IAES International Journal of Robotics and Automation (IJRA)
Vol. 15, No. 3, September 2026, pp. 678~689
ISSN: 2722-2586, DOI: 10.11591/ijra.v15i3.pp678-689

PIKER-NET: Multi-class retinal disease classification using Pied Kingfisher optimization-based improved residual network

Lissy Devasahayam[1], Ramya Devi Murugadasan[2], Anandhi Samuel Vijayalakshmi[3], Chanthiya Puhalenthi[4], Ramnath Muthusamy[3], Ahilan Appathurai[5], Natarajan Mohana Suganthi[1]
[1]Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Tamil Nadu, India
[2]Department of Computer Science and Engineering, Hindusthan College of Engineering and Technology, Coimbatore, India
[3]Department of Artificial Intelligence and Data Science, Ramco Institute of Technology, Rajapalayam, Tamil Nadu, India
[4]Department of Artificial Intelligence and Data Science, Dr. Sivanthi Aditanar College of Engineering, Thiruchendur, Tamil Nadu, India
[5]Department of Electronics and Communication Engineering, PSN College of Engineering and Technology, Tirunelveli, India

Article Info

Article history:

Received May 1, 2025
Revised June 29, 2026
Accepted July 20, 2026

Keywords:

Feature selection
Pied Kingfisher optimization
Residual multilayer perceptron
Retinal disease
Scalable range adaptive bilateral filter

ABSTRACT

Retinal diseases are vision-threatening conditions, including age-related macular degeneration (ARMD), diabetic retinopathy (DR), and glaucoma, that require early and accurate detection to prevent blindness. However, existing methods often struggle with limited feature representation, high inter-class similarity, intra-class variability, and reduced performance in handling noisy and low-quality retinal images. To address these challenges, a novel PIKER-NET framework is proposed for accurate multi-class retinal disease classification. The input fundus images from the RFMiD dataset are pre-processed using a scalable range adaptive bilateral (SCRAB) filter to enhance image clarity by preserving edges while reducing noise. The improved residual network-rescaled (ImResNet-RS) integrated with temporal attention is then employed to extract deep hierarchical features with enhanced discriminative power. Pied Kingfisher optimization (PKO) algorithm is utilized for feature selection, effectively reducing redundant information while retaining the most relevant features. Residual multilayer perceptron (ResMLP) is used to classify retinal diseases into ARMD, branch retinal vein occlusion (BRVO), diabetic neuropathy (DN), DR, healthy, and myopia (MYA). The PIKER-NET achieves an overall accuracy of 98.14% and F1-score of 97.06%. The PIKER-NET approach improves overall accuracy by 3.24%, 4.24%, 6.23%, and 2.00% compared to EyeDeep-Net, IDL-MRDD, DeepDiabetic, and VisionDeep-AI, respectively. The proposed approach has strong clinical relevance by supporting earlier disease screening, reducing misdiagnosis, and enabling faster diagnosis to assist ophthalmologists in improving patient outcomes.

Corresponding Author:

Natarajan Mohana Suganthi
Department of Computer Science and Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology
Avadi, Tamil Nadu, India
Email: mohanasuganthi154n@outlook.com

1. INTRODUCTION

Retina is one of the primary elements of the eye that sustains vision [1]. Its primary function is to convert light entering the eye into electrical signals that are transmitted to the brain via the optic nerve. It is situated in the rear of the eye [2]. Because of its nature, the retina can exhibit both eye-specific diseases and

more general physiological ailments, such as circulatory and brain disorders [3]. Age-related macular degeneration (ARMD), diabetic retinopathy (DR), and glaucoma are among the diseases that cause blindness to over 10 million individuals worldwide each year.

Retinal diseases are caused by factors such as genetics, aging, diabetes, hypertension, and lifestyle choices [4]. It is critical to detect retinal diseases early and initiate treatment as soon as possible to preserve vision and manage them [5]. An early diagnosis is achieved with the aid of Fluorescein angiography, fundus imaging, retinal imaging, and optical coherence tomography (OCT) [6].

In deep learning (DL) [7] techniques, convolutional neural networks (CNN) [8] and transformer architectures more recently. A significant amount of effort has been made to identify prevalent retinal illnesses such glaucoma, diabetic retinopathy, and age-related macular degeneration [9]. convolutional neural networks (pre-trained CNN and VGG16) [10] to diagnose DR based on lesion patch probability [11]. It created a system to categories fundus images into referable DR and non-referable DR using three CNNs: Inception-v3, ResNet152, and Inception-ResNet-v2 [12], [13].

Sengar *et al.* [14] introduced a non-invasive automated deep learning approach to identify numerous eye illnesses using color fundus images. The suggested methodology's effectiveness in classifying and identifying diseases using digital fundus images was demonstrated by a comparison with the most recent approaches. Gour and Khanna [15] proposed a transfer learning-based CNN utilizing two distinct optimizers and four pre-trained CNN architectures. They determined that the stochastic gradient descent (SGD) optimizer pre-trained on VGG16 of ocular diseases for multi-class, fundus imagine classification.

Vaiyapuri *et al.* [16] presented an intelligent deep learning-based multi-retinal disease diagnostic (IDL-MRDD) framework using fundus image. With a maximum accuracy of 0.963, the experimental values produced better results than the current methods. Rodríguez *et al.* [17] developed a multi-label retinal disease classification using transformers. CNNs (pre-trained CNN and VGG16) to diagnose DR based on lesion patch probability. It created a system to categories fundus images into referable DR and non-referable DR using three CNNs: Inception-v3, ResNet152, and Inception-ResNet-v2.

