International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I8P123

Volume 73 Issue 8, 262-272, August 2025


*Original Article*

# An Enhanced Diabetic Retinopathy Classification Using ResNet-DenseNet Hybrid Model

Ashok Kumar Kavuru[1], Rajesh Kumar Patjoshi[2], Rakhee Panigrahi[3]

[1]*Department of Electronics and Communication Engineering, Biju Patnaik University of Technology, Odisha, India.*
[2]*Department of Electronics and Communication Engineering, NIST University, Odisha, India.*
[3]*Department of Electrical Engineering, Parala Maharaja Engineering College, Odisha, India.*

[1]*Corresponding Author : ashok.kavuru@gmail.com*

Received: 25 April 2025 Revised: 12 August 2025 Accepted: 18 August 2025 Published: 30 August 2025

***Abstract -*** *Diabetic Retinopathy is one of the blood vessel problems in the eye, which develops in people with uncontrolled diabetes, causing vision impairment worldwide. Early diagnosis of DR requires urgent attention because it helps to minimize the effects of this disease. Many deep learning methods have been proposed for classifying diabetic retinopathy, but developing models that are both effective and dependable requires a lot of knowledge and computing power. The proposed approach adopts ResNet and DenseNet deep learning architectures for the purpose of DR classification from retina fundus images. The enhancement of retinal fundus image features, including microaneurysms and hemorrhages during image preprocessing, is improved by using Contrast Limited Adaptive Histogram Equalization (CLAHE) and Particle Swarm Optimization (PSO), which serves as a tool to perform hyperparameter tuning by optimizing learning rate together with dropout rate parameters to enhance the performance of the model. The combination methodology produces quicker model convergence with better generalization capabilities. The suggested model for finding diabetes retinopathy has a classification accuracy of 95.01% on the EyePACS dataset, which surpasses existing diagnostic systems. Hence, the proposed system is appropriate for large-scale DR screening.*

***Keywords -*** *Diabetic Retinopathy, ResNet, DenseNet, CLAHE, Particle Swarm Optimization, Fundus images.*

## 1. Introduction

Diabetic Retinopathy (DR) can result from uncontrolled diabetes mellitus (blood sugar) that affects the retina's light-sensitive layer [1]. It is a major contributor to vision impairment globally. If not treated early on, diabetic retinopathy can result in visual loss or full blindness. When blood sugar levels stay too high for a long time, they can hurt the small blood vessels in the retina. This can cause the blood vessels to leak, bulge and develop abnormally. In earlier times, doctors manually visually examined retinal fundus images and classified the disease. However, it requires a huge amount of time, and to avoid these limitations, there is a need for efficient, accurate, and scalable diabetic retinopathy screening systems. Moreover, high sugar levels can damage multiple organs and tissues throughout the body [2]. The retina functions as one of the most susceptible parts of the eye because it represents the light-sensitive tissue that exists at its posterior end. The scientific understanding of diabetic retinopathy as a major vision impairment disorder started at the beginning of the twentieth century along with the increasing numbers of diabetic patients worldwide. Diabetic retinopathy exists as a major preventable reason for vision loss throughout the world despite recent significant progress in diagnostic and therapeutic approaches. This problem worsens because of a continuous rise of diabetes cases worldwide [3]. The diabetic retinopathy health crisis impacts approximately 103 million people world-wide at present, of which nearly 28 million are at risk for progression to vision threatening stages of the disease [4]. An estimated 30% of people over the age of 40 with diabetes in developed countries like the United States (US) have diabetic retinopathy, and 4.4% of these individuals have macular edema with a significant risk of losing vision. Diabetic retinopathy is much common in lower and moderate-income countries due to less access to early diagnosis and correct treatment. DR contains two main types, including Non-Proliferative Diabetic Retinopathy (NPDR), which follows Proliferative Diabetic Retinopathy (PDR). The initial stage of the disease is characterized by mild changes in the retinal known as NPDR which progressively deteriorates [5]. Subsequently, the development of microaneurysms and retinal haemorrhages are important signs of NPDR and without proper and timely management, these early manifestations could progress and potentially lead to a more severe outcome and evolution to the proliferative phase. Diabetic retinopathy has a last stage characterized by typical fragile blood vessel growth on the retina's surface and it is referred to as Proliferative Diabetic Retinopathy (PDR). The newly formed vessels are susceptible to burst, and this leads

to vitreous haemorrhage and retinal detachment. If not received proper treatment by the medical person, it may lead to irreversible blindness or major visual impairment during this crucial time [6]. Considerably, recent improvement in the Artificial Intelligence (AI) along with Machine Learning (ML) technologies made the way for deep learning approaches to be introduced and delivers very effective and powerful solutions to the automatic process of medical images. Various literature deliberates that Convolutional Neural Networks (CNNs) are highly successful in disease detection using medical images and achieving more accuracy over the performance of human experts [7]. Owing to their ability to extract and learn complex hierarchical features of retinal images, CNN-based architectures are one of the widely used models in DR identification [8].

Yet these models provide several benefits; they also have some drawbacks. Hyperparameters such as learning rate, batch size, optimizer and architectural depth all have a significant impact on model performance. Tuning these manually or automatically is time-consuming and computationally expensive. Overfitting is another limitation when training on small or imbalanced datasets. Moreover, CNN architectures like ResNet and DenseNet have become highly accepted architectures owing to their amazing performance in image classification tasks.

Therefore, ResNet utilizes a residual learning framework that effectively makes it easier to train deeper networks by reducing the gradient vanishing problem. On the other hand, DenseNet uses dense connections by having direct connections between all layers, allowing for better improvement in information flow through the network and feature reuse. Such an architecture resolves gradient vanishing issues effectively and improves feature propagation, making it highly suitable for complicated image classification problems. Since effective feature extraction is vital for the classification of diabetic retinopathy. Recent research demonstrates that combining deep learning models allows separate architectures to share their respective benefits, thus achieving superior results than individual models would achieve alone.

This paper formulates an efficient classification model for DR using a hybrid ResNet-DenseNet architecture, CLAHE for image processing and PSO for hyperparameter tuning. The goal of this effort is to sort retinal images into 5 stages of diabetic retinopathy: No DR, Mild DR, Moderate DR, Severe DR and Proliferative DR. The proposed method provides a precise, scalable computer-based method for the early diagnosis of the disease. Employing high-end deep learning architectures and various suitable preprocessing and optimization, followed by evaluating the performance on the widely used fundus image dataset, this system could contribute to a significant reduction of diabetes associated visual impairment and blindness.

