Journal of Artificial Intelligence and Metaheuristic

♦ ISSN: 2833-5597 Vol: 10, No. 02, 2025 ♦ PP. 01–08

Article DOI: https://doi.org/10.54216/JAIM.100201

FULL LENGTH ARTICLE

Hybrid CNN–LSTM Architecture for OCT Retinal Disease Classification

Ehsan Khodadadi[1,*] **P. K. Dutta**[2] **Amel Ali Alhussan**[3] **Marawa Metwally**[4,5]

[1] Department of Chemistry and Biochemistry, University of Arkansas, Fayetteville, AR 72701, USA
[2] School of Engineering and Technology, Amity University Kolkata, Kolkata, India
[3] Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[4] Computer Science and Intelligent Systems Research Center, Blacksburg, VA 24060, USA
[5] Jadara Research Center, Jadara University, Irbid 21110, Jordan

Emails: Ehsan.k@uark.edu; pkdutta@kol.amity.edu; aalhussan@pnu.edu.sa; marwa@jcsis.org

Received: April 15, 2025 Revised: June 25, 2025 Accepted: September 05, 2025 ⋆ Corresponding author

ABSTRACT

Optical coherence tomography (OCT) provides detailed retinal images, but their interpretation can be difficult across clinical settings and patient groups. This paper presents a hybrid Convolutional Neural Network–Long Short-Term Memory (CNN–LSTM) architecture for classifying retinal OCT images into eight disease categories. Two parallel convolutional branches extract spatial features; their outputs are pooled, reshaped into a sequence, and processed by an LSTM layer to model inter-feature relationships before dense classification. The model is evaluated on 24,000 images and compared with ResNet50, MobileNetV2, InceptionV3, and Xception. The proposed architecture achieves 93.00% accuracy, 94% precision, 93% recall, and a 0.93 F1-score, outperforming the compared models. The results indicate that combining convolutional feature extraction with sequence modeling is a promising approach to automated retinal disease classification.

Keywords: Retinal disease classification ▪ Optical coherence tomography (OCT) ▪ CNN–LSTM ▪ Deep learning ▪ Transfer learning

1. INTRODUCTION

Optical coherence tomography (OCT) has enabled detailed, non-invasive examination of retinal structure and is used to assess conditions including diabetic retinopathy (DR) and age-related macular degeneration (AMD). In practice, diagnostic consistency remains affected by image quality, disease progression, coexisting pathologies, and differences between patient groups and clinical settings [1, 2, 3]. DR can progress without early symptoms, while AMD-related structural changes can cause substantial visual impairment; timely and accurate interpretation is therefore important [4, 5, 6, 7, 8, 9].

Healthcare systems face barriers to deploying deep-learning classifiers at scale, including the limited availability of annotated cases and the difficulty of producing standardized datasets across institutions [10]. Transfer learning offers a practical way to adapt the representations learned by pre-trained networks to diagnostic tasks with less task-specific data [11]. In this study, the target classes are AMD, choroidal neovascularization (CNV), DR, diabetic macular edema (DME), central serous retinopathy (CSR), drusen, macular hole (MH), and normal retina [12].

© 2025 The Author(s). Published by ASPG. This article is licensed under CC BY 4.0.

AI has become an increasingly common component of ophthalmic research and clinical decision-support, and deep learning can learn image representations without requiring manually engineered features [13, 14, 15]. CNNs have been applied to medical-image classification, segmentation, and detection, while network weights, biases, and hyperparameters remain important to final performance [16, 17]. Artificial neural networks and deep-learning systems differ in model depth and learning strategy; gradient-based methods are widely used, while metaheuristic optimization has also been investigated to address some of their limitations [18, 19].

CNN-based systems may shorten diagnostic workflows and support clinicians, but conventional deep-network training is data intensive. Medical datasets can be limited by privacy requirements, annotation costs, and the need for expert review [20]. Accordingly, this study fine-tunes ImageNet-pretrained Xception, InceptionV3, MobileNetV2, and ResNet50 models for OCT images, adapting general visual representations to retinal disease patterns [21]. The proposed hybrid design supplements convolutional spatial-feature extraction with LSTM sequence modeling: the CNN encodes salient image structure and the LSTM learns relationships across the resulting feature sequence. The model is trained and tested on 24,000 OCT images representing eight conditions using TensorFlow and Keras [22].

