

Enhanced diabetic retinopathy detection and classification using fundus images with ResNet50 and CLAHE-GAN

Sowmyashree Bhoopal, Mahesh Rao, Chethan Hasigala Krishnappa

Department of Electronics and Communication Engineering, Maharaja Institute of Technology Mysore,
Affiliated to Visvesvaraya Technological University, Belagavi, India

Article Info

Article history:

Received Jan 27, 2024

Revised Feb 25, 2024

Accepted Mar 20, 2024

Keywords:

APTOS-2019 dataset
CLAHE image enhancement
Diabetic retinopathy detection
Fundus image analysis
ResNet50 application

ABSTRACT

Diabetic retinopathy (DR), a progressive eye disorder, can lead to irreversible vision impairment ranging from no DR to severe DR, necessitating precise identification for early treatment. This study introduces an innovative deep learning (DL) approach, surpassing traditional methods in detecting DR stages. It evaluated two scenarios for training DL models on balanced datasets. The first employed image enhancement via contrast limited adaptive histogram equalization (CLAHE) and a generative adversarial network (GAN), while the second did not involve any image enhancement. Tested on the Asia Pacific Tele-Ophthalmology Society 2019 Blindness Detection (APTOS-2019 BD) dataset, the enhanced model (scenario 1) reached 98% accuracy and a 99% Cohen kappa score (CKS), with the non-enhanced model (scenario 2) achieving 95.4% accuracy and a 90.5% CKS. The combination of CLAHE and GAN, termed CLANG, significantly boosted the model's performance and generalizability. This advancement is pivotal for early DR detection and intervention, offering a new pathway to prevent irreversible vision loss in diabetic patients.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Sowmyashree Bhoopal

Department of Electronics and Communication Engineering, Maharaja Institute of Technology Mysore

Affiliated to Visvesvaraya Technological University

Belagavi-580018, India

Email: sowmyashreeb.kp@gmail.com

1. INTRODUCTION

The last decade has seen a marked increase in the prevalence of eye diseases, among which diabetic retinopathy (DR) poses a significant threat due to its ability to lead to severe vision impairment or even blindness. The World Health Organization (WHO) reports that over 2.2 billion individuals globally suffer from vision impairment, with nearly half of these cases being preventable or treatable [1], [2]. The link between diabetes, a condition affecting millions worldwide, and DR is particularly concerning as it leads to damage in the retinal blood vessels, often progressing unnoticed until vision loss becomes evident. The silent progression of DR in its initial stages highlights an urgent need for enhanced diagnostic methodologies capable of early and precise detection [2]. While public awareness campaigns are crucial for early intervention, the traditional diagnostic methods, such as ophthalmoscopy, which rely heavily on manual annotation and are time-consuming, fall short in early-stage detection due to a lack of precision [3]. Routine eye examinations play a significant role in early DR detection, identifying early signs such as microaneurysms (MAs), hemorrhages (HMs), and exudates (EXs) [4].

For diagnosing DR, ophthalmologists categorize retinal lesions into MAs, HEMs, soft exudates (SE), and hard exudates (HE) [5], [6]. MAs are small, red dots, indicative of early DR, while HEMs are larger, irregularly shaped dots. SEs appear as white ovals, and HEs as yellow spots on the retina [5]-[10].

Figure 1 depicts DR's varying severity stages, ranging from normal (a) to proliferate (e). Figure 1(a) shows a normal retina, without any signs of DR. Figure 1(b) represents mild DR, where minimal abnormalities such as microaneurysms begin to appear. Figure 1(c) depicts moderate DR, highlighting more evident signs such as increased microaneurysms and hemorrhages. While Figure 1(d) illustrates severe DR, where significant abnormalities are present, including severe hemorrhages and exudates. Finally, Figure 1(e) shows proliferative DR, the most advanced stage, characterized by the growth of new blood vessels on the retina, leading to severe vision impairment or blindness.

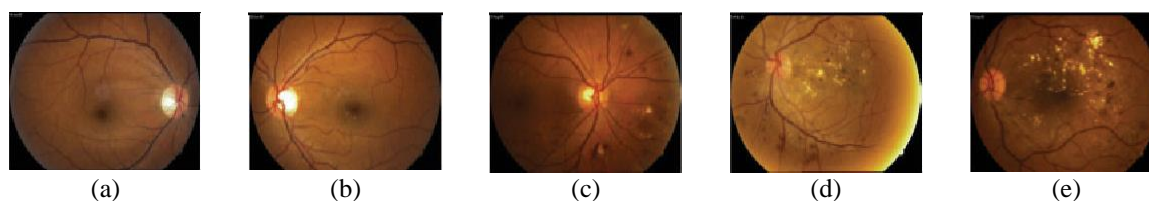


Figure 1. The five sub-classes of DR by severity: (a) normal, (b) mild, (c) moderate, (d) severe, and (e) proliferate

Prior research has laid a significant foundation by identifying the characteristic lesions associated with DR, and categorizing the severity into distinct stages. However, these efforts have been hindered by challenges such as data scarcity, variability, and the manual, labor-intensive nature of diagnostic processes. In response to these challenges, this study proposes an automated deep learning model that significantly enhances the precision of DR detection. By leveraging image enhancement methods like contrast limited adaptive histogram equalization (CLAHE) and employing ensemble learning techniques benefiting from batch normalization and fine-tuning techniques, our model aims to overcome the limitations of current diagnostic approaches. Our contributions to the field of ophthalmology and DR detection are manifold and address several gaps in the current diagnostic landscape:

- Introduces a new method for enriching training and validation datasets, which addresses the crucial issues of data scarcity and variability. This advancement ensures a more robust model training process, leading to improved diagnostic accuracy.
- Our approach incorporates ensemble learning strategies, including a batch normalization layer and fine-tuning techniques, to enhance the model's robustness and accuracy, marking a significant improvement over existing model.
- The application of CLAHE for enhanced image quality, crucial for accurate DR detection on the APTOS-2019 dataset.
- We employ a thorough evaluation framework utilizing precision, recall, Kappa score, and classification matrices to validate the effectiveness of our model comprehensively.
- The innovative application of pre-trained networks, with ResNet50 serving as the backbone, optimizes the detection efficiency, setting a new standard in DR diagnostic methodologies.