The following important points are brought out in this work:

- ResNet50 and DenseNet121 are used as dual backbones for feature extraction, utilizing their strengths in medical image analysis.
- CLAHE is the preprocessing step to enhance contrast in fundus images, improving the visibility of key diagnostic features like microaneurysms and hemorrhages.
- PSO algorithm is used for fine-tuning the hyperparameters, including learning rate, momentum and dropout rates.

The methodology was tested on the EyePACS dataset and yielded significant performance. In summary, the methodology of the research seeks to greatly enhance the diagnosis and categorization of DR disorders, offering a useful tool for healthcare practitioners and maybe resolving issues related to conventional diagnostic techniques. The rest of this study's framework is set up as follows: In Section 2, a literature review of the current methods is presented. The description of the proposed model is discussed in Section 3. Section 4 gives the results of the experiment and a comparison study for classifying diabetic retinopathy. Finally, Section 5 summarizes the study and defines the future scope of the study.

## 2. Literature Review

The current section describes a wide-ranging review of new advancements in DR classification and medical image investigation, aiming at deep learning methods. Early methods involved traditional machine learning approaches, which required manual features and struggled with scalability. The introduction of CNNs can fully change the growing accuracy of classification. The overview of existing studies to classify DR disease based on deep learning methods is summarized in Table 1. This table presents methodologies, preprocessing, architecture, datasets, performance measures, and a comparative survey of available knowledge.

These details highlight the progress and obstacles of each approach. Qureshi et. al. [8] introduced an Active Deep Learning (ADL) framework that utilized CNNs for feature extraction in conjunction with Expected Gradient Length (EGL) based active learning to decrease the burden of manual annotation for DR severity classification by using the EyePACS dataset, and showed that their model achieved high accuracy. Similarly, Das et al. proposed adaptive histogram equalization and morphological operations to improve the overall vessel segmentation and feature extraction process and their method was found to have a stable performance regarding the DIARETDB1 dataset [9]. Aware of the need for hyperparameter optimization, researchers such as Menaouer et al. stated about hybrid CNN models with VGG16 and VGG19 architectures.

**Table 1. Summary of recent surveys on diabetic retinopathy classification**

| Study | Methodology | Preprocessing Techniques | Architecture/ Model | Dataset | Performance Metrics |
|---|---|---|---|---|---|
| Menaouer et al.[10] | Hybrid Deep Learning model | None specified | CNN with VGG16 and VGG19 | 5,584 retinal images | Accuracy: 90.6%, Recall: 95%, F1-Score: 94% |
| Dai et al. [11] | Multi-Task Deep Learning (DeepDR) | None specified | Multi-task CNN | 466,247 fundus images | AUC: 0.943-0.972 across severity levels |
| Alyoubi et al. [12] | Synergic Deep Learning (SDL) | Histogram-based segmentation | SDL model | Messidor | High classification accuracy |
| Math and Fatima [13] | Segment-based learning framework | None specified | Pre-trained CNN modified for segment-level estimates | Kaggle DR | AUC: 0.963, Sensitivity: 96.37% |
| Kobat et al. [14] | Patch-based segmentation and DenseNet201 | Horizontal and vertical patch division | DenseNet201 + SVM | Custom dataset and APTOS 2019 | Accuracy: 94.06% |
| Tymchenko et al. [15] | Multi-stage transfer learning strategy | None specified | CNN with transfer learning | APTOS 2019 | Sensitivity: 0.99, Specificity: 0.99 |
| Sikder et. al. [16] | Decision tree-based ensemble learning | Gray-level intensity enhancement | Ensemble learning | APTOS 2019 | Accuracy: 94%, F-measure: 93% |
| Nazir et al. [17] | Two-phase framework with CenterNet | Feature extraction and dataset annotations | CenterNet with DenseNet-100 | APTOS 2019, IDRiD | High accuracy and cross-dataset robustness |
| Yi et al. [20] | RA-Efficient Net with residual attention blocks | Transfer learning | RA-Efficient Net | ImageNet pre-trained model | Enhanced feature extraction and lesion detection |
| Madarapu et. al. [21] | Multi-Resolution based Convolutional Attention Network (MuR-CAN) | Layer-by-layer convolution with expansion rates | MuR-CAN + SVM | Experimental evaluations | Improved performance over existing methods |
| Akhtar et al. [22] | EfficientNetB3 and VGG16 | Gaussian filtering, data augmentation | EfficientNetB, VGG16 | APTOS 2019 | Superior accuracy and robustness |
| Mushtaq and Farheen [19] | DenseNet-169-based classification | Data augmentation and preprocessing | DenseNet169 | APTOS 2019 | Accuracy: 90% |

The authors used retinal ischaemia to classify DR severity with an accuracy of 90.6% on a database of 5,584 images [10]. Dai et al. [11] Further expanded this work by developing DeepDR, a multi-task CNN-based system trained on more than 466,247 images to detect lesions and grade DR severity and reported AUC values above 0.94. Alyoubi et al. [12] proposed a Synergic Deep Learning (SDL) architecture to classify DR types obtained from pre-processed retinal images. Their method combines histogram-based segmentation with CNNs to achieve good accuracy on the Messidor dataset. Math et. al. [13] proposed a segment-based CNN framework without requiring any manual annotations, which could learn features from the data, achieving 96.37% sensitivity value on the Kaggle DR dataset for the segmentation problem. Kobat et al. [14] used a patch-based segmentation method combining DenseNet201 architecture and SVM classifiers to enhance lesion detection. This method achieves 94.06% accuracy by patch separation and employs NCA. Tymchenko et al. [15] proposed an improved transfer learning method to reduce the influence of label inconsistency and achieved impressive sensitivity and specificity scores on the APTOS2019 dataset. Transfer learning-based approaches have been introduced to tackle dataset challenges. Sikder et al. [16] proposed an ensemble learning method in which texture features and gray level intensity are extracted from the fundus images, and achieved an accuracy of 94% and F-measure of 93.5% using a decision tree-based classifier. Similarly, Nazir et al. [17] achieved simultaneous lesion classification on the APTOS 2019 and IDRiD by combining CenterNet and DenseNet-100, and attaining scalability and robustness. Employing Efficient neural architectures to classify vision-threatening DR stages, the best AUC values of 0.984 were reported over the EyePACS dataset by Chetoui and Akhloufi [18]. Preprocessing and data augmentation were conducted during model development by Mushtaq and Farheen [19], leading to the employment of attention mechanisms and ensemble learning techniques to enhance classification accuracy. Yi et al. [20] incorporated residual attention blocks

