International Journal of Medical Toxicology & Forensic Medicine
2026, Volume 16, E53260

Research Paper

A Hybrid Vgg19–Xgboost Adaptive Transfer Learning Framework for High-Accuracy Multiclass Retinal Disease Classification Using Optical Coherence Tomography

Suja Ganapathy Premalatha[1], Thanga Parvathi Balasubramanian[2], Pothumani Selvamuthukumaran[3], Dennis Ebenezer Dhanaraj[4], Palanisamy Brinda[5*], Sangita Gautam Lade[6]

1. *Department of Computer Science, Muslim Arts College, Thiruvithancode, Kanniyakumari, India.*
2. *Department of Computer Technology, Bannari Amman Institute of Technology, Sathyamangalam, India.*
3. *Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology, India.*
4. *Department of Computer Science and Engineering, Nehru Institute of Technology, Coimbatore, India.*
5. *Department of Computer Science and Engineering, Vel Tech High Tech Dr Rangarajan Dr Sakunthala Engineering College, Avadi, India.*
6. *Department of Computer Engineering, Vishwakarma Institute of Technology, Pune, India.*

Citation Ganapathy Premalatha S, Parvathi Balasubramanian Th, Pothumani S, Ebenezer Dhanaraj D, Brinda P, Gautam Lade S. A Hybrid Vgg19–Xgboost Adaptive Transfer Learning Framework for High-Accuracy Multiclass Retinal Disease Classification Using Optical Coherence Tomography. International Journal of Medical Toxicology and Forensic Medicine. 2026; 16:E53260.

https://doi.org/10.22037/ijmtfm.v16.53260

Article info:

Received: 24 August, 2026

Accepted: 14 September, 2026

Published: 30 October, 2026

Keywords:

Optical coherence tomography (OCT), Retinal disease classification, Hybrid deep learning, transfer learning, VGG-19, XGBoost, Gradient boosting, medical image analysis, Computer-aided diagnosis (CAD), Multiclass classification

ABSTRACT

Background: Early detection of retinal abnormalities is crucial to avoid irreversible visual loss and enable timely clinical interventions using Optical Coherence Tomography (OCT). Multiclass OCT classification remains difficult to automate because of inter-class similarity, limited samples for rare retinal diseases, and significant class imbalance.

Materials and Methods: This paper introduces a new adaptive transfer-learning method, HybVGG-XGB, which combines the fine-tuned VGG19 deep feature extraction with Extreme Gradient Boosting (XGBoost) to classify retinal images into multiclass categories. The proposed framework uses VGG19 to learn discriminative retinal representations and then transfers these features to XGBoost for nonlinear multiclass decision-making, unlike conventional transfer-learning methods that work directly with a fully connected or Softmax classification head.

Results: An experimental subset consisting of 1,618 OCT images from seven categories was used in the experiments: age-related macular degeneration (AMD), diabetic macular edema (DME), epiretinal membrane (ERM), normal (NO), retinal artery occlusion (RAO), retinal vein occlusion (RVO), and vitreomacular interface disease (VID). The proposed HybVGG-XGB framework achieved an accuracy of 88.48%, a macro-F1 score of 0.7890, a sensitivity of 0.7540, a specificity of 0.9782, an ROC-AUC of 0.9903, and a PR-AUC of 0.9342. The findings suggest that task-adaptable convolutional representations, combined with gradient-boosted classification, offer an innovative approach to retinal OCT classification under class imbalance and for computer-aided retinal screening.

Conclusion: The results provide preliminary evidence that combining convolutional feature extraction with gradient-boosted classification is a suitable approach for retinal OCT analysis. However, the degree of minority-class representation and issues of patient-level independence and statistical uncertainty make it difficult to draw conclusions about clinical screening utility.

*** Corresponding Author:**
P. Brinda, PhD
Department of Computer Science and Engineering, Vel Tech High Tech Dr Rangarajan Dr Sakunthala Engineering College, Avadi, India.
E-mail: *brindhaassistantprofessor@gmail.com*