Albelaihi and Ibrahim [18] introduced a DeepDiabetic approach for identification of diabetic eye diseases using of deep neural networks. The accuracy of the proposed model is 98.76%, whereas the accuracy of the earlier research was only 88.33%, 89.54%, 97.23%, and 80.33%, respectively. Joshi *et al.* [19] examined a VisionDeep-AI approach for retinal blood vessels segmentation and multi-class classification framework for eye diagnosis. Multi-modal deep feature-fused classification architecture achieved a test accuracy of 81.50%, with a specificity of 93.83%. Yang *et al.* [20] developed a CRAT approach using transformer-based DL algorithm in OCT image classification. They are 97.33%, 96.33%, 97.08%, 99.17%, and 98.74% on the external test dataset, respectively.

Existing methods in the literature review have many shortcomings that limit their effectiveness in retinal disease classification. However, existing methods often struggle with accurate classification of multi-class retinal diseases due to limited feature representation, poor generalization across diverse datasets, high inter-class similarity, and intra-class variability. Additionally, many approaches are not performed to reduce noise and low-quality images, and reduced accuracy in early-stage disease detection. To overcome these challenges, a novel PIKER-NET approach was accurately classifying retinal diseases enhancing diagnostic efficiency and improving early disease detection. The key research contributions of the PIKER-NET approach are highlighted: i) Scalable range adaptive bilateral (SCRAB) filter is utilized to preserve edges while effectively reducing noise in retinal images, improving image quality, ii) The proposed ImResNet-RS integrated with temporal attention approach to capture hierarchical features for improve generalization, and strengthen discriminative learning for multi-class disease detection, iii) Pied kingfisher optimization (PKO) is utilized to select the most discriminative features by balancing exploration and exploitation, thereby enhancing retinal disease classification accuracy, and iv) ResMLP for precise classification of multiple retinal diseases, stabilizing training with residual connections.

The remainder of this work is organized into the following sections. Section 2 presents the recent works of retinal disease classification, section 3 introduces the PIKER-NET approach in practical aspect, section 4 demonstrates the experimental arrangement and the findings of the PIKER-NET, finally conclusion and future work are included in section 5.

The proposed approach raises the following scientific questions:

SQ1: How can noise be reduced and image clarity improved in retinal fundus images for better disease detection?

SQ2: Can deep learning with feature optimization accurately classify multiple retinal diseases?

SQ3: How does the proposed PIKER-NET framework perform compared to existing retinal disease classification methods?

2. METHOD

In this research, a novel PIKER-NET approach is proposed for retinal disease classification from the gathered retinal fundus multi-disease image dataset (RFMiD). Figure 1 illustrates the overall process of the PIKER-NET method.

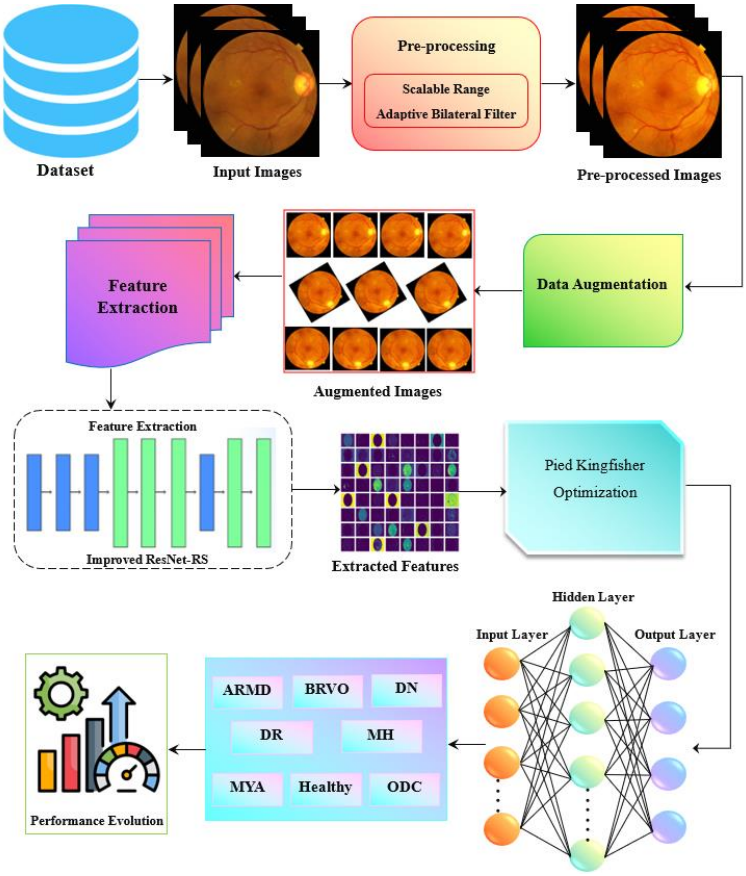


Figure 1. Proposed PIKER-NET approach

2.1. Dataset description

The RFMiD is an openly accessible dataset. It is specifically created for the categorization of retinal disorders into multiple classes. It consists of 3,200 high-resolution fundus images collected from real-world clinical settings covering 46 different retinal conditions such hypertensive retinopathy, ARMD, glaucoma, and DR. The dataset includes expert-labelled annotations, and it is highly diverse featuring a wide range of disease severities, image qualities, and variations in patient demographics. The distribution of retinal fundus images across different disease classes in the RFMiD dataset comprises a total of 3,372 images. The dataset covers 10 categories including pathological conditions such as age-related macular degeneration (ARMD), branch retinal vein occlusion (BRVO, diabetic retinopathy (DR), healthy, and myopia (MYA), along with healthy cases. The Healthy class contains the highest number of images 669, while the ODP class has the fewest 115. This dataset provides a balanced yet diverse representation of common and rare retinal diseases.