into Efficient Net, which improved lesion detection and yielded competitive AUC scores. They proposed RA-Efficient Net based on the principle of residual attention blocks at all levels for the enhanced detection of the lesions in DR images, and to accurately perform multi-class classification of DR, they improved the algorithm by adaptive transfer learning and multi-stage classifiers. Madarapu et al. [21] proposed Multi-Resolution Convolutional Attention Networks (MuR-CAN), which include dilation-based convolution layers to better capture fine spatial details for improved classification performance. The CNN architecture was further improved by non-local blocks and channel-spatial attention channels to improve contextual feature extraction and reduce noise captured in DR images, resulting in better accuracy. Similarly, Akhtar et al. [22] employed transfer learning techniques by merging EfficientNetB3 and VGG16 and showed that Gaussian filtering and augmentation could provide better accuracy and robustness. Despite these improvements, older methods suffered from overfitting, a lack of generalization capabilities, and computational inefficiencies. Considering these concerns, more contemporary methods, such as CLAHE for contrast enhancement and optimization algorithms such as PSO for hyperparameter tuning, have been presented to increase model performance and robustness in diabetic retinopathy classification. These techniques effectively prevent overfitting by extracting features and minimizing noise, thus making the model robust. A combination of CLAHE preprocessing, along with PSO optimization in a CNN-based architecture, has developed a more accurate and larger-scale model for DR classification. Specifically, the framework aims to overcome the challenges faced by data imbalance, feature inconsistency and overfitting.

## 3. Proposed Methodology

This paper suggests a hybrid deep learning architecture for the efficient classification of the stages of diabetic retinopathy. Medical image analysis with dual backbones such as ResNet50 and DenseNet121 has been applied to utilize their strengths. CLAHE improves fundus, such as microaneurysms and hemorrhages. PSO is utilised for fine-tuning of hyperparameters together with dropout rates, learning rate and momentum for optimization. This method guarantees optimal convergence and enhanced classification performance. This method was tested on Kaggle's EyePACS fundus image datasets, yielding important performance.