Copyright © 2026 The Author(s);
This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0: https://creativecommons.org/licenses/by-nc/4.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Retinal diseases such as choroidal neovascularization and diabetic macular edema threaten sight and must be identified early to prevent permanent damage to the retina. Optical coherence tomography (OCT) is a noninvasive imaging technology with high-resolution imaging of retinal microstructures that can be used to assess fluid accumulation, neovascularization, and macular thickening, but it requires expertise and time. Multimodal structural and vascular imaging shows subtle retinal changes associated with systemic and degenerative diseases [1], and molecular and clinical biomarkers emphasize the importance of diagnostic frameworks that can capture pathological multiscale variation [2,3]. Expert-level automated feature extraction and retinal disease detection have been made possible by deep learning [4], and benchmarks such as RFMiD have driven multiclass retinal classification [5]. However, generalization issues, missing annotated data, and domain variations result in class imbalance; data harmonization and cross-institutional robustness are required [6]. Continual learning, distributed training, and foundation models offer opportunities for improvement due to their adaptability and scalability, but their computational and data requirements can be challenging to apply [7]. Metaheuristic optimization is also used to study convergence and parameter-tuning problems [8]. To address retinal image analysis, Hu et al. [9] proposed a multiscale CNN based on a modified cross-entropy loss function for vessel segmentation, highlighting the importance of multiscale feature learning and loss-function optimization for this problem. Kaushik et al. [10] adopted residual networks for DR detection and implemented skip connections to improve feature flow and training stability. The residual network (ResNet) [11], introduced by He et al. in recent years, uses skip connections to optimize deeper networks. The next step towards automated diagnosis is to apply deep learning based on OCT to the diagnosis of treatable retinal diseases, as Kermany et al. [12] demonstrated, achieving very high accuracy. Khan et al. [13] also proposed a task-specific architecture called RenalSegNet for kidney tumor segmentation in computed tomography (CT) images, highlighting the versatility of task-specific architectures for segmentation. Nagel et al. [14] pointed out complementary information in two imaging modalities, OCT and scanning laser ophthalmoscopy (SLO), and the need to develop computational methods that account for the characteristics of both. Vidiveli et al. [15] tuned hyperparameters for classifying ophthalmic images, while El Harti et al. [16] compared results from deep learning and traditional machine learning approaches and found that a hybrid classification approach can be applied with limited data. These studies indicate a need to explore more architectures that leverage transferable deep representations alongside regularised non-linear classifiers in imbalanced and heterogeneous data contexts. In this context, this study proposes a Hybrid VGG19-XGB (HybVGG-XGB) Adaptive Transfer Learning Framework for multiclass classification of retinal OCT, where VGG19 is adapted to learn task-specific hierarchical features used by the gradient-boosting classifier XGB for classification across different style-representation learning and final decision modeling. These morphology-preserving preprocessing and augmentations help generalization and are compared with conventional transfer-learning baselines to test the hybridization effect (e.g., ResNet50). Evaluation includes accuracy, macro-F1, sensitivity, specificity, ROC-AUC, and PR-AUC to measure the overall discrimination and performance under class imbalance. The aim of the framework is to increase diagnostic accuracy and confidence; model transfer between datasets must be evaluated. In addition, it could enable the analysis and presentation of connected ophthalmic screening in Internet of Things healthcare workflows for OCT scans; further testing of test reliability, memory requirements, communication overhead, and inference latency would be needed to demonstrate deployability readiness.

Materials and Methods

The proposed HybVGG-XGB construct is run and tested using the OCTDL dataset [17], a publicly available repository of labeled retinal images belonging to several pathological groups and provided in Optical Coherence Tomography. The dataset includes a subset of clinically annotated OCT scans covering a wide range of retinal pathologies, facilitating robust multiclass analysis under realistic clinical variability. The proposed HybVGG-XGB is a hybrid model that extracts features from retinal OCT images using the VGG19 architecture and classifies them using XGBoost. The OCTDL dataset contains seven disability categories: age-related macular degeneration (AMD), diabetic macular edema (DME), epiretinal membrane (ERM), normal retina (NO), retinal artery occlusion (RAO), retinal vein occlusion (RVO), and Vitreomacular interface disease (VID). A total of 2,064 images of 821 patients were obtained. The class distribution is also highly unbalanced, with 1231 images of AMD and 22 images of RAO. Therefore, assessment should not rely only on overall accuracy; class-specific performance and macro-averaged metrics should also be considered. Before training the

model, an extensive preprocessing pipeline is used to ensure consistent input and enhance generalization performance. All OCT images are downsampled to a uniform spatial resolution of 224 x 224 pixels and converted to a three-channel RGB format so they fit the pre-trained VGG-19 architecture. Pixel intensity values are scaled to the range [0,1] to stabilize gradient propagation and speed up convergence. To reduce overfitting and increase robustness to intra-class variability, data augmentation methods such as controlled rotation, horizontal flipping, brightness manipulation, zoom scaling, and spatial shifting are used during training. Table 1 summarizes the OCTDL dataset used in this paper.

Table 1. Summary of OCTDL Dataset Characteristics [17].

