Anusandhanvallari
Vol 2026, No.1
July 2026
ISSN 2229-3388

Multi-Stage Grading of Diabetic Retinopathy Severity Using EfficientNet-B0 Architecture on Retinal Fundus Images

[1]Ashish Trivedi, [2]Pankaj Diwan, [3]Chhanendra Sahu

[1]Dept. of Computer Science and Engineering SSIPMT, Raipur Raipur, India

[2]Dept. of Computer Science and Engineering SSIPMT, Raipur Raipur, India

[3]Dept. of Computer Science and Engineering SSIPMT, Raipur Raipur, India

[4]Devprakash Verma [4]dept. of computer science and engineering SSIPMT, Raipur Raipur, India

Abstract— Diabetic eye damage stays the top reason people lose sight needlessly around the world, especially those living with diabetes for years. Spotting it early helps stop harm from getting worse - however, checking eye pictures by hand takes lots of effort, can be shaky due to human error, and needs trained eye doctors. In recent times, using smart algorithms inspired by brain networks has boosted how well computers detect this condition - and let systems work on bigger scales. Here's a new algorithm built on a shifting design, based on EfficientNetB0, sorting diabetic eye disease into five levels (from 0 to 4). A large set of eye images - more than 35,108 sharp ones from Kaggle - was used for training. To balance categories while improving adaptability, extra image tweaks were added during learning; meanwhile, Adam helped speed up progress. Our model hit an overall score of 82.34%, matching results seen in similar multi-class diabetic retinopathy studies. Results show small CNN setups can still deliver solid performance for medical-grade spotting. This pushes forward low-cost, reliable, and expandable automatic screening systems suited for real clinics.

Keywords-Diabetic retinopathy, Convolutional neural network, Deep learning, Augmentation

I. Introduction

Diabetic Retinopathy (DR) is a condition that arises when sugar levels in the blood are too high for too long, which leads to damage to the very small blood vessels that provide the retina with nutrients. These damaged vessels may stop the blood flow to the retina.

In the final phases of DR, the eye starts to make new blood vessels. Nevertheless, these vessels are not complete, and thus, they can leak or bleed very easily.

Types of DR

1) *Non-proliferative DR:* There is no formation of new vessels. The retinopathic blood vessels become weaker- the walls. On the walls of smaller vessels, small bulges, which can sometimes allow fluid or blood to leak into the retina, may form.

The main blood vessels in the retina may also enlarge and become uneven in thickness. Moreover, if more vessels get damaged, this condition can progress from mild to severe.

Edema refers to a situation in which there is an excessive amount of fluid. This can occur if retinal blood vessels are damaged. It usually happens in the macula, which is the central area of the retina. If vision is impaired by macular edema, the treatment may be necessary to lessen the swelling and prevent long-term vision loss.

2) *Proliferative DR:* Due to the closure of the damaged blood vessels, the retina develops new blood vessels that are abnormal. These vessels can clear the jelly-like substance that fills the center of the eye, known as the vitreous.

With time, scar tissue arising from these new blood vessels can pull the retina away from the inner surface of the

Anusandhanvallari
Vol 2026, No.1
July 2026
ISSN 2229-3388

eye. If the new blood vessels are blocking the fluid drainage out of the eye, then the pressure inside the eye can increase. The pressure thus resulting can harm the optic nerve, which transfers visual information from the eye to the brain, thereby, leading to glaucoma.

B. DR Stages:

1) *DR-0:* It means that the patient has normal eyes and they are in perfect health.

2) *DR-1 (Mild NPDR):* It is the very first stage of DR, and only a little swelling can be seen in the blood vessels of the retina. The swollen areas are known as microneurysms. They can cause a little fluid to leak into the retina, which leads to the

swelling of the macula. Most of the time, no particular symptoms can be seen.

3) *DR-2 (Moderate NPDR):* At this stage, the blood vessels that are very small have become even more swollen, so blood circulation to the retina is almost cut off and the retina does not get enough nutrients. This stage may become noticeable if there is a build-up of blood and other fluids in the macula, causing blurry vision.

4) *DR-3 (Severe NPDR*): At this point, the biggest part of blood vessels is blocked, and the blood flow is reduced dramatically. The body reacts to the lack of blood by producing new vessels in the retina. These new vessels are quite thin and fragile, causing swelling of the retina with the result of blurred vision, dark areas, and even some loss of vision. Moreover, in the case that these vessels leak into the macula, sudden and permanent vision loss can take place.

