IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Traitement du Signal
Vol. 43, No. 3, June, 2026, pp. 1435-1443
Journal homepage: http://iieta.org/journals/ts

Automated Detection of Retinal Diseases Using DenseNet201-SE Attention Mechanisms

Jayanthi Rajee Balashunmugam*, Mohamed Mansoor Roomi Sindha

Department of ECE, Thiagarajar College of Engineering, Madurai 625020, India

Corresponding Author Email: jayanthi@student.tce.edu

https://doi.org/10.18280/ts.430327

Received: 5 April 2026
Revised: 2 June 2026
Accepted: 16 June 2026
Available online: 30 June 2026

Keywords:
retinal disease classification, deep learning in ophthalmology, Squeeze-and-Excitation networks, DenseNet201-SE

ABSTRACT

The manual diagnosis of retinal diseases from fundus images is labor-intensive and prone to error. Deep learning (DL) models have received greater attention due to their accuracy and flexibility. Existing models struggle with multi-class imbalance and subtle feature extraction across eight distinct pathologies. This paper proposes a novel automated detection system for eight retinal conditions—diabetic retinopathy (DR), age-related macular degeneration (AMD), glaucoma (GA), hemorrhages (HG), epiretinal membrane (EM), hypertension (HT), myopia (MY), and normal cases. The primary novelty lies in two key contributions: first, DenseNet201 architecture is proposed based on Squeeze-and-Excitation (SE) attention blocks. The SE attention blocks are added in dense layers to perform channel-wise feature recalibration. It is used for the model to prioritize clinically relevant subtle variations. Next, a Modified Election Optimization Algorithm (MEOA) is proposed to dynamically tune the hyperparameters of a DenseNet201 backbone. Evaluated on the MRD-8 dataset, the DenseNet201-SE model optimized via MEOA achieves a superior classification accuracy of 92% when compared to existing models.

1. INTRODUCTION

Diagnosing retinal disorders is a critical task in the medical field due to the importance of vision in daily life [1]. Retinal diseases, such as age-related macular degeneration (AMD), diabetic retinopathy (DR), epiretinal membrane (EM), glaucoma (GA), myopia (MY), hemorrhages (HG), and hypertension (HT), can significantly impair vision [2, 3]. AMD is common among older people and one of the most prominent causes of loss of vision and blindness, particularly in advanced countries, due to its influence on the choroid and retina. Retinopathy caused by diabetic conditions affects the arteries of the retina and falls under four categories: mild, moderate, severe, and proliferative retinopathies. Epiretinal membrane, an abnormal formation that occurs over the retina, and GA, a disease that raises intraocular pressure, are some other conditions that might impair the retinal structure. The condition of hemorrhage associated with bleeding in the retinal vessels and HT, one of the primary risk factors of numerous diseases such as heart and kidney conditions, is another condition that affects retinal function. Recent developments in Artificial Intelligence (AI) technology enable automated identification of retinal conditions.

Despite significant advances in deep learning (DL)-based retinal disease classification, several critical research gaps remain unaddressed in the existing literature [4]. Most prior works focus on binary classification. Next, the existing models struggle with class imbalance and fail to extract clinically relevant features. Then, the current attention mechanisms primarily use spatial attention or self-attention transformers [5]. Finally, hyperparameter optimization in most retinal disease studies is based on manual tuning or grid search. This work addresses these gaps by proposing a DenseNet201-SE architecture with channel-wise attention and a Modified Election Optimization Algorithm (MEOA) for automated hyperparameter tuning across eight retinal disease classes. The main contribution of the work is

•Compilation of a comprehensive dataset with eight retinal disease classes: AMD, DR, GA, HG, EM, HT, MY, and Normal images.

•Development of a DenseNet-SE attention model for multi-disease classification.

•A MEOA is proposed to tune the parameters of the model.

•Evaluation of dataset augmentation techniques and comparison with state-of-the-art methods using pre-trained networks and attention mechanisms.

2. RELATED WORK

To reduce feature dimensionality, Arunkumar and Karthigaikumar [6] proposed a multi-class Support Vector Machine (SVM) classifier combined with DL based feature extraction. In their study, Vaiyapuri et al. [7] suggested an Intelligent Deep Learning-based Multi-Retinal Disease Diagnosis (IDL-MRDD) model for processing fundus images. This model uses Artificial Flora Algorithm with Shannon's Function (AFA-SF) for multi-level thresholding. For feature extraction, it uses SqueezeNet. The classification is done with Stacked Sparse Autoencoder (SSAE). In validation, the model shows 96.3% accuracy. Similarly, Choudhary et al. [8] developed a customized Visual Geometry Group 19-layer

(VGG-19) architecture with transfer learning for Optical Coherence Tomography (OCT) image classification into four retinal conditions. The model achieves 99.17% accuracy, 0.995 specificity, and 0.99 sensitivity. A hybrid model based on ResNet and Transformer is proposed by Zhao et al. [9]. This model uses the strategy of learnable label embeddings for better feature learning. Di Giammarco et al. [10] presented a Fully Convolutional Neural Network plus (FCNNplus) for retinal disease classification with 93.3% accuracy. Also, it uses the explainability technique of Class Activation Mapping (CAM) algorithms for disease localization. Likewise, Sahoo et al. [11] proposed DeepRetinaNet for disease identification. It involves RetiSegNet for vessel extraction and Spatio-Temporal Deep Network (STDeepNet) for disease identification. Wen et al. [12] proposed a Structure-Aware Transformer-based attention fusion Network (SAT-Net) for multi-class retinal disease detection. Hemal and Saha [13] applied transfer learning on eight DL models like MobileNetV2, InceptionV3, NASNetMobile, DenseNet169, EfficientNetB4, DenseNet121, ConvNeXt, and Xception.