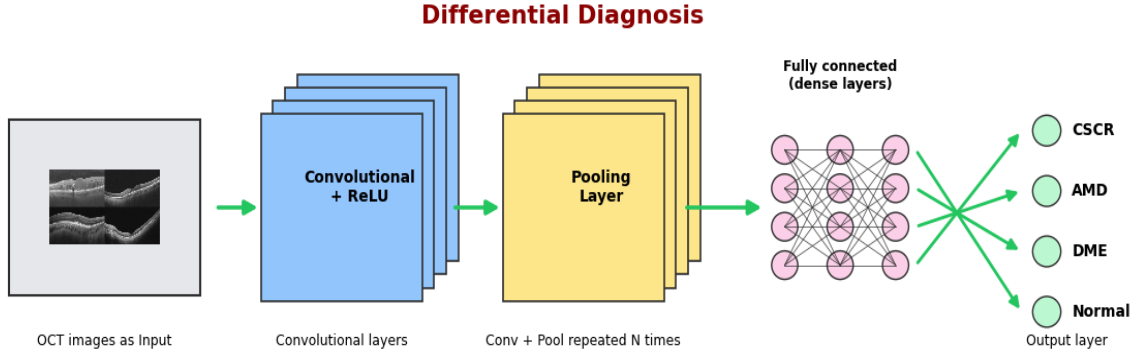
Category	Number of Images
Age-Related Macular Degeneration (AMD)	1231
Diabetic Macular Edema (DME)	147
Epiretinal Membrane (ERM)	155
Normal	332
Retinal Artery Occlusion (RAO)	22
Retinal Vein Occlusion (RVO)	101
Vitreomacular Interface Disease (VMID)	76
Total	**2064**

International Journal of
Medical Toxicology & Forensic Medicine

The original OCTDL collection was created for retinal OCT image analysis and comprises seven main diagnostic categories: Age-related macular degeneration (AMD), Diabetic macular edema (DME), Epiretinal membrane (ERM), Normal retina (NO), Retinal artery occlusion (RAO), Retinal vein occlusion (RVO), and Vitreomacular interface disease (VID). The entire OCTDL database has 2,064 OCT scans from 821 patients. The original distribution comprises 1,231 AMD, 147 DME, 155 ERM, 332 normal, 22 RAO, 101 RVO, and 76 VID images. The dataset is therefore highly skewed. AMD accounts for about 59.64% of the available scans, while RAO accounts for about 1.07%. There is a large disparity between the largest class (AMD) and the smallest class (RAO), i.e., 55.95:1, which highlights the challenge of assessing the performance of the seven-class classifier using overall accuracy. Figure 1 demonstrates the proposed model's workflow. Experiments were carried out with images from the publicly available Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods (OCTDL).

Results

The dataset used in the experiments is OCTDL, which includes seven categories: AMD, DME, ERM, NO, RAO, RVO, and VID. Experiments were reported with class-stratified training, validation, and test partitions. Class stratification is not sufficient for patient-level separation, however. OCTDL is a multi-image dataset from a single patient; patient identifiers need to be documented across partitions to ensure independence. This condition is not shown in the present reporting. In this regard, patient-level leakage is possible and cannot be ruled out; therefore, the reported performance must be considered preliminary until the partitioning procedure is verified. During training, images were pre-processed, resized to 224 × 224 pixels, normalized to the range [0,1], and augmented with rotations, flips, and zooms. HybVGG-XGB used an ImageNet-pretrained VGG19 backbone, with the upper convolutional layers unfrozen for domain adaptation purposes. We used Adam, categorical cross-entropy loss, and early stopping for training. Deep features were extracted from the images and fed to an optimized XGBoost classifier, with



International Journal of
Medical Toxicology & Forensic Medicine

Figure 1. Workflow diagram of the proposed model.

adaptive probability fusion used to obtain the final class label probabilities. Evaluation comprised ACC, macro-F1, SEN, SPE, ROC-AUC, and PR-AUC. Figure 2 shows the training loss and accuracy curves. The loss function shows smooth convergence, and the training and validation losses are close, indicating stable optimization and limited overfitting; however, the model still needs to be tested independently for generalization.