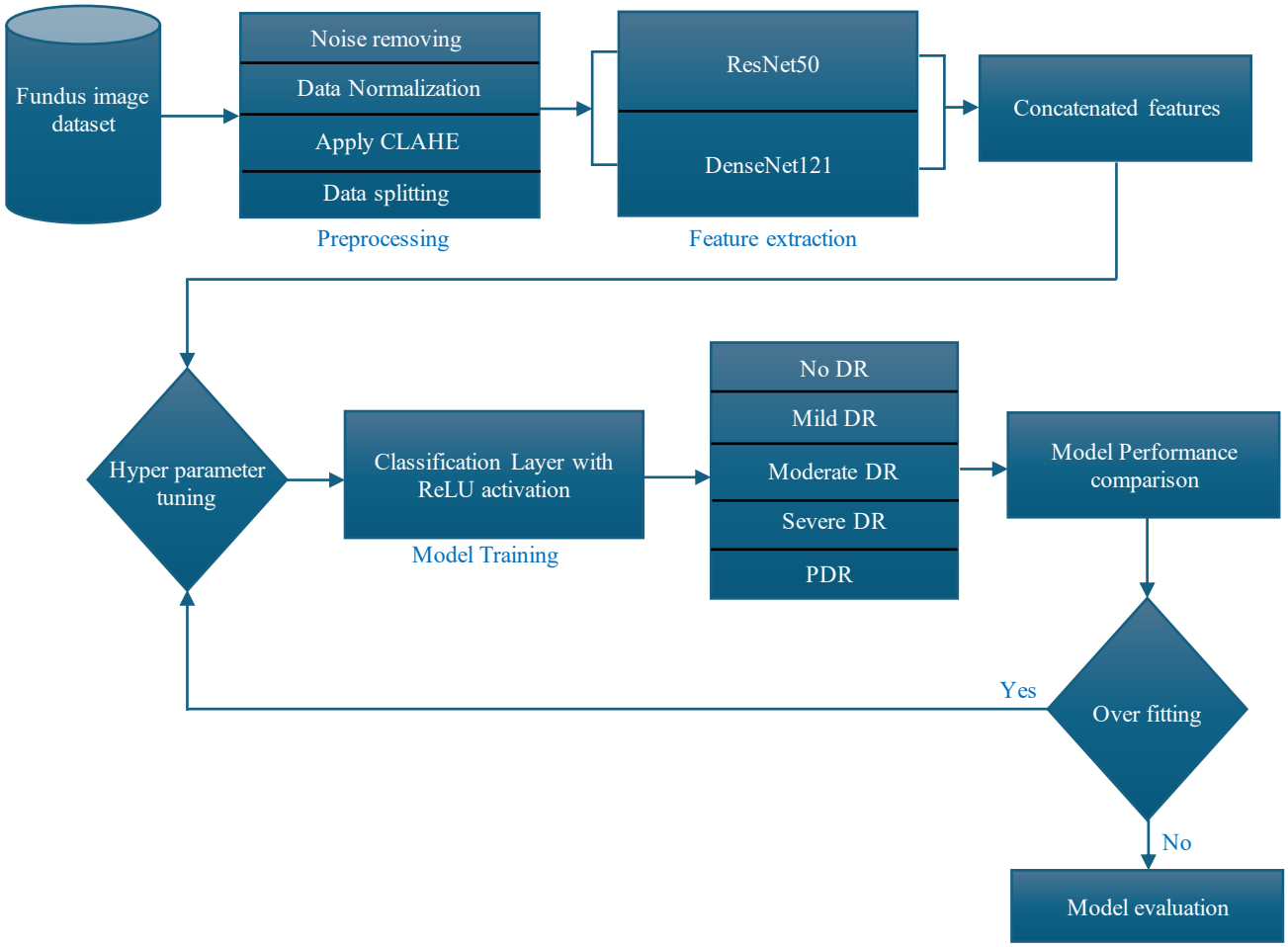


**Fig. 1 Overview of the proposed methodology**

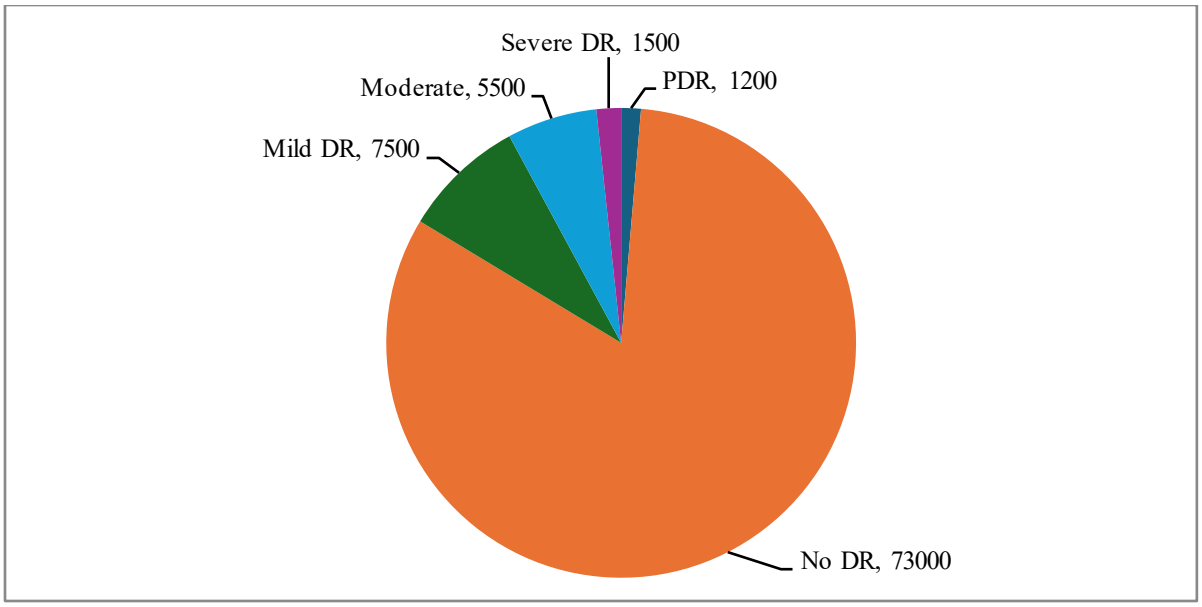


**Fig. 2 Distribution of the EyePACS dataset**

The architecture of the suggested method utilized for the classification of DR is shown in Figure 1. It describes the important phases of methodology, starting from input image preprocessing of fundus images using CLAHE. They are followed by two DL architectures, namely ResNet50 and DenseNet121, for feature extraction. PSO is applied as an optimizer to tune the hyperparameters of the CNN model image by enhancing the structures important for diagnosis.

### *3.1. Input Dataset*

Deep learning models need a lot of data for effective training. The EyePACS dataset [23] is one of the most popular publicly available fundus image datasets for research on classifying diabetic retinopathy. It has a large and varied collection of fundus images. The diabetic retinopathy levels based on severity lead to image categorization between five groups: No DR, Mild DR, Moderate DR, Severe DR and Proliferative DR. The Kaggle edition contains 88,700 fundus images and image distribution as follows in Figure 2.

### *3.2. Preprocessing*

Noise removal is the first step in preprocessing, which removes undesired artifacts and fluctuations from the image. Here, key features like blood arteries are carefully preserved using the technique of median filtering. The size of each image is normalized in the dataset. Normalization adjusts the pixel intensity values to fall within a specified range, usually between 0 and 1. This ensures uniformity across all images and helps the model train more efficiently.

Image normalization is calculated by using:

$$I_{norm} = \frac{I - I_{min}}{I_{max} - I_{min}} \quad (1)$$

I initial value of the pixel
$I_{min}$ represents the image's minimal pixel value.
$I_{max}$ represents the image's maximum pixel value.

The image quality improvement is achieved by enhancing contrast in retina images; the CLAHE technique is applied [24]. CLAHE improves blood vessels and other critical features, especially in darker regions of the retina. The CLAHE algorithm can be represented as:

$$I_{enhanced} = \text{CLAHE}(I, \text{ClipLimit}, \text{TileGridSize}) \quad (2)$$

Here, I is the original input image
ClipLimit controls the contrast enhancement.
tileGridSize defines the size of the contextual regions.

The success of CLAHE in improving fundus image quality is visibly displayed in Figure 3, which shows pre- and post-processed images for various stages of diabetic retinopathy, including mild, moderate, severe and proliferative. The use of CLAHE produces substantial contrast enhancement, allowing for improved delineation of retinal structures such as microaneurysms, hemorrhages, and neovascular formations. This augmentation is especially useful in darker parts of the retina, where diseased abnormalities can be difficult to identify. The enhanced visibility of these parameters allows for more accurate diagnosis and feature extraction in later processing phases.

### *3.3. Feature Extraction and Feature Fusion*

Two pre-trained deep learning models, namely Residual Network with 50 layers (ResNet50) [25] and Densely Connected Convolutional Network with 121 layers (DenseNet121) [26], are employed to extract the features. These models operate without enabling fully connected layers

to obtain feature maps from the convolutional layers. The key benefit behind hybridizing ResNet and DenseNet is utilisation of ResNet's deep residual learning capabilities alongside DenseNet's feature reuse mechanism.