Khan et al. [14] proposed a Densely Connected Multidilated Convolution Neural Network (DCM-CNN) for fundus image processing. The model is validated on the ODIR-5K dataset. Sivaz and Aykut [15] suggested a Swin Transformer V2 for medical image classification. In addition, it uses a Shunted Cross-Attention (SCA) classification head and Adaptive Sharpness-Aware Minimization (ASAM) optimizer to achieve a 87.60% accuracy. In their work, Ghislain et al. [16] proposed a lightweight CNN with biorthogonal wavelet transforms for better feature learning in fundus images. Similarly, Cutur and Inan [17] compared Vision Transformers (ViTs) with other DL models in terms of accuracy and precision rates. Lisha and R [18] recommended a hybrid of Improved LinkNet and SqueezeNet for retinal image classification. Also, Modified Gaussian Filtering (MGF) is applied for image enhancement. In their study, Mathews and Anzar [19] proposed an Attention-Driven Retinal Classification Model (ADRCM) for Diabetic Macular Edema (DME) grading. For accuracy improvement, the model uses a Feature Fusion and Enhancement Module (FFEM). A ResNet-50 model combined with a Generative Adversarial Network (GAN) is proposed by Jeyasri and Karthiyayini [20] for synthetic image generation and disease classification.

3. DENSENET201-SE BASED RETINAL DISEASE CLASSIFICATION

In this work, a DenseNet201 architecture is applied for automated retinal disease classification. The architecture is modified with Squeeze-and-Excitation (SE) attention modules to improve the classification accuracy. The proposed model is designed to identify eight retinal disorders from fundus images contained in the MRD-8 dataset. The overall workflow is shown in Figure 1.

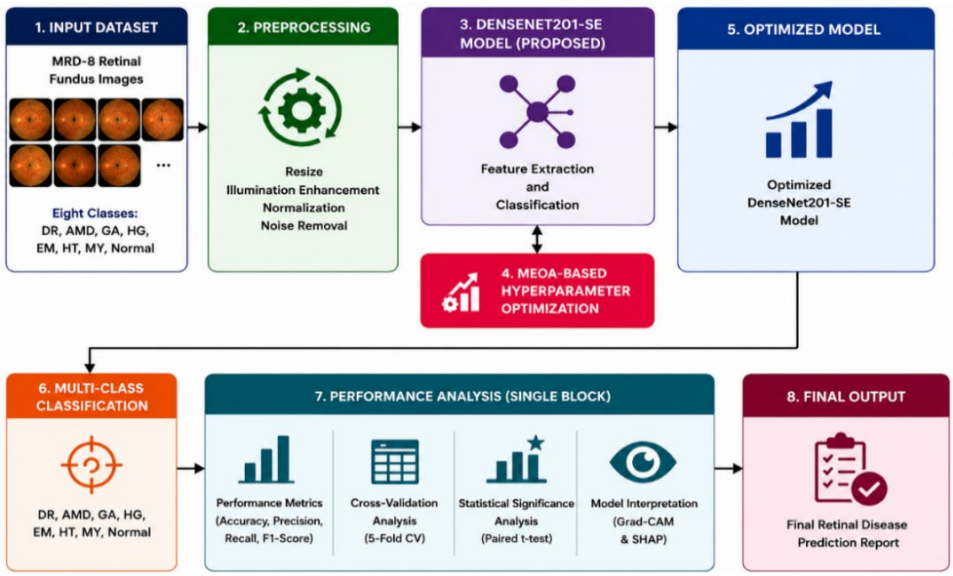

Figure 1. Proposed system workflow

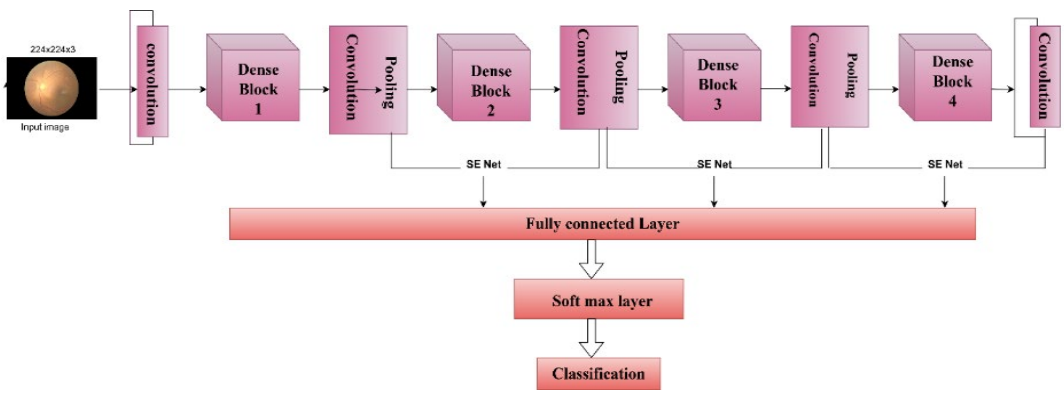

Figure 2. DenseNet201-SE model

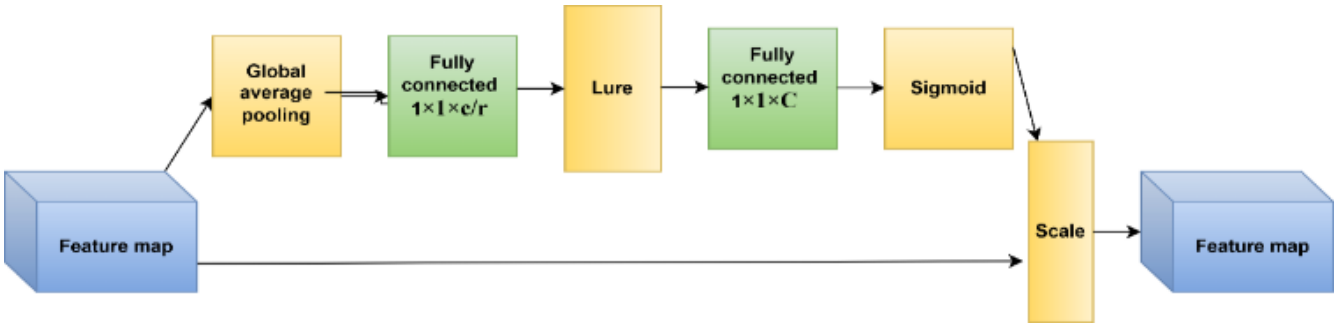

Figure 3. Squeeze-and-Excitation (SE) block architecture

Initially, all retinal images are resized to a fixed spatial resolution of 224 × 224 pixels and normalized to the range [0,1] before being supplied to the network.