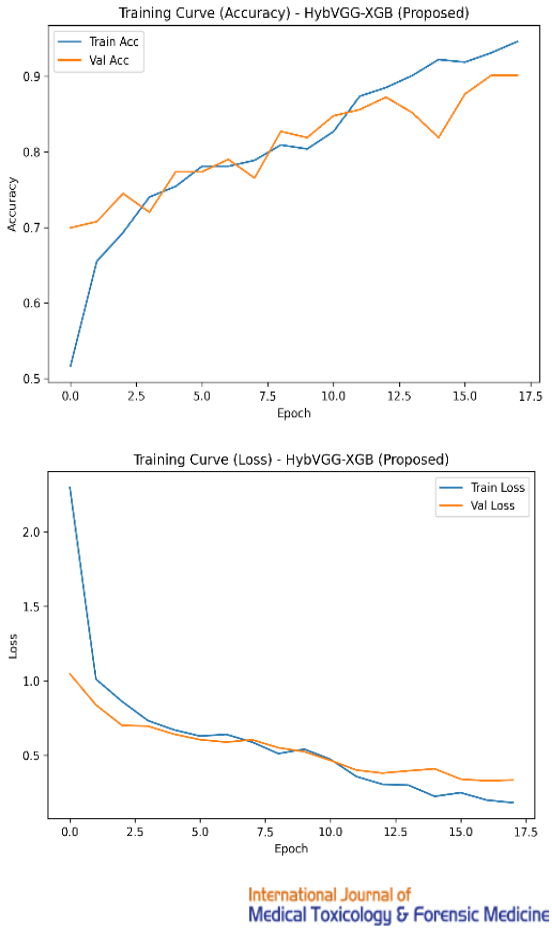


Figure 2. Accuracy and Loss curve of the proposed HYBVGG-XGB model.

Similarly, the training and validation accuracy curves increase steadily with epochs, reaching high performance levels. The error between generalization (i.e., a small generalization gap (~4-5)) is a sign of strong feature learning and good regularization. Figure 3 (Overall Metrics) represents the general performance indicators of the proposed HYBVGG-XGB model.

Performance of the proposed HybVGG-XGB model is also further considered in terms of classification in terms of confusion matrix and normalized confusion matrix as shown in Figure 4. The analysis of the confusion matrix, in general, speaks in favor of the performance of the proposed model framework and correlates with the obtained accuracy of 88.48, macro-F1 score of 78.90, sensitivity of 75.40, and specificity of 97.82, and the high ROC-AUC score of 0.9903 means that the framework is highly separable across classes. These findings indicate that the proposed approach can achieve strong classification performance with minimal false positives, especially in a clinical screening context.

There was a wide range of performance among classes. Sensitivity values were 0.9398, 0.9524, and 0.8000 for RAO, VID, and ERM, respectively, with lower values for AMD and normal images. The sensitivity of RAO was 0.5000 (two images correctly classified from four test images), and the sensitivity of VID was 0.6250 (five images correctly classified from eight test images). The small denominators make these estimates very sensitive to individual misclassifications. Although precision was 1.0000 for both RAO and VID, these values were based on only two and five positive predictions, respectively. Therefore, these results cannot be trusted to provide adequate minority-class detection and should be interpreted alongside sample sizes and patient-aware uncertainty estimates (Figure 5).

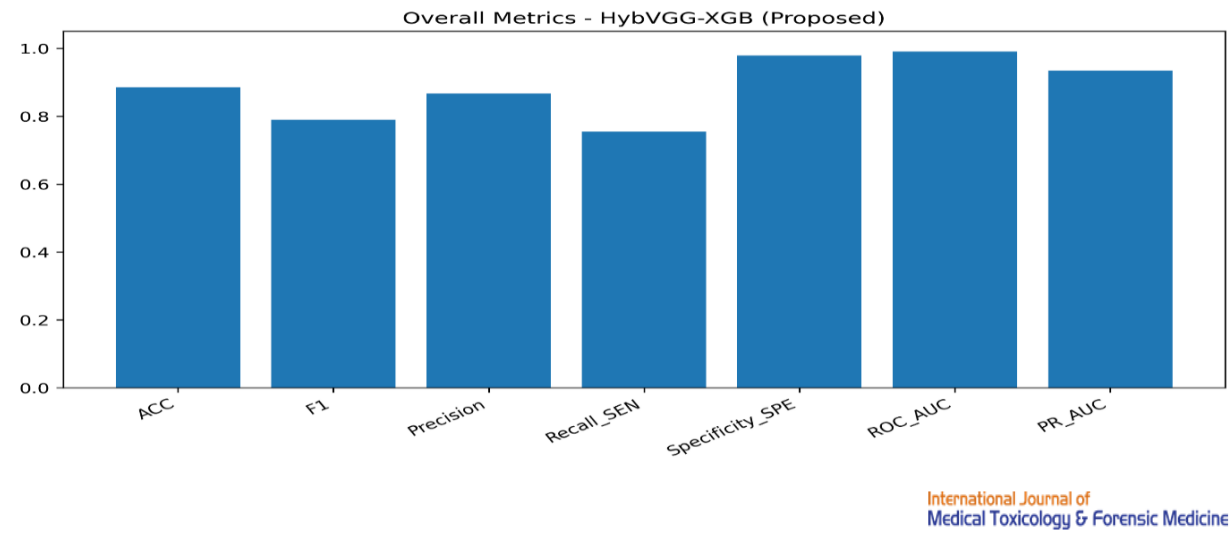


Figure 3. Overall metrics of the proposed HYBVGG-XGB model.



Figure 4. Confusion metrics of the proposed HYBVGG-XGB model.