.

2.2. Pre-processing using scalable range adaptive bilateral (SCRAB)

SCRAB filter is utilized to enhance retinal images by preserving edges while reducing noise. The bilateral filter is used for noise elimination, levelling, non-linear, and edge-preserving in retinal images. A major shortcoming of the bilateral filter is acquiring significant data about edge fluctuations in noisy images. This tricky is solved by the SCRAB filter, which is determined as follows:

$$z_r(j,k,l) = \propto exp\left(-\frac{1}{2}\left(\frac{||f(j)-f(k)-f(l)||}{\sigma_r}\right)^2\right) + \beta \quad (1)$$

$$q(k) = \begin{cases} |f(k) - mean(\Omega_y)||j - l| \leq c \\ 0 \qquad\qquad otherwise \end{cases} \tag{2}$$

where Ω_y denotes the pixel-set of $(2k+1) * (2k+1)$ pixel window when $k = 2$. The optimistic parameters are $\propto$ and β, Ω_y is the average value, c is a stable variable and $q(k)$ is a range-based functions. The three variables are utilized for controlling z_r are the scaling factor σ_r, the linear constant coefficients $\alpha = 2$ and $\beta = 1$. Since these variables σ_r guarantees a high degree of visual similarity between pixel n in the center and pixel o in the neighboring area. This SCRAB filter effectively enhances edge preservation and noise reduction in retinal images by refining intensity differences while maintaining visual similarity.

2.3. Feature extraction

Improved residual network (ImResNet-RS) integrated with temporal attention approach is used for extracting hierarchical features from retinal images. ResNet-RS is a scaled ResNet architecture that improves accuracy and efficiency by systematically adjusting depth, width, and input resolution. ImResNet-RS further enhances this design for better performance with low computational cost. Figure 2 demonstrates the architecture of proposed ResNet-RS.

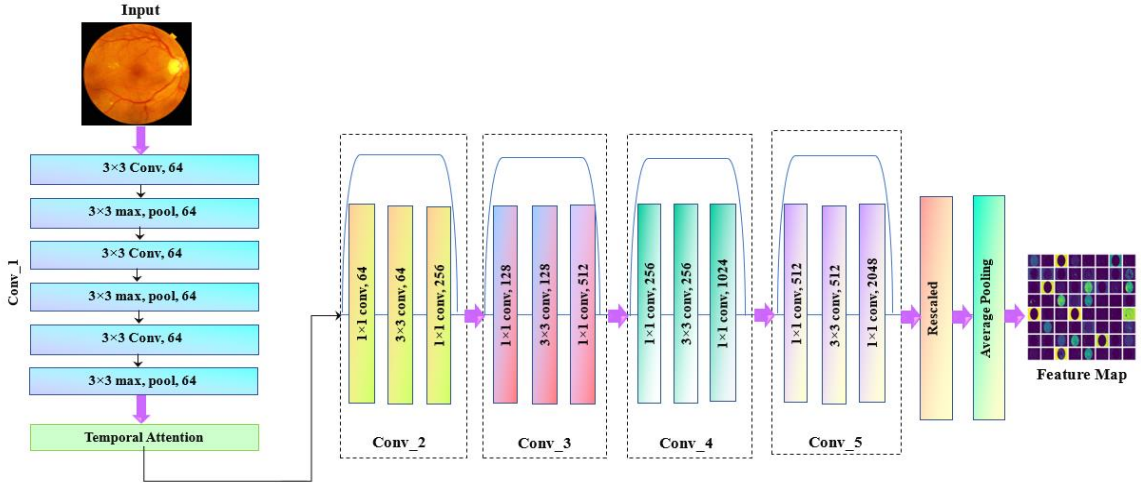


Figure 2. Architecture of the proposed ImResNet-RS

The primary contribution of ResNet-RS is its scaling rule, which adjusts depth (d), width (w), and resolution (r) according to a compound coefficient $\emptyset$. This scaling follows (3):

$$d = \alpha^{\emptyset} d_0, \quad w = \beta^{\emptyset} w_0, \quad r = \gamma^{\emptyset} r_0 \tag{3}$$

where d_0, w_0, r_0 are the initial depth, width, and resolution of the baseline ResNet model, while α, β, γ are constants determined through grid search. In the proposed framework, a temporal attention module is embedded after convolutional layers to refine the extracted features by weighting the most informative patterns. Mathematically bottleneck block expressed as:

$$y = x + f\left(R_3.\sigma\left(R_2.\sigma(R_1.x)\right)\right) \tag{4}$$

where x is the input feature map R_1, R_2, R_3 are weight matrices for the 1×1, 3×3, and 1×1 convolution, σ represents a non-linearity and $f$is the function that applies batch normalization and activation and y is the final output of the operation. ImResNet-RS enhances retinal image analysis by efficiently extracting hierarchical features while maintaining computational efficiency.