The models' training and testing was done using the APTOS-2019 dataset, demonstrating a 98% classification accuracy with enhancement techniques and 95.4% without, validated through a 70:30 hold-out method. Enhanced image analysis was performed using CLAHE and GAN (referred to as CLANG), proving effective in DR stage classification. Conversely, in non-enhanced cases, image oversampling was vital due to dataset class imbalance. The models were developed using ResNet-50, comparing their performance on the APTOS-2019 dataset.

This paper is structured into five main sections: in section 1, the background and motivation for our study is presented, a review of related work is highlighted in section 2. A detailed description of the proposed methodology is presented in section 3. In section 4, a comparative analysis of the results against existing state-of-the-art models is presented, and a conclusion that outlines future research directions is in section 5. By bridging critical gaps in early DR detection, our study not only contributes to the field of ophthalmology but also opens new pathways for future advancements in diagnostic technologies, ultimately aiming to reduce the global burden of vision impairment caused by diabetic retinopathy.

2. RELATED WORK

The detection of DR has been thoroughly investigated from two primary perspectives. It is the traditional approach and the deep learning approach. Given the focus of this research on deep learning (DL) techniques, relevant literature within this specific field is presented.

2.1. Deep learning approaches for DR diagnosis

DL algorithms, particularly convolutional neural networks (CNNs), have revolutionized DR diagnosis by their exceptional ability to classify and predict DR severity from fundus images without the need for explicit feature annotation by experts. This is a stark departure from traditional machine learning methods that rely heavily on manually annotated features, which are not only time-consuming but also prone to overlooking critical lesion features in the images. CNNs, through their hierarchical feature extraction capabilities, can discern both fine-grained details and higher-level semantic information from the full image, facilitating a more comprehensive analysis of potential DR indicators.

Pratt *et al.* [11] leveraged data augmentation and CNNs for feature extraction, successfully identifying complex lesion features that are crucial for DR diagnosis. Abbas *et al.* [12] proposed a novel approach that bypasses the need for pre- or post-processing by utilizing scale-invariant color density and gradient location direction histograms for deep visual feature extraction. This methodology enhances the efficiency and efficacy of DR detection by simplifying the feature extraction and classification process.

Kanungo *et al.* [13] emphasized the impact of hyperparameters, training data quality, and quantity on model performance, underscoring the importance of optimal dataset utilization and parameter tuning. Quellec *et al.* [14] introduced an innovative method to generate heatmaps for visualizing the pixels influencing image-level predictions, offering insights into the “black box” of CNN decision-making processes.

The integration of attention mechanisms for refining DR severity classification was exemplified by Zhao *et al.* [15], demonstrating the utility of focusing on small, significant lesions in fundus images. Orujov *et al.* [16] employed adaptive histogram equalization and a modified fuzzy rule for enhanced vascular detection, showcasing the potential for combining image processing techniques with DL for improved DR diagnosis. Then the images of the fundus are preprocessed such that a segmentation model can recover the vascular branches; the regions of interest are then adjusted using an approach proposed by Das *et al.* [17].

Fan *et al.* [18] developed a residual convolutional block attention model, utilizing adaptive weights within a MobileNetV3 network for DR severity classification. This approach, along with Liu *et al.* [19] compact bilinear pooling network model, highlights the advancements in creating efficient, accurate, and accessible DL models for DR detection. Macsik *et al.* [20] suggest a new alternative to local binary CNN deterministic filter generation that can get close to replicating the conventional CNN’s performance with less training set and memory use. Quantitative tests were done on a publicly available dataset using VGG16, ResNet50 [21], and VGG19 [22] models that had already been trained. These models were picked above dense net and inception-Resnet owing to their simpler structure. As a result of adding more data, the models were more accurate and had less overfitting. Similarly, the researches [23]–[26] investigated the use of ensemble approaches for semantic segmentation using pre-trained U-Net and VGG19 models. The success of ensemble approaches and CNN-based detectors in achieving high accuracy and precision further attests to the potent capabilities of DL in transforming DR diagnostic processes. Ikechukwu and S. Murali [27] favoured the classification of normal and abnormal patients using fundus images. To reconstruct blood vessels from images, maximal primary curvature was used. Missing regions were fixed by adjusting the histogram’s equalization and opening up the morphology. With a precision of 0.97, DIARETDB1 performed admirably.

2.2. Survey findings

The comparative analysis between DL and traditional machine learning approaches for DR detection underscores the superior accuracy and efficiency of DL models. The utilization of publicly available datasets like Kaggle DRD and DRIVE has been instrumental in training and evaluating these models. Despite the encouraging outcomes, challenges such as dataset diversity and model interpretability persist, affecting the generalization of DL models across different populations and obscuring the understanding of model predictions. This comprehensive examination of DL approaches for DR diagnosis underscores a collective move towards creating more explainable, accurate, and efficient diagnostic tools. The ongoing research aims to harness the power of DL ensembles and fine-tuning methodologies to develop an automated, explainable model capable of early DR diagnosis, addressing both the technical and practical challenges in the field.

3. METHOD

The study utilizes the APTOS-2019 dataset [28], known for its diversity in high-resolution retinal images across all five DR stages. With 3,662 images, the dataset presents challenges like data imbalance and variation in image quality due to different photographers and conditions. Figure 2 and Table 1 detail the dataset’s distribution with the number of classes in each subtype.