The remainder of this paper reviews related work in Section 2, describes the dataset and methods in Section 3, presents results and discussion in Section 4, and concludes in Section 5.

2. RELATED WORK

Recent retinal-image research spans multi-label classification, denoising, transformer architectures, biomarkers, and clinical evaluation. Deep-intronic variants and splicing defects have been investigated as sources of inherited retinal disease, while vision-transformer methods address concurrent pathologies that are not well represented by single-label classification [23, 24]. For OCT images, multi-scale denoising residual networks aim to preserve diagnostic detail under speckle noise [25].

Other work explores tear extracellular-vesicle microRNAs as potential ocular biomarkers, interpretable lightweight CNNs for OCT classification, and the clinical assessment of AI-assisted grading of epiretinal membranes [26, 27, 28]. Model-based transformers have been applied to OCT images and videos, and hybrid convolution/vision-transformer methods combine local texture with global contextual features [29, 30].

Clinical translation has also motivated work on AI-enabled telemedicine, deep-learning approaches to glaucoma screening, and the genetic and clinical heterogeneity of inherited retinal diseases [31, 32, 33]. Additional studies propose global-attention networks for OCT, scale-adaptive systems for AMD grading, approaches to dataset bias and class imbalance, and residual encoder–decoder networks for exudate segmentation [34, 35, 36, 37]. Reviews of AI in retinal care discuss potential clinical applications alongside implementation challenges [38].

The reviewed work addresses retinal diagnosis at several levels. Molecular studies of deep-intronic variants show how aberrant splicing can expand the known mutational landscape and complicate inherited-disease diagnosis; understanding those variants can inform both diagnostic tools and gene-therapy targets [23]. For acquired retinal disease, multi-label learning is relevant because patients may present with concurrent pathologies, whereas single-label classifiers do not represent that clinical complexity. Transformer-based classification has therefore been explored to model multiple conditions together [24].

OCT provides detailed views of the macula, but speckle noise can obscure disease features. The multi-scale denoising residual convolutional network described in prior work combines soft denoising, multi-scale context extraction, and feature fusion to suppress noise while retaining diagnostically useful information [25]. Complementary research investigates tear extracellular-vesicle microRNAs as non-invasive molecular biomarkers that may reflect ocular health, disease progression, or treatment response [26].

Manual OCT interpretation can be time-consuming and subject to reader variation. Lightweight CNNs have been evaluated for classification robustness and interpretability, including methods intended to explain model decisions and assess susceptibility to adversarial perturbations [27]. Clinical studies have also compared AI-assisted with unassisted evaluation of epiretinal-membrane severity. They report the potential for decision support while emphasizing careful review of model errors and patient-safety considerations [28].

Transformer methods extend this work from static images to temporal OCT data. The model-based transformer approach applies Vision Transformer and Swin Transformer models to OCT images and multi-scale transformer models to OCT video, with sparse representations used to handle high-dimensional input and capture long-range dependencies [29]. Conv-ViT takes a complementary approach, combining convolutional extraction of local texture with transformer attention to represent global shape and context [30].

AI-assisted retinal care is also connected to telemedicine and portable imaging. Remote OCT acquisition and automated analysis could enable earlier feedback in underserved settings, although these systems must be integrated into clinical workflows responsibly [31]. A review of deep-learning glaucoma screening methods compares datasets, metrics, preprocessing, anatomical localization, classification, and segmentation, emphasizing the importance of timely screening for a disease that can remain silent in its early stages [32]. In inherited retinal disease, variable expression among female carriers of X-linked conditions has been linked to mosaic gene expression following X-inactivation, with implications for risk assessment and future precision therapies [33].

Several studies focus on model design and dataset properties. The global-attention block in GABNet generates attention over feature-map height, width, and channels to emphasize diagnostically relevant regions; Grad-CAM visualizations provide additional information about the regions driving predictions [34]. Scale-adaptive autoencoder and CNN methods have been used to detect and grade AMD from fundus images by capturing subtle retinal-pigment-epithelium and drusen changes [35]. Work on unified retinal datasets highlights the need to counter class imbalance across common conditions such as DR, AMD, and glaucoma to improve

