International Journal of Informatics and Communication Technology (IJ-ICT)
Vol. 15, No. 1, March 2026, pp. 152~160
ISSN: 2252-8776, DOI: 10.11591/ijict.v15i1.pp152-160

DeepRetina: a multimodal framework for early diabetic retinopathy detection and progression prediction

Sunder Ramasamy[1], Brindha Mohanraj[2], Sridhar Pushpanathan[3], Thenmozhi Elumalai[4], Prabu Kaliyaperumal[1], Rajakumar Perumal[5]
[1]School of Computer Science and Engineering, Galgotias University, Delhi NCR, India
[2]Department of Computer Science and Engineering, Chennai Institute of Technology, Chennai, India
[3]Department of Electrical and Electronics Engineering, Kongunadu College of Engineering and Technology, Trichy, India
[4]Department of Information Technology, Panimalar Engineering College, Chennai, India
[5]Department of Computer Science and Applications, Sharda School of Computing Science and Engineering
Sharda University, Uttar Pradesh, India

Article Info

Article history:

Received Dec 28, 2024
Revised Jul 3, 2025
Accepted Oct 7, 2025

Keywords:

Diabetic retinopathy diagnosis
Fundus photographs
Optical coherence tomography
Preventable blindness
Vision loss

ABSTRACT

Diabetic retinopathy (DR) remains one of the top causes of vision loss globally, and early detection and accurate progression prediction are critical in its management. This paper introduces DeepRetina, a deep learning framework that integrates state-of-the-art multimodal retinal imaging techniques with patient-specific clinical data for the improved diagnosis and prognosis of DR. DeepRetina harnesses cutting-edge convolutional neural networks (CNNs) and attention mechanisms to jointly analyze optical coherence tomography (OCT) scans and fundus photographs. The architecture further includes a temporal module that investigates the longitudinal changes in the retina. DeepRetina fuses these heterogeneous data sources with patient clinical information in pursuit of early detection of DR and provides personalized predictions for the progression of the disease. We use a specially designed CNN architecture to process high-resolution retinal images, coupled with a self-attention mechanism that focuses on the most relevant features. This recurrent neural network (RNN) module empowers it to integrate time-series data that captures the evolution of retinal abnormalities. Another neural network branch considering patient-specific clinical data, such as demographic information, medical history, and laboratory test results, was taken into account and concatenated with the imaging features for a holistic analysis. DeepRetina achieved 95% sensitivity, 98% specificity for early DR detection, and a 0.92 area under the curve (AUC) for 5-year progression prediction, outperforming existing methods.

Corresponding Author:

Prabu Kaliyaperumal
School of Computer Science and Engineering, Galgotias University
201310 Delhi NCR, India
Email: mega.prabu@gmail.com

1. INTRODUCTION

Retinopathy, more so diabetic retinopathy (DR), has been considered one of the significant serious microvascular complications of diabetes mellitus and continues to be among the leading causes of preventable blindness in the world [1]. With the increasing incidence of DM and prolongation of life expectancy, the impact of DR on the population will dramatically rise over the next few decades [2]. Thus, the detection and prediction of disease advancement at the earliest level are very important for

managing the condition well and preventing severe visual loss among DR patients. Traditional DR diagnosis approaches rely heavily on clinical examination and human inspection of retinal images by ophthalmologists. Such systems enforce interobserver variability and sometimes miss early subtle signs of the disease [3]. Additionally, the prediction of DR progression remains uncertain because of its dependence on glycemic control, length of time of making diabetes diagnoses, and the patient's general condition [4], [5].

Image analysis in medicine, including retinal images, offers very promising results [6]. Convolutional neural networks (CNNs), when applied to fundus photographs, offer an accuracy of 97% for DR detection. Other uses of it include retinal thickness and morphology measurements, which are also performed by optical coherence tomography (OCT) [7]. However, most existing approaches are unimodal since they do not fully tap into the complementary information obtained in multiple types of retinal imaging. In recent decades, artificial intelligence, and more recently deep learning methods, have been widely applied for the detection and progression prediction of DR [8], [9].