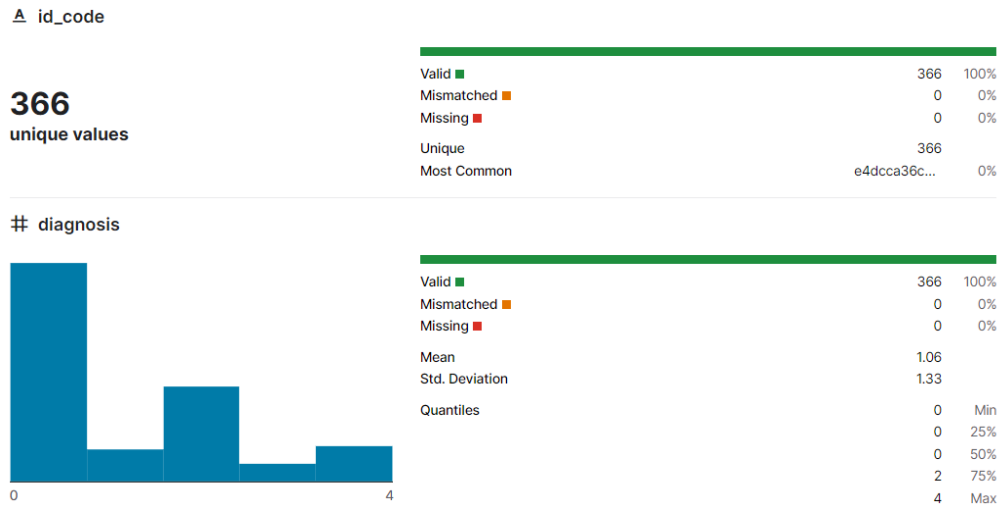


Figure 2. APTOS dataset description with more of class 0 cases

Table 1. Variations in inter-class distribution on APTOS 2019 dataset

Class index	Description	Number of samples
0	No DR	1,805
1	Mild DR	370
2	Moderate DR	999
3	Severe DR	193
4	Proliferate DR	295

The methodology comprises two scenarios: case 1 uses CLAHE [29] preprocessing followed by GAN, while case 2 involves no preprocessing. Both cases focus on preventing overfitting and maintaining image quality. The CLAHE technique is used to control contrast amplification, and GAN is applied for gradient control. ResNet50 [30] is employed for image classification, with preprocessing tasks like resizing and normalization. Figure 3 illustrates the methodology block diagram, including steps like data augmentation and image enhancement.

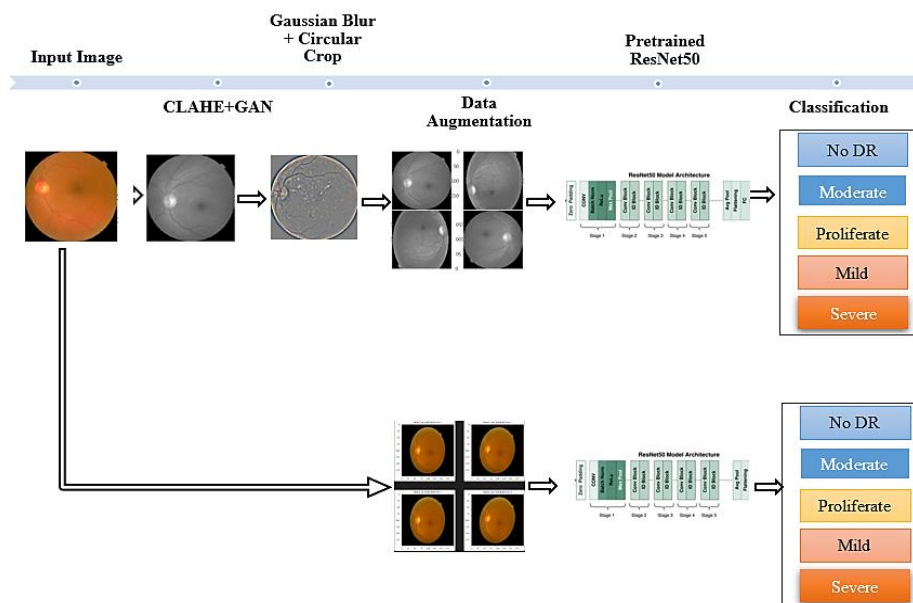


Figure 3. The block diagram of the proposed DR detection (CLANG) approach

3.1. Effect of data augmentation

Data augmentation plays a key role in addressing data scarcity and enhancing dataset diversity [31]. Techniques like horizontal and vertical flips, rescaling, zooming, rotations, and shifts are employed as detailed in Table 2. Figure 4 provides a comprehensive view of the effects of image augmentation on the dataset, enhancing our understanding of its contribution to model generalization. Specifically, Figure 4(a) illustrates the class distribution before image augmentation, while Figure 4(b) showcases the improved number of training samples subsequent to the augmentation process.

Table 2. Parameters for data augmentation

Method	Default	Augmented
Horizontal flip	None	True(p=0.5)
Vertical flip	None	True(p=0.5)
Rescale (normalization)	-	1./255
Zoom range	-	0.25
Rotation (°)	-	60, 90 & 120
x-Shift, y-Shift	None	[-0.1, +0.1]
x-Scale, y-Scale	None	[0.75, 1.25]
Adjusted image	3216 x 2136	224 x 224

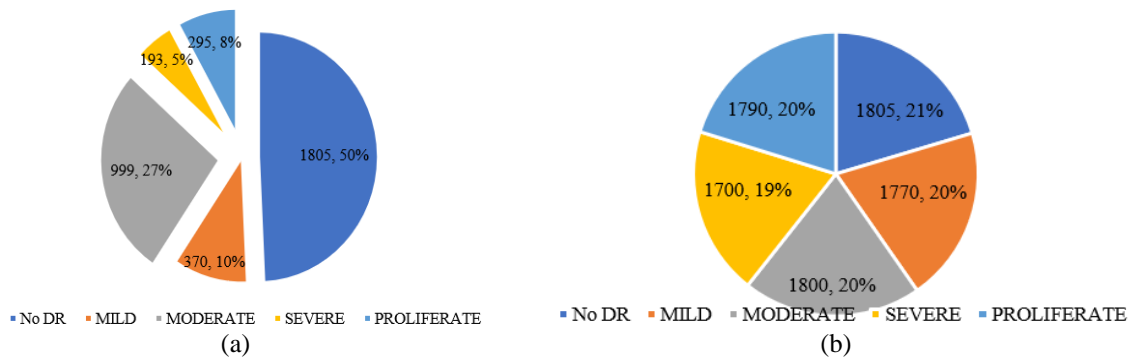


Figure 4. Comparison of training data (a) before augmentation and (b) after augmentation

3.2. Pretrained ResNet50

ResNet-50, a 50-layer CNN, is utilized for its ability to learn residuals [32], addressing vanishing and exploding gradient problems. It employs a residual function, adding the input x to the final output of layers in (1) to (5).

$$y = F(x) + x \tag{1}$$

$$F(x) = H(x) - x \tag{2}$$

$$\text{And } y = P(x) + x = H(x) - x + x = H(x) \tag{3}$$

Since it is now possible to learn identity mappings by simply setting all weights to zero, $H(x)=0$, and $F(x)=x$, the mapping $y=x$ can be learned. Next, the activation function $f()$ is applied, and the final result is $H(x)$, as illustrated in (5). Figure 5 shows the difficulty in separating overlapping class boundaries. Figure 6 shows the ResNet50 with residual blocks, and Figure 7 presents the model hyperparameters, focusing on preventing overfitting [33].