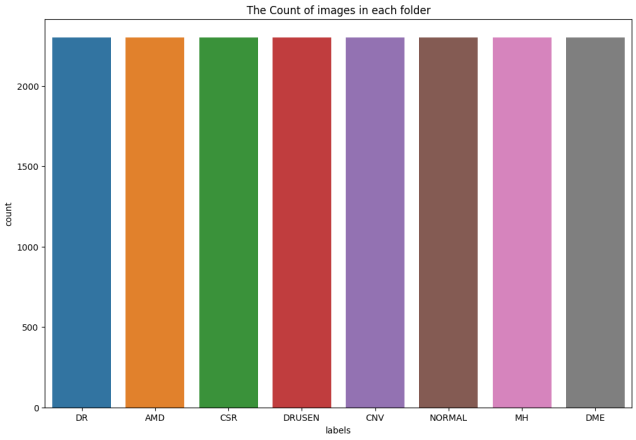


Figure 1. Dataset distribution for each class.

clinical generalization [36]. Residual encoder–decoder architectures have also been applied to automated segmentation of diabetic-retinopathy exudates, an important marker for disease assessment [37]. More broadly, clinical reviews identify opportunities for AI to improve screening speed and access while noting challenges in validation, workflow integration, and safe use [38].

Table 1 summarizes the studies cited in this review. Collectively, they reflect a shift toward robust and generalizable models, attention mechanisms, and explainable predictions. Further work is needed on multimodal data, personalized diagnostic tools, and validation in real-world clinical settings.

3. MATERIALS AND METHODS

3.1 Dataset

The dataset comprises 24,000 retinal OCT images divided among eight classes: central serous retinopathy (CSR), diabetic macular edema (DME), diabetic retinopathy (DR), age-related macular degeneration (AMD), choroidal neovascularization (CNV), drusen, macular hole (MH), and normal retina. Each class contains 3,000 images. The dataset is organized into training, validation, and test folders, with 2,300 training images per class. Horizontal flipping, padding, and cropping are used during preprocessing to augment the data and reduce overfitting. The dataset is available from Kaggle at `https://www.kaggle.com/datasets/obulisainaren/retinal-oct-c8`; Figure 1 shows its class distribution.

3.2 Compared CNN Architectures

3.2.1 ResNet50

Residual networks (ResNet) use skip connections to pass activations from earlier layers to deeper layers, helping address vanishing gradients in deep networks. Residual and identity blocks are the basic units. The ResNet50 input image shape used in this study is $224 \times 224 \times 3$, with three input channels [39].

3.2.2 MobileNetV2

MobileNet was introduced as a simplified architecture for efficient vision applications. MobileNetV2 adds inverted residual and linear bottleneck modules to improve training efficiency and accuracy. In this study, MobileNetV2 is used for feature generation and classification; the reported model has 2,261,827 parameters and 154 layers [40].

3.2.3 InceptionV3

Inception architectures use parallel convolutional operations and pooling to expand network depth and breadth. InceptionV3 replaces the initial 7×7 convolution used in earlier versions with three 3×3 convolutions and uses 299×299 input images. The implementation used here has 311 layers and 21,808,931 parameters [41, 42].

3.2.4 Xception

Xception replaces the original Inception modules with depthwise separable convolutions, combining spatial convolution per input channel with pointwise 1×1 convolution. The described model has 14 modules, 36 convolutional layers, 132 layers in total, and 20,867,627 parameters; its input size is 299×299 [43].

3.3 Proposed Hybrid CNN–LSTM Model

The proposed model has two parallel feature-extraction branches. The first extracts deep features through convolutional and identity blocks based on the ResNet structure. The second examines spatial features at a deeper level using max-pooling operations. The branch outputs are processed with global average pooling, reshaped into sequences, and passed to an LSTM layer to learn inter-feature dependencies. Dense layers then classify each image into one of the eight disease categories. The architecture combines spatial features learned by CNNs with sequence modeling by the LSTM.

Figure 2 outlines the analysis pipeline, from image loading and preprocessing through data splitting and model training to classification.

4. RESULTS AND DISCUSSION

The proposed CNN–LSTM model is evaluated against ResNet50, MobileNetV2, InceptionV3, and Xception on the retinal OCT test dataset. Performance is summarized using accuracy, precision, recall, and F1-score. Here, TP and TN denote correctly identified positive and negative cases, while FP and FN denote false positives and false negatives.

Table 3 reports the comparative test results. The hybrid CNN–LSTM achieves the highest overall accuracy (0.9300), precision (0.94), recall (0.93), and F1-score (0.93). ResNet50 is the second-best model by accuracy, while InceptionV3 and MobileNetV2 achieve moderate results. Xception has the lowest accuracy and F1-score among the compared models.

