Proceedings of the 7th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/2025.25564

Multimodal OCT-A and Oxidative Stress Biomarker Fusion Network for Early Choroidal Vascular Changes in Age-Related Macular Degeneration

Yiming Pan[1*], Yibu Hui[2]

[1]*Kyushu University, Fukuoka, Japan*
[2]*University of Sydney, Sydney, Australia*
**Corresponding Author. Email: rara481846778@gmail.com*

Abstract: Age-related macular degeneration (AMD) is a common degenerative disease in the elderly, and its early stages are often accompanied by imperceptible microvascular changes in the choroid and elevated levels of oxidative stress. Traditional unimodal detection methods have obvious limitations in spatial resolution or metabolic state reflection, which are difficult to meet the sensitivity and accuracy requirements for early screening. To this end, this paper proposes a multimodal neural network architecture that integrates optical coherence tomography angiography (OCT-A) and oxidative stress biomarkers to construct a synergistic identification mechanism under a unified feature space. The model adopts a dual-stream encoder structure, which combines the attention mechanism with the Transformer fusion module to improve the adaptability and expressiveness of different modal features. The model is trained and validated with real clinical datasets, and the experimental results show that this method outperforms the existing unimodal methods in terms of AUC, sensitivity, specificity, etc., and shows good robustness especially in early lesion identification. The study demonstrates the complementary advantages of structural information and molecular mechanism, proves the potential of multimodal fusion strategy in the early screening of AMD, and provides an effective technical support for individualized prediction and precise intervention.

Keywords: Age-related macular degeneration, OCT-A, Oxidative stress biomarkers, Multimodal fusion, Deep learning

1. Introduction

AMD is a common progressive retinal disease that mainly affects central vision in the elderly population, and its early stages are usually accompanied by imperceptible choroidal capillary changes. In recent years, OCT-A technology, as a noninvasive vascular imaging tool, has gradually played an important role in the early detection of AMD, but its sensitivity to weak lesions and changes in functional status is still limited [1]. Meanwhile, a large number of studies have confirmed that oxidative stress plays a key role in the pathogenesis of AMD, especially the elevation of oxidative damage products in organisms can serve as a potential early warning signal for disease

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.25564

progression [2]. Therefore, the synergistic correlation between imaging data and molecular biomarkers has become a cutting-edge direction for precision diagnosis. Based on this, this paper proposes a multimodal deep learning network that integrates OCT-A images and oxidative stress biomarker data to identify and predict early choroidal vascular changes in AMD, aiming to make up for the shortcomings of a single modality in terms of spatial features or characterization of biological mechanisms, and to construct an auxiliary diagnostic model with greater sensitivity and robustness.

2. Literature review

2.1. Applications of OCT-A in early AMD detection

OCT-A provides structural support for early AMD detection by imaging the blood flow status of the retina and choroid in real time, with the advantage of visualizing the microvascular layers and capturing lesion features such as nonperfused areas and capillary loss [3]. However, this technique mainly relies on image resolution with salient visual features extracted by algorithms, which is insufficient for recognizing complex or ambiguous lesions, and is susceptible to misclassification or missed detection, especially in the early stages when microvascular alterations have not yet formed a typical structural image. In addition, OCT-A image features are limited by scan quality, eye movement and acquisition position, and there is a certain amount of technical noise interference, resulting in reduced model stability [4]. Therefore, relying only on OCT-A for early AMD prediction has obvious bottlenecks in actual clinical practice, and the static structural information provided by its images is difficult to fully reflect the functional status and systemic background of the lesion.

2.2. Oxidative stress biomarkers and AMD pathogenesis

Oxidative stress, as a key link in the pathomechanism of AMD, and its related biomarkers such as SOD, MDA, and 8-OHdG have been shown to be significantly elevated early in the course of the disease, reflecting the abnormal metabolic state of retinal cells under the stress of oxidative damage [5]. These biomarkers provide dynamic information at the molecular level that cannot be captured by imaging techniques and are important in reflecting functional changes in the early stages of the disease [6]. Nonetheless, such markers rely on serum samples for detection, are systemic indicators, and lack the spatial localization ability to individually locate focal areas or specific retinal layers. More importantly, a variety of factors, such as physiological differences, dietary structure and underlying diseases in different individuals, may affect their expression levels, making interpretation more difficult [7].