The formula for:

$$P(x) = f(wx + b) \tag{4}$$

whereas the formula for:

$$P(x) = f(x) + x \tag{5}$$

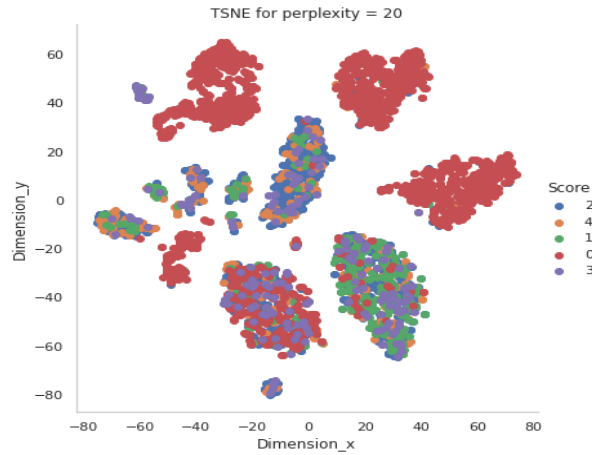


Figure 5. TSNE plot showing the challenges in inter-class separation

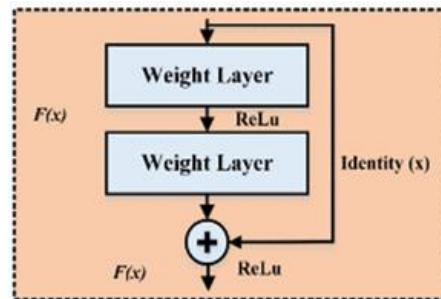


Figure 6. ResNet50 with residual blocks

Layer (type)	Output shape	Param
conv2d (Conv2D)	(None, 26, 26, 28)	784
max_pooling2d	(None, 13, 13, 28)	0
conv2d_1 (Conv2D)	(None, 11, 11, 64)	16,192
max_pooling2d_1	(None, 5, 5, 64)	0
conv2d_2 (Conv2D)	(None, 3, 3, 64)	36,928
flatten (Flatten)	(None, 576)	0
dense (Dense)	(None, 640)	369,280
dropout (Dropout)	(None, 640)	0
dense_1 (Dense)	(None, 264)	169,224
dense_2 (Dense)	(None, 64)	16,960
dense_3 (Dense)	(None, 3)	260

Figure 7. Hyperparameters used in ResNet50 on APTOS dataset

4. RESULTS AND DISCUSSION

4.1. Experimental setup and model training

The study employed the APTOS 2019 dataset, dividing it into training (70%), testing (20%), and validation (10%) subsets. Training involved resizing images to 224×224×3 pixels, using an Intel® CPU with a GTX 1060 GPU. The Adam optimizer was utilized with specific hyperparameters, including a learning rate of 0.0001, a batch size of 8, and 40 epochs. Table 3 details these parameters. The training accuracy reached 98.0%, with early stoppage after the 17th epoch due to no further improvements.

Table 3. Model parameters for training on APTOS-19

Parameter	Value
Batch size	8
Epochs	40
Learning rate	1e-4
Patience	5
Early stoppage	True
Reduce on plateau patience	3
Decay drop	0.5
Kernel size	(3,3)
Filter size	64

4.2. Performance evaluation

Performance was assessed using metrics like accuracy, precision, recall, Cohen's kappa score (CKS), and F-measure, calculated based on true/false positives/negatives in (6) to (10).

$$Accuracy\% = \left(\frac{TP+TN}{TP+FP+TN+FN} \right) * 100 \quad (6)$$

$$Precision\% = \left(\frac{TP}{TP+FP} \right) * 100 \quad (7)$$

$$Recall\% = \left(\frac{TP}{TP+FN} \right) * 100 \quad (8)$$

$$CKS\% = \left(\frac{Po-Pe}{1-Pe} \right) * 100 \quad (9)$$

$$F - Measure\% = 2 \left(\frac{Precision * Sensitivity}{Precision+ Sensitivity} \right) \quad (10)$$

Where TP =true positive, FP =false positive, TN =true negative, and FN =false negative.

The model's classification accuracy for each DR stage is displayed in Figure 8, with a focus on the correct predictions and misclassifications.

4.3. Performance of ResNet50 model

In examining the performance of the ResNet50 model, this study explored two different approaches: one with image enhancement called CLANG and one without it. The results showed a significant improvement in accuracy when using the CLANG enhancement, achieving 98.0% accuracy compared to 95.4% without it. This improvement highlights the value of using enhanced images for better model performance. The findings, including detailed performance metrics, are presented in Tables 4 and 5, focusing specifically on validation accuracy as a key measure of success.

The accuracy of the ResNet50 model, both with and without the use of an early stopping parameter, is visually demonstrated in Figures 9 and 10. Additionally, Figures 11 and 12 showcase examples of the model's predictions on the severity of diabetic retinopathy, providing a clear comparison between the two approaches. Finally, Figure 13 summarizes the performance metrics, highlighting the overall effectiveness of the model.

This comparison clearly shows that using image enhancement techniques like CLANG can lead to better outcomes. Such techniques not only improve the accuracy of models but also make them more reliable for analyzing medical images, an essential factor in healthcare settings. The study strongly supports the integration of image enhancement as a way to enhance the performance of deep learning models in medical diagnostics.