2.4. Pied kingfisher optimization (PKO)

The PKO is to select the most discriminative features by balancing exploration and exploitation, thereby enhancing retinal disease classification accuracy. PKO is a bio-inspired optimization algorithm that

mimics the perching, hovering, and diving behaviors of Pied Kingfishers. In the proposed PIKER-NET, PKO is employed for feature selection, where perching enables global exploration of the feature space, hovering refines promising feature subsets, and diving performs intensive local exploitation near the best candidates, leading to optimal and highly discriminative feature selection. This optimization improves the model's performance by reducing irrelevant features for enhancing classification accuracy while lowering computational cost.

Step 1: Initialization

The initial input parameters such as population size and feature subset length are set.

Step 2: Random generation

Candidate feature subsets are generated randomly from the extracted features.

Step 3: Fitness function

Fitness is determined by the classification performance, and it is defined as:

$$F = min(1 - Accuracy) \tag{5}$$

where accuracy represents the classification accuracy of the ImResNet-RS approach when evaluated with a particular subset of features.

Step 4: Exploration

PK's hovering and perching behavior supports to identify subsets of features with high potential. The exploratory mechanism allows PKO to scan the feature space widely and select diverse candidate subsets for evaluation. The PK's foraging activity is used in PKO to determine the locations of search agents. The updated positions are calculated as:

$$D_j(T+1) = D_j(T) + \beta * t \times \left(D_k(T) - D_k(T)\right) \tag{6}$$

where $D_j(T+1)$ denotes solution for next iteration, $D_j(T)$ represents the position of the current iteration, β and t denotes parameter. This process ensures wide exploration of feature subsets, thereby increasing the possibility of selecting the most discriminative features.

Step 5: Exploitation

The diving behavior of PK ensures refined exploitation of promising feature subsets. The search process emphasizes subsets that maximize accuracy while minimizing redundancy. Mathematically, the updated candidate feature subset positions are expressed as:

$$D_f(t+1) = D_f(t) + \alpha.\beta.\left(Best_{features} - D_f(t)\right) \tag{7}$$

where $D_f(t+1)$ is represents the updated feature subset position at iteration $t+1$, and $D_f(t)$ denotes the current feature subset position at iteration t, the α is the learning rate that controls the step size towards the best feature subset. β denotes the control parameter that balances exploration and exploitation during the optimization. Then $Best_{features}$ indicate the most discriminative feature subset identified and based on classification accuracy. $Best_{features} - D_f(t)$ symbiosis the direction of movement toward the optimal subset in the feature space.

Step 6: Updating

The best subset is updated iteratively using classification accuracy as the objective until convergence.

Step 7: Termination

The process terminates the maximum number of repetitions is reached or when no significant improvement in feature selection is observed. The final selected features are then passed to the classifier for final retinal disease prediction.

2.5. ResMLP

Residual multilayer perceptron (ResMLP) is used to classify the retinal disease ARMD, BRVO, DN, DR, healthy, MYA. ResMLP is an advanced variant of MLP that integrates residual connections to stabilize training and enhance feature learning. A ResMLP usually comprises multiple hidden layers interconnected with residual skip connections along with the output layer. Weighted connections are used to link each neuron to the neurons in the layer below. Figure 3 illustrates the ResMLP structure for Retinal disease classification.

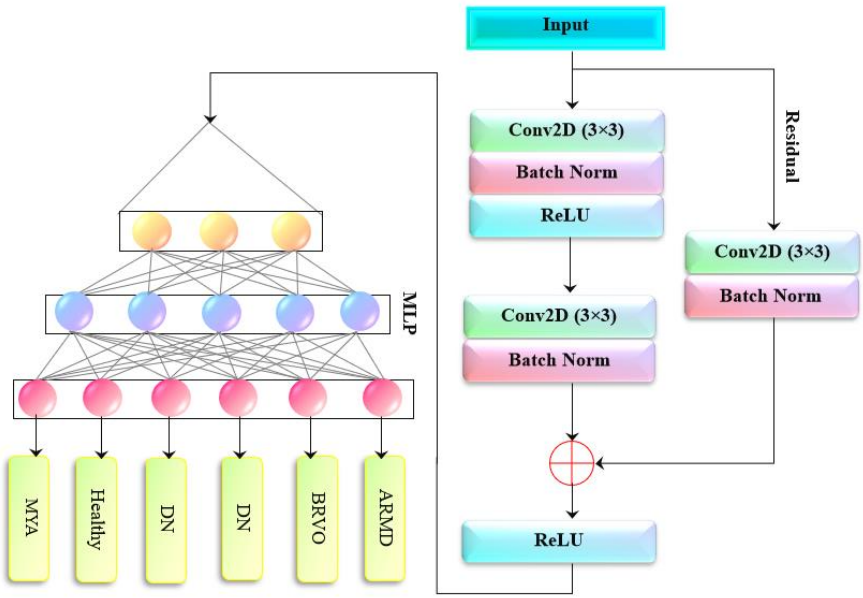


Figure 3. Structure of ResMLP

The fundamental unit of a ResMLP is the perceptron, which introduces non-linearity by applying an activation function and calculating the inputs weighted quantity. Mathematically a given neuron j, the output is given by (8).

