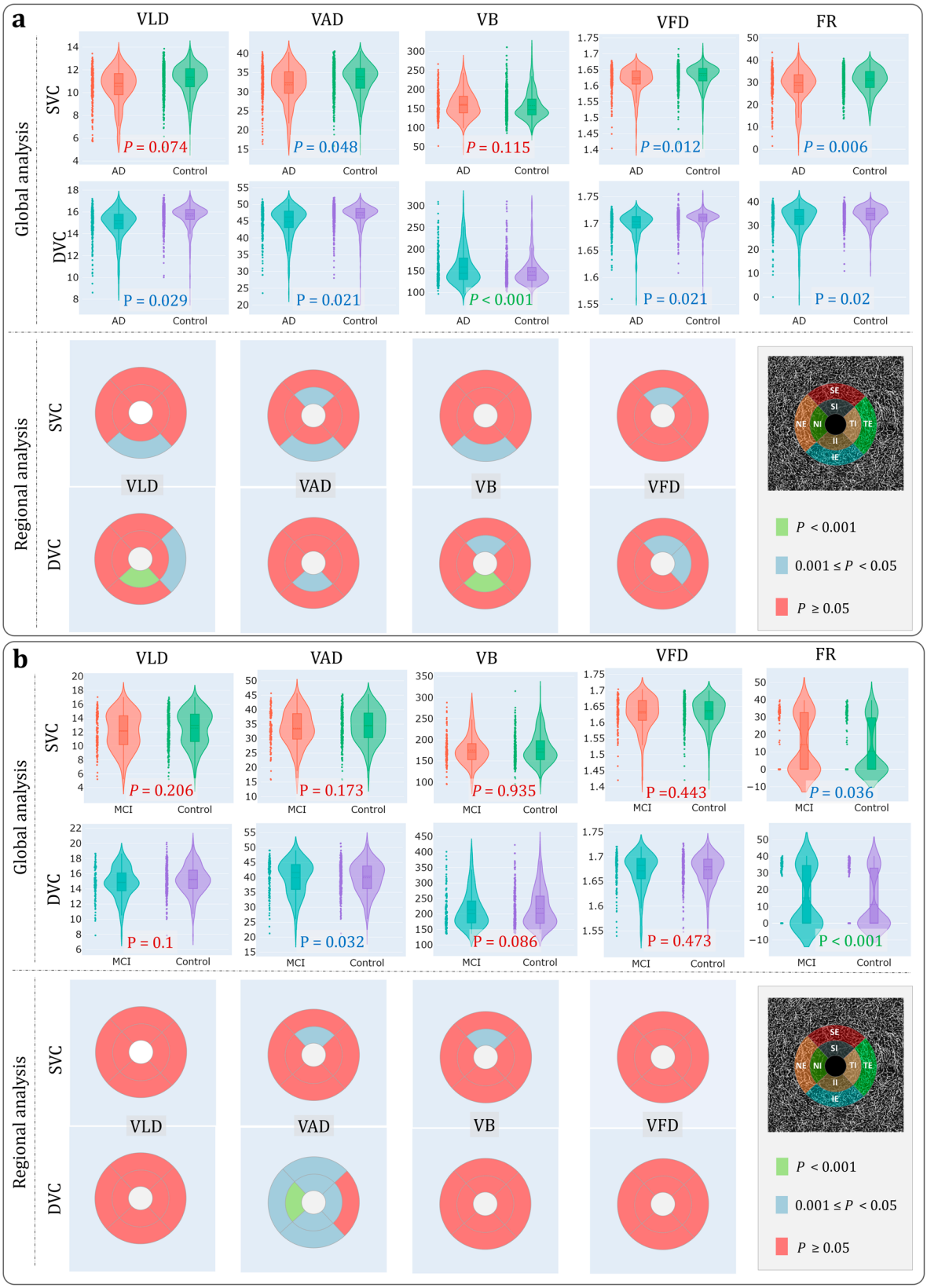


Fig. 4 | The results of the global and regional statistical analysis of the parameters, adjusted for factors such as age, gender, hypertension, diabetes, and education level, were obtained utilizing Generalized Estimation Equation. a shows the results of comparison between EOAD and the control group: **b** illustrates the result between MCI and the control group.

procedures, such as cognitive assessments, neuroimaging, and cerebrospinal fluid biomarker evidence. The distinct layers of the retina contain different types of information, together providing a relatively complete picture of retinal changes. The SVC consists of a mixture of large and small vessels and has a centripetally branching pattern that terminates at the capillary ring around the FAZ. The DVC consists of thin layers of capillaries arranged in lobular patterns without directional preference. The CC contains a wealth of vascular information on the choroidal circulation[24]. The proposed Eye-AD model incorporates a novel multilevel graph method to model and utilize the intra- and inter-instance relationship of different retinal layers and fully exploits and fuses this layer information for the improvement of the detection performance in AD. Rapid identification of potential AD patients in community settings would enable timely patient referral for further examination, and early treatments for the best outcomes. To this end, it is feasible and more valuable to identify patients with prodromal AD or MCI. In this work, we evaluated our Eye-AD on the MCI dataset. Our results suggest that the MCI detection shows satisfactory results (internal: AUC = 0.8630; external: AUC = 0.8037). However, compared with the results of the EOAD detection, MCI detection is more challenging, possibly because detectable changes in the retinal microvascular system accumulate only slowly as dementia develops. Future research could improve the Eye-AD model's sensitivity and specificity in MCI detection: combine biomarkers extracted from retina scanning with other noninvasive, accessible modalities such as blood tests or cognitive tests as a potential solution.