Table 4. Optimal performance metrics on APTOS-19 with enhancement (CLANG)

Accuracy (val)	Precision	Recall	CKS	F-measure	Top-2 acc	Top-2 acc
98.0	99.1	99.0	99.0	99.0	99.7	99.9

Table 5. Optimal performance metrics on APTOS-19 without enhancement

Accuracy(val)	Precision	Recall	CKS	F-measure	Top-2 acc	Top-2 acc
95.4	89.2	81.0	90.5	82.3	92.3	95.2

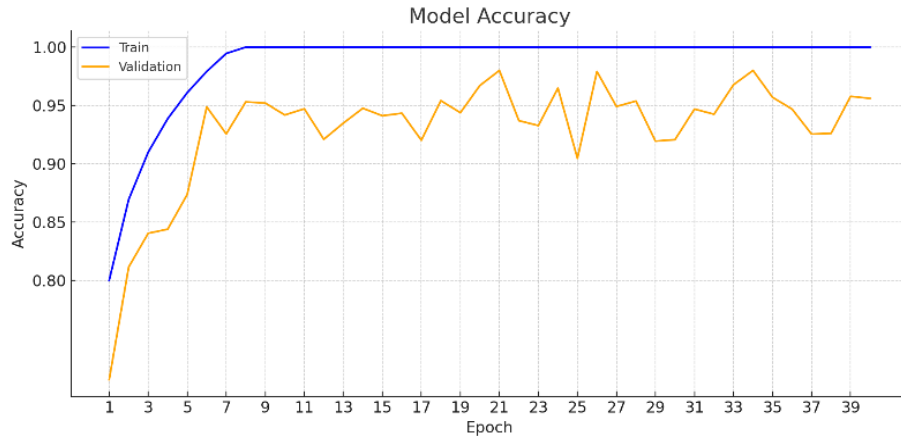


Figure 8. ResNet50 accuracy with an early stoppage factor for case 1

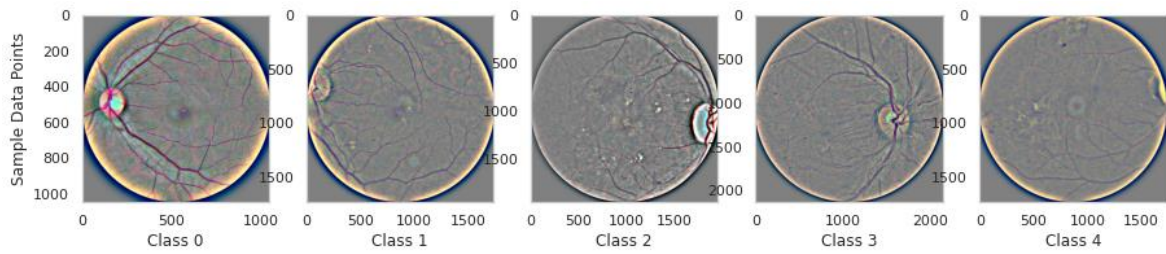


Figure 9. Predicted image corresponding to the DR severity for case 1

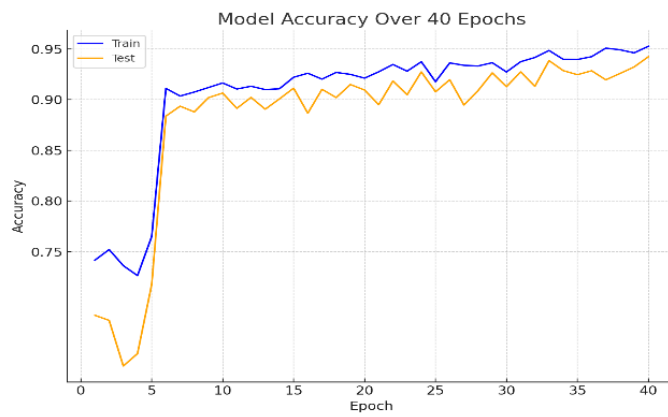


Figure 10. ResNet50 accuracy without an early stoppage factor for case 2

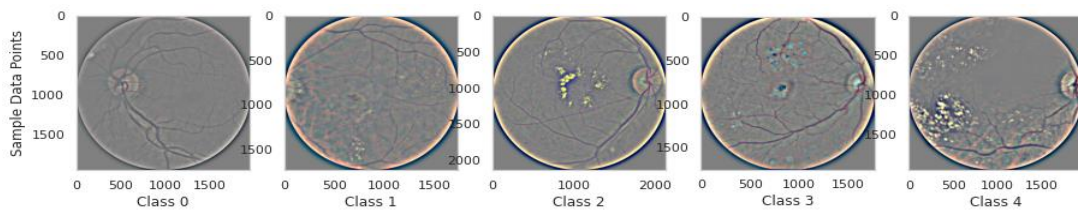


Figure 11. Predicted image corresponding to the DR severity for case 2

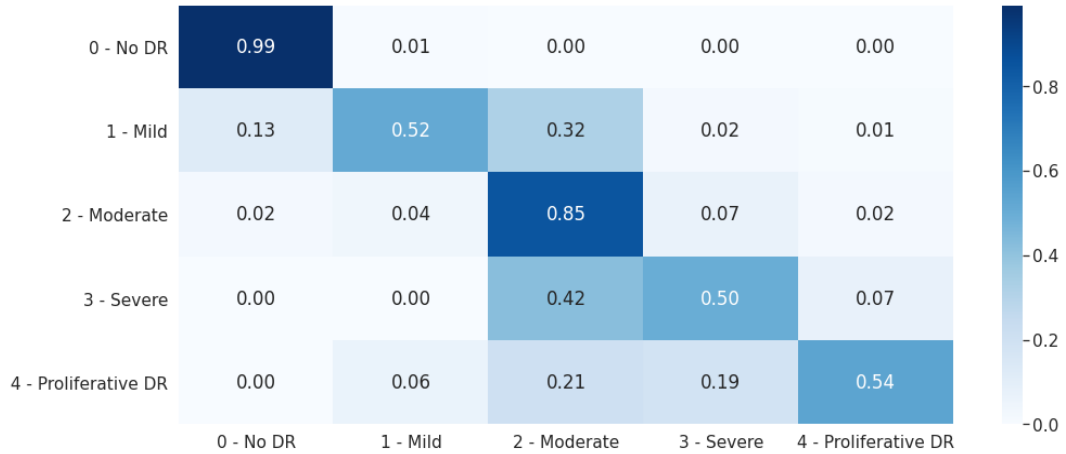


Figure 12. Confusion matrix of the validation set

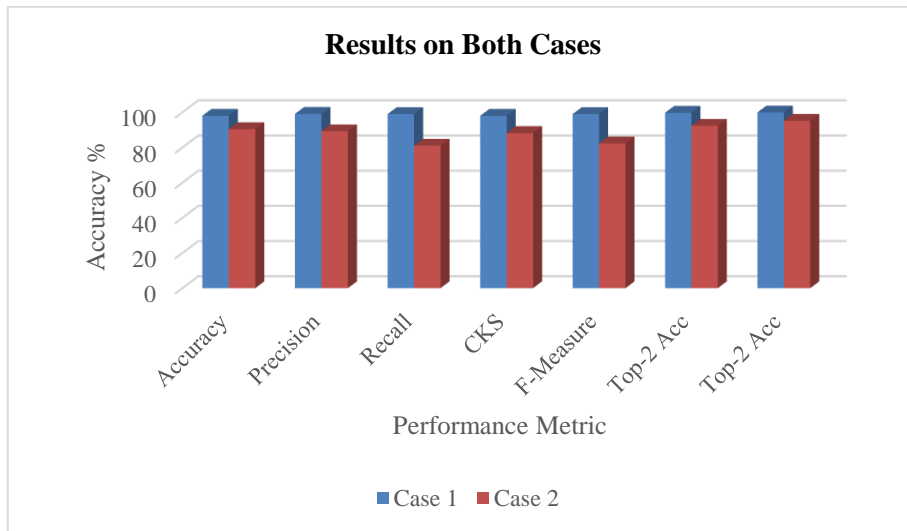


Figure 13. Performance comparison of case 1 vs case 2

4.4. Comparison with other models

The model was compared with other state-of-the-art techniques as shown in Table 6. It outperformed these models, achieving the highest accuracy with CLANG at 98.0%. The study highlighted the importance of data augmentation and CLANG in enhancing accuracy, as shown in Figure 13.

Table 6. Comparisons with relevant literatures

S. No	Method (technique)	Performance accuracy (%)
[34]	CNN	95
[35]	Inception-V2	97
[36]	VGG-16	75
[37]	EfficientNet-B6	86
[38]	Inception-V3, ResNet50	85
[39]	Inception-V2	72
[40]	MobileNetV2	93
[41]	DenseNet	96
[42]	Hybrid U-Net	94
Proposed	Case 1: (ResNet50 with CLANG)	98
	Case 2: (ResNet50 without CLANG)	95.4

The findings demonstrate a significant improvement in the accuracy of DR detection, with the enhanced model achieving a 98% classification accuracy. This surpasses existing models, which typically report accuracies below this benchmark. The integration of CLAHE and GAN for image enhancement has been pivotal. This approach outperforms traditional methods by enhancing the visibility of crucial retinal features, thus facilitating more accurate and early detection of DR. The comparison with state-of-the-art models reveals our model's superior ability to detect subtle indicators of DR at earlier stages, a key factor in preventing the progression of the disease.