Figure 3 shows the training and validation loss curves. Training loss decreases throughout learning; validation loss fluctuates before declining, with the accompanying study text reporting a minimum at epoch 17. The overall convergence of the curves was interpreted as evidence of learning without substantial overfitting.

Training and validation accuracy are shown in Figure 4. Training accuracy improves over the course of training, while the accompanying study text reports that validation accuracy reaches its peak at epoch 27.

The proposed model's ROC curves in Figure 5 show high areas under the curve across all eight retinal disease classes.

Table 1. Selected studies on retinal disease classification and diagnosis.

Ref.	Main focus	Methodology	Key finding
[23]	Deep-intronic ABCA4 splicing defects	Splicing analysis and CRISPR–Cas9 correction	Identified splicing defects and demonstrated correction in photoreceptor precursors.
[24]	Multi-label retinal disease classification	Vision-transformer model	Addressed classification of concurrent retinal disorders.
[25]	Noisy OCT disease classification	Multi-scale denoising residual CNN	Improved classification under strong image noise.
[26]	Tear extracellular-vesicle miRNAs	Profiling of tear EVs	Identified miRNA profiles potentially associated with ocular disease.
[27]	Interpretable OCT classification	Lightweight CNN with adversarial robustness evaluation	Reported a robust model for retinal disease classification.
[28]	Epiretinal membrane grading	Clinical evaluation of OCT-based AI systems	Assessed the potential benefits and risks of AI-assisted grading.
[29]	OCT image and video classification	Model-Based Transformer	Reported improved discrimination for retinal disease detection.
[30]	Retinal disease detection	CNN and Vision Transformer feature fusion	Combined texture and shape information.
[31]	AI and telemedicine in retinal care	Perspective on AI-enabled remote care	Highlighted opportunities and implementation challenges.
[32]	Automated glaucoma screening	Review of deep-learning methods	Summarized approaches and performance benchmarks.
[33]	X-linked inherited retinal disease	Review of genetics, manifestations, and therapies	Described clinical variability and therapeutic considerations.
[34]	OCT disease classification	GABNet with global attention	Reported improved classification accuracy.
[35]	AMD detection and grading	Scale-adaptive model using fundus images	Adapted feature analysis to retinal scales.
[36]	Dataset bias and class imbalance	Multi-class learning on diverse data	Targeted generalization to clinical settings.
[37]	Exudate segmentation in DR	Residual encoder–decoder	Supported automated segmentation for DR assessment.
[38]	Clinical AI for retinal disease	Review of applications and future directions	Discussed benefits and workflow limitations.

Table 2. Performance evaluation metrics.

Metric	Formula
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$
F1-score	$\frac{2\,\text{Precision}\,\text{Recall}}{\text{Precision}+\text{Recall}}$

Table 3. Comparative performance analysis of deep-learning models on the test dataset.

Model	Accuracy	Precision	Recall	F1-score
ResNet50	0.8452	0.85	0.85	0.84
MobileNetV2	0.7428	0.77	0.74	0.73
InceptionV3	0.7604	0.78	0.76	0.76
Xception	0.4711	0.52	0.47	0.37
CNN–LSTM	0.9300	0.94	0.93	0.93

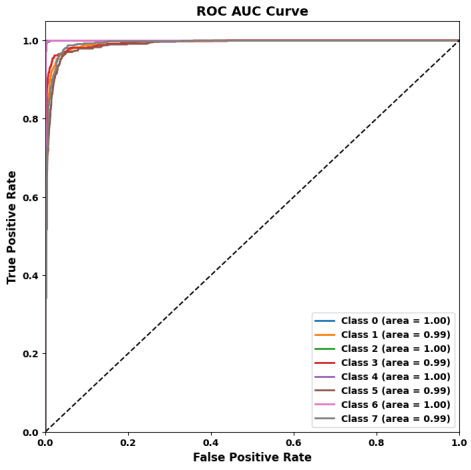


Figure 5. ROC curves for the proposed CNN–LSTM model across retinal disease categories.

The reported results attribute this discrimination to the combination of spatial representation from CNN layers and feature-sequence relationships modeled by the LSTM.

Figure 6 shows weaker ROC performance for Xception, consistent with its 47.1% test accuracy.

Figure 7 shows reasonable but variable discrimination across classes for InceptionV3, whose overall accuracy is 76.0%.