Eye-AD provides a reliable and interpretable detection result. Previous studies have explored the retinal change in AD patients by analyzing a variety of potentially relevant microvasculature parameters[5,6,8,25–27]. The conclusions drawn from these studies exhibit poor consistency and even contradict each other. For example, several studies[8,25,26] suggested that AD patients exhibited a significant decrease in the microvasculature in the SVC compared to control, while no significant difference was observed in the DVC between the two groups. Contrarily, other reports[5,6] found that the DVC showed a significant difference between AD and control groups, while no significant difference was found in the SVC. The conflicting conclusions of previous studies may be due to any of several causes. Firstly, the sample sizes employed in most studies were relatively small, with the majority utilizing fewer than 100 participants: this may have rendered the results susceptible to inherent biases within the datasets. Furthermore, these studies relied on handcrafted features, requiring the prior extraction of the retinal structures before the calculation of the relevant parameters. The chosen parameters may not be sufficient for a comprehensive assessment of changes in the microvasculature and may be affected by variations in the precision of the prior structure extraction, leading to inconsistent results between studies. Our study involved 1671 participants, forming by far the largest OCTA dataset in the investigation of OCTA's predictive capability for EOAD and MCI. Analysis of the interpretability of Eye-AD indicates that retinal changes associated with EOAD/MCI impact both the SVC and DVC, with potentially greater effects on the DVC. This is evidenced by the differing contributions of each *en face* to the model's predictions. To further strengthen the credibility of our interpretability results, we conducted a statistical analysis, which revealed that more parameters of the DVC were significantly associated with EOAD/MCI than the SVC. The interpretative results of Eye-AD suggest that the FAZ and its surrounding microvasculature are more sensitive biomarkers for EOAD diagnosis. This finding is further supported by the statistical analysis of the sub-sector parameters of the microvasculature. Regarding image-level importance, our findings show that the DVC consistently exhibited higher importance scores (40% for EOAD and 49% for MCI) compared to the SVC and CC. This observation aligns with our parameter analysis, where more parameters in the DVC showed significant differences between the EOAD/MCI and control groups than those in the SVC. In terms of region-level importance, attention maps reveal that EOAD cases display more activation in smaller vessels within the SVC and larger activation areas in the DVC around the FAZ compared to control cases. Additionally, a sub-sector analysis of the microvasculature around the FAZ revealed significant differences between the EOAD and control groups, particularly in the DVC, supporting our model's interpretive results. These results underscore the strong and consistent association between the interpretive outputs of Eye-AD and traditional morphological parameter analysis, highlighting the DVC's crucial role in detecting EOAD and MCI.

Our study provides an aid to understanding the effects of AD on the retina. The interpretable results obtained from Eye-AD offer valuable novel insights. Our findings suggest that changes in the DVC were more significantly associated with EOAD/MCI compared to changes in either the SVC or the CC. The DVC is made up of capillaries and has a thin cross-section, which makes it more sensitive to the progression of the disease cascade[28]. Importantly, the DVC is responsible for the venular circulation in the retina[29]. Recent reports[30–32] have highlighted the role of cerebral venular degeneration in AD; given the retinal microvasculature reflects the cerebral microcirculation, it is plausible to suggest that these retinal changes reflect underlying cerebral changes. Parameter statistics on both the ROAD and ROMCI datasets reveal fewer parameters with significant differences between MCI and control groups than between EOAD and controls. Specifically, five parameters in DVC (VLD, VAD, VB, VFD, and FR) show significant differences between EOAD and control, while only two parameters (VAD and FR) exhibit a significant difference between MCI and controls. These findings suggest that the vascular alterations in the retina may be subtle during the early stages of cognitive decline, and the deterioration of the retinal microvasculature increases with the progression of dementia. In addition, the ablation experiments demonstrate that by harnessing complementary information of SVC, DVC, and CC, we may achieve more accurate detection of both EOAD and MCI. Moreover, our study identifies a critical association between EOAD and the choroid, as indicated by the differences in the pixel-level importance distribution of the CC between EOAD cases and controls. Although earlier studies were limited by image resolution and did not fully explore the role of the choroid in AD, our study reveals that the choroid could serve as a potential diagnostic biomarker for the disease, as previously suggested[33]. With advances in OCTA technology, future studies must place greater emphasis on investigating the role of the choroid in facilitating rapid screening for AD. Such research has the potential to enable early detection of AD, providing significant benefits in the field of AD diagnosis and intervention.

The proposed Eye-AD model not only provides a solution for the rapid detection and screening of EOAD and MCI but also offers a proof-of-concept, interpretable deep learning model for understanding the association between biomarkers in the retina and dementia. However, it is important to acknowledge the limitations of our study. First, although our dataset is larger than previous OCTA studies on EOAD and MCI, the sample size is still relatively small for deep learning. In particular, our dataset lacked diversity in ethnic groups. As our study was limited to Chinese populations, its generalizability to other ethnicities is unclear. The difficulty of obtaining retinal imaging from EOAD patients currently limits sample sizes. Additionally, our exclusion of patients unable to cooperate could further bias results. Because recognizing EOAD from MCI is a valuable task, we also performed a multi-label recognition task on samples from EOAD, MCI, and controls using our model and dataset. The results were as follows: Accuracy: 0.8758 and AUC: 0.9320. However, as the reviewer pointed out, detecting MCI and EOAD is a more difficult task and we believe that these results obtained may be largely due to the heterogeneity of the data, specifically the different models of OCTA devices used to detect the MCI and EOAD cohorts. This discrepancy may allow the model to learn disease-irrelevant information. In future work, we will establish more homogeneous cohorts to further investigate the effects of different stages of dementia on the fundus. Second, the current model focuses solely on retinal OCTA images, which, while valuable, may not capture the complete picture of AD pathology. Integrating multimodal data, such as blood biomarkers, genetic information, and cognitive assessments, could enhance the model's predictive power and generalizability. Third, the real-life deployment of Eye-AD in clinical settings faces several challenges. These include the standardization of imaging protocols across different devices and institutions,

ensuring data privacy and security, and obtaining regulatory approvals. Addressing these issues requires collaborative efforts among researchers, clinicians, and policymakers. Lastly, longitudinal studies are necessary to validate the utility of retinal biomarkers for predicting the progression of cognitive decline over time. Such studies would provide insights into the temporal dynamics of retinal changes in AD and MCI and could inform the development of intervention strategies aimed at halting or slowing disease progression. In short, larger longitudinal studies are needed to assess the utility of retinal biomarkers for diverse populations. Future work should focus on standardizing retinal imaging and cognitive assessments for early AD detection.

In conclusion, we have proposed a novel multilevel graph-based deep learning model, Eye-AD, to detect EOAD and MCI through the unique characteristics of OCTA images. This model allows for significant exploration and understanding of both intra- and inter-instance relationships across diverse retinal layers. Eye-AD achieves state-of-the-art performance in distinguishing AD patients from healthy controls. The model also shows its generalizability when validated on an external dataset. Our interpretability analysis reveals that Eye-AD can identify meaningful patterns related to EOAD/MCI. We find that changes in the DVC are more significantly associated with EOAD/MCI compared to those in either the SVC or the CC. We also find that vascular alterations in the retina may be subtle during the early stages of cognitive decline, and the deterioration of the retinal microvasculature increases with the progression of dementia. Our findings provide evidence that retinal OCTA imaging coupled with artificial intelligence may serve as a rapid and noninvasive approach for EOAD and MCI detection and screening. Although our study has made progress in understanding the association between retinal changes and EOAD/MCI, further research is still needed. By combining multiple data modalities and expanding datasets, the proposed model has the potential for use in large-scale community screening and longitudinal monitoring of AD.