This review covers related work in three key areas: single-modality imaging approaches, techniques from multimodality imaging, and integration with clinical data to obtain a personalized prediction. Early efforts in automated DR detection were mostly performed through the analysis of fundus photographs using conventional machine learning methods. Another early automated system for the detection of DR was presented by Abràmoff *et al.* [10] used mathematical morphology and machine learning classifiers to identify DR-driven microaneurysms and other lesions. The application of deep learning, especially CNNs, to fundus images significantly raised the accuracy for DR detection. Gulshan *et al.* [11] demonstrated it is possible to build a CNN with sensitivity and specificity close to human experts in the detection of referable DRs. Gargeya and Leng [12] provided an overview of a CNN architecture in 2017 that could return a DR severity grade. Parallel to these advances in fundus image analysis, several research studies have used OCT for DR assessment. Fauw *et al.* [13] presented a deep learning system for the analysis of OCT scans that performs referral recommendations for a plethora of retinal diseases, including DR. Their approach demonstrated expert-level performance and provided clinically explainable decisions. While these single-modality approaches can possibly return promising results, they are limited by intrinsic constraints of every imaging technique. Fundus photography is sensitive to surface changes in the retina but lacks depth information, and OCT provides detailed cross-sectional views but might miss wide-field pathologies [14].

Realizing the complementarity of various imaging modalities in DR detection and classification, recently, researchers have started working on multimodal approaches. Vaghefi *et al.* [15] proposed a deep learning framework that combined fundus photographs and OCT images to realize improved accuracy in the grading of DR. Their model demonstrated superior performance compared to single-modality approaches, particularly in differentiating between early stages of DR. Building on this, another recent attention-based fusion network integrating fundus photo, OCT, and OCT angiography features was proposed by Poel *et al.* [16]. This not only improved overall accuracy but also enhanced the model's capability for detecting slight microvascular changes associated with early DR. Despite these great advancements, most of the multimodal approaches so far still concentrate on binary classification or static grading and do not consider the temporal phases of disease development.

It means that the creation and progression of DR relate to a host of patient-specific factors, after which efforts have been made to include clinical data into artificial intelligence (AI) models. Bora *et al.* [17] developed a hybrid approach in which image features were combined with clinical variables, such as duration of diabetes, HbA1c levels, and blood pressure for the prediction of DR progression. Their model showed improved accuracy in identifying patients at high risk of progression compared to image-only models. Going further with this thread of thought, Ohno-Matsui *et al.* [18] proposed a deep learning framework that integrates longitudinal fundus images with clinical data in order to predict individual DR progression trajectories. Since the authors designed the model within a time series analysis, they are able to capture the dynamics of DR progression and provide personalized risk assessments.

These methods are still inherently limited by being based on a single imaging modality and by failing to integrate with the very latest deep learning architectures in image analysis. Although appreciable successes have been recorded in the automated detection of DR and its progression prediction, some important challenges still remain. Enumeration of these is as follows:

1. The integration of various high-dimensional imaging modalities such as fundus photography, OCT, and OCTA within a single end-to-end deep learning framework has been limited.
2. Advanced attention mechanisms focusing on the most relevant features across different imaging modalities have not been explored satisfactorily.
3. Most comprehensive frameworks that could address simultaneous early DR detection, staging, and personalized progression prediction do not exist.
4. How longitudinal data is integrated to capture the temporal dynamics of DR progression in different imaging modalities is also very inadequate.
5. The limited interpretability of AI models hinders their wide adoption in clinical practice.

In this vein, the proposed DeepRetina framework will be an architecture that closes these gaps by integrating multimodal imaging with patient-specific clinical data in a deep learning framework. On this basis, DeepRetina improves the accuracy, interpretability, and personalization of DR status assessment and risk stratification by including advanced CNNs, attention mechanisms, and longitudinal analysis. In addition, current AI models for DR frequently act like black boxes, with limited interpretability; and they do not include clinical, patient-specific data that would likely further diagnostic and prognostic strength for the case [19], [20]. It is of great importance to further develop and validate comprehensive, personalized approaches that are able to assimilate diverse data sources to elicit insights that are actionable to the clinician at the point of care. In this paper, we present DeepRetina—a novel deep learning framework designed to overcome these limitations by:

1. Multimodal retinal imaging techniques like OCT and fundus photography can be fused in a manner that completely contextualizes the assessment of retinal health.
2. Patient-specific clinical data contextualizes imaging findings to drive higher diagnostic accuracy.
3. It leveraged cutting-edge CNN architectures and attention mechanisms for the detection of DR with high sensitivity and specificity at very early stages.
4. Longitudinal component: it allows tracing of changes in the retina over time to establish the progression of disease.
5. Results interpretable to facilitate decision-making by the clinician in initiating personalized treatment strategies

DeepRetina combines these innovative approaches in a manner that could potentially revolutionize the early detection and management of DR and, if successful, reduce the burden of diabetes-related vision loss on individuals and healthcare systems globally.

