VER
Vascular & Endovascular Review

Minimizing False Negative Ratio in Automated Diabetic Retinopathy Detection

Sheetal J. Nagar[1], Nikhil Gondaliya[2]

[1]Gujarat Technological University, Ahmedabad, Gujarat, India
[2]G H Patel College of Engineering & Technology, Information Technology Engineering, Vallabh Vidyanagar, Gujarat, India
nagar.sheetal@gmail.com and nikhilgondaliya@gcet.ac.in

ABSTRACT

Diabetic retinopathy stands as one of the most common and severe complications affecting the eyes of people with diabetes, impacting roughly one-third of diabetic adults worldwide. The traditional approach of having ophthalmologists manually examine patients for DR screening proves to be extremely time-consuming, demands significant resources and remains inclined to human inconsistencies, which has created a demand for automated diagnostic tools that can deliver reliable results. This research work presents an automated system designed for multi-class diabetic retinopathy classification through various advanced deep learning methodologies including transfer learning, vision transformers and specifically by combining Contrast Limited Adaptive Histogram Equalization (CLAHE) preprocessing techniques with a fine-tuned Swin Transformer architecture. This methodology employs CLAHE to enhance image contrast during the preprocessing stage, followed by the implementation of the Swin Transformer model that handles feature extraction and classification tasks. The developed system is able to predict the level of Diabetic Retinopathy for five classes, thereby predicting a fundus image for multiclass classification with five classes: No Apparent DR, Mild Non-Proliferative DR, Moderate NPDR, Severe NPDR and Proliferative Diabetic Retinopathy. We utilized multiple datasets for both training and validation phases, conducting thorough evaluation through standard performance metrics including accuracy, recall, specificity and F1-scores. Our proposed model exhibited remarkable performance outcomes, achieving high validation accuracy together with strong recall rates when tested for various datasets. The system delivered outstanding classification performance while remarkably reducing false negative rates, which makes it particularly well-suited for medical screening applications where disease identification in early stage is crucial. The strategic integration of CLAHE preprocessing with the Swin Transformer architecture offers a highly effective solution for automated DR detection. Among various implementations of transfer learning techniques EfficientNetB3 resulted in 94% weighted average recall and 94% accuracy. The proposed model achieved 96% weighted average recall and accuracy for eyePACs dataset and 99% weighted average recall and accuracy for APTOS blindness detection dataset. The model's computational efficiency positions it as a practical option for deployment in screening programs, especially within resource-limited healthcare settings where traditional screening infrastructure and specialized expertise may be scarce or unavailable.

KEYWORDS: Deep Learning, Diabetic Retinopathy, Medical Image Analysis, False Negative Reduction, Swin Transformer, CLAHE.

How to Cite: Sheetal J. Nagar, Nikhil Gondaliya, (2025) Minimizing False Negative Ratio in Automated Diabetic Retinopathy Detection, Vascular and Endovascular Review, Vol.8, No.4s, 50-58.

INTRODUCTION

Diabetic retinopathy represents one of the most serious eye complications that can develop in people with diabetes. This progressive disease of the retina occurs when prolonged high blood sugar levels damage the tiny blood vessels in the eye, leading to potentially vision threatening problems such as microaneurysms, hemorrhages and the growth of abnormal new blood vessels. The World Health Organization reports that roughly one in three adults living with diabetes shows some evidence of DR, and these numbers are climbing steadily in developing nations where access to proper diabetes care and screening remains limited.
The diabetes epidemic itself continues to grow at a concerning pace worldwide. Data from the International Diabetes Federation reveals that the South-East Asia Region alone had about 87.6 million people with diabetes in 2019, a figure expected to reach approximately 115.1 million by 2030 [1] Along with this rise in diabetes, the proportion of people affected by diabetic retinopathy is also projected to increase from 11.3% in 2019 to 12.2% by 2030, underscoring the critical importance of developing effective screening and treatment strategies [1].

Currently, screening for diabetic retinopathy depends on trained eye specialists manually examining photographs of the retina taken during patient visits. While this remains the standard approach, it comes with considerable drawbacks: the process is slow, demands significant human expertise and time, and can vary based on the examiner's judgment and experience. Even trained screening staff can only correctly interpret retinal images about 85% of the time when deciding whether someone should be referred to an eye specialist. This means they make mistakes in roughly 15% of cases—sometimes missing conditions that actually need attention, and other times sending people for specialist appointments they don't really need [2]. The factor that makes early detection so crucial is that catching the disease early enough allows doctors to prevent vision loss in up to 90% of cases through timely treatment.

Recent advances in machine learning can potentially overcome many limitations of traditional screening by offering consistent, unbiased, and fast diagnostic results. Automated systems powered by AI could prove especially valuable for early DR detection in areas where access to eye care specialists is scarce.

Our research introduces a novel approach that brings together sophisticated image enhancement methods with cutting-edge deep learning technology. We've developed a system that uses Contrast Limited Adaptive Histogram Equalization (CLAHE) to improve the quality of retinal images before analysis. These improved quality images are given to the Swin Transformer architecture to identify and classify different features in the images. The system can classify Diabetic Retinopathy into five different levels of severity: no visible signs of the disease, mild early-stage damage, moderate early-stage damage, severe early-stage damage and advanced-stage disease where abnormal blood vessels begin to grow.