Methods

Study data

The study data consists of data for Retinal OCTA-based EOAD detection (ROAD-I, ROAD-II), and Retinal OCTA-based MCI detection (ROMCI-I and ROMCI-II). This study was approved by the Ethics Committee of the Laboratory of Advanced Theranostic Materials and Technology, Ningbo Institute of Materials Technology and Engineering, Chinese Academy of Sciences, and West China Hospital of Sichuan University Ethics Committee (Ethics number 2020[922]). All the OCTA images acquired are centered on the fovea, and the *en face* images of the superficial vascular complex (SVC), deep vascular complex (DVC), and choriocapillaris (CC) are used for the analysis[24]. Images of any subject will only be used either in the training or test sets to avoid information leakage. The OCTA imaging device includes built-in correction algorithms that remove artifacts from the images. To address potential data bias, we ensured that the same models of collection equipment were used for the same tasks. Additionally, we applied the same data inclusion and exclusion criteria across different institutions when building the cohort.

Retinal OCTA data of the EOAD study (ROAD) contains an internal (ROAD-I) and an external (ROAD-II) subset. ROAD-I was used for the development and internal test of the proposed Eye-AD model for the detection of EOAD, and ROAD-II for its external test only. **ROAD-I** comprises 810 OCTA volumes from the Affiliated People's Hospital of Ningbo University, China, including 199 OCTA volumes of EOAD subjects and 611 OCTA volumes of controls (one volume per eye per patient). EOAD patients met the National Institute on Aging and Alzheimer's Association (NIA-AA) criteria[34] and were diagnosed before the age of 65. The inclusion criteria are as follows: have the capacity to consent, or, if cognitively impaired, have a legally authorized representative who can provide consent; can complete OCTA imaging; and be willing and able to complete all study procedures. The exclusion criteria are as follows: not be pregnant or lactating; not have a history of other brain disorders (both neurologic and psychiatric, except for seizures thought to be related to EOAD or headaches); have not participated in therapeutic trials; have no history of substance abuse, or suicidal behaviors in the past 12 months. Individuals with MRI evidence of infection, focal lesions such as strokes, multiple or strategic lacunae, and/or space-occupying lesions are also excluded. Controls are subjects without dementia or cognitive impairment, defined as no objective impairment on the neuropsychological assessment. These participants do not have any neurological diseases. We use a five-fold cross-validation approach to train and evaluate the models. Each individual fold of the data (20%) is used in turn as the internal testing set, whilst the other 4 folds (80%) are set as the training set. **ROAD-II** is an independent dataset acquired from a different hospital (the West China Hospital, Sichuan University, Chengdu, China), with a view to further evaluation of the generalization ability of our Eye-AD model. The inclusion/exclusion criteria were the same as those in ROAD-I. This dataset contained 382 OCTA volumes, including 150 OCTA volumes of EOAD subjects and 232 OCTA volumes of controls.

Retinal OCTA data of MCI study (ROMCI) also comprises an internal (ROMCI-I) and an external subset (ROMCI-II) for the validation of the proposed model for the detection of MCI. **ROMCI-I** participants were selected from the Second Affiliated Hospital of Zhejiang University, Hangzhou, China. Initially, MCI participants were evaluated and diagnosed clinically by an experienced neurologist, following the diagnostic guidelines and recommendations of the Petersen criteria[35]. Clinical history, cognitive testing, and neuroimaging were reviewed for diagnostic accuracy by an experienced neurologist with a specialization in memory disorders. Exclusion criteria were: (1) significant sensory impairment, e.g., verbal or hearing impairment, etc; and (2) presence of major depression or other psychiatric disorders. A total of 545 OCTA volumes were used to construct the dataset, including 104 OCTA volumes of subjects of MCI and 441 OCTA volumes of control. We use a five-fold cross-validation approach for the development and evaluation of the performance of the model. **ROMCI-II** subjects are selected from the Affiliated People's Hospital of Ningbo University, Ningbo, China. Their inclusion criteria for MCI patients and controls are the same as those used in ROMCI-I. This second dataset contained 180 OCTA volumes, including 35 OCTA volumes of MCI subjects and 145 OCTA volumes of control.

Model architecture

Eye-AD includes two main components: CNN for embedding extraction, and multilevel GNN for final prediction, as illustrated in Fig. 5. We apply ResNet18[36] as the encoders, which retain the first four blocks. The input *en face* image $X^k \in X$ of size $h \times w$ is partitioned into $n \times n$ non-overlapping regions. $k \in [1, K]$, and K is the total number of *en face* images of a patient.

We devise a hierarchical multilevel graph, consisting of instance-level and subject-level graphs to model the inter- and intra-instance relationship of different *en face* angiograms, as shown in Fig. 6. We construct the graph set from bottom-level (i.e., level 1 in Fig. 6) to top-level (i.e., level 2 in Fig. 6). An instance-level graph is constructed based on an *en face* image, and a subject-level graph is constructed based on the instance-level graph. Each region is considered as a node of an instance-level graph in which node features are obtained from the last layer of the encoder, which is the output of the fourth block of ResNet18. Undirected edges are assigned between node pairs according to the adjacency between regions in the image. We describe $\mathbb{G} = (\mathbb{V}, \mathbb{E})$ as a instance-level graph, with $|\mathbb{V}| = n^2$ nodes and $|\mathbb{E}|$ edges. $|\cdot|$ represents the cardinality of a set. For each $v_i \in \mathbb{V}$, h_i is the corresponding F-dimensional feature vector. Let $H \in \mathbb{R}^{n^2 \times F}$ be the node feature matrix, and $A^{adj} \in \mathbb{R}^{n^2 \times n^2}$ be the sparse adjacency matrix, encoding the edge connections between the nodes. The instance-level graph is fed into the importance-aware graph attention layer to obtain the aggregated embeddings, which are used to construct subject-level graphs. The subject-level graph includes the relationship among different *en face* images. It is represented as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$. The subject-level graph is a fully connected graph, where each node represents an *en face* image instance, and there is an