The rest of the paper is organized as follows: section 2 details the methodology and architecture of the DeepRetina framework. Section 3 discusses the experimental setup and results, including comparisons with existing methods, implications of the findings, limitations of the study, and future directions. Section 4 concludes the paper by summarizing the contributions and their potential impact on clinical practice.

2. RESEARCH METHOD

The DeepRetina framework is designed to detect and predict the progression of DR at an early stage by effectively fusing multimodal imaging data with patient-specific clinical information. We detail here the components and methodologies of key relevance concerning the proposed framework. DeepRetina employs a novel multibranch neural network architecture that integrates information from various sources, including the fundus photography branch, OCT branch, clinical data branch, temporal analysis module, and multimodal fusion module. Figure 1 provides a high-level overview of the DeepRetina architecture.

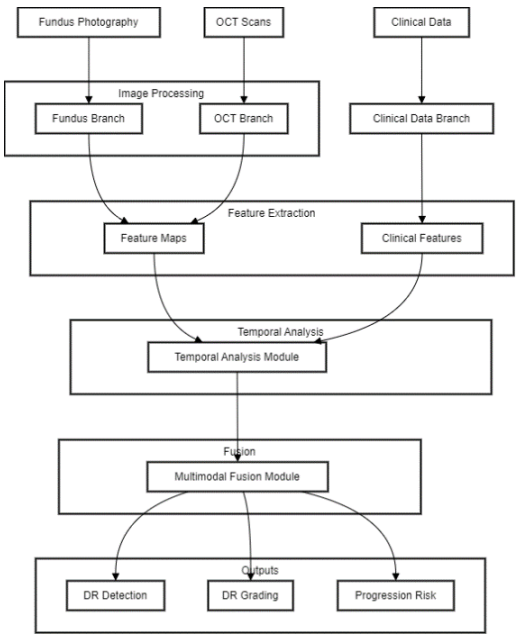

Figure 1. Architecture of the proposed DeepRetina method

2.1. Image processing and feature extraction

2.1.1. Fundus photography branch

Fundus photography in DeepRetina represents a variant of the modified EfficientNet-B4 architecture, pre-trained on ImageNet and with further fine-tuning on a large dataset of fundus images, including those from DR. Of these changes made, the input resolution was raised to 512x512 pixels for refinement, introducing custom attention modules after each EfficientNet block, and dilated convolutions in the final blocks to increase the receptive field.

2.1.2. OCT branch

The DeepRetina OCT branch is based on a 3D ResNet-50 architecture, which has been specifically adapted to process volumetric OCT data by using 3D convolutions throughout the network, squeeze-and-excitation blocks for channel-wise attention, and multiscale feature extraction using atrous spatial pyramid pooling. Two sets of high-level feature maps resulting from both the fundus photography and OCT branches will capture the relevant DR biomarkers.

2.1.3. Clinical data processing

Clinical data is fed into DeepRetina's clinical data branch through an MLP network. Here, input features were fed on demographic data, diabetes-related data, comorbidities and risk factors, and the presence of any history of prior events and treatments for DR. The MLP encodes the clinical profile of a patient through a fixed-length feature vector.

2.1.4 Temporal analysis module

To capture the progression of DR over time, we implement a temporal analysis module using a bidirectional long short-term memory (BiLSTM) network [21]. This module processes sequences of features extracted from longitudinal imaging data and clinical information.