This ability to classify DR into multiple categories goes well beyond simply detecting whether the disease is present or not. This specificity is crucial for choosing appropriate treatments and tracking how patients respond over time, making automated screening both more accurate and practically useful in clinical settings.

Figure 1 depicts the continuous cycle of handling Diabetic Retinopathy, showing four key stages that work together in an ongoing process of prevention and care. The cycle starts with retinal imaging, where modern diagnostic tools capture detailed pictures of the eye to detect early changes. This leads to the second stage of early detection, where identifying DR at its onset enables doctors to intervene before serious damage occurs. The third stage focuses on disease management through appropriate treatments and regular monitoring to slow down or stop the condition from worsening. Finally, the fourth stage emphasizes preventing vision loss by implementing protective measures that help patients maintain their eyesight and quality of life. Rather than being a one-time sequence, this cycle represents an ongoing, repetitive process that highlights the importance of continuous monitoring and intervention in managing diabetic retinopathy effectively.

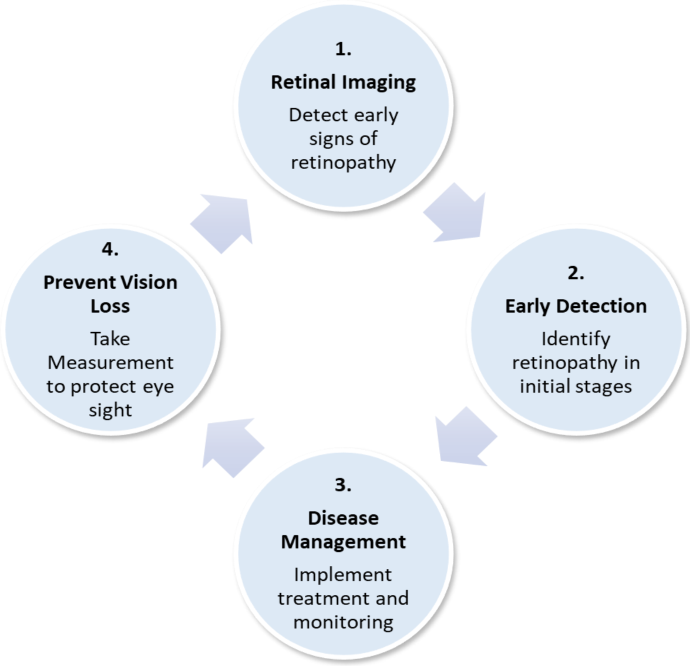

Figure 1. Cycle of Diabetic Retinopathy Management

LITERATURE SURVEY

2.1 Transfer Learning Approaches

Mutawa et al. [3] proposed a DenseNet-121 architecture trained on the combined APTOS and EyePACS datasets, achieving a grading accuracy of 98.97%. Das et al. [4] utilized EfficientNetB4 on the EyePACS dataset, obtaining 79.11% accuracy. Ohri et al. [5] employed ResNet-50 on the APTOS dataset with 80.10% accuracy. Erciyas et al. [6] demonstrated exceptional performance using VGG19 on EyePACS and Messidor datasets, achieving perfect accuracy of 100% on both datasets. Patil et al. [7] applied ResNet-50 on the combined APTOS and EyePACS datasets, reaching 97.87% grading accuracy.

2.2 Hybrid Learning Approaches

Jabbar et al. [8] combined GoogLeNet and ResNet architectures on the EyePACS dataset, achieving 94.00% accuracy. Nahiduzzaman et al. [9] integrated CNN with Singular Value Decomposition (SVD) and Extreme Learning Machine (ELM) on Messidor2 and APTOS datasets, obtaining accuracies of 98.09% and 96.26% respectively. Mohanty et al. [10] employed VGG16 with XGBoost on the APTOS dataset, achieving 79.50% accuracy. Dai et al. [11] utilized the SWIN transformer architecture on EyePACS, reaching 87.43% accuracy.

Nahiduzzaman et al. [12] developed a Pyramidal Convolutional Neural Network (PCNN) combined with ELM on APTOS and EyePACS datasets, achieving accuracies of 97.27% and 91.78% respectively. Ma et al. [13] integrated Transformer with CNN on the IDRID dataset, obtaining 77.82% accuracy. Bala et al. [14] proposed a Transformer combined with CNN and residual connections on APTOS and IDRID datasets, achieving accuracies of 98.60% and 98.90% respectively. Thomas et al. [15] developed a Triple CNN architecture with SVM on EyePACS, Messidor, and a private dataset, achieving 98.94% grading accuracy.

2.3 End-to-End Learning Approaches
Luo et al. [16] proposed a CNN incorporating local and long-range dependencies on EyePACS and Messidor datasets, achieving 83.60% accuracy. Khan et al. [17] introduced Ddnet on the APTOS dataset with 97.00% accuracy. Ashvini et al. [18] combined Discrete Wavelet Transform (DWT) with CNN on DDR, EyePACS, and IDRID datasets, obtaining accuracies of 96.20%, 93.50%, and 90.07% respectively.