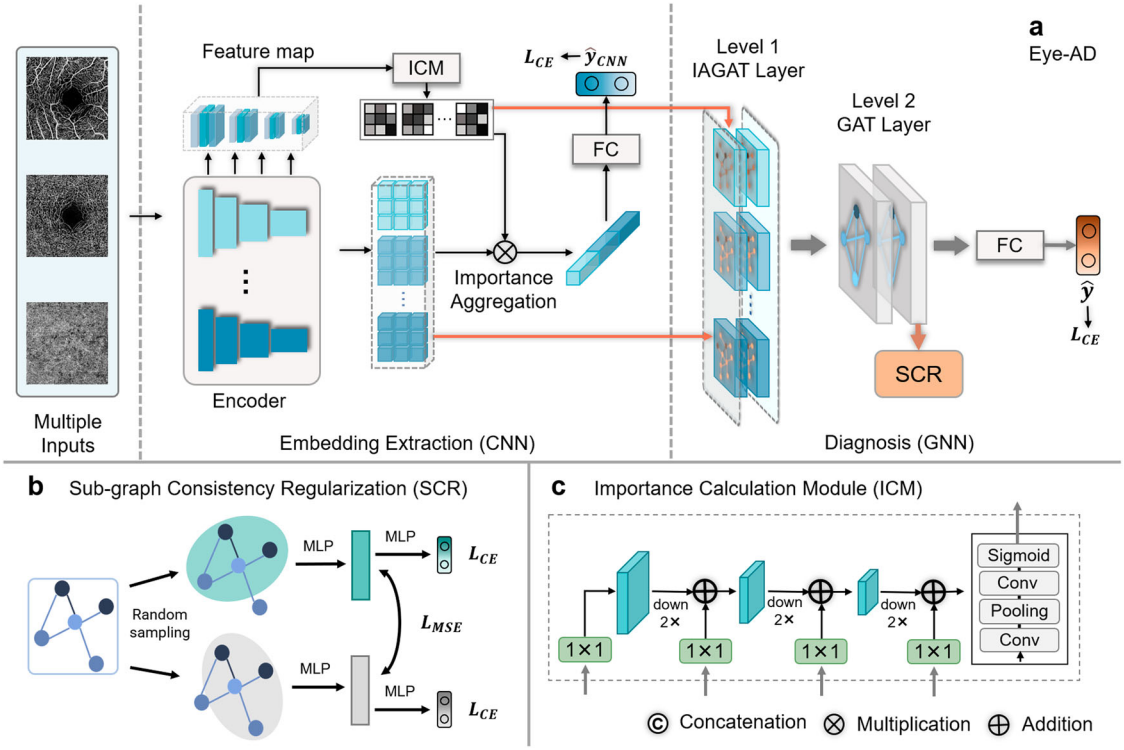
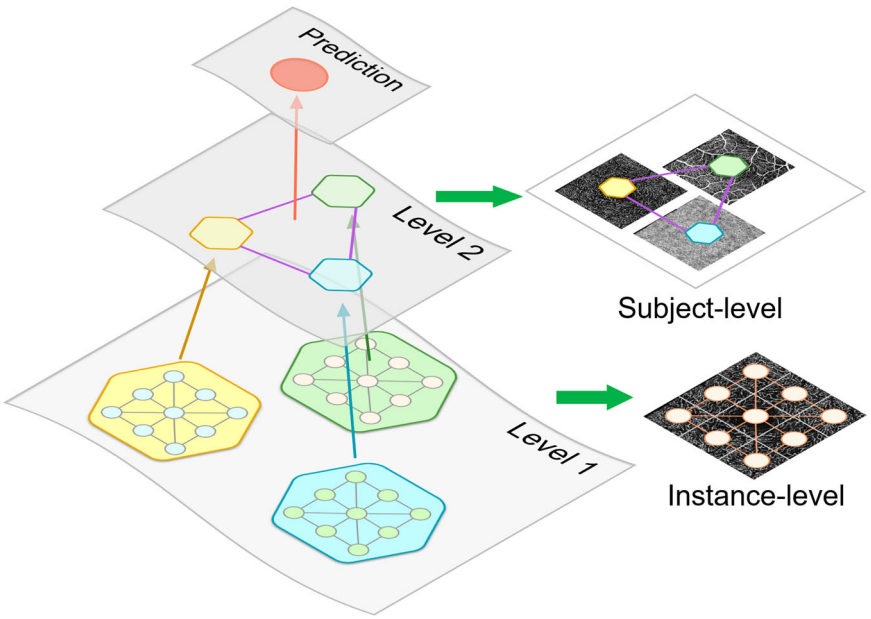


Fig. 5 | The design of Eye-AD. a Architecture illustration of Eye-AD, consisting of an Embedding Extraction module and a Diagnosis module. **b** Details regarding the regularization of sub-graph consistency. **c** The neural network structure of the Importance Calculation Module (ICM).

Fig. 6 | Construction of multilevel graph. Level 1 is the instance-level graph, and level 2 is the subject-level graph. An instance-level graph is constructed based on an *en face* image, and a subject-level graph is constructed based on the instance-level graph.



edge between every two instances, the k_{th} node feature of $\mathbb{G}$ is the aggregated embedding of $\mathbb{G}^k$.

The ICM consists of multiple convolutional layers with down-sampling operations, a final convolutional layer with a pooling, and a sigmoid activation layer, used to map the features into the form of a probability map with a size of $n \times n \times K$ that can be treated as the importance. The ICM takes the features of the different layers from the encoders as input, and the corresponding output is the obtained importance for each *en face* image. As shown in Fig. 5, the outputs of the ICM are first applied for importance aggregation during the CNN module pretraining process. Specifically, the different importance maps are separately multiplied with the corresponding encoded features, and the results are then fed into a fully connected layer for classification. This operation ensures that the ICM can be trained during the CNN pre-training stage. Additionally, the importance matrix is then used during the graph convolution operation, as shown in Equations (1)–(4).

Specifically, given an instance-level graph $\mathbb{G}^k = (\mathbb{V}, \mathbb{E})$ with a set of node features $h = \left\{\vec{h}_1, \vec{h}_2, ..., \vec{h}_{n^2}\right\}, \vec{h}_i \in \mathbb{R}^F$, the IAGAT layer updates the node features and obtains the new embedding, i.e., $h' = \left\{\vec{h}'_1, \vec{h}'_2, ..., \vec{h}'_{n^2}\right\}$,

$\vec{h}_i' \in \mathbb{R}^{F'}$, and F' is the dimension of the updated node feature. The details for the update are as follows. Firstly, a linear transformation parameterized by a shared weight matrix $W \in \mathbb{R}^{F' \times F}$ is employed for each node. Then we calculate the attention coefficients e_{ij} using the self-attention operation:

$$e_{ij} = LeakyReLU(\vec{a}^T[W\vec{h}_i \parallel W\vec{h}_j]), \quad (1)$$

where $e_{ij} \in E$ indicates the attention value of node j to node i, and $E \in \mathbb{R}^{n^2 \times n^2}$ is the attention coefficient matrix. $\parallel$ is the concatenation operation: $\vec{a}^T \in \mathbb{R}^{2F'}$ is a learnable weight vector implemented by a fully connected layer, followed by the LeakyReLU activation (with a negative input slope $\alpha = 0.2$).