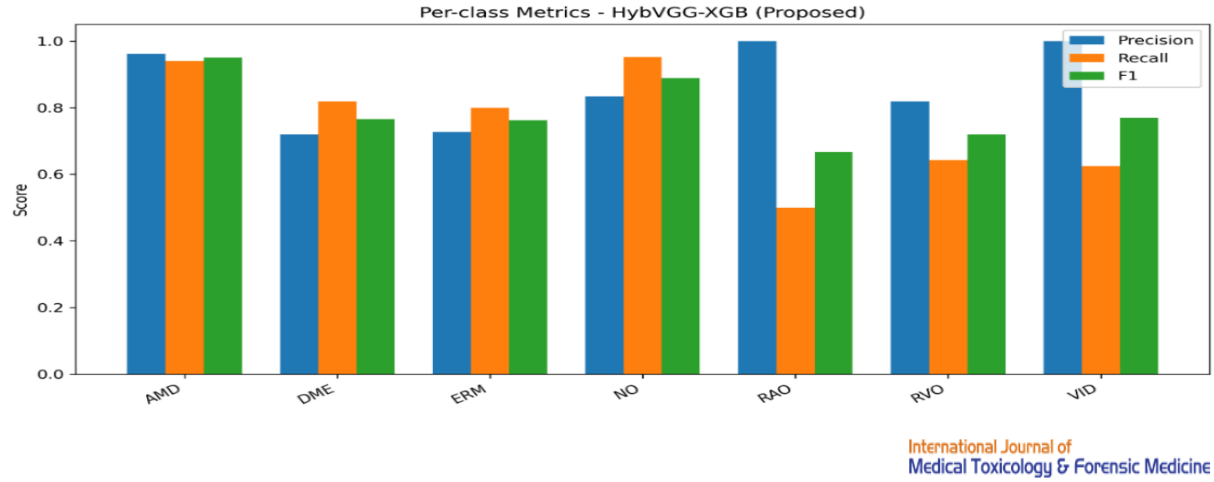


Figure 5. Per-class comparisons of precision, Recall, and F1 score of the proposed HYBVGG-XGB model.

ROC-AUC and macro-F1 are two different measures of model performance. In a 1 vs rest analysis, the continuous prediction score is used to compare each diagnostic category to all the other diagnostic categories. Macro-averaged ROC-AUC gives equal weight to the class-specific ROC AUC; micro-averaged ROC-AUC contains the image–class label and score pairs. Sensitivity and F1, on the other hand, are based on final class assignment. Thus, it is possible to have positive ranking while maintaining low sensitivity and macro-F1. The reported AUC should not be used as an