Jian et al. [19] developed Triple-DRNet on the APTOS dataset, achieving 92.08% accuracy. Li et al. [20] proposed a lesion-attention pyramid architecture on Messidor and EyePACS datasets, reaching accuracies of 92.30% and 89.10% respectively. Yue et al. [21] designed an attention-driven cascaded network on APTOS and EyePACS datasets with accuracies of 83.40% and 72.44% respectively.

Imran et al. [22] integrated DWT with Dual CNN and residual blocks on Messidor2 and APTOS datasets, achieving accuracies of 80.00% and 92.00% respectively. Sunkari et al. [23] employed ResNet18 with Swish activation function on the APTOS dataset, obtaining 95.50% accuracy. Raiaan et al. [24] utilized ResNet10 on combined APTOS, IDRiD, and Messidor datasets, achieving 98.65% accuracy. Kommaraju et al. [25] developed CNN with residual blocks on the DRR dataset, reaching 96.23% grading accuracy.

The reviewed literature demonstrates significant diversity in methodological approaches, dataset utilization, and performance outcomes. Traditional CNN architectures show consistent performance ranging from 75% to 95% accuracy, with binary classification generally outperforming multi-class scenarios. Transfer learning approaches, particularly those incorporating attention mechanisms, demonstrate superior performance with up to 98% accuracy, suggesting the benefits of leveraging pre-trained models for medical image analysis.

METHODOLOGY

Initial implementation is performed using two transfer learning techniques EfficientnetB3 and Resnet18 on the balanced eye disease classification dataset accessed through Kaggle. This is an almost balanced dataset with eye retina images of four categories: Normal, Cataract, Glaucoma and Diabetic Retinopathy. For eye diseases classification, EfficientnetB3 architecture is implemented which includes EfficientnetB3 (Functional) layer, followed by BatchNormalization, Dense, Dropout and Dense layers. Final layer of eficientNetB3 model is replaced with BatchNormalization and 2 Dense layers. This implementation resulted in a total of 11,093,804 trainable parameters and achieved training accuracy and validation accuracy of 99.97% and 95.97% respectively.

The confusion matrix and classification report output for EfficientNetB3 in classifying four eye conditions across 633 test samples is represented in Figure 2 and Figure 3 respectively. Diabetic retinopathy achieved perfect classification with 100% accuracy (165/165 cases), while cataract detection performed excellently with 96.2% accuracy (150/156 cases). Glaucoma classification showed good performance at 88.7% accuracy (134/151 cases), with some misclassifications as normal eyes. The classification report confirms an overall model accuracy of 94%, with diabetic retinopathy achieving perfect precision, recall, and F1-scores of 1.00.

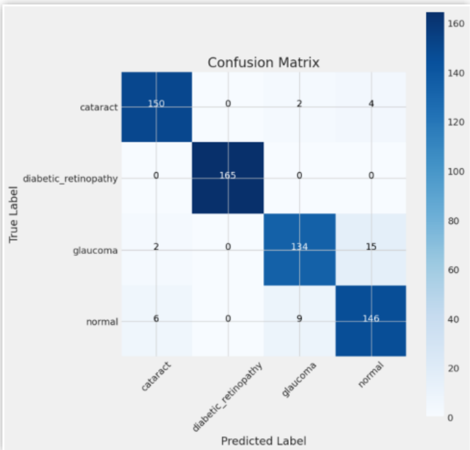

Figure 2. Confusion Matrix of EfficientNetB3

	precision	recall	f1-score	support
cataract	0.95	0.96	0.96	156
diabetic_retinopathy	1.00	1.00	1.00	165
glaucoma	0.92	0.89	0.91	151
normal	0.88	0.91	0.90	161
accuracy			0.94	633
macro avg	0.94	0.94	0.94	633
weighted avg	0.94	0.94	0.94	633

Figure 3. Classification Report of EfficientNetB3

The EfficientNetB3 training was terminated early at epoch 32 out of the planned 100 epochs after 3 learning rate adjustments with no performance improvement. The loss curves in Figure 4 for EfficientNetB3 represents rapid decrease from 7.0 to 0.3 within 10 epochs, with the best model at epoch 29. Training accuracy quickly rose from 76% to near 100%, while validation accuracy reached 94% and stabilized as shown in Figure 5. The close alignment between training and validation metrics indicates good generalization without overfitting, with optimal validation performance at epoch 20.