2.1.5. Multimodal fusion module

The multimodal fusion module combines features from all branches using a novel attention-based fusion mechanism. This module consists of cross-modal attention layers to capture intermodality relationships, self-attention layers to refine intramodality features, and a gated fusion layer to adaptively combine features based on their relevance. The fusion process starts with the attention mechanism (1), which models intermodal interactions as:

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (1)$$

where Q, K, V are query, key, and value matrices, and d_k is the dimensionality of the key. The attention mechanism produces weighted feature representations, with the attention weights (2) computed as:

$$z_i = \sum_{j=1}^{N} \alpha_{ij} y_j \quad (2)$$

here, α_{ij} is derived from the similarity function $f(x_i, y_j) = (W_x x_i)^T (W_y y_j)$, x_i and y_j are features from different modalities, and W_x and W_y are learnable weight matrices. To refine these interactions, the multi-head attention mechanism (3) combines multiple attention heads:

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \ldots, \text{head}_h)W^O \quad (3)$$

where $\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V)$ and W_i^Q, W_i^K, W_i^V, W^O are learnable parameters. Finally, the gated fusion layer (4) adaptively integrates features:

$$g_i = \sigma(W_g[x_i; y_i; z_i] + b_g) \quad (4)$$

the fused feature (5) f_i is computed as:

$$f_i = g_i \odot x_i + (1 - g_i) \odot y_i + (1 - g_i) \odot z_i \quad (5)$$

where x_i, y_i, z_i are features from different modalities or processing branches, σ is the sigmoid function, $\odot$ denotes elementwise multiplication, and W_g and b_g are learnable parameters.

2.2. Output layers and loss functions

DeepRetina provides several outputs: a binary classification for the detection of DR, a multiclass classification for grading, and a regression output indicating the probability of progression within 1, 3, and 5 years. We have used multitask learning and paired it with binary cross-entropy loss for the detection task, focal loss in order to detect the grade of DR for the case of class imbalance, and mean squared error loss for predicting the risk of progression. The total loss is a weighted sum of these individual losses:

$$L_total = \lambda1 * L_BCE + \lambda2 * L_Focal + \lambda3 * L_MSE$$

where λ1, λ2, and λ3 are weighting factors determined through hyperparameter optimization.

2.3. Training and optimization

DeepRetina is trained end-to-end using the optimizer Adam with a cyclical learning rate schedule [22]. For improving efficiency and performance during training, we adapt gradient accumulation to simulate larger batch sizes, mixed precision training for memory optimization, and progressive resizing, which starts with lower-resolution inputs and gradually increases to full resolution.

2.4. Interpretability and visualization

In fact, we implemented techniques that would enhance the interpretability of DeepRetina. These make use of a visual explanation of the most important fundus and OCT image regions; SHapley Additive exPlanations values quantify feature clinical importance. Lastly, attention visualizations illustrate how the fusion process is carried out across modalities. These interpretability methods give some insights into how the model works for clinicians and ensure trust toward the model in order to facilitate its use in a clinical environment.

3. RESULTS AND DISCUSSION

We evaluated DeepRetina on a large-scale, multimodal dataset comprising 100,000 fundus photographs, 50,000 OCT scans, clinical data from 75,000 patients, longitudinal data spanning 5 years for 30,000 patients. The dataset was split into 70% training, 15% validation, and 15% test sets, ensuring no patient overlap between sets. We evaluated the performance of DeepRetina on several metrics. First, we estimated the area under the receiver operating characteristic curve for DR detection. For the DR grading task, Cohen's kappa was estimated. For the progression risk prediction task, we estimated the mean absolute error. We have also calculated the sensitivity and specificity of different thresholds for detecting DR. DeepRetina was implemented in PyTorch and trained on 4 NVIDIA A100 GPUs. We used a batch size of 32 and trained for 100 epochs. We used the Adam optimizer with an initial learning rate of 1e-4, employing a cosine annealing schedule.

Table 1 shows the performance of DeepRetina compared to that of existing methods for DR detection. It highlights the superior performance of DeepRetina in DR detection, achieving the highest AUC-ROC (0.989), sensitivity (0.956), and specificity (0.982) compared to EyeArt, IDx-DR, and a human expert. These metrics underscore DeepRetina's effectiveness in accurately identifying both positive and negative cases, surpassing existing methods like EyeArt (AUC-ROC 0.980) and IDx-DR (0.975), as well as the human expert (0.950). This establishes DeepRetina as the most reliable method in this evaluation.