Transfer learning is adopted by initializing DenseNet201 with weights pretrained on the ImageNet dataset. It is used for effective feature extraction from retinal images. The architecture of the DenseNet SE model is shown in Figure 2.

The DenseNet201 consists of multiple dense blocks interconnected through transition layers. Compared to conventional convolutional neural networks, each layer receives feature maps from all preceding layers. This strategy supports feature reuse and alleviates the vanishing gradient problem. For the l^{th} layer, the feature representation is expressed as:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \tag{1}$$

where, x_l denotes the output feature map of the l^{th} layer, $[x_0, x_1, \dots, x_{l-1}]$ represents the concatenation of all preceding feature maps, and $H_l(\cdot)$ denotes a composite operation consisting of Batch Normalization (BN), Rectified Linear Unit (ReLU), and convolution. The transformation performed by a convolutional layer is defined as

$$z^{(l)} = W^{(l)} * x^{(l-1)} + b^{(l)} \tag{2}$$

To improve channel-wise feature discrimination, three SE blocks are inserted after selected dense blocks of DenseNet201. The architecture of SE is shown in Figure 3. The SE mechanism performs a squeeze operation using global average pooling to aggregate spatial information from each channel. For channel c, the squeeze operation is given by

$$Z_{CH} = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} R_c \tag{3}$$

where, $R_c(i,j)$ denotes the activation value at the spatial location (i, j) of channel c, while H and W represent the height and width of the feature map, respectively. The resulting vector $z = [z_1, z_2, \dots, z_C]$ captures global channel statistics, where C denotes the number of channels. The excitation stage models channel interdependencies and generates channel attention weights as

$$s = \sigma(W_2\, \delta(W_1 z)) \tag{4}$$

where, W_1 and W_2 are the learnable parameters of two fully connected layers, $\delta(\cdot)$ denotes the ReLU activation function, and $\sigma(\cdot)$ represents the sigmoid activation function. The output vector $s = [s_1, s_2, \dots, s_C]$ contains importance weights for all feature channels. Finally, the recalibration stage scales each feature channel according to its learned importance:

$$\hat{R}_c = s_c \cdot R_c \tag{5}$$

where, s_c is the attention weight corresponding to the channel c, R_c is the original feature map, and $\hat{R}_c$ denotes the recalibrated feature map. Through this adaptive weighting process, the network learns diagnostically relevant retinal structures and suppresses the less informative features.

After feature extraction, a Global Average Pooling layer is used to reduce spatial dimensions. The SoftMax output layer with eight neurons corresponding to the retinal disease categories is used. The network is trained using the categorical cross-entropy loss function, and model parameters are optimized using the Adam optimizer. The optimal hyperparameters are obtained using the MEOA. During training, the model learns discriminative retinal representations. The SE modules dynamically recalibrate feature responses.

3.1 Election Optimization Algorithm

The Election Optimization Algorithm (EOA) is a population-based metaheuristic algorithm [21]. It is inspired by the democratic presidential election process. Each candidate represents a potential solution. The election activities mimic search mechanisms to explore and exploit the solution space. The party nomination phase performs global exploration. The presidential election phase performs local exploitation to converge on the optimum solution. EOA has been successfully applied to complex engineering optimization problems due to its balance between diversity and convergence. The main drawbacks of EOA are premature convergence and its heavy dependence on random factors. To address these drawbacks, a MEOA is proposed in this work. A new candidate update strategy and adaptive control factors are added to improve the optimizer's performance. The initialization phase generates candidates uniformly within the search bounds:

$$x_{i,j} = LB_j + (UB_j - LB_j) \cdot rand, i = 1, \dots, N, j = 1, \dots, Dim \tag{6}$$

where, N is the number of candidates, Dim is the problem dimension, and $rand \in [0,1]$.

3.1.1 Party nomination phase (exploration)

The party nomination phase simulates internal party elections. In this phase, candidates refine their strategies by

learning from fellow party members. Also, it incorporates innovative suggestions from their staff teams.

Strategy Reference (SR). It is used for the candidates to adjust strategies based on high-performing members.

$$X_1(t+1) = X_R(t) + (X(t) - X_{best}(t)) \cdot AF_{adapt} \quad (7)$$

$$AF_{adapt} = \alpha \cdot r_3 \cdot e^{(1-t/T)} \cdot (1 + \beta \cdot rand) \quad (8)$$

where, $X_R(t)$ is a randomly chosen reference candidate, $X_{best}(t)$ is the best candidate at iteration t, α is a sensitivity factor, and β is an adaptive diversity factor. This adaptive strategy reference encourages candidates to explore underrepresented regions while gradually converging toward promising areas.

Innovative Suggestions of the Staff Team (ISST).

$$X_2(t+1) = w_1 \cdot X(t) + w_2 \cdot X_S(t) \quad (9)$$

$$X_S(t) = (UB - LB) \cdot r_6 + LB \quad (10)$$

Here, $X_S(t)$ represents the staff-proposed innovation. The weights w_1 and w_2 are dynamically adjusted according to candidate performance. This mechanism uses fresh candidate positions and reduces the risk of premature convergence.