$$z_j = \sum_{i=1}^{n} w_{ji}\, x_i + b_j \tag{8}$$

where x_i represents the input w_{ji} represents the weight assigned to every input connection and a_j is the bias term. This weighted sum goes through an activation function after that f like the tanh function ReLU or sigmoid.

$$a_j = f(z_j) \tag{9}$$

A ResMLP with multiple layers and residual links is mathematically represented using matrix notation. Given an input vector X and weight matrices W, the output of a layer written as (10).

$$M^{(l)} = f\left(W^{(l)} M^{(l-1)} + N^{(l)}\right) + M^{(l-1)} \tag{10}$$

where $M^{(l)}$ symbolizes the layer activations at layer $l, W^{(l)}$ is the weight matrix for the layer l, and $N^{(l)}$ is the bias vector. The residual connection $M^{(l-1)}$ supports gradients and accelerates convergence. This ResMLP-based approach enhances classification of retinal disease types by residual learning from extracted features to achieve precise classification.

3. RESULTS AND DISCUSSION

This section details the experimental results of the PIKER-NET approach was implemented using MATLAB 2022b. The proposed PIKER-NET approach was assessed using a number of metrics in this section including specificity, accuracy, F1 score, recall, and precision based on images taken from the RFMiD dataset. The experimental results of the PIKER-NET approach are illustrated in Figure 4 for retinal disease classification.

The experimental result of the PIKER-NET approach is shown in Figure 4. In column 1, the input retinal images are collected from the RFMiD dataset. In column 2, SCRAB to enhance retinal images by preserving edges while reducing noise followed by data augmentation in column 3. Feature extraction using ImResNet-RS to extract the feature images in column 4. Finally, the ResMLP is used to classify the retinal disease namely ARMD, BRVO, DN, DR, Healthy, and MYA in column 4.

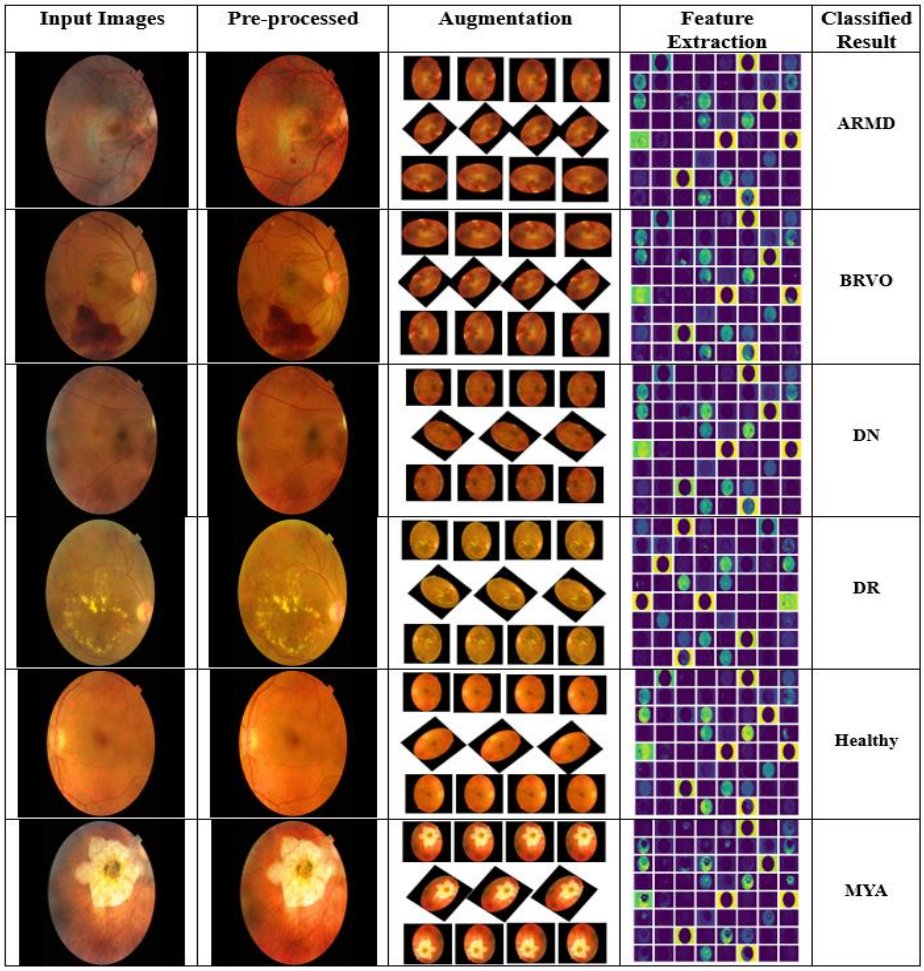


Figure 4. Experimental result of the PIKER-NET approach

3.1. Performance analysis

The PIKER-NET was assessed in this section using a number of metrics, including specificity (Sp), F1 score, accuracy (Ac), precision (Pr), and recall (Rc), on the gathered RFMiD dataset.