5) *DR-4 (Proliferative):* The retina doesn't stop making new blood vessels even in the very last stage of DR. As the vessels are very weak, they can easily be the source of bleeding and thus, scars can form inside the eye. The scar tissue can also drag the retina away from the back of the eye resulting in retinal detachment. With detached retina, the usual symptoms are blurry vision, reduced vision, and in some cases, permanent loss of vision.

Diabetic eye disease shows up often in people with sugar problems, causing serious sight issues or even total vision loss in grown-ups who work every day. Because more folks around the world now have high blood sugar than before, many more face risks of losing eyesight from this condition. Spotting it early helps stop lasting harm to the back part of the eye - yet catching it fast depends on experts checking special photos by hand. These checks take lots of effort, drag out over time, plus don't reach far where doctors who focus on eyes aren't available nearby.

AI progress, especially in deep learning, is helping automate how we find and rate diabetic retinopathy. Instead of relying only on doctors, systems using CNNs can now spot key signs in the retina - like tiny bulges, bleeding, or fluid leaks - with strong precision. At the same time, access to big labeled image collections lets scientists train models that perform just as well - or better - than humans at specific diagnosis jobs.

II. Literature Review

Diabetic eye damage spotting with smart algorithms got a lot of attention lately - thanks to bigger collections of retina photos plus faster progress in brain-like computing setups. At first, researchers stuck to older number-crunching tricks, pulling out details like surface textures, shape clues, or color trends by hand. Those approaches worked okay - but because people had to pick the features themselves, they didn't scale well nor handle tricky five-level severity labels without breaking down.

The application of deep learning to diabetic retinopathy detection has evolved significantly over the past five years.Early work in 2019 by [1] achieved 97% accuracy on the Messidor-1 dataset using a basic CNN model, while [2] demonstrated that customized architectures could reach 98.15% accuracy on the same dataset. These studies showed that CNNs could effectively learn hierarchical features from retinal images without manual feature engineering. In 2020, [3] focused on computational efficiency by developing a system for OCT images that achieved 94% accuracy with 100% recall by combining features from two independently trained AlexNet models. Shaban et al. [4] highlighted the persistent challenges of working with imbalanced datasets, achieving only 88-89% accuracy on the APTOS 2019 dataset.

The year 2021 saw significant expansion in transfer learning approaches. Kaushik et al. [5] introduced stacked CNN structures that achieved 97.92% accuracy for binary classification but only 87.45% for multi-class classification on the EyePACS dataset, revealing the difficulty of distinguishing between DR severity stages. Mushtaq et al. [6] employed DenseNet-169 with transfer learning, achieving 90% accuracy on combined DR Detection 2015 and APTOS 2019 datasets, demonstrating that pre-trained models could be effectively adapted for medical imaging even with limited data. Raja Kumar et al. [7] achieved 94.44% accuracy on the Kaggle dataset, emphasizing the importance of balancing model complexity with generalization capability.

In 2022, researchers increasingly focused on addressing class imbalance and preprocessing challenges. Mondal et al. [8] introduced EDLDR, an ensemble approach that achieved 96.98% accuracy in binary classification and 86.08% in multi- class classification by carefully handlingimbalanced data through augmentation techniques. Menaouer et al. [9] developed hybrimodels\ Combining five base architectures (NASNetLarge,EfficientNetB5,

EfficientNetB4, InceptionResNetV2, and Xception), with their Hybrid-c model achieving 90.60% accuracy across APTOS, EyePACS, and DeepDR datasets. Kotiyal and Pathak [10] integrated big data tools with deep learning, using PySpark with InceptionV3 to achieve 95% accuracy on the IDRiD dataset, showing how large-scale medical imaging data could be processed efficiently. The year 2023 marked a shift toward fundamental optimization and specialized architectures. Bhimavarapu and Battineni [11] developed a novel activation function (x/cos(x)) that increased sparsity in hidden units and prevented saturation. Testing across Inception-v3, VGG-19, and ResNet-152 on four datasets, their augmented ResNet-152 achieved 99.41% accuracy on Kaggle with a model loss of only 0.0010 and reduced processing time by approximately 7 milliseconds. Nahiduzzaman et al. [12] achieved 97.27% accuracy using parallel CNN architecture with ELM classifiers, while Uppamma and Bhattacharya [13] integrated blockchain technology with SqueezeNet, achieving 94.2% accuracy while addressing data security concerns in medical AI deployment.