3.1.2 Presidential election phase (exploitation)

After securing a party nomination, elite candidates compete in the presidential election phase. This stage is used for the exploitation process to refine solutions toward global optima while maintaining population diversity.

Televised Debate (TD) simulates interaction between candidates from rival parties.

$$X_3(t+1) = X_{best}(t) + r_7 \cdot (X_{best}(t) - TF \cdot X_R(t)) + \gamma \cdot (X_R(t) - X(t)) \quad (11)$$

Here, $TF \in [0,2]$ controls the influence of the rival candidate X_R, and γ is an additional diversity factor to prevent stagnation. TD is used for candidates to exploit the knowledge of competitors**.**

Campaign Speech (CS) models interaction with voters.

$$X_4(t+1) = X_{best}(t) + (X_M(t) - X_{best}(t)) \cdot SF_{adapt} \cdot r_8 \quad (12)$$

$$SF_{adapt} = \left((t+1) - \frac{2r_9 + 1}{(T+1)^2}\right) \cdot (1 + \delta \cdot rand) \quad (13)$$

$$X_M(t) = \frac{1}{N}\sum_{i=1}^{N} X_i(t) \quad (14)$$

The control factor SF_{adapt} dynamically balances exploitation and population diversity. By moving toward the mean solution, candidates avoid being trapped in narrow optima and encourage convergence across the population.

(C) Termination and Best Solution. The process iterates through party nomination and presidential election phases until a maximum number of iterations. T or the convergence criterion is reached. The best candidate X_{best} at termination represents the optimal solution of the problem. Overall, the adaptive control factors (β, γ, δ) dynamically balance exploration and exploitation.

Algorithm 1: Hyperparameter Tuning of DenseNet201-SE using MEOA

Input:

- MRD-8 training dataset D_{train}, validation dataset D_{val}
- DenseNet201-SE architecture M
- Hyperparameter search bounds
- Population size N
- Maximum iterations T_{max}

Output:

- Optimal hyperparameter set X_{best}
- Trained DenseNet201-SE model

1. Initialize candidate population $P = \{X_1, X_2, \dots, X_N\}$ using the initialization equation.
2. Evaluate the fitness of each candidate using the Fitness Function Equation.
3. Determine the current best candidate X_{best}.
4. For $t = 1$to T_{max}do
5. Party Nomination Phase (Exploration)
6. For each candidate $X_i$do
7. Generate candidate position using SR Equation.
8. Generate staff innovation using the ISST Equation.
9. Combine SR and ISST solutions using the Combined Exploration Update Equation.
10. Evaluate the fitness of the updated candidate using the Fitness Function Equation.
11. Replace the current candidate if the fitness improves.
12. End For
13. Presidential Election Phase (Exploitation)
14. Select elite candidates from the population.
15. For each elite candidate, do
16. Update candidate position using TD Equation.
17. Refine candidate position using CS Equation.
18. Combine TD and CS solutions using the Combined Exploitation Update Equation.
19. Evaluate the fitness of the updated elite candidate using the Fitness Function Equation.
20. Replace the elite candidate if the fitness improves.
21. End For
22. Update global best candidate X_{best}.
23. Adjust adaptive parameters using the Adaptive Parameter Update Equation.
24. End For
25. Return X_{best} and train the final DenseNet201-SE model using the optimal hyperparameters.

3.2 Hyperparameter tuning of DenseNet201-SE using Modified Election Optimization Algorithm

The DenseNet201-SE architecture achieves robust feature extraction and channel-wise recalibration, but the classification performance is highly sensitive to the selection of training hyperparameters. Manual tuning is time-consuming and suboptimal. To address this, the MEOA is applied to automatically determine the optimal hyperparameter configuration for the DenseNet201-SE model, thereby maximizing validation accuracy and minimizing overfitting. In this work, each candidate solution in MEOA represents a specific set of hyperparameters (Algorithm 1).

$$X_i = [lr_i, bs_i, dr_i, ep_i] \quad (15)$$

where, lr_i is the initial learning rate (searched in $[10^{-5}, 10^{-2}]$), bs_i is the batch size (searched in $\{16, 32, 64, 128\}$), dr_i is the dropout rate (searched in $[0.2, 0.7]$), and ep_i is the number of epochs (searched in

$\{10, 15, 20, 25, 30\}$). The fitness function for each candidate is defined as the validation accuracy of the DenseNet201-SE model trained with that hyperparameter set on the MRD-8 dataset.

$$Fitness(X_i) = Accuracy_{val}(DenseNet201\text{-}SE(X_i)) \quad (16)$$

The MEOA iteratively optimizes this fitness over multiple generations. The pseudocode for tuning the DenseNet201-SE model using MEOA is provided in Algorithm 1 (Pseudocode).

Initially, a population of candidate solutions is randomly generated within the predefined search space. The fitness is evaluated using the model validation performance. During each iteration, the Party Nomination Phase focuses on exploration by updating candidate positions through the SR and ISST mechanisms. The generated solutions are combined and evaluated. Subsequently, the Presidential Election Phase emphasizes exploitation by selecting elite candidates. The best-performing solutions are preserved, and adaptive parameters are updated dynamically. After reaching the maximum number of iterations, the optimal hyperparameter set is obtained and used to train the final DenseNet201-SE model for MRD-8 classification.

4. RESULTS AND DISCUSSION

The DenseNet201-SE model is coded in Python and simulated using Google Colab. The MRD-8 dataset has been created to support research and clinical applications in retinal disease diagnosis. The images are collected from different sources and merged to solve data imbalance issues. Specifically, the AMD, DR, EM, HG, GA, and MY classes are collected from sources [22-24]. The HT and MY classes are obtained from the ODIR-5K dataset [25]. The MRD-8 dataset provides a comprehensive resource for multi-disease retinal diagnosis. It represents eight different classes of retinal disorders and is selected based on expert recommendations from ophthalmologists. This careful selection enhances the dataset's utility for both medical research and automated diagnostic model development. The dataset contains a total of 13200 images with an equal number of samples in each class. The model is trained with 70% of the images and tested with 30% of the images. The number of samples in the test set is given in Table 1.