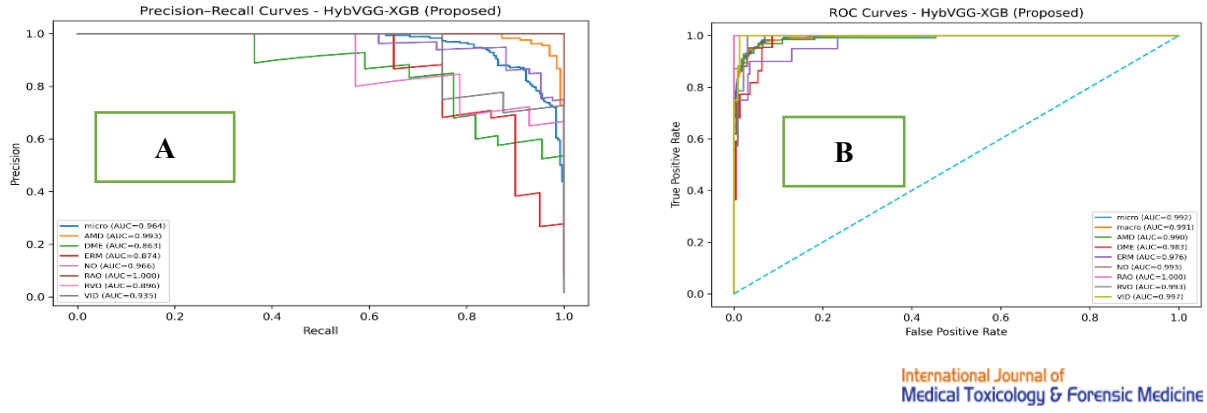
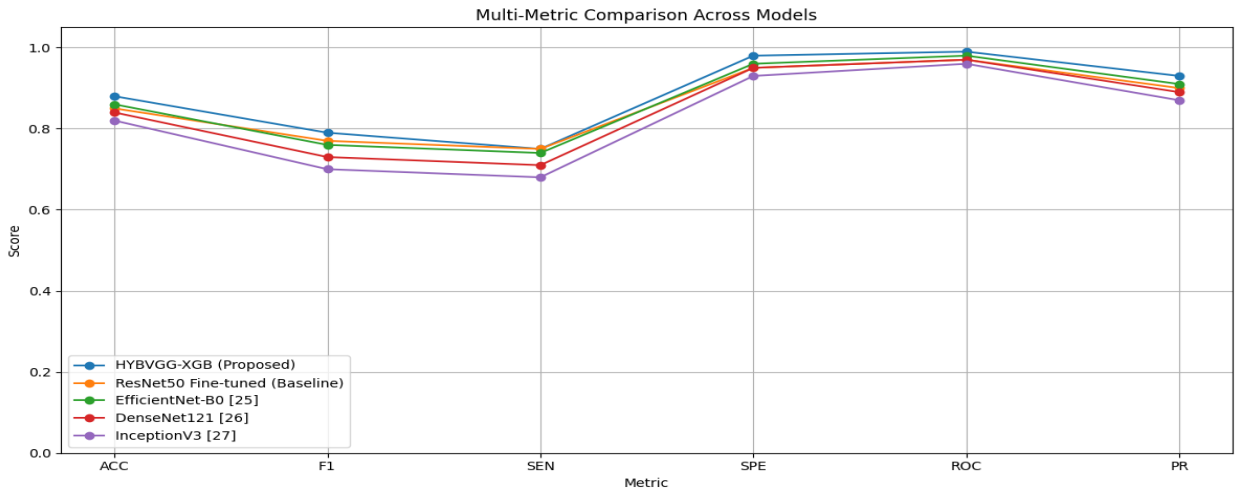
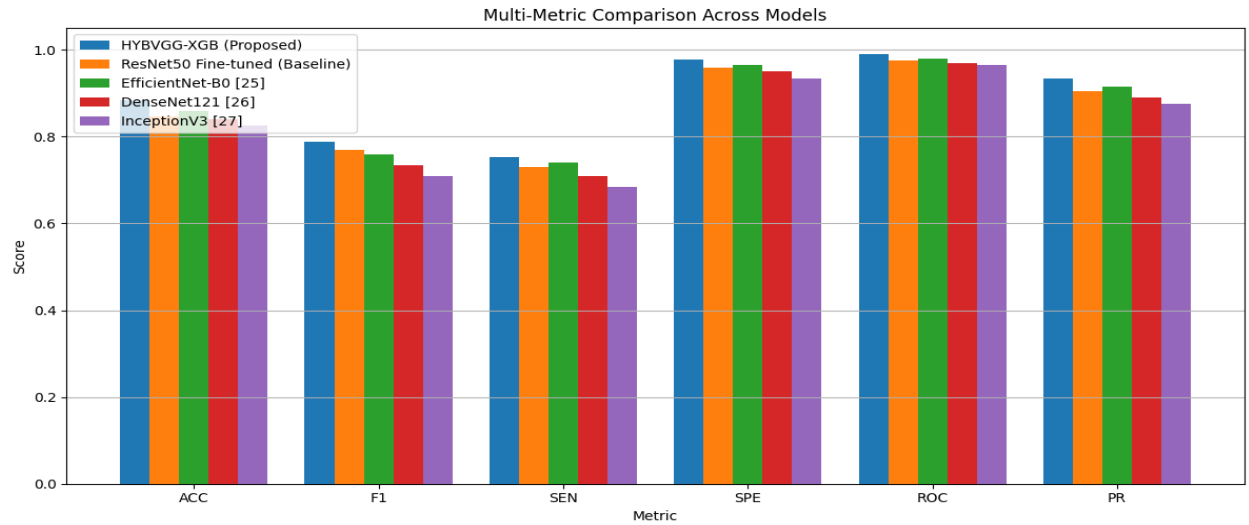


Figure 6. Precision–Recall (PR) and Receiver Operating Characteristic (ROC) curves of the proposed HYBVGG-XGB model.



International Journal of Medical Toxicology & Forensic Medicine

Figure 7. Multi-model comparisons across models.



International Journal of Medical Toxicology & Forensic Medicine

Figure 8. Grouped comparison of all existing methods.

indication of consistently high detection rates for all seven categories, especially for minority-class sample sizes**.** Figure 6 (A) shows that the ROC analysis yields micro-AUC and macro-AUC of 0.992 and 0.991, respectively, indicating good separability among the classes. Similarly, the Precision-Recall curves in Figure 6 (B) also show high performance, especially in the AMD and NO classes, demonstrating the model's ability to address class imbalance. The high ROC and PR values across all categories prove the strength and credibility of the proposed model in differentiating the classes.