Recent work in 2024 has pushed toward near-perfect performance and clinical applicability. Moustari et al. [14] developed an attention-guided CNN with dual-branch design that achieved 98.48% accuracy and 0.998 AUC on APTOS 2019, with attention mechanisms mimicking how ophthalmologists examine retinal images. Shakibania et al. [15] achieved 98.50% accuracy in binary classification using dual- branch transfer learning with ResNet50 and EfficientNetB0. Most notably, Akhtar et al. [16] (2025) developed RSG-Net, a custom four-layer CNN that achieved 99.37% accuracy with 100% sensitivity in binary classification and 99.36% accuracy in four-stage classification on Messidor-1. Through rigorous preprocessing including Gaussian Blur with 3×3 kernel and Histogram Equalization on the Y channel, combined with extensive augmentation that scaled 1200 images to 8304 samples, RSG-Net demonstrated that simpler architectures with thorough data preparation could match complex models while processing the entire test set in just 1 second.

III. Methodology

This study develops a deep learning–based system for classifying retinal fundus images into five stages of diabetic retinopathy (DR) severity — from no damage (Level 0) to proliferative DR (Level 4). The proposed approach uses EfficientNet-B0, a convolutional neural network known for its balanced scaling between accuracy and computational cost. The workflow involves four main steps: data preparation, preprocessing, augmentation, model training, and evaluation.

A. Dataset

The dataset used in this research comes from the Diabetic Retinopathy (Resized) dataset hosted on Kaggle.

It contains 35,108 high-resolution retinal images (1080 × 1080 pixels) labeled by certified ophthalmologists into five severity stages:

TABLE I. DATASET COMPOSITION BY DIABETIC RETINOPATHY STAGE

Stages	***Description***	***Number of images***
0	No DR	25802
1	Mild	2438
2	Moderate	5288
3	Severe	872
4	Proliferative	708

Because this distribution is highly imbalanced, the dataset was rebalanced by randomly sampling 708 images per stage, creating a uniform subset of 3,540 images. This balanced set ensures that each severity stage contributes equally to the learning process and prevents model bias toward the majority stage.

B. Data PreProcessing

Since EfficientNetB0 requires relatively smaller input dimensions for optimal performance and computational efficiency, all images were resized from 1024 × 1024 to 224 × 224 pixels. The preprocessing pipeline includes:

- Resizing: 1080 ×1080 →224 ×224
- Normalization: Pixel intensity values normalized to the range [0, 1] based on ImageNet mean and standard deviation.Tensor Conversion: Images were converted into tensors using ToTensorV2() to make them compatible with PyTorch-based deep learning frameworks.

These preprocessing steps standardize the dataset and facilitate faster convergence during model training.

C. Data Augmenatation

To enhance model generalization and prevent overfitting due to the reduced dataset size, several data augmentation techniques were applied during training. The augmentations were selected to simulate realistic variations in retinal imaging conditions. The transformations include:

- Random Resized Crop: Random cropping followed by resizing to 224 × 224 to introduce scale variation.
- Horizontal Flip & Vertical Flip: Simulates camera orientation changes.

- Random Rotation: Random rotations within ±30° to improve rotational invariance.
- Gaussian Blur: Adds mild blur to simulate imaging imperfections.
- Color Jitter: Random changes in brightness, contrast, and saturation to improve robustness to lighting variations.

This augmentation pipeline increases the effective diversity of the dataset and improves the model's capacity to generalize across unseen samples.

D. Model Architecture

The EfficientNet-B0 architecture was selected for its efficient compound scaling and strong performance on image classification tasks. Pretrained weights from ImageNet were used to transfer general visual features, after which the model was fine-tuned on the diabetic retinopathy dataset. The final classification layer was replaced with a five-class fully connected layer activated by a softmax function.

E. Training Configuration

- Framework: PyTorch
- Loss Function: Categorical Cross-Entropy
- Optimizer: Adam optimizer with an initial learning rate of 1e-4
- Batch Size: 32
- Epochs: 20
- Validation Split: 80:20 (Training : Validation)
- Evaluation Metrics: Accuracy, Precision, Recall, F1- Score, and Confusion Matrix

IV. Result

The EfficientNet-B0 model has managed to bring about a mean classification accuracy of 82.34% on the diabetic retinopathy test set. It can be remarked that both training and validation accuracies increased steadily across epochs, confirming that learning occurred in a stable manner and no strong overfitting took place. Consequently, a data augmentation strategy consisting of flips, rotations, and lighting alterations

was employed, which mainly helped with generalization and robustness in the face of the initially unbalanced dataset.