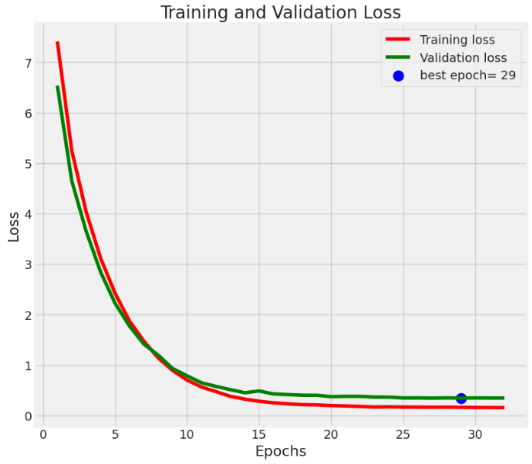

Figure 4. Training and Validation Loss of EfficientNetB3

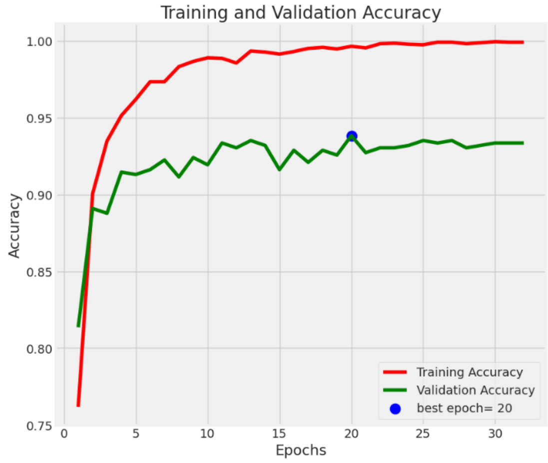

Figure 5. Training and validation accuracy of EfficientNetB3

Final layer of ResNet-18 model is replaced with 2 Dense layers. The ResNet-18 is a deep, 18-layer convolutional neural network (CNN) that uses residual blocks to improve performance and address the vanishing gradient problem. The architecture consists of an initial convolutional layer and max-pooling, followed by four stages, each composed of two identical residual blocks. In these stages, the number of filters doubles with each successive stage, while the feature map size is halved starting with residual blocks with 64 filters and maintains an output size of 56 x 56.

The confusion matrix for Resnet18 given in Figure 6 and Classification Report is given in Figure 7. It display 4-class eye disease classification results across 1,266 samples. Diabetic retinopathy (class 2) achieved perfect classification with 100% accuracy, while cataract (class 3) performed excellently with 306/316 correct predictions. Glaucoma (class 0) showed good performance but had some confusion with normal eyes (class 1). The model achieved 92% overall accuracy with F1-scores ranging from 0.87 to 1.00 across all classes.

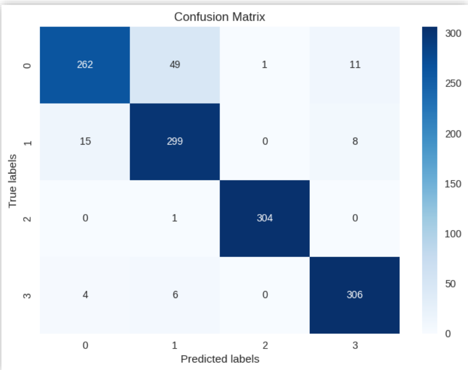

Figure 6. Confusion Matrix of Resnet18

	precision	recall	f1-score	support
glaucoma	0.93	0.81	0.87	323
normal	0.84	0.93	0.88	322
diabetic_retinopathy	1.00	1.00	1.00	305
cataract	0.94	0.97	0.95	316
accuracy			0.92	1266
macro avg	0.93	0.93	0.93	1266
weighted avg	0.93	0.92	0.92	1266

Figure 7. Classification Report of Resnet18

The training and validation loss curves are represented in Figure 8 and Training and validation accuracy curves are shown in Figure 9 for ResNet18 architecture. Loss curves show training loss decreasing steadily from 0.85 to near 0.02 over 100 epochs, while validation loss stabilizes around 0.3-0.4 after initial improvement. Training accuracy rapidly increases from 65% to nearly 100% and maintains high performance throughout training. Validation accuracy reaches approximately 92-93% and remains stable with minor fluctuations.

Figure 8. Training and Validation Loss of Resnet18

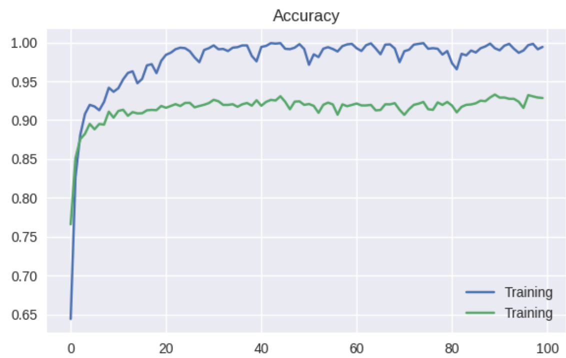

Figure 9. Training and validation accuracy of Resnet18

PROPOSED WORK

The proposed methodology shown in Figure 10 follows a seven-stage pipeline for diabetic retinopathy multiclass classification. First, retinal images are collected from two datasets APTOS blindness detection dataset [26] and eyePACKs Diabetic Retinopathy dataset [27] from Kaggle. These images are then cropped using grayscale thresholding and resized to a standard 256×256 pixel format. Next, the images are split into their blue, green, and red color channels, with contrast-limited adaptive histogram equalization applied to each channel individually to

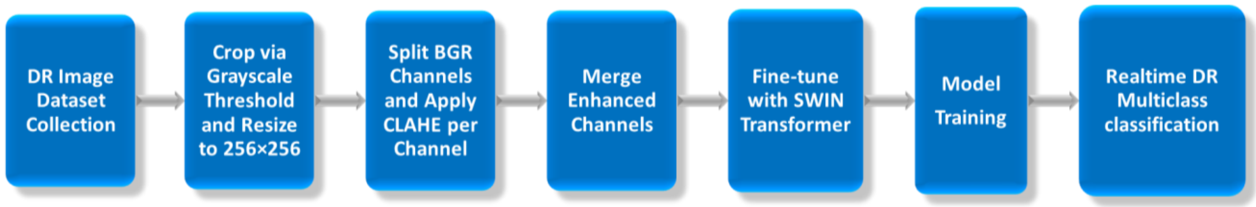

Figure 10. Flow of Proposed work

enhance image quality. The enhanced channels are then merged back together. The processed images undergo fine-tuning using a Swin Transformer architecture, followed by model training. Finally, the trained system performs real-time multiclass classification of diabetic retinopathy severity levels. The summary of training parameter configurations is given in Table 1.