MobileNetV2 provides intermediate ROC performance and 74.3% accuracy (Figure 8); its lightweight design offers computational efficiency, although the reported curves are less

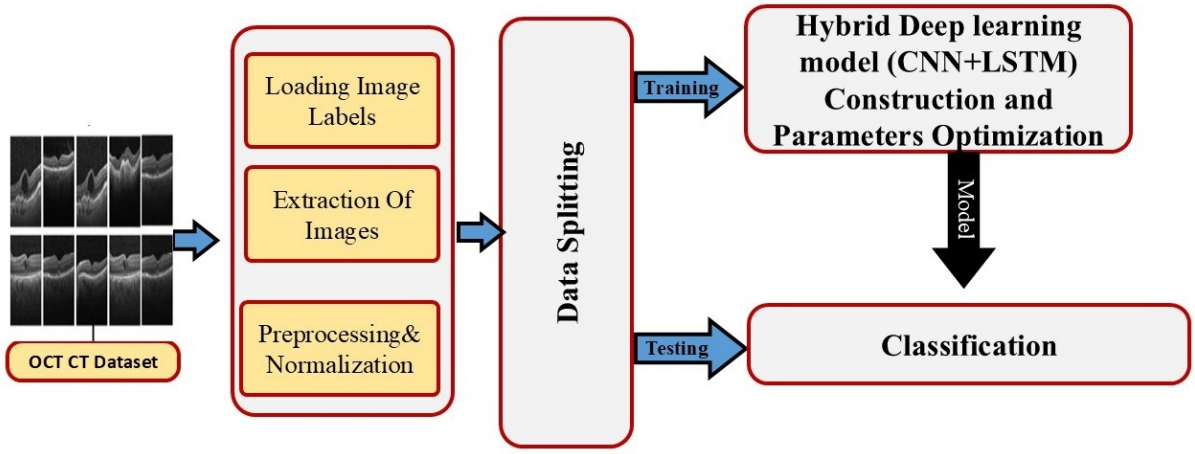


Figure 2. Overview of the proposed methodology.



Figure 3. Training and validation loss curves for the CNN–LSTM model.

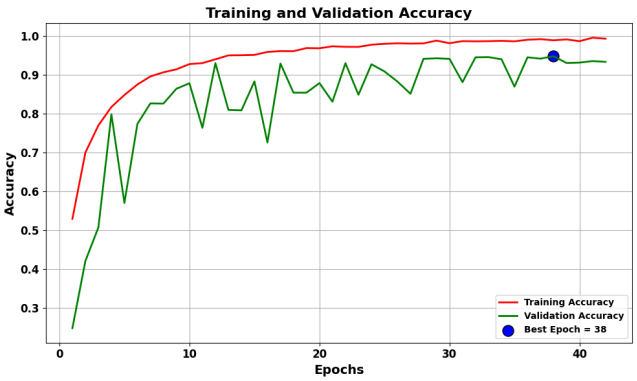


Figure 4. Training and validation accuracy curves for the CNN–LSTM model.

consistent across classes than those of the hybrid model.

ResNet50 has the strongest aggregate result among the comparison architectures, reaching 84.5% accuracy, and its ROC curves show generally good class discrimination (Figure 9). The proposed CNN–LSTM nevertheless outperforms all four comparison models on the reported aggregate metrics. The study attributes this result to the hybrid architecture's task-specific design and the combination of spatial feature extraction with sequence modeling.

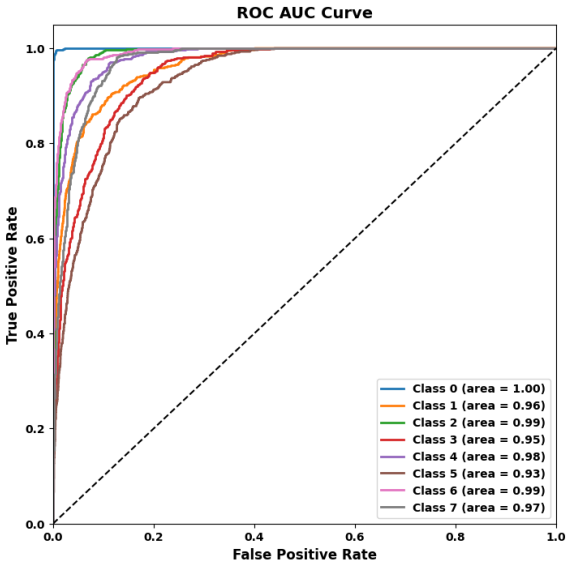


Figure 6. ROC curves for Xception on retinal OCT images.