Table 1. Dataset description

Disease	Samples Count
Age-related Macular Degeneration (AMD)	494
Diabetic Retinopathy (DR)	481
Epiretinal Membrane (EM)	468
Hypertensive Retinopathy / Hemorrhage (HG)	421
Glaucoma (GA)	482
Myopia (MY)	432
Hypertension (HT)	428
Normal	483
Total	3689

The candidate population size (N) of 30 is assigned with 50 iterations. The tuned values are given in Table 2. The optimiser selects a smaller learning rate and prevents the model from overshooting fine feature gradients during backpropagation. The batch size of 64 is used to smooth out the gradient updates per step. The performance of the proposed DenseNet201-SE is evaluated using training and validation accuracy and loss curves, as illustrated in Figure 4. In the case of the accuracy curve, it can be seen that there is a steady increase in the training accuracy, with an approximate value of 96%, and stability in the validation accuracy at about 93%. This indicates that the model has a strong ability to generalize. In the graph of loss, there is an immediate drop in the loss values of both training and validation in the early epochs, and later, there is convergence.

Table 2. Model hyperparameter configurations before and after Modified Election Optimization Algorithm (MEOA) optimization

Hyperparameter Category	Search Space Bounds	Baseline Setting (Before Tuning)	Optimized Setting (After MEOA Tuning)
Initial Learning Rate (α)	$[10^{-5}, 10^{-2}]$	0.001 (1×10^{-3})	0.00025 (2.5×10^{-4})
Batch Size	{16, 32, 64, 128}	32	64
Dropout Rate	[0.1, 0.5]	0.20	0.35
Number of Epochs	[20, 100]	50	75
Backbone Optimizer	Fixed	Adam	Adam (MEOA-optimized)

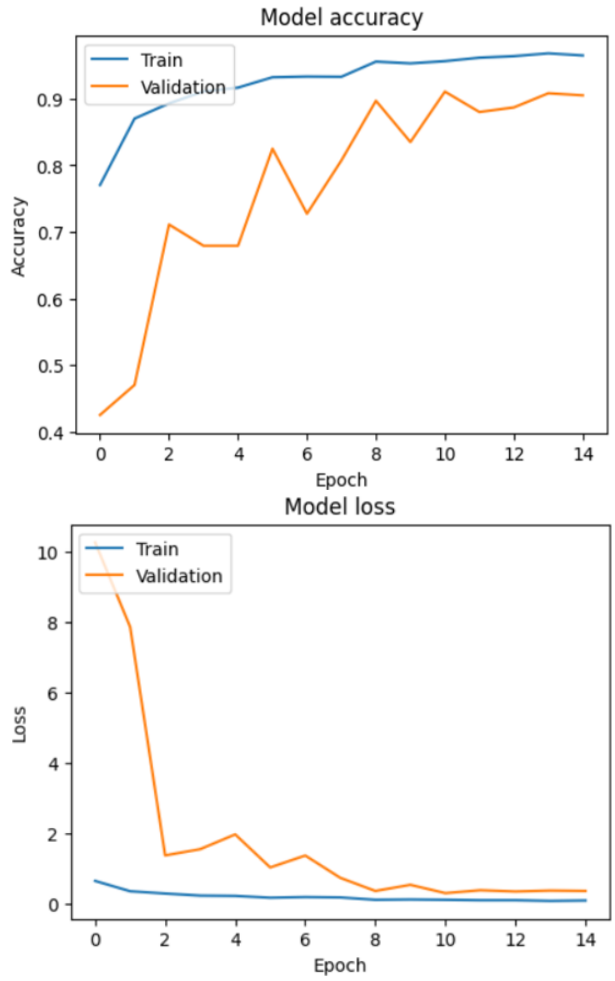

Figure 4. Training and validation accuracy and loss curves of DenseNet201-SE

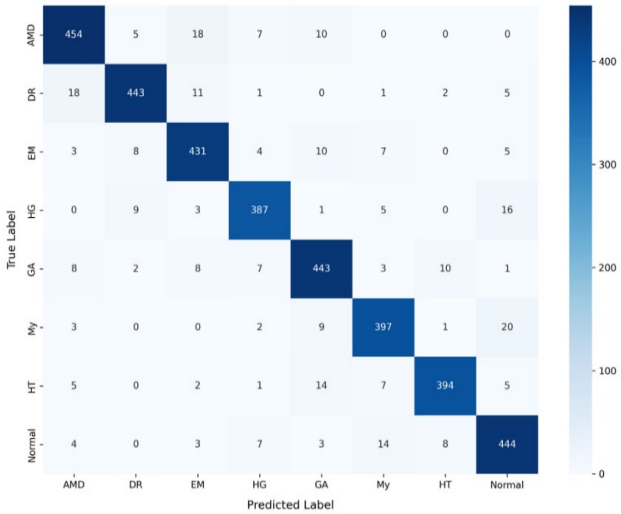

Figure 5. Confusion matrix

The confusion matrix obtained for the DenseNet201-SE model is shown in Figure 5. The higher diagonal values prove the strong capability of the model in distinguishing different retinal disorders. The confusion matrix proves that the proposed DenseNet201-SE achieves high classification accuracy with minimal inter-class confusion for multi-class retinal disease diagnosis. The classwise performance of the model is given in Table 3.

Table 3. Classwise performance of DenseNet201-SE

Class Label	Precision (%)	Recall (%)	F1-Score (%)
AMD	91.72	91.90	91.81
DR	94.86	92.10	93.46
EM	90.55	92.09	91.31
HG	93.03	91.92	92.47
GA	90.41	91.91	91.15
MY	91.47	91.90	91.69
HT	94.94	92.06	93.48
Normal	89.52	91.93	90.71

Note: AMD = Age-related Macular Degeneration; DR = Diabetic Retinopathy; EM = Epiretinal Membrane; GA = Glaucoma; HG = Hemorrhage; HT = Hypertension; MY = Myopia; Normal = Normal retinal image.