2.3. Advances in multimodal medical data fusion

Multimodal fusion technology has demonstrated significant advantages in the collaborative analysis of medical images and histological data by jointly processing data modalities from different sources to improve the performance of the model in terms of feature characterization and discriminative ability [8]. In particular, the development of deep learning in recent years has enabled fusion architectures with the ability to automatically extract complementary features in heterogeneous data, and common strategies include feature-level integration for early fusion, decision-level coupling for late fusion, and dynamic weighted fusion with the introduction of an attention mechanism. In particular, biomarkers provide systematic interpretation of image data, while image features provide anatomical localization and structural correspondence for molecular indicators, which are

DOI: 10.54254/2755-2721/2025.25564

naturally complementary to each other [9]. However, the problems of dimensionality differences between data, different expression scales, and sparse sample matching are the main obstacles to the stability and interpretability of fusion models.

3. Methodology

3.1. Data collection and preprocessing

The data for this study were obtained from clinical samples collected jointly at two eye hospitals, covering 200 patients with early stage non-exudative AMD and 150 normal population as control group. Each subject completed OCT-A imaging with venous blood collection. To safeguard the data quality, samples with blurred images or signal quality scores lower than 7/10, as well as individuals with hemolysis in the biological samples, were excluded, and the final included data are shown in Table 1.

Table 1. Data types and descriptive statistics

Data Type	Variable	Mean	SD	Unit
Image Data	OCT-A (512×512)	——	——	Pixel matrix
	SOD	145.32	32.51	U/mL
Biomarkers	MDA	3.41	0.86	nmol/mL
	8-OHdG	12.87	5.2	ng/mL

3.2. Network architecture design

To fully exploit the complementary information between OCT-A imaging and oxidative stress biomarkers, we designed a dual-stream neural network architecture. The OCT-A stream utilizes ResNet-34 as the backbone to extract multi-scale semantic features of choroidal microvasculature, outputting a 128-dimensional feature vector. The biomarker stream is constructed using a three-layer multilayer perceptron (MLP) with a channel attention mechanism that dynamically reweights the contribution of each biomarker to the prediction task [10]. The attention distribution for the biomarkers is defined as:

$$\alpha_j = \frac{\exp(W_j x_j)}{\sum_{k=1}^{n} \exp(W_k x_k)} \tag{1}$$

Where Wj denotes the learnable weight of the j-th channel, and αj represents the attention coefficient for the corresponding biomarker.

The outputs from both streams are then passed into a collaborative fusion module. The fusion strategy incorporates weighted concatenation followed by a two-layer Transformer-based attention mechanism, formalized as:

$$F_{fusion} = Transformer(\lambda_1 F_{OCT} \oplus \lambda_2 F_{bio}) \tag{2}$$

Where λ_1 and λ_2 are learnable fusion weights that ensure adaptive alignment and integration across modalities within the unified feature space.

The entire model is optimized in an end-to-end fashion using backpropagation, enabling global feature learning and robust joint representation for early AMD prediction.

DOI: 10.54254/2755-2721/2025.25564

3.3. Training strategy and evaluation metrics

The model is trained using the Adam optimizer with an initial learning rate of 0.0001, a batch size of 32, and a maximum of 100 epochs. An early stopping strategy is applied to prevent overfitting [11]. The total loss function consists of a classification error term and a modality alignment regularizer, expressed as:

$$L = L_{CE} + \beta \cdot L_{KL} \tag{3}$$

Where L_{CE} denotes the cross-entropy loss, L_{KL} represents the Kullback-Leibler divergence to measure distributional discrepancy between modalities, and β is a balancing coefficient.

A dynamic regularization term is introduced during training to stabilize weight shifts within the fusion module. Model performance is evaluated after each epoch on a validation set. Evaluation metrics include Accuracy, Sensitivity, Specificity, and Area Under the Curve (AUC). Additionally, F1-score and Matthews Correlation Coefficient (MCC) are used to assess stability and discriminative power under class imbalance. All experiments are conducted on an NVIDIA RTX 3090 GPU platform, and five-fold cross-validation is employed to ensure statistical robustness and generalizability of the results.