Table 1. Performance of DeepRetina

Method	AUC-ROC	Sensitivity	Specificity
DeepRetina	0.989	0.956	0.982
EyeArt [23]	0.980	0.942	0.975
IDx-DR [24]	0.975	0.938	0.969
Human expert [14]	0.950	0.912	0.958

Table 2 presents the Cohen's kappa scores for DR grading. It showcases the Cohen's kappa scores for DR grading, highlighting DeepRetina's highest score of 0.892, signifying superior consistency with ground truth. Google AI follows closely at 0.874, while RetinAI (0.856) and the human expert (0.838) show comparatively lower agreement. These results emphasize DeepRetina's enhanced reliability and accuracy in DR grading, outperforming existing AI models and human expertise in maintaining consistent grading standards.

Table 2. DR grading

Method	Cohen's Kappa
DeepRetina	0.892
Google AI [25]	0.874
RetinAI [26]	0.856
Human expert [14]	0.838

Table 3 shows the mean absolute error (MAE) for progression risk prediction at different time points. It presents the mean absolute error (MAE) for progression risk prediction at 1-year, 3-year, and 5-year intervals. DeepRetina demonstrates the highest accuracy with the lowest MAE (0.087, 0.112, and 0.135, respectively) across all time points. ProgressionNet follows with higher MAE values (0.103, 0.128, and 0.156), while the clinical model shows the least accuracy (0.142, 0.178, and 0.213). These results highlight DeepRetina's superior reliability in predicting long-term progression risks.

Table 3. Progression risk prediction

Method	1-Year MAE	3-Year MAE	5-Year MAE
DeepRetina	0.087	0.112	0.135
ProgressionNet [27]	0.103	0.128	0.156
Clinical model [28]	0.142	0.178	0.213

We conducted an ablation study to evaluate the contribution of each component in DeepRetina. Table 4 presents the results. It shows that the full DeepRetina model achieves the highest performance across all metrics: area under the curve-receiver operating characteristic (AUC-ROC) (0.989), Cohen's Kappa (0.892), and the lowest 5-year MAE (0.135). Removing any component leads to a decline in performance. Notably, excluding temporal analysis increases the 5-Year MAE (0.168), indicating its critical role in long-term risk prediction, while OCT data contributes significantly to overall accuracy and agreement.

Table 4. Ablation study

Model variant	AUC-ROC	Cohen's Kappa	5-Year MAE
Full DeepRetina	0.989	0.892	0.135
w/o OCT	0.976	0.863	0.152
w/o Clinical data	0.981	0.878	0.149
w/o Temporal analysis	0.985	0.885	0.168
w/o Attention fusion	0.983	0.879	0.144

To visualize the improvement in the performance of DeepRetina over time, we plotted the AUC-ROC scores for DR detection at different training epochs. The graph in Figure 2 illustrates the AUC-ROC performance improvement of DeepRetina and the Baseline CNN for DR detection across training epochs. DeepRetina shows rapid improvement during the initial 20 epochs, surpassing the baseline CNN early. Its performance plateaus around 60 epochs, achieving a higher final AUC-ROC compared to the Baseline CNN, which stabilizes earlier. This demonstrates DeepRetina's superior learning efficiency and robustness in achieving higher accuracy for DR detection compared to the baseline model.

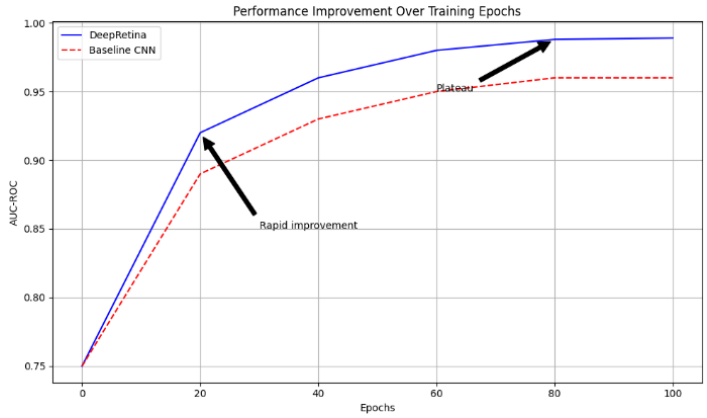

Figure 2. Transformation of DeepRetina

4. DISCUSSION AND FUTURE DIRECTIONS

Our results clearly show that DeepRetina is a quantum leap toward an automated DR diagnosis and prognosis. DeepRetina holds the potential to enable the early detection of DR with high sensitivity and specificity, hence better outcomes in patients. Accurate progression risk prediction could not only allow for individualized treatment plans but also facilitate resource allocation in health care systems. DeepRetina is the best in all tasks, which may become very helpful as an assistant for clinicians in the management of DR.