DenseNet201-SE model is compared with other models. All models are evaluated on the MRD-8 retinal disease dataset with identical preprocessing, image resizing, and training settings. The results are given in Table 4. The proposed DenseNet201-SE model shows the highest overall performance in F1-score. It maintains a strong balance between precision and recall rates. It is the most reliable model for retinal disease classification in this comparison. The other models, like ResNet, VGG, CNNs, Transformers, and GAN-based architectures, perform well but slightly under the proposed model.

To assess the robustness and generalization capability of the proposed DenseNet201-SE model, a 5-fold cross-validation experiment is conducted. The dataset is randomly divided into five equal subsets. During each iteration, four folds are used for training. The remaining fold is used for testing. The measured results are given in Table 5. The results demonstrate stable performance across all folds with very low variance. The model does not suffer from overfitting and maintains consistent predictive capability on unseen retinal images. The average accuracy obtained across the five folds is close to the overall test accuracy. This proves the reliability and robustness of the proposed framework.

Table 4. Comparison of DenseNet201-SE with other models

Method	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
IDL-MRDD + SSAE	90.00	89.50	89.75	90.00
Swin Transformer V2	91.00	90.50	90.75	91.00
ResNet + Transformer	89.50	89.00	89.25	89.50
FCNNplus	88.50	88.00	88.25	88.50
DeepRetinaNet	88.00	87.50	87.75	88.00
SAT-Net	87.50	87.00	87.25	87.50
Transfer Learning Models	88.50	88.00	87.25	88.50
DCM-CNN	85.50	85.00	85.25	85.50
Customized VGG-19	84.00	84.50	84.25	84.00
Lightweight CNN + Wavelet	83.50	83.00	83.25	83.50
Vision Transformer (ViTs)	89.00	90.50	89.74	91.00
Improved LinkNet + SqueezeNet	82.50	82.00	82.25	82.50
ADRCM	82.00	81.50	81.75	82.00
ResNet-50 + GAN	92.50	92.00	9.25	91.50

Table 5. Five-fold cross-validation results of the proposed DenseNet201-SE model

Fold	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Fold-1	91.84	91.72	91.78	91.80
Fold-2	92.11	91.95	92.03	92.00
Fold-3	92.32	92.18	92.25	92.30
Fold-4	91.96	91.89	91.92	91.90
Fold-5	92.07	92.14	92.10	92.00
Average	92.06	91.98	92.01	92.00
Std. Deviation	0.18	0.18	0.18	0.19

To verify whether the performance improvement achieved by the proposed DenseNet201-SE model is statistically significant, a paired t-test is performed against the competing retinal disease classification models using the cross-validation accuracies. The measured results are given in Table 6. The null hypothesis assumes that there is no significant difference between the compared methods. A significance level of 0.05 is adopted. The obtained p-values are lower than 0.05 for all baseline methods. This shows that the improvements achieved by the proposed model are statistically significant and not due to random variations in the dataset. To evaluate the contribution of each component in the proposed DenseNet201-SE model, a comprehensive ablation study is conducted. The measured results are given in Table 7. The study systematically removes or modifies key architectural and optimization components to quantify their individual impact on classification performance.

Table 6. Statistical significance (paired t-test) analysis

Compared Method	Mean Accuracy (%)	*P*-Value	Significance
IDL-MRDD + SSAE	90.00	0.012	Significant
Swin Transformer V2	91.00	0.028	Significant
ResNet + Transformer	89.50	0.008	Significant
FCNNplus	88.50	0.005	Significant
DeepRetinaNet	88.00	0.004	Significant
SAT-Net	87.50	0.003	Significant
Transfer Learning Models	88.50	0.006	Significant
DCM-CNN	85.50	0.001	Significant
Customized VGG-19	84.00	<0.001	Significant
Lightweight CNN + Wavelet	83.50	<0.001	Significant
Vision Transformer (ViTs)	91.00	0.031	Significant
Improved LinkNet + SqueezeNet	82.50	<0.001	Significant
ADRCM	82.00	<0.001	Significant
ResNet-50 + GAN	91.50	0.041	Significant
DenseNet201-SE (Proposed)	92.00	—	Reference Model

Table 7. Ablation study results on MRD-8 dataset

Model Variant	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)	Improvement over Baseline
Baseline DenseNet201	84.23	83.91	84.07	84.00	Reference
DenseNet201 + SE (without MEOA)	88.45	88.12	88.28	88.30	+4.30%
DenseNet201 + MEOA (without SE)	89.12	88.76	88.94	89.00	+5.00%
DenseNet201-SE + Random Search	90.34	90.01	90.17	90.20	+6.20%
DenseNet201-SE + MEOA (Proposed)	92.06	91.98	92.01	92.00	+8.00%

Table 8. Performance comparison of optimization algorithms for DenseNet201-SE tuning

Optimization Algorithm	Best Accuracy (%)	Mean Accuracy (%)	Std. Deviation (%)	Convergence Iterations	Time per Iteration (s)	Final Loss
Grid Search	88.50	88.50	0.00	N/A	1245.0 (total)	0.42
Random Search	90.20	89.34	0.52	N/A	18.5	0.35
Bayesian Optimization	90.80	90.15	0.38	28	22.3	0.32
Genetic (GA) Algorithm	90.50	89.87	0.45	35	19.8	0.34
Particle Swarm Optimization (PSO)	90.70	90.02	0.41	32	17.2	0.33
Election Optimization Algorithm (EOA)	91.20	90.56	0.35	26	16.5	0.30
Proposed MEOA	92.00	91.45	0.28	22	16.8	0.26