Considering the importance of each region, i.e., the node of the instance-level graph, during the disease diagnosis, we re-weight the attention coefficients using the importance value obtained by the ICM. We denote the importance matrix of $\mathbb{G}^k$ as M^k, whose size is $n^2 \times 1$. The re-weighting operation is defined as:

$$E' = E(Diag(M^k)), \quad (2)$$

where $Diag(\cdot)$ is the diagonalization operation. Then we normalize the coefficient e'_{ij} using the softmax function to make it comparable across different nodes:

$$\alpha_{ij} = Softmax_j(e'_{ij}) = \frac{exp(e'_{ij})}{\sum_{l \in N_i} exp(e'_{il})}, \quad (3)$$

where N_i denotes the neighborhood of node i in the graph, showing that only e'_{ij} for neighboring node $j \in N_i$ is considered during the update of the node feature to avoid involving irrelevant nodes. Finally, we use the normalized attention coefficients α_{ij} to calculate a weighted sum of the involved node features to obtain the updated features for each node:

$$\vec{h}_i' = ELU\left(\sum_{j \in N_i} \alpha_{ij} W \vec{h}_j\right), \quad (4)$$

where ELU represents the exponential linear unit (ELU) nonlinearity. In this way, the proposed IAGAT layer can extract instance embedding by cascading the updated node features of $\mathbb{G}^k$, which are used to construct the subject-level graph $\mathcal{G}$. Then, we employ the GAT layer[19], a popular GNN method with an attention mechanism, as the extractor to obtain the subject features. Finally, the subject feature is fed into a fully connected layer followed by a softmax activation layer for the final subject classification.

To improve generalization and capture the subtle differences between different subgraphs of the same subject, we design a sub-graph consistency regularization (SCR)-based loss function. The SCR randomly samples two subgraphs from the same subject-level graph and minimizes their squared L_2 distance after passing through a two-layer MLP network. Specifically, for each input, a graph consisting of different projection layers is generated with n nodes. During training, two graphs consisting of different n-1 nodes, i.e., subgraphs, are sampled from n nodes through random sampling. We aim for the features related to disease prediction represented by these two subgraphs to be consistent, achieved through the consistency constraints of the SCR loss function. This SCR encourages the model to learn more robust and discriminative representations that can capture the subtle differences between different subgraphs of the same subject, thus improving the generalization ability of the model. This regularization term is particularly effective for small datasets, where overfitting is a major concern, and has been shown to improve classification performance in our experiments.

Training strategy

Eye-AD is an end-to-end framework. We have found in practice that a proper training strategy is critical to obtaining a well-trained model. To extract more discriminative disease-related embeddings and speed up convergence, we add an intermediate objective $\mathcal{L}_{CE}(\hat{y}_{CNN}, y)$ for the CNN component, as shown in Fig. 5. The total loss in Eye-AD includes CNN loss $\mathcal{L}_{CE}(\hat{y}_{CNN}, y)$, consistency loss $\mathcal{L}_{SCR}$ and final prediction loss $\mathcal{L}_{CE}(\hat{y}, y)$. Since CNN and GNN have different convergence properties, simultaneously training CNN and GNN by minimizing the sum of these three losses may lead the model to fall into a local optimum or even cause gradient vanishing. Therefore, we adopt an alternating strategy to update the model. To accelerate convergence, we first pre-train the CNN and then perform the joint alternating training. In the first step, we only optimize the CNN by minimizing $\mathcal{L}_{CE}(\hat{y}_{CNN}, y)$. Then we update the parameters of the GNN model by minimizing the sum of $\mathcal{L}_{CE}(\hat{y}, y)$ and $\mathcal{L}_{SCR}$. Finally, CNN is optimized by minimizing the final prediction loss $\mathcal{L}_{CE}(\hat{y}, y)$. The above alternating process is repeated with training iterations. We summarize the training strategy in Supplementary Algorithm 1. The effectiveness of the proposed training strategy is shown in Fig. 7.

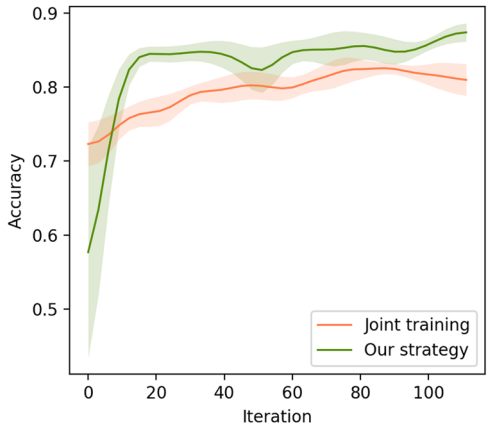


Fig. 7 | Effectiveness of the proposed training strategy for Eye-AD. By using the proposed training strategy, the model can more easily converge to the global optimal solution, thereby achieving better performance.

Implementation details

All the experiments were implemented using PyTorch and run on a workstation with 4 NVIDIA GeForce 3090 GPUs. We used ResNet18 as the backbone of the embedding extractor in the CNN branch. The level-1 graph block consists of an IAGAT layer and a GAT layer. The level-2 graph block consists of a GAT layer and a fully connected layer, followed by a softmax activation layer for subject classification. The IAGAT and GAT layers in the GNN branch both have two attention heads computing 512-dimensional features.

An Adam optimizer is used to optimize the model, with a weight decay of 0.0005 and a batch size of 16. The initial learning rates of CNN and GNN are set to 0.0001 and 0.00001 respectively, which are gradually decayed to zero using a cosine annealing scheduler. The trade-off factor λ is set to 2. The iteration numbers for different steps, i.e., K_1, K_2, and K_3, are set to 1, 4, and 1 respectively. In addition, data augmentation is employed including random horizontal and vertical flips to expand the training dataset. For the trained model, the image-level importance is calculated as in ref. 37, while the pixel level is calculated using Grad-cam[38].