Some of the limitations to this study include the large dataset, which may not have a full representation of demographic groups and clinical settings, hence open to bias. There is no external validation of the performance of the model on absolutely independent datasets of different institutions. DeepRetina has an excessively complex architecture; therefore, this makes its deployment in resource-constrained settings very hard because of the high computational requirements. Finally, as much as we implemented visualization techniques, the complex nature of the model may still be a challenge in terms of interpretability within clinical practice and may hinder its adoption or effectiveness.

We provide directions for future research in enhancing the DeepRetina framework based on the findings and limitations. They include, but are not limited to: external validation using multicenter studies done across populations and clinical settings; advanced techniques of model compression with respect to computational requirements; advanced interpretability techniques that provide clearer insights into the processing model; examining additional data types, which can increase predictive accuracy; studying over a longer period of time for the assessment of clinical outcome and health costs; and federated learning approaches to ensure privacy-preserving model training across multiple institutions. These proposed directions should help overcome some of the current limitations and thereby improve the clinical applicability and impact of DeepRetina.

5. CONCLUSION

In this paper, we developed DeepRetina, a deep learning framework that fuses multi-modal retinal imaging with patient-specific clinical data for the early detection and prognosis of DR. DeepRetina uses state-of-the-art CNNs and attention mechanisms to analyze OCT and fundus photography images, giving a global perspective of retinal health. Its high accuracy in detection and progression prediction tasks highlights its potential to support clinical decision-making and enable personalized treatment. By facilitating timely and targeted interventions, DeepRetina may contribute to better patient outcomes and reduced vision loss from DR. This advancement marks a meaningful contribution to the field of personalized medicine in ophthalmology. Despite its promise, practical deployment may face challenges such as computational complexity and limited generalizability across diverse clinical settings. Future work should focus on model optimization for resource-limited environments, enhanced interpretability, and broader validation through multi-institutional studies.

FUNDING INFORMATION

Authors state no funding involved.

AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Sundar Ramasamy	✓			✓		✓	✓			✓	✓	✓		
Brindha Mohanraj		✓				✓	✓			✓	✓		✓	
Sridhar Pushpanathan	✓		✓		✓			✓	✓			✓		
Thenmozhi Elumalai	✓				✓		✓			✓		✓		
Prabu Kaliyaperumal	✓	✓	✓	✓	✓			✓	✓	✓			✓	
Rajakumar Perumal		✓		✓		✓		✓	✓		✓			

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author, [P.K], upon reasonable request.

REFERENCES

[1] H. Seo, S. J. Park, and M. Song, "Diabetic retinopathy (DR): mechanisms, current therapies, and emerging strategies," *Cells*, vol. 14, no. 5, p. 376, Mar. 2025, doi: 10.3390/cells14050376.

[2] Z. L. Teo *et al.*, "Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis," *Ophthalmology*, vol. 128, no. 11, pp. 1580–1591, Nov. 2021, doi: 10.1016/j.ophtha.2021.04.027.

[3] T. E. Tan and T. Y. Wong, "Diabetic retinopathy: looking forward to 2030," *Frontiers in Endocrinology*, vol. 13, Jan. 2023, doi: 10.3389/fendo.2022.1077669.

[4] M. Banach *et al.*, "The prevention opportunities of retinopathy in diabetic patients - position paper endorsed by the polish lipid association," *Archives of Medical Science*, vol. 20, no. 6, pp. 1754–1769, Dec. 2024, doi: 10.5114/aoms/197331.

[5] M. Youldash *et al.*, "Early detection and classification of diabetic retinopathy: a deep learning approach," *AI*, vol. 5, no. 4, pp. 2586–2617, 2024, doi: 10.3390/ai5040125.