$$Sp = \frac{TN}{TN+FP} \tag{11}$$

$$Ac = \frac{TP+TN}{TP+TN+FP+FN} \tag{12}$$

$$Pr = \frac{TP}{TP+FP} \tag{13}$$

$$Rc = \frac{TP}{TP+FN} \tag{14}$$

$$f1score = 2(\frac{precision*recall}{precision+recall}) \tag{15}$$

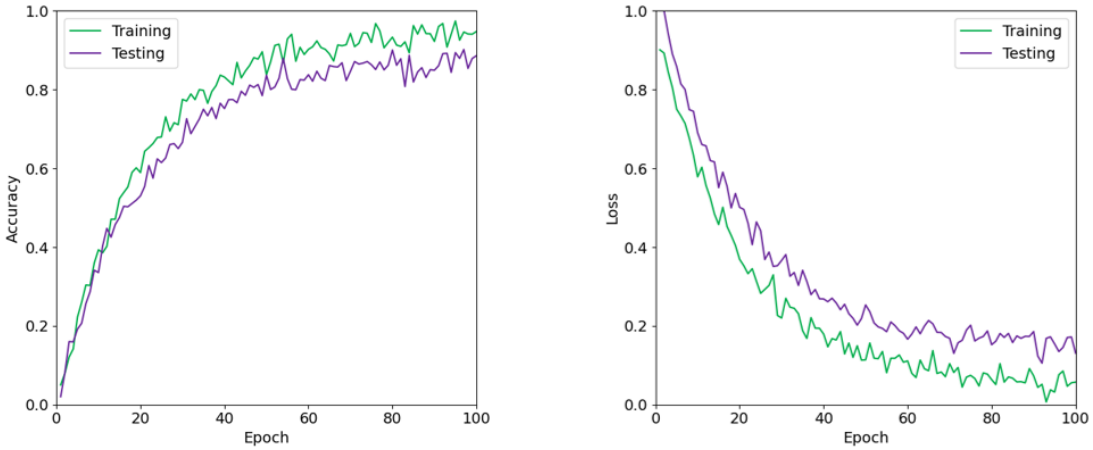
where TP and TN indicate true positives and negatives of the sample images, FP and FN indicates false positives and negatives of the sample images.

Table 1 presents the specificity, F1 score, recall, accuracy, and precision of the proposed method for classes with multiple levels. Table 1 displays the efficiency metrics of the proposed method for the classes listed below: ARMD, BRVO, DN, DR, Healthy, and MYA. The accuracy of the proposed approach is 96.37% for ARMD, 98.56% for BRVO, 97.81% for DN, 99.11% for DR, 99.87 for Healthy, and 97.16% for MYA respectively. The proposed approach achieves a total accuracy rate of 98.14%.

The performance evaluation of the proposed PIKER-NET approach across 100 epochs is illustrated in Figure 5, showing both the accuracy and loss curves during the training and testing stages. Figure 5(a) demonstrates the accuracy curve of the PIKER-NET approach based on the gathered RFMiD dataset. The PIKER-NET training and testing accuracy curve displays a high accuracy level of 98.14%. The PIKER-NET approach loss curve, which displays 100 epochs, is illustrated in Figure 5(b). It achieves a low loss of 1.86%, indicating that the PIKER-NET model performs effectively during both the testing and training stages.

Table 1. Performance analysis of PIKER-NET method

Classes	Accuracy%	Specificity%	Precision%	Recall%	F1score%
ARMD	96.37	97.40	95.32	96.38	98.02
BRVO	98.56	98.11	94.65	98.30	97.21
DN	97.81	96.32	94.22	95.81	96.89
DR	99.11	98.97	96.37	98.42	99.04
Healthy	99.87	99.21	97.94	95.39	98.27
MYA	97.16	91.80	94.37	93.94	92.95



(a) (b)

Figure 5. Performance curves of the proposed PIKER-NET approach on the RFMiD dataset, showing (a) the training and testing accuracy curve and (b) the training and testing loss curve

3.2. Comparative analysis

This section includes a comparative assessment of traditional neural networks and the PIKER-NET approach. It has been demonstrated to be more productive than existing methods as demonstrated by the comparison of performance metrics including precision, specificity, F1 score, recall, and accuracy. The comparative evaluation involves the existing DL Network and proposed model.

Table 2 demonstrates the comparison of existing approaches and proposed approach. The accuracy obtained by the PIKER-NET approach and existing methods such as CNN, VGG16, DBN, and MLP. The proposed model performs 7.70%, 5.21%, 3.39%, and 2.98% better than the conventional approaches, such as CNN, VGG16, DBN and MLP.

Table 2. Comparison of existing models with proposed model

Techniques	Accuracy%	Specificity%	Precision%	Recall%	F1score%
CNN	91.12	89.62	88.07	91.63	89.23
VGG16	93.28	91.74	90.96	92.95	92.67
DBN	94.92	93.31	92.35	94.02	93.91
MLP	95.30	94.75	93.69	95.93	94.38
Proposed ResMLP	98.14	95.32	96.23	97.14	98.97

Table 3 shows the comparison of existing models and the proposed PIKER-NET approach. The evidence includes a comprehensive evaluation of the proposed PIKER-NET approach using multiple performance metrics on the RFMiD dataset. The PIKER-NET approach achieves a higher accuracy rate of

98.14% in classifying retinal diseases, outperforming existing methods by 3.24%, 4.24%, 6.23%, and 2.00% compared to EyeDeep-Net, IDL-MRDD, DeepDiabetic, and VisionDeep-AI respectively. The proposed PIKER-NET model shows that the comparison of the accuracy is superior to the existing approach.