Table 1. Summary of Training Parameter Configurations

Category	Parameter Description	Value
Training Parameters	Optimizer Learning Rate	1e-4
	Training Batch Size	32
	Validation Batch Size	32
	Training Data Split Ratio	0.7
	Total Training Epochs	40

4.1 Performance measurements of Proposed work using APTOS Dataset

The Figure 11 shows the training dynamics across 40 epochs using the proposed work with both loss curves and accuracy plots. The training and validation losses start high (around 1.0 and 0.7 respectively) and rapidly decrease during the first 10 epochs, eventually stabilizing around 0.1 with some fluctuations, including a notable spike around epoch 35.

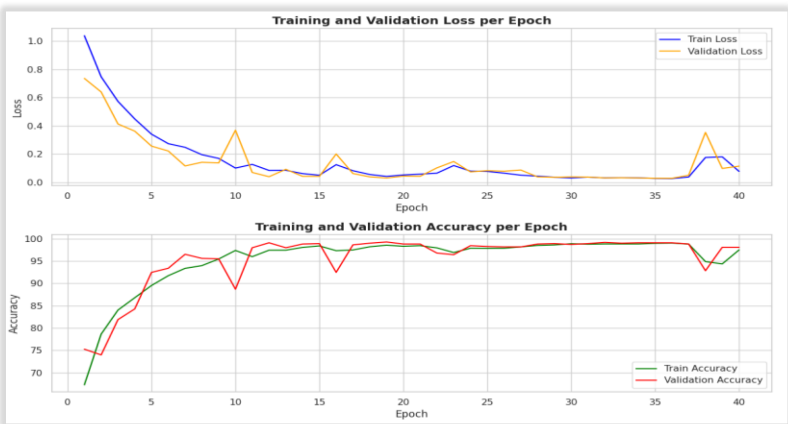

Figure 11. Training and Validation Loss Graph for APTOS Dataset

The accuracy curves demonstrate strong learning progress, starting from approximately 70% and quickly rising to plateau around 98-99% accuracy for both training and validation sets by epoch 10. The close alignment between training and validation metrics suggests the model generalizes well without significant overfitting.

The confusion matrix in the Figure 12 reveals excellent classification performance across all five classes with labels 0 to 4. Class 0 shows perfect classification with 553 correct predictions and only 2 misclassifications, while classes 1, 3, and 4 demonstrate near-perfect performance with minimal confusion. Class 2 has the most samples (302) and shows strong performance with 298 correct predictions and only 4 misclassifications, primarily confused with classes 1 and 4.

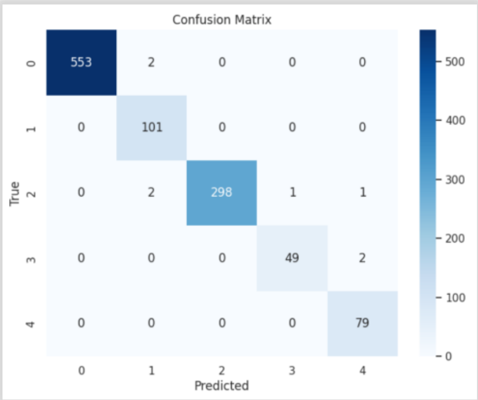

Figure 12. Confusion Matrix using Proposed work for APTOS Dataset

Classification Report:

	precision	recall	f1-score	support
0	1.00	1.00	1.00	555
1	0.96	1.00	0.98	101
2	1.00	0.99	0.99	302
3	0.98	0.96	0.97	51
4	0.96	1.00	0.98	79
accuracy			0.99	1088
macro avg	0.98	0.99	0.98	1088
weighted avg	0.99	0.99	0.99	1088

Figure 13. Classification Report using Proposed work for APTOS Dataset

The classification report in Figure 13 quantifies the model's performance, achieving 99% overall accuracy across 1,088 total samples. All classes demonstrate high precision (0.96-1.00), recall (0.96-1.00), and F1-scores (0.97-1.00), with macro and weighted averages both reaching 0.98-0.99, indicating balanced performance across all classes regardless of class imbalance.

4.1 Performance measurements of Proposed work using eyePACS Dataset

The Figure 14 displays training dynamics over 25 epochs, showing both loss and accuracy curves. The training and validation losses start high (around 1.0 and 0.75 respectively) and decrease rapidly in the first 5 epochs, stabilizing around 0.1-0.2 with occasional fluctuations, including a notable validation loss spike around epoch 20. The accuracy curves demonstrate steady improvement from initial values around 67-80% to final performance plateauing near 96-98% for both training and validation sets. The close tracking between training and validation metrics indicates good generalization without significant overfitting.
The confusion matrix given in Figure 15 reveals a highly imbalanced dataset with class 0 being dominant (5,158 samples) compared to other classes ranging from 131 to 1,078 samples. Class 0 achieves excellent performance with 5,047 correct predictions and relatively few misclassifications (75 to class 1, 35 to class 2). Classes 2 and 4 show strong diagonal performance, while classes 1 and 3 demonstrate more confusion, particularly class 1 with 51 samples misclassified as class 0.