We calculated the parameter size and FLOPS of different models. From the Table 3, it is evident that the parameter sizes and FLOPs vary significantly across different methods. For instance, the MCC method has the highest parameter size at 94.37 M, indicating a more complex model with potentially higher computational demands. In contrast, the Early Fusion method has the lowest parameter size at 11.87 M, which suggests a simpler model with lower computational requirements. When it comes to FLOPs, UG-GAT exhibits the highest computational complexity with 54.19 GMacs, while Early Fusion again shows the least complexity with 5.36 GMacs. Our proposed method has a parameter size of 66.32 M and 33.83 GMacs, which

Table 3 | Comparison of Parameters and FLOPs for Different Methods

	Early fusion	Middle fusion	Late fusion	MCC	MUCO	GCN	GAT	UG-GAT	Eye-AD
Parameters (M)	11.87	34.22	35.61	94.37	60.18	65.53	65.43	35.97	66.32
FLOPs (GMac)	5.36	16.09	19.08	13.31	24.5	33.05	32.96	54.19	33.83

Table 4 | Performance of Eye-AD on the ROAD-I dataset for various patch sizes (i.e., 3 × 3, 5 × 5, 7 × 7 and 9 × 9)

Methods	Accuracy	Precision	F1-score	Kappa	AUC
3 × 3	0.8859	0.8835	0.8840	0.6948	0.9341
5 × 5	0.8885	0.8862	0.8867	0.7018	0.9355
7 × 7	0.8938	0.8915	0.8914	0.7130	0.9302
9 × 9	0.8872	0.8856	0.8823	0.6851	0.9123

Table 5 | Analysis of the efficiency of different components in Eye-AD on the ROAD-I dataset

Methods	Accuracy	Precision	F1-score	Kappa	AUC
Eye-AD w/o ICM	0.8584	0.8537	0.8539	0.6948	0.9018
Eye-AD w/o SCR	0.8715	0.8703	0.8637	0.7018	0.9297
Eye-AD	0.8885	0.8862	0.8867	0.7018	0.9355

Table 6 | Impact of different inputs on classification performance on the ROAD-I dataset

Methods	Accuracy	Precision	F1-score	Kappa	AUC
SVC	0.8440	0.8459	0.8448	0.6001	0.9056
DVC	0.8610	0.8578	0.8588	0.6288	0.9119
CC	0.8532	0.8606	0.8366	0.5523	0.8966
Multi-input	0.8885	0.8862	0.8867	0.7018	0.9355

positions it as a balanced approach between model complexity and computational efficiency. This balance is crucial for applications requiring both high performance and manageable resource usage.

Ablation study

We carry out ablation studies to investigate the effects of different components over the internal ROAD dataset. In Eye-AD, an instance-level graph is constructed by dividing an image into $n \times n$ patches. To investigate the influence of the patch size on the results, we divided the image into 3 × 3, 5 × 5, 7 × 7, and 9 × 9 patches, respectively, corresponding to 9, 25, 49, and 81 nodes in the instance-level graph. As shown in Table 4, it can be seen that the 5 × 5 configuration achieves the highest AUC, while the 9 × 9 configuration results in a slightly lower score compared to the other configurations. One possible reason is that if a region in one node is too small, it will contain fewer pixels, and thus yield less information for the detection of the disease.

Table 5 shows the results of the proposed Eye-AD with and without some proposed components. It can be seen that the use of an ICM improves F1-score from 0.8539 to 0.884. We also investigate the effectiveness of the SCR. After it was removed, Eye-AD performance declined across all the metrics, especially in terms of kappa, from 0.6943 to 0.6322.

We then investigate the importance of exploiting the correlation and complementary between different *en face* images. Specifically, we compare the performances when using each of the SVC, DVC, and CC only, by replacing different inputs with three identical ones. Table 6 shows the performance of Eye-AD with different inputs from the ROAD dataset. It is easy to observe that multi-input outperforms every single input in terms of all the metrics, indicating that multiple inputs can provide more information to the model, leading to more accurate and reliable results. In addition, we can see that the DVC, when used as the only input, produces overall better results than either the SVC or the CC.

Furthermore, we conducted experiments to investigate the influence of different encoders on the performance, including ResNet18, ResNet50, DenseNet, EfficientNet, ConvNeXt, and ViT. As shown in Table 7, the performance gap between the different CNN-based feature extractors is not significant, with ConvNeXt leading by a narrow margin. This suggests that the key to classification performance lies in the fusion of features between different layers, and general feature extractors are sufficient for our needs. Additionally, the ViT-based encoder performed poorly, likely because ViT models require substantial training data to reach peak performance. With our current dataset, ViT was prone to overfitting and achieved inferior results compared to CNNs.

Statistical analysis

To analyze the parameters of interest of microvasculature and FAZ, we first extract the retinal structures, i.e., microvasculature, FAZ, and vascular bifurcations, using a multi-task learning method[39]. The final vessel and FAZ detection results were obtained using the OTSU automatic thresholding method. Then eight parameters were calculated, including vascular length density (VLD), vascular area density (VAD), vascular bifurcation number (VB), vascular fractal dimension (VFD), FAZ Area (FA), FAZ Circularity (FC), FAZ roundness (FR) and FAZ Solidity (FS).

- **VAD** is defined as the total length in millimeters of perfused retinal microvasculature per unit area of the analyzed image.
- **VLD** is defined as the ratio of the total number of pixels on microvascular centerlines to the measurement area.
- **VFD** is a measure of the global branching complexity of the vasculature.
- **VB** is the total number of bifurcations in the analyzed area.
- **FA** is defined as the total number of pixels in the FAZ region.
- **FC** measures the degree of roundness of the FAZ, calculated as: $FC = 4\pi^{*}FA/FP^{2}$. FP is the perimeter of FAZ. A larger FC indicates a more circular shape. A value of 1.0 denotes a perfect circle.
- **FR** is similar to FC but is less sensitive to irregular borders along the perimeter of the FAZ, and is calculated as: $FR = 4\pi * FA/L^{2}_{major}$. $/L_{major}$ is the major axial length of FAZ.
- **FS** describes the extent to which the FAZ is convex or concave: it is defined as the ratio between the area of the FAZ, and the area of the convex hull covering the FAZ. The farther the solidity deviates from 1, the greater the concavity of the structure.

We explore the differences of these parameters between AD and healthy control groups in the whole image as well as in sub-sectors. The sub-sectors include the superior inner, temporal inner, inferior inner, nasal inner, superior outer, temporal outer, inferior outer, and nasal outer[8]. We use a multiple linear regression model with generalized estimating equations to correlate AD and retinal microvasculature and FAZ measurements, adjusting for age, gender, hypertension, diabetes, and education level. We perform the analyses using standard statistical software (SPSS, v.24.0, IBM, US). The results are presented in Table 8.