4.5. Implications and future directions

The implications of our study extend beyond academic circles, offering a blueprint for the integration of advanced artificial intelligent (AI) models into clinical practice. By automating the detection of DR, our model has the potential to streamline diagnostic processes, enabling earlier interventions and reducing the burden on healthcare professionals. However, challenges such as data privacy, model interpretability, and integration with existing medical records systems must be addressed. Future research should focus on enhancing the model's generalizability, exploring its applicability to other eye diseases, and developing real-time diagnostic systems. This research direction holds promise for revolutionizing ophthalmic care, making early and accurate diagnosis more accessible worldwide.

5. CONCLUSION




The study presents a novel approach for diagnosing DR using the APTOS dataset. Two scenarios are explored: case 1 with CLANG (CLAHE and GAN) image enhancement, and case 2 without it. The enhanced image technique in case 1 significantly improves image quality, contributing to a 98% accuracy rate, compared to 95.4% in case 2. This approach, particularly with CLANG, shows potential to match expert diagnostic accuracy. This limitations include its reliance on a specific dataset, which may not fully represent the global diversity of DR cases. Future work would aim to validate and refine the model using a broader array of datasets to ensure its effectiveness across different populations. Despite these challenges, our contributions significantly advance the field of DR detection, offering a new paradigm for the use of AI in medical diagnostics. Our findings underscore the potential of AI to enhance healthcare delivery, urging continued exploration and development of innovative diagnostic tools. Moving forward, it is imperative to focus on the ethical and practical aspects of AI implementation in healthcare to fully realize its benefits for patient care.