Table 3. Comparison of the existing and proposed approaches

Authors	Methods	Accuracy
Sengar *et al.* [14]	EyeDeep-Net	95.06%
Vaiyapuri *et al.* [16]	IDL-MRDD	94.14%
Albelaihi and Ibrahim [18]	DeepDiabetic	92.38%
Joshi *et al.* [19]	VisionDeep-AI	96.21%
Proposed	MCRNET-RS	98.14%

4. CONCLUSION

This research proposed a novel PIKER-NET approach for accurate multi-class retinal disease classification. The input fundus images from the RFMiD dataset are pre-processed using a SCRAB filter for enhances image clarity by preserving edges while reducing noise. The ImResNet-RS integrated with Temporal Attention is then employed to extract deep hierarchical features with enhanced discriminative power. PKO algorithm is utilized for feature selection, effectively reducing redundant information while retaining the most relevant features. ResMLP is used to classify retinal diseases into ARMD, BRVO, DN, DR, Healthy, and MYA. The effectiveness of the proposed PIKER-NET approach is evaluated using performance metrics such as accuracy, specificity, recall, precision, and F1-score. The PIKER-NET approach achieves an overall accuracy of 98.14%, and F1score of 97.06%. The PIKER-NET approach improves an overall accuracy by 3.24%, 4.24%, 6.23%, and 2.00% compared to EyeDeep-Net, IDL-MRDD, DeepDiabetic, and VisionDeep-AI respectively. In the future, the proposed approach can extend to larger multimodal retinal datasets to improve strong and generalization.

ACKNOWLEDGMENTS

The author would like to express his heartfelt gratitude to the supervisor for his guidance and unwavering support during this research for his guidance and support.

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
Lissy Devasahayam		✓		✓	✓			✓	✓				✓	
Ramya Devi Murugadasan		✓	✓			✓		✓		✓	✓	✓		
Anandhi Samuel Vijayalakshmi	✓			✓	✓		✓		✓			✓	✓	
Chanthiya Puhalenthi		✓	✓			✓	✓		✓		✓	✓		
Ramnath Muthusamy	✓		✓			✓		✓		✓	✓		✓	
Ahilan Appathurai	✓			✓	✓		✓		✓			✓		
Natarajan Mohana Suganthi	✓		✓	✓	✓		✓			✓	✓		✓	

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

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

INFORMED CONSENT

Authors certify that authors have explained the nature and purpose of this study to the above-named individual, and authors have discussed the potential benefits of this study participation. The questions the individual had about this study have been answered, and we will always be available to address future questions.

ETHICAL APPROVAL

Our research guide reviewed and ethically approved this manuscript for publishing in this journal.

DATA AVAILABILITY

Data sharing not applicable to this article as no datasets we regenerated or analysed during the current study.

REFERENCES

[1] M. Elsharkawy *et al.*, "OCT-Trans: a novel transformer backbone with multimodal feature extraction in OCT-based retinal disease classification," in *2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*, Apr. 2025, pp. 1–4, doi: 10.1109/ISBI60581.2025.10980790.

[2] G. R. Hemalakshmi, M. Murugappan, M. Y. Sikkandar, S. S. Begum, and N. B. Prakash, "Automated retinal disease classification using hybrid transformer model (SViT) using optical coherence tomography images," *Neural Computing and Applications*, vol. 36, no. 16, pp. 9171–9188, Jun. 2024, doi: 10.1007/s00521-024-09564-7.

[3] M. Elkholy and M. A. Marzouk, "Deep learning-based classification of eye diseases using convolutional neural network for OCT images," *Frontiers in Computer Science*, vol. 5, Jan. 2024, doi: 10.3389/fcomp.2023.1252295.

[4] T. D. Nguyen, D.-T. Le, J. Bum, S. Kim, S. J. Song, and H. Choo, "Retinal disease diagnosis using deep learning on ultra-wide-field fundus images," *Diagnostics*, vol. 14, no. 1, p. 105, Jan. 2024, doi: 10.3390/diagnostics14010105.

[5] S. D. Khan, S. Basalamah, and A. Lbath, "A novel deep learning framework for retinal disease detection leveraging contextual and local features cues from retinal images," *Medical & Biological Engineering & Computing*, vol. 63, no. 7, pp. 2029–2046, Jul. 2025, doi: 10.1007/s11517-025-03314-0.

[6] C. Venkataiah *et al.*, "A novel eye disease segmentation and classification model using advanced deep learning network," *Biomedical Signal Processing and Control*, vol. 105, p. 107565, Jul. 2025, doi: 10.1016/j.bspc.2025.107565.

[7] M. Berrimi and A. Moussaoui, "Deep learning for identifying and classifying retinal diseases," in *2020 2nd International Conference on Computer and Information Sciences (ICCIS)*, Oct. 2020, pp. 1–6, doi: 10.1109/ICCIS49240.2020.9257674.

[8] B. R.M., K. B. Vardhan, M. Nidhish, S. Kiran C., D. Nahid Shameem, and V. Sai Charan, "Eye disease detection using deep learning models with transfer learning techniques," *ICST Transactions on Scalable Information Systems*, vol. 11, Jul. 2024, doi: 10.4108/eetsis.5971.