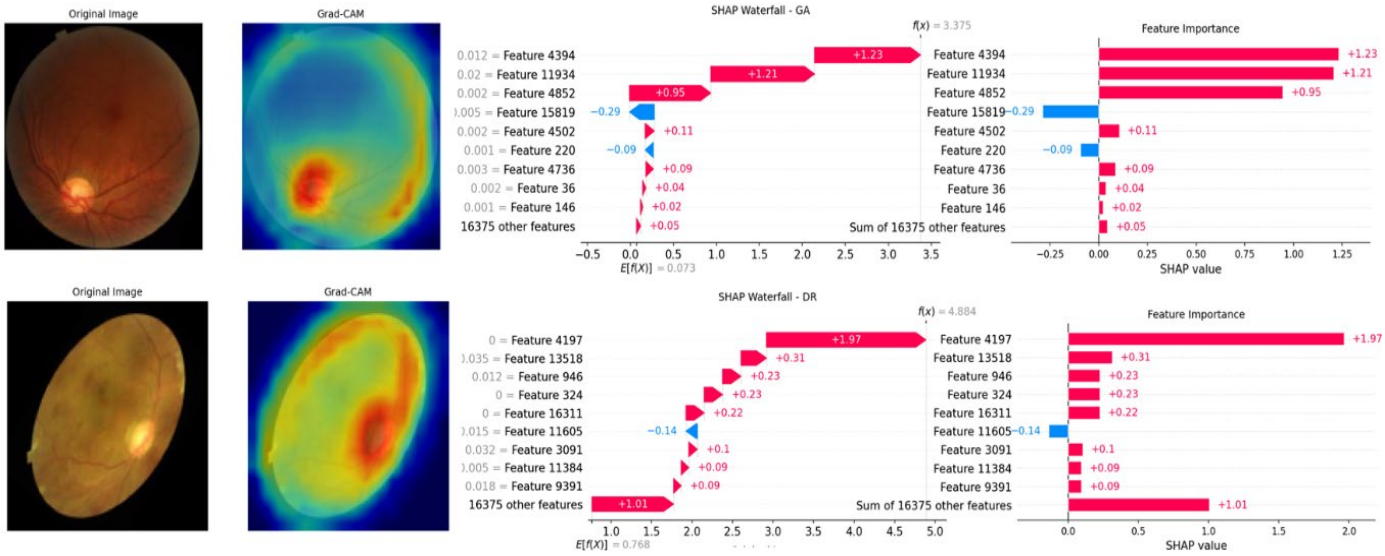

Figure 6. Gradient-weighted Class Activation Mapping (Grad-CAM) and SHapley Additive exPlanations (SHAP) results

All variants are evaluated on the MRD-8 test dataset under identical experimental conditions. The SE attention blocks to the baseline DenseNet201 improve accuracy by 4.30%. This proves that channel-wise feature recalibration is used for the model to focus on clinically relevant retinal structures. The use of MEOA for hyperparameter tuning without SE attention achieves a 5.00% accuracy improvement. The full proposed model achieves 92.00% accuracy. Overall, the SE attention

and MEOA optimization complement each other and support increasing accuracy.

To validate the effectiveness of the proposed MEOA for hyperparameter tuning, a comparative study is conducted against several established optimization algorithms as given in Table 8. The proposed MEOA converges to its optimal solution in 22 iterations. It is faster than the original EOA. This improved convergence rate is attributed to the adaptive control factors. Also, the MEOA achieves the best results when compared to all other optimizers.

For model interpretation, Gradient-weighted Class Activation Mapping (Grad-CAM) and SHapley Additive exPlanations (SHAP) analyses are performed. The plots are shown in Figure 6. It visualizes the discriminative regions and quantify the contribution of learned features responsible for disease classification. The generated heatmaps indicate that the proposed DenseNet201-SE model focuses on clinically relevant structures within the retinal fundus images rather than irrelevant background regions. This behavior shows that the network learns meaningful pathological characteristics associated with each retinal disease category. The localization capability of Grad-CAM proves that the model effectively captures disease-specific visual patterns.

SHAP values reveal how each extracted feature either increases or decreases the probability of a particular disease class. The SHAP waterfall plots quantify the contribution of individual features to the final model output. For GA, features 4394, 11934, and 4852 contributed the most positively to the prediction. This shows their strong influence in identifying glaucomatous patterns. A few features, such as 15819 and 220, had minor negative contributions. This feature slightly reduces the predicted probability. Similarly, for DR, feature 4197 had the highest positive impact. It denotes the model's reliance on these key features to correctly identify DR.

5. CONCLUSION

The proposed method uses attention mechanisms to enhance the classification of retinal diseases. Specifically, DenseNet201 with SE blocks has been utilized to classify eight retinal disorders using a compiled dataset. The performance of the DenseNet201-SE model has been assessed using the MRD-8 databases. The efficiency of this model is illustrated by comparison with other pre-trained models and attention mechanisms. The technique allows decisions on common retinal diseases, even without ophthalmologists. In future studies, an attempt will be made to expand its scope to cover more disorders and concentrate on identifying retinal features.