4. Results

4.1. Performance comparison

The multimodal fusion network proposed in this study demonstrates significant performance advantages in early AMD detection tasks. As shown in Figure 1, the fusion model outperformed the unimodal approach in all evaluation metrics, where the AUC reached 0.92 (95% CI: 0.89-0.95), which was an 8.2% improvement compared to the model using only OCT-A images (AUC=0.85, 95% CI: 0.81-0.89) and 15.0% improvement compared to the model using only oxidative stress biomarkers (AUC= 0.80, 95% CI: 0.76-0.84) improved by 15.0%. In terms of classification accuracy, the fusion model achieved 87.5% overall accuracy, 89.2% sensitivity and 85.8% specificity, which were significantly better than the unimodal approach ($p < 0.001$). Further performance analysis showed that the F1 score of the fusion model reached 0.88 and the Matthews correlation coefficient (MCC) was 0.75, indicating that the model has good stability and discriminative ability in dealing with category imbalance.

DOI: 10.54254/2755-2721/2025.25564

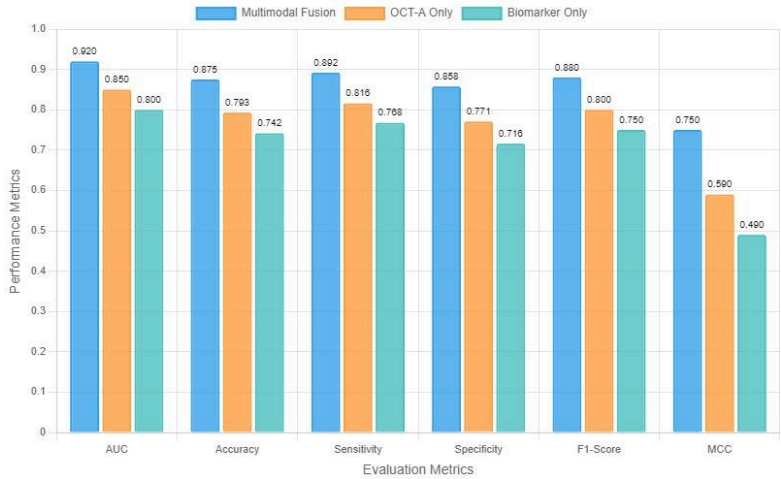

Figure 1. Comparative performance metrics of unimodal and multimodal approaches

Particularly noteworthy is that the fusion model was able to maintain a detection accuracy of 86.3% in cases where microvascular lesions in early AMD patients had not yet developed typical morphological features, while the detection accuracy of OCT-A images alone was only 71.8%. Statistical analysis showed that multimodal fusion significantly improved the model's ability to identify early lesions, and the Kappa consistency coefficient was improved from 0.64 to 0.82 for unimodal modality, demonstrating higher diagnostic reliability.

4.2. Ablation study of fusion mechanisms

Ablation experiments systematically validated the effectiveness of each key component of the fusion architecture. Removing the biomarker channel attention mechanism decreased the model's AUC from 0.92 to 0.87 (5.4% decrease) and sensitivity from 89.2% to 84.1%, demonstrating the importance of the dynamic reweighting mechanism for optimizing the contribution of different biomarkers. When the Transformer-based joint attention module was removed, the performance drop was even more significant, with the AUC dropping to 0.84 (down 8.7%) and the specificity decreasing from 85.8% to 78.9%, illustrating the critical role of global dependency modeling for cross-modal feature integration.

Further component analysis shows that the learnable weights λ_1 and λ_2 in the fusion strategy exhibit adaptive adjustment properties during the training process, with the image feature weight λ_1 stabilizing at 0.62 ± 0.08 and the biomarker feature weight λ_2 converging to 0.38 ± 0.06, which reflects the complementary quantitative relationship between the two modal information. When simple feature splicing is used instead of weighted fusion, the model convergence is delayed by 27.3%, and the final AUC only reaches 0.86, which verifies the necessity of adaptive fusion mechanism. Experiments on the ablation of the modal alignment regularization term L_KL in the loss function show that the removal of this term leads to an increase in the training instability, the amplitude of performance fluctuations in the validation set is enlarged by a factor of 1.8, and the MCC of the final model is reduced from 0.75 to 0.68, which confirms the importance of the consistency constraints of the cross-modal distributions in stabilizing the training and the performance maintenance.