REFERENCES




- [1] T. R. Fricke *et al.*, "Global prevalence of presbyopia and vision impairment from uncorrected presbyopia: systematic review, meta-analysis, and modelling," *Ophthalmology*, vol. 125, no. 10, pp. 1492–1499, Oct. 2018, doi: 10.1016/j.ophtha.2018.04.013.
- [2] Tedros Adhanom Ghebreyesus, "World report on vision," *World Health Organisation*, vol. 214, no. 14, pp. 180–235, 2019.
- [3] R. Cheloni, S. A. Gandolfi, C. Signorelli, and A. Odone, "Global prevalence of diabetic retinopathy: protocol for a systematic review and meta-analysis," *BMJ Open*, vol. 9, no. 3, p. e022188, Mar. 2019, doi: 10.1136/bmjopen-2018-022188.
- [4] A. D. Association, "Microvascular complications and foot care: standards of medical care in diabetes-2021," *Diabetes Care*, vol. 44, pp. S151–S167, 2021, doi: 10.2337/dc21-S011.
- [5] M. Z. Atwany, A. H. Sahyoun, and M. Yaqub, "Deep learning techniques for diabetic retinopathy classification: a survey," *IEEE Access*, vol. 10, pp. 28642–28655, 2022, doi: 10.1109/ACCESS.2022.3157632.
- [6] M. Dubow *et al.*, "Classification of human retinal microaneurysms using adaptive optics scanning light ophthalmoscope fluorescein angiography," *Investigative Ophthalmology and Visual Science*, vol. 55, no. 3, pp. 1299–1309, 2014, doi: 10.1167/iovs.13-13122.
- [7] K. Mazhar, R. Varma, F. Choudhury, R. McKean-Cowdin, C. J. Shtir, and S. P. Azen, "Severity of diabetic retinopathy and health-related quality of life: The Los Angeles Latino eye study," *Ophthalmology*, vol. 118, no. 4, pp. 649–655, 2011, doi: 10.1016/j.ophtha.2010.08.003.
- [8] J. R. Willis *et al.*, "Vision-related functional burden of diabetic retinopathy across severity levels in the United States," *JAMA Ophthalmology*, vol. 135, no. 9, pp. 926–932, Sep. 2017, doi: 10.1001/jamaophthalmol.2017.2553.
- [9] P. Vora and S. Shrestha, "Detecting diabetic retinopathy using embedded computer vision," *Applied Sciences (Switzerland)*, vol. 10, no. 20, pp. 1–10, Oct. 2020, doi: 10.3390/app10207274.
- [10] J. Amin, M. Sharif, and M. Yasmin, "A review on recent developments for detection of diabetic retinopathy," *Scientifica*, vol. 2016, pp. 1–20, 2016, doi: 10.1155/2016/6838976.
- [11] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Procedia Computer Science*, vol. 90, pp. 200–205, 2016, doi: 10.1016/j.procs.2016.07.014.
- [12] Q. Abbas, I. Fondon, A. Sarmiento, S. Jiménez, and P. Alemany, "Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features," *Medical and Biological Engineering and Computing*, vol. 55, no. 11, pp. 1959–1974, Nov. 2017, doi: 10.1007/s11517-017-1638-6.
- [13] Y. S. Kanungo, B. Srinivasan, and S. Choudhary, "Detecting diabetic retinopathy using deep learning," in *RTEICT 2017 - 2nd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, Proceedings*, vol. 2018-January, pp. 801–804, May 2017, doi: 10.1109/RTEICT.2017.8256708.