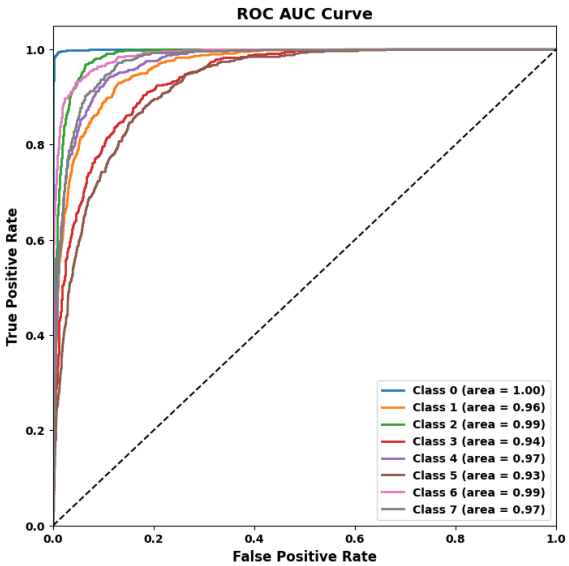


Figure 7. ROC curves for InceptionV3 across retinal disease categories.

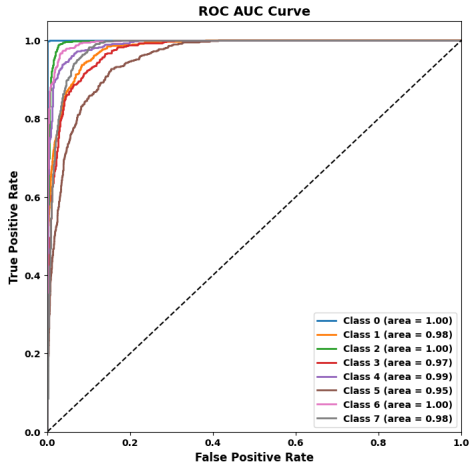


Figure 8. ROC curves for MobileNetV2 across retinal disease categories.

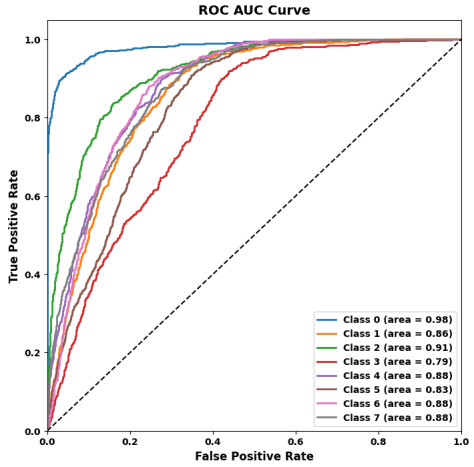


Figure 9. ROC curves for ResNet50 across retinal disease categories.

5. CONCLUSION

This paper presents a hybrid CNN–LSTM architecture for automated classification of retinal disease from OCT images. CNN branches learn spatial features, while the LSTM models relationships within the extracted feature sequence. Evaluation on a dataset of 24,000 images across eight classes yields 93% accuracy, 94% precision, 93% recall, and a 0.93 F1-score, exceeding the reported results for ResNet50, InceptionV3, MobileNetV2, and Xception.

The approach combines transfer learning with sequence modeling. Fine-tuning pretrained image models on OCT data provides a way to adapt existing visual representations to retinal texture and structural patterns, while the hybrid network performs automated feature learning without handcrafted descriptors. The reported classification scope includes AMD, CNV, DR, DME, CSR, drusen, MH, and normal retina.

The authors identify potential clinical value in settings with limited access to ophthalmological expertise, where automated analysis may support faster and more standardized assessment. These results should be interpreted as an evaluation on the reported dataset; broader external validation across diverse imaging devices, patient populations, and clinical environments is needed to establish generalizability and practical utility.

REFERENCES

[1] S. Pang, B. Zou, X. Xiao, Q. Peng, J. Yan, W. Zhang, and K. Yue, "A novel approach for automatic classification of macular degeneration oct images," *Scientific Reports*, vol. 14, p. 19285, Aug. 2024.