The performance of HybVGG-XGB and the comparison models is summarized in Figures 7 and 8, respectively. These numerical comparisons are descriptive; the current reporting does not imply that the models used have the same patient-level partitions or development and evaluation processes. The results shown may vary because of differences in sample selection, processing, augmentation, and hyperparameter optimization. Therefore, these comparisons are not meant to be statistically superior. Direct comparative conclusions must be based on the same held-out patients, report the processing requirements unique to each model, and provide paired estimates of model-specific performance differences with their attendant uncertainty (Table 2).

Discussion

The reported results provide an exploratory

Table 2. Classification report of the proposed model.

Class	Precision	Recall (SEN)	F1-Score	Support
AMD	0.9615	0.9398	0.9506	133
DME	0.7200	0.8182	0.7660	22
ERM	0.7273	0.8000	0.7619	20
NO	0.8333	0.9524	0.8889	42
RAO	1.0000	0.5000	0.6667	4
RVO	0.8182	0.6429	0.7200	14
VID	1.0000	0.6250	0.7692	8
Accuracy	—	—	**0.8848**	243
Macro Avg	0.8658	0.7540	0.7890	243
Weighted Avg	0.8919	0.8848	0.8837	243

International Journal of
Medical Toxicology & Forensic Medicine

evaluation of using VGG19 features with XGBoost to classify retinal OCT images. They don't show statistically significant superiority over competing methods or show clinical screening utility. The gap between overall accuracy and macro-F1 scores suggests that overall performance is not a good measure of performance across unbalanced diagnostic categories. The study did not assess operating thresholds, clinically justified thresholds, or probability calibration. The study therefore does not allow determination of the reliability of the predicted probabilities for clinical decision-making or whether care based on these models would be of net benefit. For clinical use, it is necessary to have a clear, easily understood intent, a defined target population, and external validation across relevant institutions and imaging devices. Combining deep feature fine-tuning with gradient-boosting optimization leads to higher classification accuracy, better detection of minority classes, and strong generalization. These findings confirm the usefulness of the proposed architecture for multiclass retinal OCT disease classification and for an automated clinical decision support system. The high specificity suggests a low false-negative rate across all disease categories, and the ROC-AUC confirms strong overall discriminative ability. The differences in overall accuracy and macro-F1 also highlight that class distribution is important to consider. Since the disease class distribution is uneven, accuracy can be more strongly influenced by larger classes, while macro-F1 gives each class equal weight. Thus, these three values, macro-F1, sensitivity, and PR-AUC, are significant complementary metrics for performance on less well-represented retinal diseases. The results will be evaluated in relation to a retrospective study (on a selected subset of the OCTDL). Results on this dataset are not necessarily representative of performance at other institutions, with other imaging devices, or in different screening populations. In some disease classes, the number of test samples is small, which diminishes the accuracy of class performance estimates. The internal test set contained 243 images, including only four RAO images and eight VID images. For these classes, changing the classification of one image results in 25.0 and 12.5 percentage-point changes in sensitivity, respectively. Estimates are therefore subject to change by class and are considered exploratory. Repeated patient-grouped cross-validation would be useful to gauge sensitivity to data partitioning, with the caveat that all images for a patient are in a single fold and that model selection is limited to the images in the appropriate development folds. The number of folds should be proportional to the number of unique patients in the least represented categories. This re-testing does not add more independent patients and will not reduce the uncertainty resulting from the limited minority-class representation.

Conclusion

In this study, we proposed a hybrid adaptive transfer-learning framework for seven-class classification of retinal diseases based on OCT images, named HybVGG-XGB. The framework features fine-tuned VGG19 deep representation learning and XGBoost nonlinear multiclass classification. In contrast, the conventional transfer learning model with a CNN classification head is replaced by HybVGG-XGB, where the deep network and the gradient-boosted classifier play complementary roles: the deep network provides features, and the gradient-boosted classifier provides the final classification. The proposed model achieved an overall accuracy of 88.48%, a macro-F1 score of 0.7890, a sensitivity of 0.7540, a specificity of 0.9782, a ROC-AUC of 0.9903, and a PR-AUC of 0.9342. The ensemble evaluation highlights the need to incorporate class-sensitive metrics when evaluating retinal disease classification tasks under an imbalanced multiclass distribution. The results suggest that deep-feature extraction with gradient boosting is promising for computer-aided OCT screening and retinal disease decision support. However, the present study has some limitations: evaluation in a single dataset and a relatively small number of samples in certain retinal disease categories. Larger, multicenter, multi-device OCT datasets are needed to further investigate and establish clinical generalizability. Future work will also focus on enhancing minority-class learning, model interpretability, calibration, and prospective clinical evaluation. HybVGG-XGB should thus be considered

as a potentially useful computer-aided screening system to complement, but not supplant, specialist ophthalmological screening.

Ethical Approval

Not applicable.