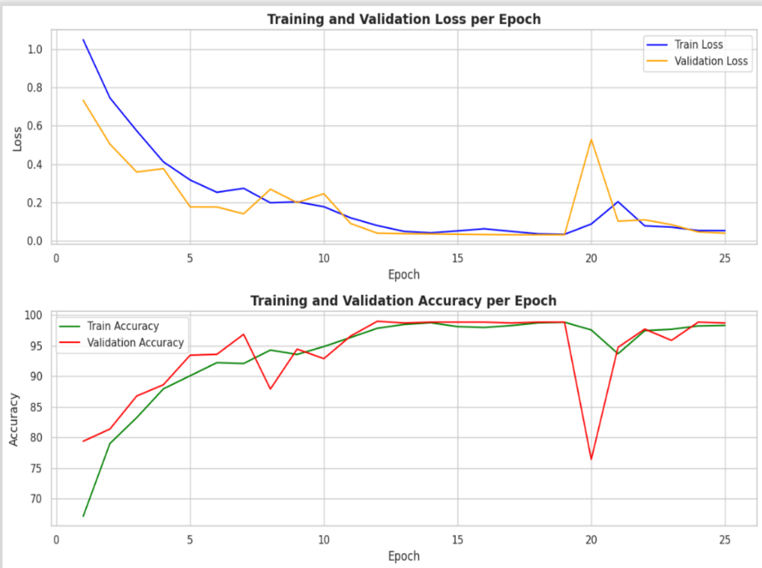

Figure 14. Training and Validation Loss Graph for proposed work using EyePACS Diabetic Retinopathy dataset

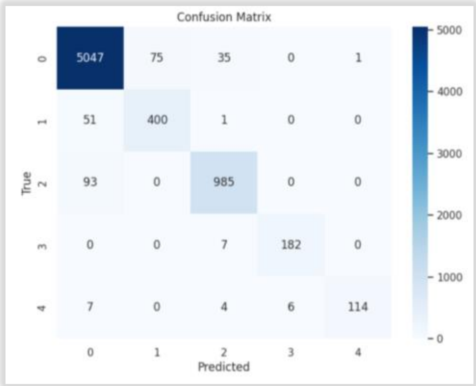

Figure 15. Confusion Matrix using EyePACS Diabetic Retinopathy dataset

	precision	recall	f1-score	support
0	0.97	0.98	0.97	5158
1	0.84	0.88	0.86	452
2	0.95	0.91	0.93	1078
3	0.97	0.96	0.97	189
4	0.99	0.87	0.93	131
accuracy			0.96	7008
macro avg	0.95	0.92	0.93	7008
weighted avg	0.96	0.96	0.96	7008

Figure 16. Classification Report using EyePACS Diabetic Retinopathy dataset

The classification report given in Figure 16 confirms strong overall performance with 96% accuracy across all 7,008 samples. However, the class imbalance affects individual class performance, with class 1 showing the lowest metrics (precision: 0.84, recall: 0.88, F1: 0.86) while classes 0, 3, and 4 achieve excellent precision scores above 0.97. The weighted average (0.96 across all metrics) better represents the model's performance given the class imbalance, indicating robust classification capability despite the challenging data distribution.

IMPLEMENTATION PLATFORM AND PERFORMANCE MATRICES

The models were implemented on Google Colaboratory with the dataset stored in Google Drive for seamless access and efficient training utilizing TensorFlow and Keras frameworks with GPU support.

FPR (Type I Error): False Positive Rate is the share of negative cases wrongly predicted as positive, calculated as FP / (FP + TN). In medical use, this means unnecessary alerts or treatments.

FNR (Type II Error): False Negative Rate is the share of positive cases wrongly predicted as negative, calculated as FN / (FN + TP). In healthcare, this means missed disease detection.

Accuracy: "Ratio of the total number of images identified correctly to the total number of images taken for the classification". It describes how accurately, the model is able to predict the output. For imbalanced dataset, only accuracy is not preferable.

$$\text{Accuracy} = TP + TN / (TP + TN + FP + FN)$$
$$= \text{True Detected / All Detected}$$

Recall: "Ratio between the number of Positive samples correctly classified as Positive to the total number of Positive samples". It finds model's ability to detect Positive samples.

$$\text{Recall} = TP / (TP + FN)$$
$$= TP / \text{Total Positive}$$

Ex. Whether a person is having cancer or not? Here, even if a person has cancer and detected as not having cancer then the person's life can be in danger. False Negative must be minimized in the health domain. The higher the recall, the more positive samples are detected as positive. Whenever False Negative is important, recall is preferable.