[2] Z. Tan, Q. Zhang, G. Lan, J. Xu, C. Ou, L. An, J. Qin, and Y. Huang, "Oct retinopathy classification via a semi-supervised pseudo-label sub-domain adaptation and fine-tuning method," *Mathematics*, vol. 12, no. 2, p. 347, Jan. 2024.

[3] K. Du, S. Shah, S. C. Bollepalli, M. N. Ibrahim, A. Gadari, S. Sutharahan, J.-A. Sahel, J. Chhablani, and K. K. Vupparaboina, "Inter-rater reliability in labeling quality and pathological features of retinal oct scans: A customized annotation software approach," *PLOS ONE*, vol. 19, no. 12, p. e0314707, Dec. 2024.

[4] C. Ou *et al.*, "A deep learning network for accurate retinal multidisease diagnosis using multiview fusion of en face and b-scan images: A multicenter study," *Translational Vision Science & Technology*, 2024.

[5] M. Guo, D. Gong, and W. Yang, "In-depth analysis of research hotspots and emerging trends in ai for retinal diseases over the past decade," *Frontiers in Medicine*, 2024.

[6] A. M. Alenezi *et al.*, "Multiscale attention-over-attention network for retinal disease recognition in oct radiology images," *Frontiers in Medicine*, 2024.

[7] J. Bell *et al.*, "Validation of inter-reader agreement/consistency for quantification of ellipsoid zone integrity and sub-rpe compartmental features across retinal diseases," *Diagnostics*, 2024.

[8] Q. Zuo *et al.*, "Multi-resolution visual mamba with multi-directional selective mechanism for retinal disease detection," *Frontiers in Cell and Developmental Biology*, 2024.

[9] J. Yang *et al.*, "Dry age-related macular degeneration classification from optical coherence tomography images based on ensemble deep learning architecture," *Frontiers in Medicine*, 2024.

[10] A. Domalpally *et al.*, "Data harmonization, standardization, and collaboration for diabetic retinal disease research: Report from the 2024 mary tyler moore vision initiative workshop on data," *Translational Vision Science & Technology*, 2024.

[11] M. Kulyabin *et al.*, "Segment anything in optical coherence tomography: Sam 2 for volumetric segmentation of retinal biomarkers," *Bioengineering*, 2024.

[12] J. Zhang, P. Zhao, Y. Zhao, C. Li, and D. Hu, "Few-shot class-incremental learning for retinal disease recognition," *IEEE Journal of Biomedical and Health Informatics*, 2024.

[13] D. J. Williamson *et al.*, "Artificial intelligence to facilitate clinical trial recruitment in age-related macular degeneration," *Ophthalmology Science*, 2024.

[14] F. H. Rizk, M. E. Mohamed, B. Sameh, A. M. Zaki, M. M. Eid, and E.-S. M. El-Kenawy, "Predictive modeling of portuguese student performance: Comparative machine learning analysis," in *2024 International Telecommunications Conference (ITC-Egypt)*, pp. 26–31, 2024.

[15] J. R. Taylor, J. Drinkwater, D. C. Sousa, V. Shah, and A. W. Turner, "Real-world evaluation of retcad deep-learning system for the detection of referable diabetic retinopathy and age-related macular degeneration," *Clinical and Experimental Optometry*, 2024.

[16] A. A. A. El-Khalek *et al.*, "A comprehensive review of ai diagnosis strategies for age-related macular degeneration," *Bioengineering*, 2024.

[17] A. Laouarem, C. Kara-Mohamed, E.-B. Bourennane, and A. Hamdi-Cherif, "Htc-retina: A hybrid retinal diseases classification model using transformer-convolutional neural network from optical coherence tomography images," *Computers in Biology and Medicine*, 2024.

[18] J. Zhao, J. Zhu, J. He, G. Cao, and C. Dai, "Multi-label classification of retinal diseases based on fundus images using resnet and transformer," *Medical and Biological Engineering and Computing*, 2024.

[19] S. V. Priyan, R. V. Kumar, C. Moorthy, and V. S. Nishok, "A fusion of deep neural networks and game theory for retinal disease diagnosis with oct images," *Journal of X-ray Science and Technology*, 2024.

[20] J. Wang, H. Peng, S. Chen, and S. Ren, "Ensemble learning for retinal disease recognition under limited resources," *Medical and Biological Engineering and Computing*, 2024.