Table 7 | Detection results on EOAD-I and ROMCI-I using different encoders

Dataset	Encoder	Accuracy	Precision	F1-score	Kappa	AUC
ROAD-I	ResNet18	0.8885	0.8862	0.8867	0.7018	0.9355
	ResNet50	0.8921	0.8902	0.8883	0.7027	0.9365
	DenseNet	0.8914	0.8891	0.8881	0.7018	0.9471
	EfficientNet	0.8790	0.8833	0.8806	0.6954	0.9330
	ConvNeXt	0.8914	0.8891	0.8881	0.7018	0.9471
	ViT	0.7744	0.7524	0.7514	0.3109	0.6902
ROMCI-I	ResNet18	0.8487	0.8506	0.8410	0.6229	0.8630
	ResNet50	0.8527	0.8702	0.8405	0.6254	0.8607
	DenseNet	0.8320	0.8314	0.8237	0.5836	0.8785
	EfficientNet	0.8208	0.8192	0.8119	0.5564	0.8124
	ConvNeXt	0.8671	0.8747	0.8588	0.6608	0.8732
	ViT	0.7946	0.7984	0.7964	0.3381	0.7399

Table 8 | The results of the global statistical analysis of the parameters, adjusted for factors such as age, gender, hypertension, diabetes, and education level, were obtained utilizing Generalized Estimation Equation

	SVC			DVC		
Variable	**EOAD/MCI**	**Control**	***P*-value**	**EOAD/MCI**	**Control**	***P*-value**
ROAD						
VAD	31.82 (4.67)	33.19 (4.03)	**0.048**	44.90 (4.53)	46.81 (3.36)	**0.021**
VLD	10.54 (1.56)	11.17 (1.30)	0.074	14.92 (1.38)	15.64 (1.03)	**0.029**
VFD	1.61 (0.04)	1.63 (0.03)	**0.012**	1.69 (0.02)	1.70 (0.01)	**0.021**
VB	161.49 (34.42)	158.00 (36.71)	0.115	158.83 (39.87)	148.22 (32.53)	**<0.001**
FR	28.68 (0.289)	30.88 (5.23)	**0.006**	32.81 (5.87)	34.40 (3.98)	**0.02**
FA	0.67 (0.53)	0.60 (0.33)	0.086	0.35 (0.17)	0.33 (0.13)	0.167
FC	0.61 (0.18)	0.65 (0.17)	0.06	0.83 (0.17)	0.86 (0.19)	0.118
FS	0.82 (0.07)	0.83 (0.07)	0.134	0.90 (0.05)	0.91 (0.06)	0.226
ROMCI						
VAD	32.12 (9.09)	32.22 (9.97)	0.173	38.72 (8.79)	38.66 (8.60)	**0.032**
VLD	11.57 (3.57)	11.78 (3.87)	0.206	14.36 (3.19)	14.73 (3.22)	0.1
VFD	1.62 (0.04)	1.63 (0.04)	0.443	1.67 (0.03)	1.67 (0.03)	0.473
VB	167.90 (49.55)	167.36 (54.92)	0.935	206.55 (67.07)	211.30 (70.22)	0.086
FR	13.56 (16.15)	9.83 (15.00)	**0.036**	14.67 (17.51)	10.90 (16.44)	**<0.001**
FA	1.323 (0.289)	1.294 (0.320)	0.143	1.639 (0.180)	1.612 (0.170)	0.609
FC	0.60 (0.20)	0.57 (0.22)	0.066	0.73 (0.25)	0.69 (0.24)	0.623
FS	0.80 (0.18)	0.78 (0.21)	0.156	0.86 (0.16)	0.84 (0.16)	0.269

The values in bold indicate statistically significant differences between the disease group and the control group for the corresponding parameters.

Data availability

Restrictions apply to the availability of the developmental and validation datasets, which were used with the permission of the participants for the current study. De-identified data may be available for research purposes from the corresponding authors upon reasonable request.

Code availability

All algorithms used in this study were developed using libraries and scripts in PyTorch. The source code is publicly available at https://github.com/iMED-Lab/Eye-AD.

Received: 22 May 2024; Accepted: 4 October 2024;
Published online: 20 October 2024

References

1. Nichols, E. et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the global burden of disease study 2019. *Lancet Public Health* **7**, e105–e125 (2022).
2. Abajobir, A. et al. Neurological disorders collaborator group, global, regional, and national burden of neurological disorders during 1990–2015: a systematic analysis for the global burden of disease study 2015. *Lancet Neurol.* **16**, 877–897 (2017).
3. Wu, Y.-T. et al. The changing prevalence and incidence of dementia over time-current evidence. *Nat. Rev. Neurol.* **13**, 327–339 (2017).
4. Koronyo, Y. et al. Retinal pathological features and proteome signatures of Alzheimer's disease. *Acta Neuropathol.* **145**, 409–438 (2023).