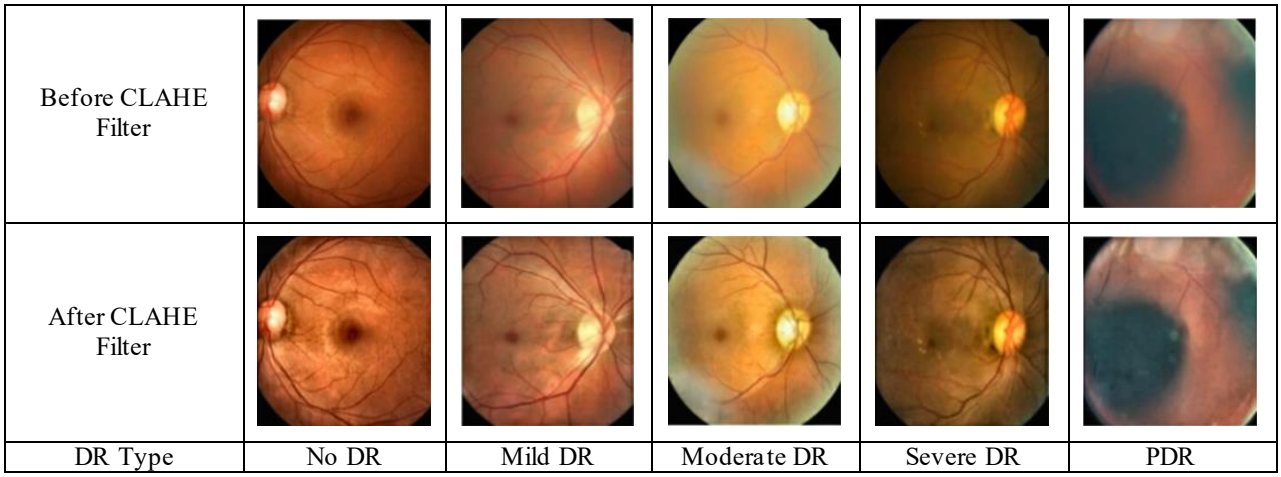


**Fig. 3 Images before and after CLAHE operation**

Let $f_{ResNet}$ and $f_{DenseNet}$ indicate the feature maps extracted from ResNet50 and DenseNet121 models, respectively.

$$F_{ResNet} = f_{ResNet}(I) \quad (3)$$

$$F_{DenseNet} = f_{DenseNet}(I) \quad (4)$$

Where I is the Input Image

The features from ResNet50 and DenseNet121 are extracted and combined to make a hybrid feature vector. This combination produces a more robust feature set by capturing the complementary qualities of both models.

Let the fused feature vector be represented as:

$$F_{fused} = [F_{ResNet} || F_{DenseNet}] \quad (5)$$

Where '||' represents the concatenation process

### *3.4. Classification Layer*

The concatenated feature vector is passed through a series of fully connected (dense) layers to perform the classification task. In these layers, every input neuron is coupled to all output neurons, making a structure that is highly connected. Each of these dense layers uses a Rectified Linear Unit (ReLU) non-linear activation function. The output of the $i^{th}$ hidden neuron

$$h_i = ReLU(W_i F_{fused} + b_i) \quad (6)$$

Where $F_{fused}$ is the input feature vector

$W_i$ and $b_i$ represent the weight matrix and bias for layer i.

The final classification layer employs a softmax function to output class probabilities for the five DR categories:

$$P(y = j|x) = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \quad (7)$$

Where $Z_j$ is the output of class j

y is the target label

x is the input features, and the total number of classes is K.

### *3.5. Optimization Using Particle Swarm Optimization (PSO)*

The proposed method uses Particle Swarm Optimization (PSO) to automate hyperparameter tuning [28], with each particle representing a possible set of hyperparameter values. These particles explore the search space by adjusting their positions depending on both their own experience (personal best) and the swarm's overall performance (global best). The optimization procedure is led by a fitness function, which is defined as the model's validation accuracy. By iteratively refining particle locations and velocities, PSO effectively explores the hyperparameter space, decreasing the need for manual tuning and the risk of becoming trapped in local optima. As a result, the model performs better in classification, with greater accuracy, F1-score, and AUC than configurations that use default or user-chosen parameters. Each particle $P_i$ in the swarm is characterized by:

Position $X_i(t)$, which represents a possible solution (set of hyperparameters)

Velocity $V_i(t)$ represents the rate of change of the particle's position. Personal best $P_i^*$ the best solution found by the particle.

Global best G is the best solution found by the entire swarm.

The velocity and position update rules in PSO are updated as follows in Equations (8) and (9).

$$V_i(t+1) = \omega V_i(t) + b_1\, r_1\left(P_i^* - X_i(t)\right) + b_2\, r_2(G - X_i(t)) \quad (8)$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \quad (9)$$

Where ω denotes the inertia weight
$b_1$ and $b_2$ are the coefficients of acceleration
$r_1$ and $r_2$ are the random values [0,1]
PSO is applied to reduce the categorical cross-entropy measure of classification error on the validation set.

Cross entropy is defined as:

$$L = -\sum_{i=1}^{N} \sum_{j=1}^{K} y_{ij} log\ (\widehat{y_{ij}}) \quad (10)$$

Here, N is the total number of samples, K represents the number of classes, $y_{ij}$ and $\hat{y}_{ij}$ are real and the predicted probability that the $i^{th}$ sample belongs to class j.

### *3.6. Training and Evaluation*

Applying the Adam optimizer, the model is initially trained using a learning rate of 0.001. To avoid overfitting, training is done over 100 epochs with early stopping. The categorical cross-entropy loss function helps to reduce training loss.

The test set is used to evaluate the model's performance, and the classification results are then compared to the most advanced diabetic retinopathy classification techniques to determine the performance of the proposed model.

## 4. Results and Discussion

The experimental setup was built on a system powered by an Intel i7 processor with Windows 11 operating system, featuring a 2.4 GHz processing speed and 16 GB RAM supported by an NVIDIA GeForce RTX 4060 Laptop GPU.

The simulations were executed using Python in Visual Studio Code.

### *4.1. Performance Metrics*

The performance of the model is assessed using a few parameters [29].

Accuracy: Accuracy is the percentage of right predictions out of all the predictions.

$$Accuracy = \left(\frac{TP+TN}{TP+TN+FP+FN}\right) \quad (11)$$

Sensitivity(recall): It measures the ability of a model to correctly identify the total positive instances

$$Sensitivity = \frac{TP}{TP+FN} \quad (12)$$

Area Under the Curve (AUC): It shows how successfully the model differentiates positive and negative classes; larger values mean better performance. The Area Under the Curve (AUC) is calculated using the Receiver Operating Characteristic (ROC) curve, which indicates the connection between the True Positive Rate (TPR) and the False Positive Rate (FPR) at various categorization thresholds.

Precision: Determines how many of the anticipated positive situations are true positives.

$$Precision = \frac{TP}{TP+FP} \quad (13)$$

F1-Score: A measure that shows the trade-off between precision and recall in tasks that involve classifying things into two or more classes

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Re\,call} \quad (14)$$

### *4.2. Performance on the EyePACS Dataset*

Ablation research is carried out to learn more about how each parameter influences the model's operation. The hyperparameters in Table 2 are tuned for their effect on the final model performance. CLAHE is used to preprocess fundus images of the EyePACS dataset as it enhances local contrast to allow better differentiation in the local neighborhood of retinal features such as hemorrhages and microaneurysms.