- [14] G. Quellec, K. Charrière, Y. Boudi, B. Cochener, and M. Lamard, "Deep image mining for diabetic retinopathy screening," *Medical Image Analysis*, vol. 39, pp. 178–193, Jul. 2017, doi: 10.1016/j.media.2017.04.012.
- [15] Z. Zhao *et al.*, "BiRA-Net: bilinear attention net for diabetic retinopathy grading," in *Proceedings - International Conference on Image Processing, ICIP*, vol. 2019-September, pp. 1385–1389, Sep. 2019, doi: 10.1109/ICIP.2019.8803074.
- [16] F. Orujov, R. Maskeliūnas, R. Damaševičius, and W. Wei, "Fuzzy based image edge detection algorithm for blood vessel detection in retinal images," *Applied Soft Computing Journal*, vol. 94, p. 106452, Sep. 2020, doi: 10.1016/j.asoc.2020.106452.
- [17] S. Das, K. Kharbanda, S. M. R. Raman, and E. D. D., "Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy," *Biomedical Signal Processing and Control*, vol. 68, p. 102600, Jul. 2021, doi: 10.1016/j.bspc.2021.102600.
- [18] R. Fan, Y. Liu, and R. Zhang, "Multi-scale feature fusion with adaptive weighting for diabetic retinopathy severity classification," *Electronics (Switzerland)*, vol. 10, no. 12, p. 1369, Jun. 2021, doi: 10.3390/electronics10121369.
- [19] P. Liu, X. Yang, B. Jin, and Q. Zhou, "Diabetic retinal grading using attention-based bilinear convolutional neural network and complement cross entropy," *Entropy*, vol. 23, no. 7, p. 816, Jun. 2021, doi: 10.3390/e23070816.
- [20] P. Macsik, J. Pavlovicova, J. Goga, and S. Kajan, "Local binary CNN for diabetic retinopathy classification on fundus images," *Acta Polytechnica Hungarica*, vol. 19, no. 7, pp. 27–45, 2022, doi: 10.12700/aph.19.7.2022.7.2.
- [21] A. V. Ikechukwu, S. Murali, R. Deepu, and R. C. Shivamurthy, "ResNet-50 vs VGG-19 vs training from scratch: a comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 375–381, Nov. 2021, doi: 10.1016/j.glt.2021.08.027.
- [22] S. Bhimsheyy and A. V. Ikechukwu, "Energy-efficient deep Q-network: reinforcement learning for efficient routing protocol in wireless internet of things," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 33, no. 2, pp. 971–980, Feb. 2024, doi: 10.11591/ijeecs.v33.i2.pp971-980.
- [23] A. V. Ikechukwu and S. Murali, "I-Net: a deep CNN model for white blood cancer segmentation and classification," *International Journal of Advanced Technology and Engineering Exploration*, vol. 9, no. 95, pp. 1448–1464, Oct. 2022, doi: 10.19101/IJATEE.2021.875564.
- [24] A. Victor Ikechukwu and S. Murali, "CX-Net: an efficient ensemble semantic deep neural network for RoI identification from chest-x-ray images for COPD diagnosis," *Machine Learning: Science and Technology*, vol. 4, no. 2, p. 025021, Jun. 2023, doi: 10.1088/2632-2153/acd2a5.
- [25] V. I. Agughasi and M. Srinivasiah, "Semi-supervised labelling of chest x-ray images using unsupervised clustering for ground-truth generation," *Applied Engineering and Technology*, vol. 2, no. 3, pp. 188–202, Sep. 2023, doi: 10.31763/aet.v2i3.1143.
- [26] A. V. Ikechukwu, M. S., and H. B., "COPDNet: an explainable ResNet50 model for the diagnosis of COPD from CXR images," in *2023 IEEE 4th Annual Flagship India Council International Subsections Conference (INDISCON)*, pp. 1–7, Aug. 2023, doi: 10.1109/INDISCON58499.2023.10270604.
- [27] A. V. Ikechukwu and S. Murali, "XAI: an explainable AI model for the diagnosis of COPD from CXR images," in *2023 IEEE 2nd International Conference on Data, Decision and Systems (ICDDS)*, pp. 1–6, Dec. 2023, doi: 10.1109/ICDDS59137.2023.10434619.
- [28] P. Prakash, S. K. Y. Hanumanthaiah, and S. B. Mayigowda, "CRNN model for text detection and classification from natural scenes," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 13, no. 1, p. 839, Mar. 2024, doi: 10.11591/ijai.v13.i1.pp839-849.
- [29] A. Torralba, R. Fergus, and W. T. Freeman, "80 million tiny images: a large data set for nonparametric object and scene recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 1958–1970, Nov. 2008, doi: 10.1109/TPAMI.2008.128.
- [30] N. Papandrianos, E. Papageorgiou, A. Anagnostis, and A. Feleki, "A deep-learning approach for diagnosis of metastatic breast cancer in bones from whole-body scans," *Applied Sciences (Switzerland)*, vol. 10, no. 3, p. 997, Feb. 2020, doi: 10.3390/app10030997.
- [31] M. Salehi, R. Mohammadi, H. Ghaffari, N. Sadighi, and R. Reiazi, "Automated detection of pneumonia cases using deep transfer learning with paediatric chest X-ray images," *British Journal of Radiology*, vol. 94, no. 1121, p. 20201263, May 2021, doi: 10.1259/bjr.20201263.
- [32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-December, pp. 770–778, Jun. 2016, doi: 10.1109/CVPR.2016.90.
- [33] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L. C. Chen, "MobileNetV2: inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 4510–4520, Jun. 2018, doi: 10.1109/CVPR.2018.00474.
- [34] N. M. Thomas and S. Albert Jerome, "Grading and classification of retinal images for detecting diabetic retinopathy using convolutional neural network," *Lecture Notes in Electrical Engineering*, vol. 881, pp. 607–614, 2022, doi: 10.1007/978-981-19-1111-8_45.
- [35] A. Crane and M. Dastjerdi, "Effect of simulated cataract on the accuracy of an artificial intelligence algorithm in detecting diabetic retinopathy in color fundus photos," *Investigative Ophthalmology and Visual Science*, vol. 63, no. 7, pp. 2100-F0089, 2022.
- [36] A. Deshpande and J. Pardhi, "Automated detection of diabetic retinopathy using VGG-16 architecture," *International Research Journal of Engineering and Technology*, 2021, [Online]. Available: www.irjet.net
- [37] Z. Maqsood and M. K. Gupta, "Automatic detection of diabetic retinopathy on the edge," *Lecture Notes in Networks and Systems*, vol. 370, pp. 129–139, 2022, doi: 10.1007/978-981-16-8664-1_12.
- [38] M. Oulhadj *et al.*, "Diabetic retinopathy prediction based on deep learning and deformable registration," *Multimedia Tools and Applications*, vol. 81, no. 20, pp. 28709–28727, 2022, doi: 10.1007/s11042-022-12968-z.
- [39] A. K. Gangwar and V. Ravi, "Diabetic retinopathy detection using transfer learning and deep learning," *Advances in Intelligent Systems and Computing*, vol. 1176, pp. 679–689, 2021, doi: 10.1007/978-981-15-5788-0_64.
- [40] C. Lahmar and A. Idri, "On the value of deep learning for diagnosing diabetic retinopathy," *Health and Technology*, vol. 12, no. 1, pp. 89–105, 2022, doi: 10.1007/s12553-021-00606-x.
- [41] M. Canayaz, "Classification of diabetic retinopathy with feature selection over deep features using nature-inspired wrapper methods," *Applied Soft Computing*, vol. 128, 2022, doi: 10.1016/j.asoc.2022.109462.
- [42] V. K. K. D.K. Salluri, V. Sistla, "HRUNET: Hybrid Residual U-Net for automatic severity prediction of diabetic retinopathy," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 11, no. 3, pp. 530–540, 2023.




BIOGRAPHIES OF AUTHORS

Sowmyashree Bhoopal    holds a Bachelor of Engineering in Electronics and Communication Engineering, Master of Technology in VLSI and Embedded Systems. Currently pursuing her Ph.D. in Electronics and Communication Engineering, besides several professional certificates and skills, and she is working as Assistant Professor in the Department of ECE at Maharaja Institute of Technology Mysore. She is a member in Indian Society for Technical Education (ISTE), International Association of Engineers (IAENG) and Institute of Electrical and Electronics Sight (IEEE sight). Her research area of interest is in image processing and AI and ML applications. She can be contacted at email: sowmyashreeb.kp@gmail.com.



Mahesh Rao    is an expert in Electrical Engineering, earned his Ph.D. from the University of Wyoming in 1988, focusing on 3-D actuator design for robotics. He started with a BE in Electronics and Communication from SJCE, University of Mysore in 1981, followed by a Masters from the University of Windsor in 1984. Rao served as an Assistant Professor at the University of Reno, Nevada, before joining Intel in 1992, advancing to Validation Manager for Chipsets and Program Manager for Integrated Microprocessors. In 2002, he founded Aspire Communication, a successful embedded product company. Returning to academia, he's now a Professor and HOD at MIT Mysore's ECE Department. His research, with over 50 papers and numerous patents, focuses on image processing, AI, and ML. He can be contacted at email: raomahesh08@gmail.com.



Chethan Hasigala Krishnappa    holds a Ph.D. in Computer Science from the University of Mysore, India. He is a Professor in the Department of Computer Science and Engineering at Maharaja Institute of Technology, India. His expertise spans a wide range of disciplines, including algorithms, artificial intelligence, and software engineering. He is highly skilled in areas such as pattern recognition, computer vision, document analysis, image processing, feature extraction, image segmentation, digital image processing, and software engineering, making significant contributions to these fields. He can be contacted at email: hkchethan@gmail.com.