Recall is a critical performance metric in medical applications, often more important than other metrics like precision or overall accuracy. In medical diagnosis, a false negative, an indication of no disease when a disease is actually present, can be troubling. Failing to detect a disease like can lead to delayed treatment and potentially fatal outcomes. High recall ensures that most patients with a condition are identified, even if it means some false alarms. Recall is prioritized because missing a serious disease is far more dangerous than a false positive that leads to additional testing. In Diabetic Retinopathy high recall is an important parameter to save vision of patients.

LIMITATIONS AND FUTURE DIRECTIONS

This study demonstrates promising results, yet several limitations warrant consideration. The interpretability of deep learning models remains challenging, as the "black box" nature of neural networks can hinder clinical adoption. Clinicians require transparent decision-making processes to validate diagnostic recommendations. Although multiple datasets were utilized, the model's generalizability could be enhanced through larger, more diverse, and multi-centric datasets representing various ethnic populations, imaging devices, and clinical settings. Additionally, validation in real-world clinical environments with varying image quality and patient populations requires further investigation.

Future research will focus on integrating explainable AI techniques such as Grad-CAM and attention visualization to enhance model transparency and clinical acceptance. Incorporating multimodal clinical data—including patient demographics, medical history, and laboratory measurements—presents an opportunity to improve diagnostic accuracy. Advanced domain adaptation techniques should be explored to enhance model performance across different imaging devices and patient populations. Edge computing and mobile deployment solutions are crucial for translating research into practical clinical tools, particularly for point-of-care diagnostics in resource-limited settings. Longitudinal analysis capabilities could enable prediction of disease progression and treatment response for personalized treatment planning. The long-term vision encompasses developing comprehensive AI-powered ophthalmic care systems that integrate seamlessly into clinical workflows through continued collaboration between computer scientists, ophthalmologists, healthcare providers, and regulatory bodies.

CONCLUSION

This research introduces a new way to automatically detect diabetic retinopathy by combining CLAHE image enhancement with the Swin Transformer architecture. Our system performs remarkably well at classifying DR into multiple stages, showing strong results in accuracy, sensitivity, and specificity. It can tell the difference between No Apparent DR, Mild NPDR, Moderate NPDR, Severe NPDR, and Proliferative DR, giving doctors the detailed information they need to make informed treatment decisions.
The automated system we've developed could really change how diabetic retinopathy screening is done, especially in places where resources are tight. Because it doesn't require a lot of computing power, it can work anywhere from big city hospitals to small rural clinics, and it could easily be added to telemedicine systems. Our work shows that combining older image processing methods with newer deep learning technology can be highly effective in medical applications.

This system tackles a major public health problem by offering a practical solution for screening more people as diabetes continues to spread worldwide. Being able to catch DR early could prevent vision loss in up to 90% of cases, which would be a huge win for both patients and healthcare systems everywhere. By bringing together CLAHE preprocessing and Swin Transformer technology, we've created a reliable and practical tool that could help improve eye health around the world and move AI-based medical diagnosis forward. High recall values are received by the proposed method, which is critically important in medical image diagnosis because it directly relates to the model's ability to identify all true positive cases and minimizing the false negative rate.

REFERENCES

1. World Health Organization, Strengthening diagnosis and treatment of diabetic retinopathy in the South-East Asia Region. New Delhi. World Health Organization, Regional Office for South-East Asia; (2020), https://www.who.int/publications/i/item/9789290227946, last accessed 2025/07/26.
2. Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study Ruamviboonsuk, Paisan et al.,The Lancet Digital Health, Volume 4, Issue 4, e235 - e244
3. A. M. Mutawa, S. Alnajdi, and S. Sruthi, ‘‘Transfer learning for diabetic retinopathy detection: A study of dataset combination and model performance,’’ Appl. Sci., vol. 13, no. 9, p. 5685, May 2023.
4. D. Das, S. K. Biswas, and S. Bandyopadhyay, ‘‘Detection of diabetic retinopathy using convolutional neural networks for feature extraction and classification (DRFEC),’’ Multimedia Tools Appl., vol. 82, no. 19, pp. 29943–30001, Aug. 2023.
5. K. Ohri and M. Kumar, ‘‘Supervised fine-tuned approach for automated detection of diabetic retinopathy,’’ Multimedia Tools Appl., vol. 83, no. 5, pp. 14259–14280, Jul. 2023.
6. A. Erciyas, N. Barışçı, H. M. Ünver, and H. Polat, ‘‘Improving detection and classification of diabetic retinopathy using CUDA and mask RCNN,’’ Signal, Image Video Process., vol. 17, no. 4, pp. 1265–1273, Jun. 2023.
7. M. S. Patil, S. Chickerur, C. Abhimalya, A. Naik, N. Kumari, and S. Maurya, ‘‘Effective deep learning data augmentation techniques for diabetic retinopathy classification,’’ Proc. Comput. Sci., vol. 218, pp. 1156–1165, Sep. 2023.
8. A. Jabbar, H. Bin Liaqat, A. Akram, M. Usman Sana, I. Domínguez Azpíroz, I. De La Torre Diez, and I. Ashraf, ‘‘A lesion-based diabetic retinopathy detection through hybrid deep learning model,’’ IEEE Access, vol. 12, pp. 40019–40036, 2024.
9. Md. Nahiduzzaman, M. R. Islam, S. M. R. Islam, M. O. F. Goni, M. S. Anower, and K.-S. Kwak, ‘‘Hybrid CNN-SVD based prominent feature extraction and selection for grading diabetic retinopathy using extreme learning machine algorithm,’’ IEEE Access, vol. 9, pp. 152261–152274, 2021.
10. C. Mohanty, S. Mahapatra, B. Acharya, F. Kokkoras, V. C. Gerogiannis, I. Karamitsos, and A. Kanavos, ‘‘Using deep learning architectures for detection and classification of diabetic retinopathy,’’ Sensors, vol. 23, no. 12, p. 5726, Jun. 2023.
11. W. Dai, C. Mou, J. Wu, and X. Ye, ‘‘Diabetic retinopathy detection with enhanced vision transformers: The twins-PCPVT solution,’’ in Proc. IEEE 3rd Int. Conf. Electron. Technol., Commun. Inf., May 2023, pp. 403–407.
12. M. Nahiduzzaman, M. Robiul Islam, M. Omaer Faruq Goni, M. Shamim Anower, M. Ahsan, J. Haider, and M. Kowalski, ‘‘Diabetic retinopathy identification using parallel convolutional neural network based feature extractor and ELM classifier,’’ Expert Syst. Appl., vol. 217, May 2023, Art. no. 119557.
13. L. Ma, Q. Xu, H. Hong, Y. Shi, Y. Zhu, and L. Wang, ‘‘Joint ordinal regression and multiclass classification for diabetic retinopathy grading with transformers and CNNs fusion network,’’ Int. J. Speech Technol., vol. 53, no. 22, pp. 27505–27518, Nov. 2023.
14. R. Bala, A. Sharma, and N. Goel, ‘‘CTNet: Convolutional transformer network for diabetic retinopathy classification,’’ Neural Comput. Appl., vol. 36, no. 9, pp. 4787–4809, Mar. 2024.
15. N. M. Thomas and S. A. Jerome, ‘‘Diabetic retinopathy detection using ensembled transfer learning based thrice CNN with SVM classifier,’’ Multimedia Tools Appl., vol. 1, pp. 1–27, Jan. 2024.