**Table 1. Hyperparameter settings**

| S. No. | Name of the Parameter | Value |
|---|---|---|
| 1 | Learning Rate | 0.001 |
| 2 | Momentum | 0.9 |
| 3 | Dropout Rate | 0.3 |

The preprocessing step significantly enhances the feature visibility of the model and enables it to discriminate among various types of diabetic retinopathy. CLAHE improves the contrast of the model trained on the EyePACS dataset without adding noise. This helps the model find important features and increases accuracy, precision and recall. The model performance is greatly enhanced on the EyePACS dataset, with the application of CLAHE and PSO for hyperparameter optimization.

The iterative search of PSO enables the model to reach convergence faster whilst having good generalization capabilities. Obtaining more precise hyperparameter values from the PSO implementation on the EyePACS dataset leads

to significantly higher performance metrics and overall model accuracy. This can be seen from the ablation study on the EyePACS dataset, as reported in Table 3. When CLAHE is applied without PSO, the accuracy comes out to be 93.26%; however, contrast is enhanced, but optimization does not yield a positive impact. Using PSO without CLAHE raises the accuracy to 94.37% by fine-tuning the hyperparameter.

**Table 3. Performance of model variants**

| Parameter | Model with CLAHE and without PSO | Model without CLAHE and with PSO | Model with CLAHE and PSO |
|---|---|---|---|
| Accuracy | 93.26% | 94.37% | 95.01% |
| Precision | 94.39% | 95.43% | 96.28% |
| Recall | 91.80% | 93.46% | 94.21% |
| F1-Score | 93.08% | 94.24% | 94.84% |
| AUC | 93% | 94% | 95% |

The proposed model's training and validation accuracy have changed over 100 epochs, as shown in Figure 4. This chart shows that the model performs well and consistently on both the training and validation datasets. The accuracy curve indicates that the model has not overfit and is well-optimized for the classification task. Throughout the training phase, the training (black) and validation (red) curves remain tightly matched, indicating that the model maintains strong generalization capability while overfitting is negligible. The absence of a significant gap between the curves indicates robust learning and suitable model complexity in relation to the dataset. After a few epochs, both the training and validation accuracy curves plat area, with a steady performance range of 90% to 95.01%. The model ultimately achieves a high validation accuracy of 95.01%, validating the effectiveness of the training technique, model architecture, and hyperparameter selection.

A comparative evaluation of the proposed model against several existing deep-learning approaches for effective DR classification is presented in Table 4. The model's performance is measured using conventional classification measures such as Accuracy, Precision, Recall, F1-Score, and Area Under the Curve (AUC). The Proposed Model (enhanced with CLAHE for contrast improvement and PSO for hyperparameter tuning) outperforms all the other methods across all five evaluation metrics. The model attained the greatest accuracy of 95.01% when categorizing Diabetic Retinopathy (DR) images across five evaluation categories.

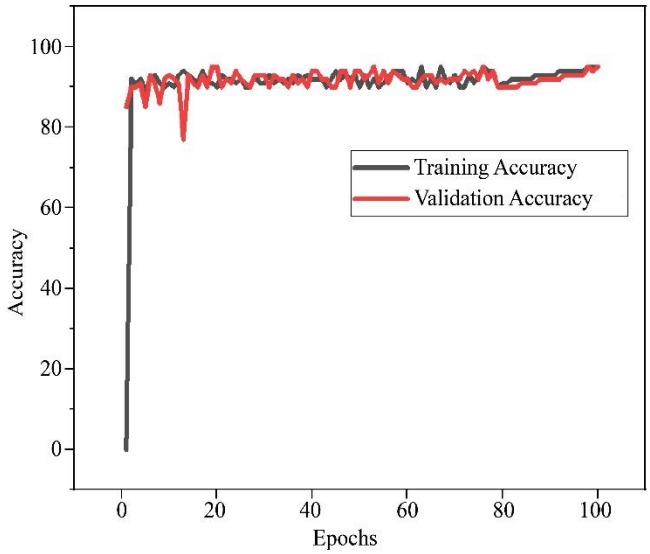


**Fig. 4 Training and validation accuracy**

**Table 4. Performance evaluation of proposed and existing metrics**

| Model | Accuracy | Precision | Recall | F1-Score | AUC |
|---|---|---|---|---|---|
| CNN [30] | 94.50 | 91.2 | 93.5 | 92.3 | 0.91 |
| CNN ResNet-101 [31] | 87.37 | 77.8 | 93.88 | 85.1 | 0.84 |
| 3-Headed CNN (TTA) [32] | 91.9 | 90.2 | 92.1 | 91.1 | 0.92 |
| CABNet [33] | 84 | 84.7 | 87.3 | 85.9 | 0.86 |
| Proposed Model with CLAHE and PSO | 95.01 | 93.7 | 94.9 | 94.3 | 0.94 |

Its precision (93.7%) and recall (94.9%) reflect an excellent balance between minimizing false positives and false negatives, which is critical in medical diagnosis scenarios. The F1-score of 94.3% confirms this balance, and an AUC of 0.94 signifies strong discriminative ability. Among the compared models, the baseline CNN model [30] also performs well, with 94.50% accuracy and an AUC of 0.91, but it still falls slightly short of the proposed model in all metrics. The 3-Headed CNN (TTA) [32] achieves a good trade-off with 91.9% accuracy and an AUC of 0.92, indicating robustness through test-time augmentation. However, its performance is marginally lower than the proposed model. In contrast, CNN ResNet-101 [31] and CABNet [33] show relatively lower performance. ResNet-101 has the lowest accuracy (87.37%) and AUC (0.84), despite a high recall (93.88%), which suggests it detects positives well but at the cost of more false positives. CABNet shows moderate results across metrics but is notably less effective than the proposed method, especially in accuracy and AUC. The proposed model with CLAHE and PSO has maximum accuracy (95.01%), indicating higher performance as shown in Figure 5. This graphical representation clearly demonstrates the effectiveness of the suggested strategy when compared to existing techniques.