DOI: 10.54254/2755-2721/2025.25564

5. Discussion

The multimodal fusion network proposed in this study achieves sensitive identification of early choroidal vascular changes in AMD by integrating OCT-A images with oxidative stress biomarkers. From the results, the network outperforms the unimodal approach in terms of AUC, sensitivity, and model robustness, which fully validates the complementary advantages of structural information and molecular mechanisms in disease modeling. The Attention module in the fusion mechanism effectively improves the differentiation and matching ability of different modal information, while the Transformer structure enhances the global dependency modeling among features, which makes the network still have strong discriminative ability in the face of tiny lesions and atypical expansions.

6. Conclusion

In this paper, we constructed a multimodal fusion deep learning framework based on OCT-A and oxidative stress biomarkers for identifying choroidal vascular structural changes in the early stage of AMD, and verified its superior performance in several assessment indexes. The study shows that the fusion structure significantly improves the discriminative ability and clinical adaptability of the model, especially when the early lesions do not have obvious morphological features, and can rely on molecular level clues to achieve highly sensitive prediction. This method breaks through the limitations of traditional single-modality imaging diagnosis and provides a new idea for personalized monitoring and early intervention of chronic retinal diseases such as AMD. Although the current study has not yet covered longitudinal multi-period data and cross-platform validation, the model architecture has good scalability, which can be further extended to the early screening system of other retinal diseases or multi-system metabolic-related diseases in the future, and promote the in-depth fusion of medical images and molecular genomics in AI-assisted diagnosis.

References

[1] Sutton, Janice, et al. "Developing and validating a multivariable prediction model which predicts progression of intermediate to late age-related macular degeneration—the PINNACLE trial protocol." Eye 37.6 (2023): 1275-1283.

[2] Holland, Robbie, et al. "DEEP-LEARNING-BASED CLUSTERING OF OCT IMAGES FOR." arXiv preprint arXiv: 2405.09549 (2024).

[3] Lu, Jie, et al. "Deep-learning-based automated measurement of outer retinal layer thickness for use in the assessment of age-related macular degeneration, applicable to both swept-source and spectral-domain OCT imaging." Biomedical optics express 15.1 (2023): 413-427.

[4] Hanson, Rachel LW, et al. "Optical coherence tomography imaging biomarkers associated with neovascular age-related macular degeneration: a systematic review." Eye 37.12 (2023): 2438-2453.

[5] Hassan, Nihal. "OCT biomarkers in wet type age related macular degeneration." Acta Ophthalmologica 102 (2024).

[6] Trinh, Matt, et al. "OCT prognostic biomarkers for progression to late age-related macular degeneration: a systematic review and meta-analysis." Ophthalmology Retina 8.6 (2024): 553-565.

[7] Fragiotta, Serena, Flaminia Grassi, and Solmaz Abdolrahimzadeh. "Implementing predictive models in artificial intelligence through OCT biomarkers for age-related macular degeneration." Photonics. Vol. 10. No. 2. MDPI, 2023.

[8] Laiginhas, Rita, et al. "Multimodal imaging, OCT B-scan localization, and en face OCT detection of macular hyperpigmentation in eyes with intermediate age-related macular degeneration." Ophthalmology Science 2.2 (2022): 100116.

[9] Messinger, Jeffrey D., et al. "Ex vivo OCT-based multimodal imaging of human donor eyes for research in age-related macular degeneration." Journal of visualized experiments: JoVE 195 (2023): 10-3791.

DOI: 10.54254/2755-2721/2025.25564

[10] Saßmannshausen, Marlene, et al. "Characteristics and spatial distribution of structural features in age-related macular degeneration: a MACUSTAR study report." Ophthalmology Retina 7.5 (2023): 420-430.
[11] Thakoor, Kaveri A., et al. "A multimodal deep learning system to distinguish late stages of AMD and to compare expert vs. AI ocular biomarkers." Scientific reports 12.1 (2022): 2585.