[21] J. Carmichael, E. Costanza, A. Blandford, R. Struyven, P. A. Keane, and K. Balaskas, "Diagnostic decisions of specialist optometrists exposed to ambiguous deep-learning outputs," *Scientific Reports*, 2024.

[22] A. J. Prabha, C. Venkatesan, M. S. Fathimal, K. K. Nithiyanantham, and S. P. A. Kirubha, "Rd-oct net: Hybrid learning system for automated diagnosis of macular diseases from oct retinal images," *Biomedical Physics and Engineering Express*, 2024.

[23] P. D. Angeli *et al.*, "Splicing defects and crispr-cas9 correction in isogenic homozygous photoreceptor precursors harboring clustered deep-intronic abca4 variants," *Molecular Therapy – Nucleic Acids*, 2024.

[24] D. Wang, J. Lian, and W. Jiao, "Multi-label classification of retinal disease via a novel vision transformer model," *Frontiers in Neuroscience*, 2024.

[25] J. Peng, J. Lu, J. Zhuo, and P. Li, "Multi-scale-denoising residual convolutional network for retinal disease classification using oct," *Sensors*, 2024.

[26] A. Torimura *et al.*, "Profiling mirnas in tear extracellular vesicles: A pilot study with implications for diagnosis of ocular diseases," *Japanese Journal of Ophthalmology*, 2023.

[27] M. Bhandari, T. B. Shahi, and A. Neupane, "Evaluating retinal disease diagnosis with an interpretable lightweight cnn model resistant to adversarial attacks," *Journal of Imaging*, 2023.

[28] Y. Yan *et al.*, "Clinical evaluation of deep learning systems for assisting in the diagnosis of the epiretinal membrane grade in general ophthalmologists," *Eye*, 2023.

[29] B. A. Hammou, F. Antaki, M.-C. Boucher, and R. Duval, "Mbt: Model-based transformer for retinal optical coherence tomography image and video multi-classification," *International Journal of Medical Informatics*, 2023.

[30] P. Dutta, K. A. Sathi, M. A. Hossain, and M. A. A. Dewan, "Conv-vit: A convolution and vision transformer-based hybrid feature extraction method for retinal disease detection," *Journal of Imaging*, 2023.

[31] C.-T. Wu, T.-Y. Lin, C.-J. Lin, and D.-K. Hwang, "The future application of artificial intelligence and telemedicine in the retina: A perspective," *Taiwan Journal of Ophthalmology*, 2023.

[32] M. J. M. Zedan, M. A. Zulkifley, A. A. Ibrahim, A. M. Moubark, N. A. M. Kamari, and S. R. Abdani, "Automated glaucoma screening and diagnosis based on retinal fundus images using deep learning approaches: A comprehensive review," *Diagnostics*, 2023.

[33] S. A. Gocuk, J. K. Jolly, T. L. Edwards, and L. N. Ayton, "Female carriers of x-linked inherited retinal diseases: Genetics, diagnosis, and potential therapies," *Progress in Retinal and Eye Research*, 2023.

[34] X. Huang *et al.*, "Gabnet: Global attention block for retinal oct disease classification," *Frontiers in Neuroscience*, 2023.

[35] N. N. El-Den *et al.*, "Scale-adaptive model for detection and grading of age-related macular degeneration from color retinal fundus images," *Scientific Reports*, 2023.

[36] A. R. Chłopowiec, K. Karanowski, T. Skrzypczak, M. Grzesiuk, A. B. Chłopowiec, and M. Tabakov, "Counteracting data bias and class imbalance – towards a useful and reliable retinal disease recognition system," *Diagnostics*, 2023.

[37] M. A. Manan, F. Jinchao, T. M. Khan, M. Yaqub, S. Ahmed, and I. S. Chuhan, "Semantic segmentation of retinal exudates using a residual encoder-decoder architecture in diabetic retinopathy," *Microscopy Research and Technique*, 2023.

[38] M. D. Varela *et al.*, "Artificial intelligence in retinal disease: Clinical application, challenges, and future directions," *Graefe's Archive for Clinical and Experimental Ophthalmology*, 2023.

[39] R. Greeshma and P. Simon, "Deep learning approach based plant seedlings classification with xception model," in *International Conference on IoT Based Control Networks and Intelligent Systems*. Springer, pp. 65–76, 2023.

[40] A. G. Howard *et al.*, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," arXiv preprint, 2017.

[41] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, 2015.

[42] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826, 2016.

[43] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1251–1258, 2017.