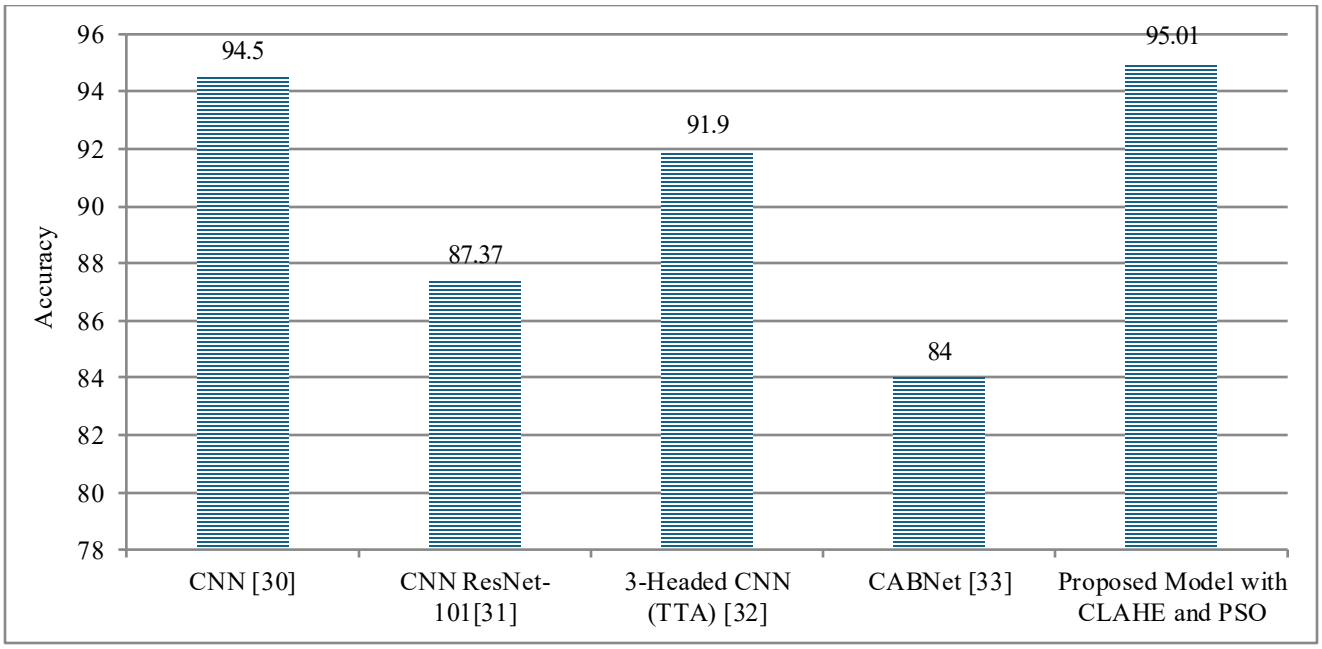


**Fig. 5 Comparison of accuracy across different models**

This model's better classification performance can be attributed to merging ResNet and DenseNet designs. ResNet contributes residual learning, which minimizes the vanishing gradient issue and enables more advanced network training without degradation. This enables the model to efficiently capture global semantic information. In contrast, DenseNet offers dense connections, in which each layer receives inputs from all previous layers, enabling feature reuse and increasing gradient flow, hence improving the learning of fine-grained local features. The hybrid model, which combines these two powerful architectures, benefits from both depth and dense feature propagation, making it well-suited for the diabetic retinopathy classification task, where both global retinal patterns and local lesions (such as microaneurysms, exudates, and hemorrhages) must be identified.

## 5. Conclusion and Future Scope

A hybrid DL model for Diabetic Retinopathy classification is proposed and implemented in this study, achieving considerable advancements with respect to existing models by extending the framework with CLAHE preprocessing and PSO. CLAHE and PSO enable the capture and enhancement of diagnostic characteristics from retinal images, adapting the model's hyperparameters in favour of maximum efficiency. CLAHE emerged as an essential preprocessing step for enhancing local contrast in fundus images without magnifying noise and showing fine features like microaneurysms, hemorrhages and exudates. CLAHE makes these tiny retinal structures more distinguishable, which is essential for diagnosing the severity of DR, thus significantly improving feature extraction by the model. As seen through the outcomes across the EyePACS dataset, CLAHE provides the model with a substantial advantage for distinguishing between different DR levels and thus yields more précised classifications. Additionally, the employment of PSO aids the success of the model in terms of hyperparameter optimization by adjusting the learning rate, momentum and dropout rate. PSO is different from traditional gradient-based optimization methods; every particle in the swarm is like a separate candidate solution moving through hyperparameter space. PSO converges to a set of hyperparameters that maximizes model performance in terms of accuracy via iterative updates informed by individual and neighborhood experiences. This leads to better convergence rates, minimized overfitting and improved generalization, as depicted from the performance metrics achieved on unseen test data. The maximum accuracy of this proposed model on the EyePACS dataset is 95.01%.

Moreover, the model has excellent precision, recall, F1-score, and AUC values, suggesting a fair capability for evaluation metrics. These results confirm the effectiveness of employing advanced techniques such as CLAHE and PSO in diabetic retinopathy classification, as these improve both the feature extraction and generalize the capacity of the model. Although the current model was trained and assessed using the EyePACS dataset, its performance can be further validated by applying it to other publicly available datasets such as Messidor and APTOS2019. This would assess its ability to generalize across various imaging environments. Furthermore, the model, incorporated into cloud-based diagnostic systems or mobile screening tools, could offer significant advantages, particularly in resource-limited regions, by facilitating early and large-scale detection of diabetic retinopathy with minimal reliance on clinical personnel.