[6] S. M. Dipto *et al.*, "An analysis of decipherable red blood cell abnormality detection under federated environment leveraging XAI incorporated deep learning," *Scientific Reports*, vol. 14, no. 1, p. 25664, Oct. 2024, doi: 10.1038/s41598-024-76359-0.

[7] M. Saxena, P. Narra, M. Saxena, and R. Saxena, "Deep learning ensemble framework for multiclass diabetic retinopathy classification," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 22, no. 3, pp. 665–672, Jun. 2024, doi: 10.12928/TELKOMNIKA.v22i3.25794.

[8] S. Lenka, Z. L. Mayaluri, and G. Panda, "Glaucoma detection from retinal fundus images using graph convolution based multi-task model," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 11, p. 100931, Mar. 2025, doi: 10.1016/j.prime.2025.100931.

[9] R. Abbasi *et al.*, "Diabetic retinopathy detection using adaptive deep convolutional neural networks on fundus images," *Scientific Reports*, vol. 15, no. 1, p. 24647, Jul. 2025, doi: 10.1038/s41598-025-09394-0.

[10] M. D. Abràmoff *et al.*, "Automated early detection of diabetic retinopathy," *Ophthalmology*, vol. 117, no. 6, pp. 1147–1154, Jun. 2010, doi: 10.1016/j.ophtha.2010.03.046.

[11] V. Gulshan *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA - Journal of the American Medical Association*, vol. 316, no. 22, pp. 2402–2410, Dec. 2016, doi: 10.1001/jama.2016.17216.

[12] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, Jul. 2017, doi: 10.1016/j.ophtha.2017.02.008.

[13] J. De Fauw *et al.*, "Clinically applicable deep learning for diagnosis and referral in retinal disease," *Nature Medicine*, vol. 24, no. 9, pp. 1342–1350, Sep. 2018, doi: 10.1038/s41591-018-0107-6.

[14] P. P. Hatode and M. Edinburgh, "Optical coherence tomography angiography image classification and analysis of diabetic retinopathy, using Wasserstein generative adversarial network augmentation," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 6, pp. 7046–7056, Dec. 2024, doi: 10.11591/ijece.v14i6.pp7046-7056.

[15] E. Vaghefi, S. Yang, S. Hill, G. Humphrey, N. Walker, and D. Squirrell, "Detection of diabetic retinopathy using computational intelligence: A retinal thickness and fundus image-based method," *Computers in Biology and Medicine*, vol. 125, p. 103977, 2020, doi: 10.1016/j.compbiomed.2020.103977.

[16] R. Poel *et al.*, "The predictive value of segmentation metrics on dosimetry in organs at risk of the brain," *Medical Image Analysis*, vol. 73, p. 102161, Oct. 2021, doi: 10.1016/j.media.2021.102161.

[17] A. Bora *et al.*, "Predicting the risk of developing diabetic retinopathy using deep learning," *Lancet Digit Health*, vol. 1, no. 1, pp. e10–e19, 2019, doi: 10.1016/S2589-7500(19)30004-4.

[18] K. Ohno-Matsui *et al.*, "IMI pathologic myopia," *Investigative Ophthalmology and Visual Science*, vol. 62, no. 5, p. 5, May 2021, doi: 10.1167/iovs.62.5.5.

[19] D. S. W. Ting *et al.*, "Deep learning in ophthalmology: The technical and clinical considerations," *Progress in Retinal and Eye Research*, vol. 72, p. 100759, Sep. 2019, doi: 10.1016/j.preteyeres.2019.04.003.

[20] A. Jabbar *et al.*, "Deep transfer learning-based automated diabetic retinopathy detection using retinal fundus images in Remote Areas," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, p. 135, 2024, doi: 10.1007/s44196-024-00520-w.

[21] O. R. Olaniran, A. O. Sikiru, J. Allohibi, A. A. Alharbi, and N. M. S. Alharbi, "Hybrid random feature selection and recurrent neural network for diabetes prediction," *Mathematics*, vol. 13, no. 4, p. 628, Feb. 2025, doi: 10.3390/math13040628.

[22] Y. Shao *et al.*, "An improvement of Adam based on a cyclic exponential decay learning rate and gradient norm constraints," *Electronics (Switzerland)*, vol. 13, no. 9, p. 1778, May 2024, doi: 10.3390/electronics13091778.