5. Zabel, P. et al. Comparison of retinal microvasculature in patients with Alzheimer's disease and primary open-angle glaucoma by optical coherence tomography angiography. *Investig. Ophthalmol. Vis. Sci.* **60**, 3447–3455 (2019).
6. Wu, J. et al. Retinal microvascular attenuation in mental cognitive impairment and Alzheimer's disease by optical coherence tomography angiography. *Acta ophthalmol.* **98**, e781–e787 (2020).
7. Cheung, C. Y. et al. A deep learning model for detection of Alzheimer's disease based on retinal photographs: a retrospective, multicentre case-control study. *Lancet Digit. Health* **4**, e806–e815 (2022).
8. Xie, J. et al. Deep segmentation of octa for evaluation and association of changes of retinal microvasculature with alzheimer's disease and mild cognitive impairment. *Br. J. Ophthalmol.* **108**, 432–439 (2024).
9. Chua, J. et al. Retinal microvasculature dysfunction is associated with alzheimer's disease and mild cognitive impairment. *Alzheimer's Res. Ther.* **12**, 1–13 (2020).
10. Cheung, C. Y.-l et al. Microvascular network alterations in the retina of patients with alzheimer's disease. *Alzheimer's Dement.* **10**, 135–142 (2014).
11. Tian, J. et al. Modular machine learning for alzheimer's disease classification from retinal vasculature. *Sci. Rep.* **11**, 238 (2021).
12. Lim, Y. J. et al. Efficient deep learning algorithm for alzheimer's disease diagnosis using retinal images. In *2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS)* 254–257 (IEEE, 2022).
13. Hermessi, H., Mourali, O. & Zagrouba, E. Multimodal medical image fusion review: Theoretical background and recent advances. *Signal Process.* **183**, 108036 (2021).
14. Zhou, T. et al. Deep multi-modal latent representation learning for automated dementia diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 629–638 (Springer, 2019).
15. Heisler, M. et al. Ensemble deep learning for diabetic retinopathy detection using optical coherence tomography angiography. *Transl. Vis. Sci. Technol.* **9**, 20 (2020).
16. Zhou, T., Canu, S., Vera, P. & Ruan, S. 3D medical multi-modal segmentation network guided by multi-source correlation constraint. In *2020 25th International Conference on Pattern Recognition (ICPR)* 10243–10250 (IEEE, 2021).
17. Wang, X. et al. Screening of dementia on octa images via multi-projection consistency and complementarity. In *International Conference on Medical Image Computing and Computer-Assisted Intervention* 688–698 (Springer, 2022).
18. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations* (2022).
19. Velickovic, P. et al. Graph attention networks. *Statistics* **1050**, 10–48550 (2017).
20. Hao, J. et al. Uncertainty-guided graph attention network for parapneumonic effusion diagnosis. *Med. Image Anal.* **75**, 102217 (2022).
21. Dosovitskiy, A. et al. An image is worth 16x16 words: Transformers for image recognition at scale. In *International Conference on Learning Representations* (2020).
22. Liu, Z. et al. Swin transformer v2: Scaling up capacity and resolution. In *Proc. IEEE/CVF conference on computer vision and pattern recognition*, 12009–12019 (IEEE, 2022).
23. Katsimpris, A. et al. Optical coherence tomography angiography in alzheimer's disease: a systematic review and meta-analysis. *Eye* **36**, 1419–1426 (2022).
24. Campbell, J. et al. Detailed vascular anatomy of the human retina by projection-resolved optical coherence tomography angiography. *Sci. Rep.* **7**, 42201 (2017).
25. Lahme, L. et al. Evaluation of ocular perfusion in alzheimer's disease using optical coherence tomography angiography. *J. Alzheimer's Dis.* **66**, 1745–1752 (2018).
26. Zhang, Y. S. et al. Parafoveal vessel loss and correlation between peripapillary vessel density and cognitive performance in amnestic mild cognitive impairment and early alzheimer's disease on optical coherence tomography angiography. *PLoS ONE* **14**, e0214685 (2019).
27. Jiang, H. et al. Altered macular microvasculature in mild cognitive impairment and alzheimer disease. *J. Neuro-Ophthalmol. Off. J. North Am. Neuro-Ophthalmol. Soc.* **38**, 292 (2018).
28. Liang, W., Olwen, M., Gonzalez, C. N., Calabresi, P. A. & Shiv, S. Emerging applications of optical coherence tomography angiography (octa) in neurological research. *Eye Vis.* **5**, 11– (2018).
29. Hormel, T. T., Jia, Y., Jian, Y., Hwang, T. S. & Huang, D. Plexus-specific retinal vascular anatomy and pathologies as seen by projection-resolved optical coherence tomographic angiography. *Prog. Retin. Eye Res.* **80**, 100878 (2020).
30. Lai, A. Y. et al. Venular degeneration leads to vascular dysfunction in a transgenic model of alzheimer's disease. *Brain* **138**, 1046–1058 (2015).
31. Fisher, R. A., Miners, J. S. & Love, S. Pathological changes within the cerebral vasculature in alzheimer's disease: New perspectives. *Brain Pathol.* **32**, e13061 (2022).
32. He, L. et al. Effect of dynamic cerebral autoregulation on the association between deep medullary vein changes and cerebral small vessel disease. *Front. Physiol.* **14**, 1037871 (2023).
33. Asanad, S. et al. The retinal choroid as an oculovascular biomarker for alzheimer's dementia: a histopathological study in severe disease. *Alzheimer's Dement. Diagn. Assess. Dis. Monit.* **11**, 775–783 (2019).
34. McKhann, G. M. et al. The diagnosis of dementia due to alzheimer's disease: Recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's Dement.* **7**, 263–269 (2011).
35. Albert, M. S. et al. The diagnosis of mild cognitive impairment due to alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimer's Dement.* **7**, 270–279 (2011).
36. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proc. IEEE conference on computer vision and pattern recognition* 770–778 (IEEE, 2016).
37. Chen, Z., Liu, J., Zhu, M., Woo, P. Y. & Yuan, Y. Instance importance-aware graph convolutional network for 3d medical diagnosis. *Med. Image Anal.* **78**, 102421 (2022).
38. Selvaraju, R. R. et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proc. IEEE international conference on computer vision* 618–626 (IEEE, 2017).
39. Hao, J. et al. Retinal structure detection in octa image via voting-based multitask learning. *IEEE Trans. Med. Imaging* **41**, 3969–3980 (2022).

Acknowledgements

This work was supported in part by the National Science Foundation Program of China (62422122, 62272444, 62371442, 62302488), in part by the Youth Innovation Promotion Association CAS (2021298), in part by the Zhejiang Provincial Natural Science Foundation of China (LR22F020008, LQ23F010007, LR24F010002, LZ23F010002), in part by Key research and development program of Zhejiang Province (2024C03101, 2024C03204) and Key Project of Ningbo Public Welfare Science and Technology (2023S012). AFF acknowledges the support of the Royal Academy of Engineering Chair INSILEX (CiET1819\9), the UKRI Frontier Research Guarantee INSILICO (EP\Y030494\1). The research of AFF was carried out at the National Institute for Health and Care Research (NIHR) Manchester Biomedical Research Centre (BRC) (NIHR203308).

Author contributions

J.H. designed and executed all the experiments conducted all the subsequent statistical analyses, and drafted the manuscript. W.R.K. and

A.F. helped design the experiments, the writing, and the data interpretation and made substantial revisions and edits of the draft manuscript. S.Z., Q.L., and T.S. contributed to the data analysis and data cleaning. S.L. and J.L. helped design the experiments and contributed to the data analysis. Y.L., Y.Z., and Y.Z. contributed to the external validation of the proposed method. H.Q., Y.Z. conceived the methodology, designed the experiments, and contributed to the writing. All authors contributed to the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41746-024-01292-5.

Correspondence and requests for materials should be addressed to Shuting Zhang, Hong Qi or Yitian Zhao.

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) 2024

https://dspace.org/

Early detection of dementia through retinal imaging and trustworthy AI

Hao, Jinkui; Kwapong, William R.; Shen, Ting; Fu, Huazhu; Xu, Yanwu ...

2024-10-04
Springer
https://doi.org/10.1038/s41746-024-01292-5
Hao J, Kwapong WR, Shen T, et al., (2024) Early detection of dementia through retinal imaging and trustworthy AI. npj Digital Medicine, Volume 7, October 2024, Article number 294

Downloaded from a DSpace repository. https://dspace.org/