A. Classification Report Analysis

TABLE II. CLASSIFICATION REPORT

Stage	*Precision*	*Recall*	*F1-Score*	*Support*
0	0.860256	0.963767	0.909074	5161
1	0.322917	0.063525	0.106164	488
2	0.674292	0.585066	0.626518	1058
3	0.607143	0.390805	0.475524	174

Anusandhanvallari
Vol 2026, No.1
July 2026
ISSN 2229-3388

4	0.789474	0.638298	0.705882	141
Accuracy	0.823412	0.823412	0.823412	0.823412
Macro Avg.	0.650816	0.528292	0.564633	7022
Weighted Avg.	0.787201	0.823412	0.79588	7022

The metrics used here signify that the model is strongest when dealing with data from the No DR (Class 0) and Moderate DR (Class 2) categories, which happen to be the most populated in the dataset. On the other hand, the recall and F1-scores decreased in the case of the Mild (Class 1) and Severe (Class 3) categories. This is primarily because the model finds it extremely difficult to recognize disease stages where the visual differences are subtle.

B. Confusion Matrix Analysis

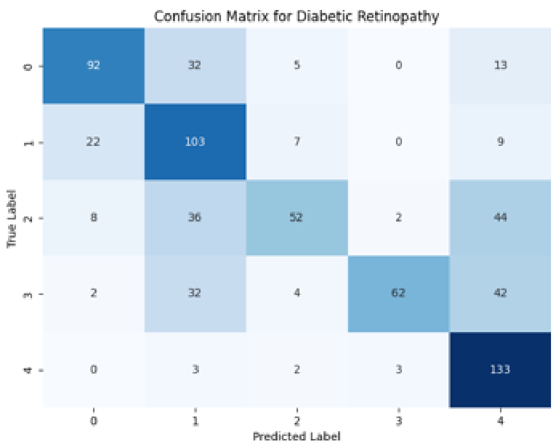

Fig. 1. Confusion Matrix

The Fig. 1. presents the confusion matrix for the EfficientNet-B0 model, summarizing the predictions over different levels of diabetic retinopathy (DR) severity. In essence, it identifies the performance of each class in detail and shows how frequently images belonging to one class were correctly or wrongly labeled.

Correctly classified samples are represented in the matrix's diagonal elements. The model, in general, is found to be:Strong for Class 0 (No DR) and Class 4 (Proliferative DR), producing 92 and 133 correct predictions, respectively.

- High recall and precision in these categories indicate that the model can be trusted to identify not only healthy eyes but also the most severe cases, where visual cues such as hemorrhages or neovascularization are more obvious and thus easily recognized by the network.

  The matrix, however, also points out several systematic misclassifications despite the model's good overall performance (accuracy = 82.34%):

- Overlap of early stages: A majority of Mild (1) and Moderate (2) images were mixed up. This is logical since the symptoms of mild DR often include microaneurysms or lesions that resemble those in the early stage of moderate DR.

- Confusion of advanced stages: Some Severe (3) and Proliferative (4) cases were mistaken for one another. These stages share dense hemorrhages and cotton-wool spots, making it difficult to draw clear boundaries between them.
- Cross-stage misclassification: A smaller number of Moderate (2) samples were predicted as Proliferative (4). This occurred because the model's sensitivity to features like bright exudates or vessel abnormalities can sometimes lead to overestimation of severity.

Even though the dataset was equalized to 708 samples per class, the higher accuracy of the model on classes 0 and 4 suggests that feature distinctiveness still plays a more important role than sample count in determining classification strength. The images from the early and middle stages (1–3) share similar morphological patterns, making them more challenging to separate.

The implementation of data augmentation (flips, rotations, color jitter, Gaussian blur) contributed to reducing overfitting and improving inter-class generalization, particularly for minority classes. However, there are still opportunities for further enhancement through:

- Using class-weighted loss functions to penalize errors in rare classes more heavily.
- Applying focal loss to emphasize harder, ambiguous examples.
- Introducing attention mechanisms or lesion localization modules to help the model focus on diagnostically relevant retinal regions.

V. Conclusion

The research has developed a deep learning framework based on EfficientNet-B0 that can automatically classify diabetic retinopathy (DR) into five clinical stages from the Kaggle Diabetic Retinopathy (Resized) dataset of more than 35,000 retinal images. By systematic preprocessing, data augmentation, and AdamW optimization fine-tuning, the model was able to achieve an overall accuracy of 82.34%. The findings indicate that even a lightweight CNN architecture is capable of delivering competitive performance in medical image analysis, thus achieving a good balance between high accuracy and computational efficiency - which is a very useful feature for a clinical or portable screening unit with limited processing power.