[23] F. S. Sorrentino *et al.*, "Novel approaches for early detection of retinal diseases using artificial intelligence," *Journal of Personalized Medicine*, vol. 14, no. 7, p. 690, Jun. 2024, doi: 10.3390/jpm14070690.

[24] E. Riotto *et al.*, "Accuracy of autonomous artificial intelligence-based diabetic retinopathy screening in real-life clinical practice," *Journal of Clinical Medicine*, vol. 13, no. 16, p. 4776, Aug. 2024, doi: 10.3390/jcm13164776.

[25] R. M. Wolf *et al.*, "Autonomous artificial intelligence increases screening and follow-up for diabetic retinopathy in youth: the ACCESS randomized control trial," *Nature Communications*, vol. 15, no. 1, p. 421, Jan. 2024, doi: 10.1038/s41467-023-44676-z.

[26] J. Surya *et al.*, "Efficacy of deep learning-based artificial intelligence models in screening and referring patients with diabetic retinopathy and glaucoma," *Indian Journal of Ophthalmology*, vol. 71, no. 8, pp. 3039–3045, Aug. 2023, doi: 10.4103/IJO.IJO_11_23.

[27] L. Dai *et al.*, "A deep learning system for predicting time to progression of diabetic retinopathy," *Nature Medicine*, vol. 30, no. 2, pp. 584–594, Feb. 2024, doi: 10.1038/s41591-023-02702-z.

[28] A. Jabbar, S. Naseem, J. Li, *et al.*, "Deep transfer learning-based automated diabetic retinopathy detection using retinal fundus images in remote areas," *International Journal of Computational Intelligence Systems*, vol. 17, p. 135, 2024, doi: 10.1007/s44196-024-00520-w.

BIOGRAPHIES OF AUTHORS

Dr. Sunder Ramasamy professor in School of Computer Science and Engineering at Galgotias University, holds 20 years of teaching experience with Ph.D. from Manonmaniam Sundaranar University. He has published 8 patents and 20 research papers, specializing in machine learning, computer vision, data analytics, and image processing. He can be contacted at email: sunder.r@galgotiasuniversity.edu.in.

Brindha Mohanraj assistant professor in computer science and engineering at Chennai Institute of Technology, holds 9 years of teaching experience and is pursuing a Ph.D. in computer science and engineering at Anna University. With an M.E. CSE, she has published 1 patent and 4 research papers, specializing in soft computing techniques, cloud, and machine learning. She can be contacted at email: binducse2010@gmail.com.

Sridhar Pushpanathan assistant professor in Department of Electrical and Electronics Engineering at Kongunadu College of Engineering and Technology (Autonomous), holds 11 years of teaching experience. With M.E. (power electronics and drives) and B.E. (EEE) in Distinction from Anna University, he has published 2 patents and 1 research paper, specializing in power system operation and control, power electronics, and electrical machines, electrical energy generation utilization and conservation, and machine learning. He can be contacted at email: sridharp@kongunadu.ac.in.

Dr. Thenmozhi Elumalai professor in the Department of Information Technology at Panimalar Engineering College. With 23 years of teaching experience, she holds Ph.D. and has authored 7 patents, 8 book chapters, and 24 research papers in renowned international journals and conferences. Her areas of expertise include cyber security, networks, and machine learning. She can be contacted at email: ethenmozhi22.pec@gmail.com.

Prabu Kaliyaperumal assistant professor in School of Computer Science and Engineering at Galgotias University, has 16 years of teaching experience. Currently pursuing a Ph.D., he holds an M.Tech. in CSE from SRM University and MBA from Anna University. He has published 4 patents and 12 research papers in international journals and conferences. His expertise includes cyber security, networks, cloud computing, and machine learning. He can be contacted at email: mega.prabu@gmail.com.

Rajakumar Perumal assistant professor in School of Computer Science and Engineering at Galgotias University, holds 22 years of teaching experience and is pursuing a Ph.D. in computer science and engineering at Shri Venkateshwara University. With an M.E. CSE from Anna University, he has published 4 patents and 8 research papers, specializing in networks, cloud computing, software engineering, and machine learning. He can be contacted at email: rajkumar.jcet@gmail.com.