Acknowledgment

None.

Funding

No external funding was received for this research.

Conflicts of Interest

The authors report there are no competing interests to declare.

References

[1] Agra CLM, Ventura CV, Da Fonte LD, Oliveira MRV, Leal M, Assunção MELSM, et al. Structural and vascular assessment of the optic nerve head and macula in chronic Chagas disease. Sci Rep. 2025;15(1). [DOI: 10.1038/s41598-025-86582-y]

[2] Bruno P, Quarta A, Calimeri F. Continual learning in medicine: a systematic literature review. Neural Process Lett. 2025;57(1). [DOI: 10.1007/s11063-024-11709-7]

[3] Camacho P, Ribeiro E, Pereira B, Nascimento J, Rosa PC, Henriques J, et al. DNA methyltransferase expression (DNMT1, DNMT3A, and DNMT3B) as a potential biomarker in age-related macular degeneration. J Clin Med. 2025;14(2):559. [DOI: 10.3390/jcm14020559]

[4] Cen L, Ji J, Lin J, Ju S, Lin H, Li T, et al. Automatic detection of 39 fundus diseases and conditions in retinal photographs using deep neural networks. Nat Commun. 2021;12(1). [DOI: 10.1038/s41467-021-25138-w]

[5] Pachade S, Porwal P, Kokare M, Deshmukh G, Sahasrabuddhe V, Luo Z, et al. RFMiD: retinal image analysis for multi-disease detection challenge. Med Image Anal. 2025;99:103365. [DOI: 10.1016/j.media.2024.103365]

[6] Domalpally A, Fickweiler W, Levine SR, Goetz KE, VanderBeek BL, Lee A, et al. Data harmonization, standardization, and collaboration for diabetic retinal disease (DRD) research: report from the 2024 Mary Tyler Moore Vision Initiative Workshop on Data. Transl Vis Sci Technol. 2024;13(10):4. [DOI: 10.1167/tvst.13.10.4]

[7] Gharehchopogh FS, Ucan A, Ibrikci T, Arasteh B, Isik G. Slime Mould Algorithm: a comprehensive survey of its variants and applications. Arch Comput Methods Eng. 2023;30(4):2683-723. [DOI: 10.1007/s11831-023-09883-3]

[8] Gholami S, Jannat F, Thompson AC, Ong SSY, Lim JI, Leng T, et al. Distributed training of foundation models for ophthalmic diagnosis. Commun Eng. 2025;4(1). [DOI: 10.1038/s44172-025-00341-5]

[9] Hu K, Zhang Z, Niu X, Zhang Y, Cao C, Xiao F, et al. Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function. Neurocomputing. 2018;309:179-91. [DOI: 10.1016/j.neucom.2018.05.011]

[10] Kaushik K, Bhardwaj A, Cheng X, Dahiya S, Shankar A, Kumar M, et al. Residual network-based deep learning framework for diabetic retinopathy detection. J Database Manag. 2025;36(1):1-21. [DOI: 10.4018/JDM.368006]

[11] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016. p. 770-8. [DOI: 10.1109/CVPR.2016.90]

[12] Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell. 2018;172(5):1122-31. [DOI: 10.1016/j.cell.2018.02.010]

[13] Khan R, Chen C, Zaman A, Wu J, Mai H, Su L, et al. RenalSegNet: automated segmentation of renal tumor, veins, and arteries in contrast-enhanced CT scans. Complex Intell Syst. 2025;11(2). [DOI: 10.1007/s40747-024-01751-2]

[14] Nagel ID, Heinke A, Agnihotri AP, Yassin S, Cheng L, Camp AS, et al. Comparison of a novel ultra-widefield three-color scanning laser

ophthalmoscope to other retinal imaging modalities in chorioretinal lesion imaging. Transl Vis Sci Technol. 2025;14(1):11. [DOI: 10.1167/tvst.14.1.11]

[15] Vidivelli S, Padmakumari P, Parthiban C, DharunBalaji A, Manikandan R, Gandomi AH. Optimizing deep learning models for ophthalmological disorder classification. Sci Rep. 2025;15(1). [DOI: 10.1038/s41598-024-75867-3]

[16] El Harti M, Zaamoun S, Jai-Andaloussi S, Ouchetto O. Classification of eye disorders using deep learning and machine learning models. In: International Conference on Advances in Computing Research; 2024. p. 184-194. Cham: Springer Nature Switzerland. [DOI: 10.1007/978-3-031-56950-0_16]

[17] Kulyabin M, Zhdanov A, Nikiforova A, Stepichev A, Kuznetsova A, Borisov V, et al. OCTDL: Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods [dataset]. Mendeley Data. 2024;Version 4. [DOI: 10.17632/sncdhf53xc.4]