## References


[1] James M. Breen et al., "501-P: Differences in Disease Severity and Treatment for Diabetes-Related Retinopathy between Type 1 and 2 Diabetes Mellitus," *Diabetes*, vol. 72, no. 1, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Nicole Eichert et al., "60-OR: High Prevalence of Multiorgan Steatosis and Fibro Inflammation, Identified by Multiparametric Magnetic Resonance Imaging, in People with Type 2 Diabetes," *Diabetes*, vol. 71, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Moien Abdul Basith Khan et al., "Epidemiology of Type 2 Diabetes - Global Burden of Disease and Forecasted Trends," *Journal of Epidemiology and Global Health*, vol. 10, no. 1, pp. 107-111, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[4] Shi-Chue Hsing et al., "The Severity of Diabetic Retinopathy Is an Independent Factor for the Progression of Diabetic Nephropathy," *Journal of Clinical Medicine*, vol. 10, no. 1, pp. 1-13, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Hamid Safi et al., "Early Detection of Diabetic Retinopathy," *Survey of Ophthalmology*, vol. 63, no. 5, pp. 601-608, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[6] Yijin Huang et al., "Identifying the Key Components in ResNet-50 for Diabetic Retinopathy Grading from Fundus Images: A Systematic Investigation," *arXiv Preprint*, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Mohammad Z. Atwany, Abdulwahab H. Sahyoun, and Mohammad Yaqub, "Deep Learning Techniques for Diabetic Retinopathy Classification: A Survey," *IEEE Access*, vol. 10, pp. 28642-28655, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[8] Imran Qureshi, Jun Ma, and Qaisar Abbas, "Diabetic Retinopathy Detection and Stage Classification in Eye Fundus Images using Active Deep Learning," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 11691-11721, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[9] Sraddha Das et al., "Deep Learning Architecture based on Segmented Fundus Image Features for Classification of Diabetic Retinopathy," *Biomedical Signal Processing and Control*, vol. 68, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Brahami Menaouer et al., "Diabetic Retinopathy Classification using Hybrid Deep Learning Approach," *SN Computer Science*, vol. 3, no. 5, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Ling Dai et al., "A Deep Learning System for Detecting Diabetic Retinopathy Across the Disease Spectrum," *Nature Communications*, vol. 12, no. 1, pp. 1-11, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[12] Wejdan L. Alyoubi, Maysoon F. Abulkhair, and Wafaa M. Shalash, "Diabetic Retinopathy Fundus Image Classification and Lesions Localization System Using Deep Learning," *Sensors*, vol. 21, no. 11, pp. 1-22, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] Laxmi Math, and Ruksar Fatima, "Adaptive Machine Learning Classification for Diabetic Retinopathy," *Multimedia Tools and Applications*, vol. 80, no. 4, pp. 5173-5186, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] S.G. Kobat et al., "Automated Diabetic Retinopathy Detection Using Horizontal and Vertical Patch Division-Based Pre-Trained DenseNET with Digital Fundus Images," *Diagnostics*, vol. 12, no. 8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[15] Borys Tymchenko, Philip Marchenko, and Dmitry Spodarets, "Deep Learning Approach to Diabetic Retinopathy Detection," *arXiv Preprint*, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] Niloy Sikder et al., "Severity Classification of Diabetic Retinopathy using an Ensemble Learning Algorithm through Analyzing Retinal Images," *Symmetry*, vol. 13, no. 4, pp. 1-26, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[17] Tahira Nazir et al., "Detection of Diabetic Eye Disease from Retinal Images Using a Deep Learning Based CenterNet Model," *Sensors*, vol. 21, no. 16, pp. 1-18, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Mohamed Chetoui, and Moulay A. Akhloufi, "Explainable Diabetic Retinopathy using EfficientNET," *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, Montreal, QC, Canada, pp. 1966-1969, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Gazala Mushtaq, and Farheen Siddiqui, "Detection of Diabetic Retinopathy using Deep Learning Methodology," *IOP Conference Series: Materials Science and Engineering*, Tamil Nadu, India, vol. 1070, no. 1, pp. 1-13, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] San-Li Yi et al., "Diabetic Retinopathy Diagnosis Based on RA-EfficientNet," *Applied Science*, vol. 11, no. 22, pp. 1-14, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[21] Sandeep Madarapu, Samit Ari, and Kamalakanta Mahapatra, "A Multi-Resolution Convolutional Attention Network for Efficient Diabetic Retinopathy Classification," *Computers and Electrical Engineering*, vol. 117, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Samia Akhtar et al., "A Classification Framework for Diabetic Retinopathy Detection Using Transfer Learning," *2024 2nd International Conference on Cyber Resilience (ICCR)*, Dubai, United Arab Emirates, pp. 1-5, 2024. [CrossRef] [Publisher Link]

[23] Diabetic Retinopathy Detection, 2025. [Online]. Available: https://www.kaggle.com/competitions/diabetic-retinopathy-detection

[24] Yuki Yoshimi et al., "Image Preprocessing with Contrast-Limited Adaptive Histogram Equalization Improves the Segmentation Performance of Deep Learning for the Articular Disk of the Temporomandibular Joint on Magnetic Resonance Images," *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, vol. 138, no. 1, pp. 128-141, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[25] Brett Koonce, *Convolutional Neural Networks with Swift for Tensorflow*, 1st ed., Apress Berkeley, CA, pp. 63-72, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[26] Marwa M. Emam et al., "Breast Cancer Diagnosis using Optimized Deep Convolutional Neural Network based on Transfer Learning Technique and Improved Coati Optimization Algorithm," *Expert Systems with Applications*, vol. 255, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[27] Yuhan Bai, "RELU-Function and Derived Function Review," *2022 International Conference on Science and Technology Ethics and Human Future (STEHF 2022)*, vol. 144, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Meetu Jain et al., "An Overview of Variants and Advancements of PSO Algorithm," *Applied Science*, vol. 12, no. 17, pp. 1-21, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[29] Juan Terven et al., "A Comprehensive Survey of Loss Functions and Metrics in Deep Learning," *Artificial Intelligence Review*, vol. 58, no. 7, pp. 1-172, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[30] Kele Xu, Dawei Feng, and Haibo Mi, "Deep Convolutional Neural Network-Based Early Automated Detection of Diabetic Retinopathy Using Fundus Image," *Molecules*, vol. 22, no. 12, pp. 1-7, 2017 [CrossRef] [Google Scholar] [Publisher Link]

[31] Shu-I Pao et al., "Detection of Diabetic Retinopathy using Bichannel Convolutional Neural Network," *Journal of Ophthalmology*, vol. 2020, no. 1, pp. 1-7, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[32] Borys Tymchenko, Philip Marchenko, and Dmitry Spodarets, "Deep Learning Approach to Diabetic Retinopathy Detection," *arXiv Preprint*, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[33] Along He et al., "CABNet: Category Attention Block for Imbalanced Diabetic Retinopathy Grading," *IEEE Transactions on Medical Imaging*, vol. 40, no. 1, pp. 143-153, 2021. [CrossRef] [Google Scholar] [Publisher Link]