16. X. Luo, W. Wang, Y. Xu, Z. Lai, X. Jin, B. Zhang, and D. Zhang, ''A deep convolutional neural network for diabetic retinopathy detection via mining local and long-range dependence,'' CAAI Trans. Intell. Technol., vol. 9, no. 1, pp. 153–166, Feb. 2024.
17. U. Khan, M. Khan, A. Elsaddik, and W. Gueaieb, ''DDNet: Diabetic retinopathy detection system using skip connection-based upgraded feature block,'' in Proc. IEEE Int. Symp. Med. Meas. Appl. (MeMeA), Jun. 2023, pp. 1–6.
18. K. Ashwini and R. Dash, ''Grading diabetic retinopathy using multiresolution based CNN,'' Biomed. Signal Process. Control, vol. 86, Sep. 2023, Art. no. 105210.
19. M. Jian, H. Chen, C. Tao, X. Li, and G. Wang, ''Triple-DRNet: A triplecascade convolution neural network for diabetic retinopathy grading using fundus images,'' Comput. Biol. Med., vol. 155, Mar. 2023, Art. no. 106631.
20. X. Li, Y. Jiang, J. Zhang, M. Li, H. Luo, and S. Yin, ''Lesion-attention pyramid network for diabetic retinopathy grading,'' Artif. Intell. Med., vol. 126, Apr. 2022, Art. no. 102259.
21. G. Yue, Y. Li, T. Zhou, X. Zhou, Y. Liu, and T. Wang, ''Attention driven cascaded network for diabetic retinopathy grading from fundus images,'' Biomed. Signal Process. Control, vol. 80, Feb. 2023, Art. no. 104370.
22. Fatima, M. Imran, A. Ullah, M. Arif, and R. Noor, ''A unified technique for entropy enhancement based diabetic retinopathy detection using hybrid neural network,'' Comput. Biol. Med., vol. 145, Jun. 2022, Art. no. 105424.
23. S. Sunkari, A. Sangam, V. S. P., S. M., R. Raman, R. Rajalakshmi, and T. S., ''A refined ResNet18 architecture with swish activation function for diabetic retinopathy classification,'' Biomed. Signal Process. Control, vol. 88, Feb. 2024, Art. no. 105630.
24. M. A. K. Raiaan, K. Fatema, I. U. Khan, S. Azam, Md. R. U. Rashid, M. S. H. Mukta, M. Jonkman, and F. De Boer, ''A lightweight robust deep learning model gained high accuracy in classifying a wide range of diabetic retinopathy images,'' IEEE Access, vol. 11, pp. 42361–42388, 2023.
25. R. Kommaraju and M. S. Anbarasi, ''Diabetic retinopathy detection using convolutional neural network with residual blocks,'' Biomed. Signal Process. Control, vol. 87, Jan. 2024, Art. no. 105494.
26. APTOS : Eye Preprocessing in Diabetic Retinopathy https://www.kaggle.com/code/ratthachat/aptos-eye-preprocessing-in-diabetic-retinopathy/data, last accessed 2025/07/26.
27. Diabetic Retinopathy Detection Competition Dataset Resized/Cropped, https://www.kaggle.com/tanlikesmath/diabetic-retinopathy-resized, last accessed 2025/07/26.