REFERENCES

[1] Abramoff, M.D., Fort, P.E., Han, I.C., Jayasundera, K.T., Sohn, E.H., Gardner, T.W. (2018). Approach for a clinically useful comprehensive classification of vascular and neural aspects of diabetic retinal disease. Investigative Ophthalmology & Visual Science, 59(1): 519-527. https://doi.org/10.1167/iovs.17-21873

[2] Muchuchuti, S., Viriri, S. (2023). Retinal disease detection using deep learning techniques: A comprehensive review. Journal of Imaging, 9(4): 84. https://doi.org/10.3390/jimaging9040084

[3] Sheet, S.S.M., Tan, T.S., As'ari, M.A., Hitam, W.H.W., Sia, J.S.Y. (2022). Retinal disease identification using upgraded CLAHE filter and transfer convolution neural network. ICT Express, 8(1): 142-150. https://doi.org/10.1016/j.icte.2021.05.002

[4] Altaf, F., Islam, S.M.S., Akhtar, N., Janjua, N.K. (2019). Going deep in medical image analysis: Concepts, methods, challenges, and future directions. IEEE Access, 7: 99540-99572. https://doi.org/10.1109/ACCESS.2019.2929365

[5] Burlina, P.M., Joshi, N., Pekala, M., Pacheco, K.D., Freund, D.E., Bressler, N.M. (2017). Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. JAMA Ophthalmology, 135(11): 1170-1176. https://doi.org/10.1001/jamaophthalmol.2017.3782

[6] Arunkumar, R., Karthigaikumar, P. (2017). Multi-retinal disease classification by reduced deep learning features. Neural Computing and Applications, 28: 329-334. https://doi.org/10.1007/s00521-015-2059-9

[7] Vaiyapuri, T., Srinivasan, S., Sikkandar, M.Y., et al. (2022). Intelligent deep learning-based multi-retinal disease diagnosis and classification framework. Computers, Materials and Continua, 73(3): 5543-5557. https://doi.org/10.32604/cmc.2022.023919

[8] Choudhary, A., Ahlawat, S., Urooj, S., Pathak, N., Lay-Ekuakille, A., Sharma, N. (2023). A deep learning-based framework for retinal disease classification. Healthcare, 11(2): 212. https://doi.org/10.3390/healthcare11020212

[9] Zhao, J.Q., Zhu, J.F., He, J.N., Cao, G.G., Dai, C.X. (2024). Multi-label classification of retinal diseases based on fundus images using ResNet and Transformer. Medical & Biological Engineering & Computing, 62: 3459-3469. https://doi.org/10.1007/s11517-024-03144-6

[10] Di Giammarco, M., Santone, A., Cesarelli, M., Martinelli, F., Mercaldo, F. (2025). Explainable retinal disease classification and localization through convolutional neural networks. Image and Vision Computing, 162: 105667. https://doi.org/10.1016/j.imavis.2025.105667

[11] Sahoo, A.K., Parida, P., Panda, M.K., Nayak, C., Mohankumar, N. (2025). DeepRetinaNet: An automated AI-based framework for retinal disease diagnosis. IEEE Latin America Transactions, 23(8): 718-728. https://doi.org/10.1109/TLA.2025.11072496

[12] Wen, Y., Luo, B., Shi, W.Z., Ji, J.H., Cao, W.M., Yang, X.K. (2025). SAT-Net: Structure-aware transformer-based attention fusion network for low-quality retinal fundus images enhancement. IEEE Transactions on Multimedia, 27: 6198-6210. https://doi.org/10.1109/TMM.2025.3565935

[13] Hemal, M.M., Saha, S. (2025). Explainable deep learning-based meta-classifier approach for multi-label classification of retinal diseases. Array, 26: 100402. https://doi.org/10.1016/j.array.2025.100402

[14] Khan, S.D., Basalamah, S., Lbath, A. (2025). A novel deep learning framework for retinal disease detection leveraging contextual and local feature cues from retinal images. Medical & Biological Engineering & Computing, 63: 2029-2046. https://doi.org/10.1007/s11517-025-03314-0

[15] Sivaz, O., Aykut, M. (2025). Image-level multi-label retinal disease classification with a novel classification head. Computers and Electrical Engineering, 124: 110410.

https://doi.org/10.1016/j.compeleceng.2025.110410
[16] Ghislain, F., Beaudelaire, S.T., Atangana, R., Daniel, T. (2025). Real-time deep learning for multi-label retinal disease diagnosis with embedded system. Computational and Structural Biotechnology Reports, 2: 100035. https://doi.org/10.1016/j.csbr.2025.100035
[17] Cutur, E.S., Inan, N.G. (2025). Multi-class classification of retinal eye diseases from ophthalmoscopy images using transfer learning-based vision transformers. Journal of Imaging Informatics in Medicine, 38: 3110-3124. https://doi.org/10.1007/s10278-025-01416-7
[18] Lisha, L.B., R, S.V.N.S. (2025). Retinal image-based disease classification using hybrid deep architecture with improved image features. International Ophthalmology, 45: 324. https://doi.org/10.1007/s10792-025-03660-w
[19] Mathews, M.R., Anzar, S.M. (2025). An attention driven retinal classification model for the joint grading of diabetic retinopathy and macular edema. SN Computer Science, 6: 743. https://doi.org/10.1007/s42979-025-04266-w
[20] Jeyasri, G., Karthiyayini, R. (2025). Deep learning-based retinal disease classification using an autoencoder and generative adversarial network. Biomedical Signal Processing and Control, 108: 107852. https://doi.org/10.1016/j.bspc.2025.107852
[21] Zhou, S., Shi, Y., Wang, D.J., Xu, X.Z., Xu, M.M., Deng, Y. (2024). Election optimizer algorithm: A new meta-heuristic optimization algorithm for solving industrial engineering design problems. Mathematics, 12(10): 1513. https://doi.org/10.3390/math12101513
[22] Pachade, S., Porwal, P., Thulkar, D., et al. (2021). Retinal fundus multi-disease image dataset (RFMiD). IEEE DataPort. https://ieee-dataport.org/open-access/retinal-fundus-multi-disease-image-dataset-rfmid.
[23] MedicMind. (n.d.). Retinal image databases. https://www.medicmind.tech/retinal-image-databases.
[24] Rivera, M.A.R., Al-Marzouqi, H., Liatsis, P. (2022). Multi-Label Retinal Diseases (MuReD) Dataset. Mendeley Data, V1. https://doi.org/10.17632/pc4mb3h8hz.1
[25] Ahamed, T. (2022). ODIR5K_Classification. Kaggle. https://www.kaggle.com/datasets/tanjemahamed/odir5k-classification.