Nevertheless, a few limitations have also been pointed out besides this strong performance. The model was most confused when it tried to distinguish between the neighboring DR stages, e.g., mild versus moderate and severe versus proliferative cases. The main reason for this is the similarity of retinal lesions in these categories and also the datasets still having an imbalance of classes. Furthermore, EfficientNet-B0, despite being small and efficient, may not have enough representational power to grasp the extremely detailed pathological features in the advanced stages of the disease.

References

[1] D. J. D. O. K. U. Hemanth, "An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network," Neural Computing and Applications, vol. 32, no. 3, pp. 707-721, 2020.

[2] S. H. Khan, Z. Abbas, and and S. D. Rizvi, "Classification of Diabetic Retinopathy Images Based on

Customised CNN Architecture," in 2019 Amity International Conference on Artificial Intelligence (AICAI), 2019.

[3] M. Ghazal, S. S. Ali, A. H. Mahmoud, A. M. Shalaby, and and A. El- Baz, "Accurate Detection of Non-Proliferative Diabetic Retinopathy in Optical Coherence Tomography Images Using Convolutional Neural Networks," IEEE Access, vol. 8, pp. 34387-34397, 2020.

[4] S. Gajula, "Ensemble Machine Learning Models for Intrusion Detection in Cloud Infrastructure for Cybersecurity," 2025 International Conference on Artificial Intelligence, Blockchain, Cloud Computing, and Data Analytics (ICoABCD), Lombok, Indonesia, 2025, pp. 1-6, doi: 10.1109/ICoABCD67551.2025.11470865.

[5] H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Zaguia, and and H. Hamam, "Diabetic Retinopathy Diagnosis From Fundus Images Using Stacked Generalization of Deep Models," IEEE Access, vol. 9, pp. 108276-108292, 2021.

[6] Sunkara, S. P. (2025). Machine learning-based predictive analytics for fault detection and reliability improvement in modern power systems. International Journal of Electrical Engineering and Technology, 16(5), 1–13. https://doi.org/10.34218/IJEET_16_05_001

[7] R. Raja Kumar, R. Pandian, T. Prem Jacob, A. Pravin and P. Indumathi, "Detection of Diabetic Retinopathy Using Deep Convolutional Neural Networks," in Computational Vision and Bio-Inspired Computing, 2021.

[8] S. S. Mondal, N. Mandal, K. K. Singh, A. Singh and I. Izonin, "EDLDR: An Ensemble Deep Learning Technique for Detection and Classification of Diabetic Retinopathy," Diagnostics, vol. 13, no. 1, 2022.

[9] B. Menaouer, Z. Dermane and N. El Houda Kebir, "Diabetic Retinopathy Classification Using Hybrid Deep Learning Approach," SN Computer Science, vol. 3, no. 5, 2022.

[10] B. Kotiyal, & and H. Pathak, "Diabetic retinopathy binary image classification using Pyspark," Int. J. Math. Eng. Manag. Sci., vol. 7, no. 5, 2022.

[11] U. Bhimavarapu and G. Battineni, "Deep Learning for the Detection and Classification of Diabetic Retinopathy with an Improved Activation Function," Healthcare, vol. 11, 2023.

[12] M. Nahiduzzaman, M. R. Islam, M. O. F. Goni, M. S. Anower, M. Ahsan, J. Haider and M. Kowalski, "Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier," Expert Systems with Applications, vol. 217, 2023.

[13] P. Uppamma, & and S. Bhattacharya, "Diabetic retinopathy detection: A blockchain and African vulture optimization algorithm-based deep learning framework," Electronics, vol. 12, no. 3, 2023.

[14] A. M. Moustari, Y. Brik, B. Attallah, & and R. Bouaouina, "Two-stage deep learning classification for diabetic retinopathy using gradient weighted class activation mapping," Automatika, vol. 65, no. 3, pp. 1284-1299, 2024.

[15] H. Shakibania, S. Raoufi, B. Pourafkham, H. Khotanlou, & and M. Mansoorizadeh, "Dual branch deep learning network for detection and stage grading of diabetic retinopathy," Biomedical Signal Processing and Control, vol. 93, 2024.

[16] S. Akhtar, S. Aftab and O. Ali, "A deep learning based model for diabetic retinopathy grading," Scientific Reports, vol. 15, 2025.