[9] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, "Automated detection of mild and multi-class diabetic eye diseases using deep learning," *Health Information Science and Systems*, vol. 8, no. 1, p. 32, Dec. 2020, doi: 10.1007/s13755-020-00125-5.

[10] K. R. Akhila, N. Muthukumaran, and A. Ahilan, "Classification of cervical cancer using an autoencoder and cascaded multilayer perceptron," *IETE Journal of Research*, vol. 70, no. 1, pp. 26–36, Jan. 2024, doi: 10.1080/03772063.2022.2142859.

[11] R. Shobana, A. V. K. Lija, and S. S. Marice, "GD2-EFFINET: Glaucoma disease detection and segmentation of deep learning models using EfficientNetB0," *International Journal of Current Bio-Medical Engineering (IJCBE)*, vol. 1, no. 2, pp. 45–51, 2023.

[12] J. Jency and S. Shunmugan, "Deep learning-based diabetic retinopathy for classifying retinal images," *International Journal of Current Bio-Medical Engineering (IJCBE)*, vol. 3, no. 3, pp. 58–64, 2025.

[13] A. Ahilan, M. Anlin Sahaya Tinu, A. Jasmine Gnana Malar, and B. Muthu Kumar, "Stationary wavelet-oriented luminance enhancement approach for brain tumor detection with multi-modality images," in *Evolution in Computational Intelligence. FICTA 2023. Smart Innovation, Systems and Technologies (SIST,volume 370)*, Springer, Singapore, 2023, pp. 461–473, doi: 10.1007/978-981-99-6702-5_38.

[14] N. Sengar, R. C. Joshi, M. K. Dutta, and R. Burget, "EyeDeep-Net: a multi-class diagnosis of retinal diseases using deep neural network," *Neural Computing and Applications*, vol. 35, no. 14, pp. 10551–10571, May 2023, doi: 10.1007/s00521-023-08249-x.

[15] N. Gour and P. Khanna, "Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network," *Biomedical Signal Processing and Control*, vol. 66, p. 102329, Apr. 2021, doi: 10.1016/j.bspc.2020.102329.

[16] T. Vaiyapuri *et al.*, "Intelligent deep learning based multi-retinal disease diagnosis and classification framework," *Computers, Materials & Continua*, vol. 73, no. 3, pp. 5543–5557, 2022, doi: 10.32604/cmc.2022.023919.

[17] M. A. Rodríguez, H. AlMarzouqi, and P. Liatsis, "Multi-label retinal disease classification using transformers," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 6, pp. 2739–2750, Jun. 2023, doi: 10.1109/JBHI.2022.3214086.

[18] A. Albelaihi and D. M. Ibrahim, "DeepDiabetic: an identification system of diabetic eye diseases using deep neural networks," *IEEE Access*, vol. 12, pp. 10769–10789, 2024, doi: 10.1109/ACCESS.2024.3354854.

[19] R. Chandra Joshi, A. Kumar Sharma, and M. Kishore Dutta, "VisionDeep-AI: Deep learning-based retinal blood vessels segmentation and multi-class classification framework for eye diagnosis," *Biomedical Signal Processing and Control*, vol. 94, p. 106273, Aug. 2024, doi: 10.1016/j.bspc.2024.106273.

[20] M. Yang, J. Du, and R. Lv, "CRAT: Advanced transformer-based deep learning algorithms in OCT image classification," *Biomedical Signal Processing and Control*, vol. 104, p. 107544, Jun. 2025, doi: 10.1016/j.bspc.2025.107544.

BIOGRAPHIES OF AUTHORS

Lissy Devasahayam completed her PhD degree from Bharathiar University. Previously she was working as an Assistant Professor in Thiruthangal Nadar College, Chennai-51 for the past 13.5 years. Now she is working as Assistant Professor in Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology. She can be contacted at lissydevasahayam@gmail.com.

Ramya Devi Murugadasan is an Associate Professor in the Department of Computer Science and Engineering at Hindusthan College of Engineering and Technology, Coimbatore. She can be contacted at ramyamurugadasan@gmail.com.

Anandhi Samuel Vijayalakshmi received her B.E degree in computer science and engineering in 2006 from Noorul Islam College of Engineering, India and M.E degree in computer science and engineering in 2010 from Dr. Sivanthi Aditanar College of Engineering, India. She can be contacted at anandhi374@outlook.com.

Chanthiya Puhalenthi is an Associate Professor in the Department of Artificial Intelligence and Data Science at Dr. Sivanthi Aditanar College of Engineering, Tiruchendur. She has over 14 years of teaching experience in the field of computer science and engineering. She can be contacted at chanthiyapuhal@gmail.com.

Ramnath Muthusamy obtained her bachelor's degree in information technology from Francis Xavier Engineering College, Anna University, Chennai in 2009. Then he obtained his master's degree in network engineering from VelTech MultiTech Dr. Rangarajan Dr. Sakunthala Engineering College, Anna University, Chennai in 2011. He can be contacted at ramnath25@gmail.com.

Ahilan Appathurai received Ph.D. from Anna University, India, and working as an Associate Professor in the Department of Electronics and Communication Engineering at PSN College of Engineering and Technology, India. He can be contacted at listentoahil@gmail.com.

Natarajan Mohana Suganthi currently holds the position of Assistant Professor in the Department of Computer Science and Engineering at Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, nestled in Avadi, Chennai, Tamil Nadu. She can be contacted at mohanasuganthin@veltech.edu.in.