SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V11I9P124

Volume 11 Issue 9, 280-289, September 2024

Original Article

Enhancing Mild Diabetic Retinopathy Detection: A Comparative Study of CLAHE-Preprocessed and Unprocessed Fundus Images Using a Minimal CNN Model

P.M. Ebin[1], P. Ranjana[2]

[1,2]*CSE, Hindustan University, Tamilnadu, India.*

[1]*Corresponding Author : pmebin195@gmail.com*

Received: 15 July 2024 Revised: 24 August 2024 Accepted: 14 September 2024 Published: 30 September 2024

Abstract *- Early identification of Diabetic Retinopathy (DR) is crucial for preserving vision in individuals with diabetes, as this condition is a leading cause of sight loss among diabetic patients. Timely detection enables effective treatment interventions. Hence, we proposed a Minimal Convolutional Neural Network (MCNN) model for detecting mild DR symptoms using fundus images. Utilizing publicly available datasets from Kaggle and Messidor, the research applies Contrast Limited Adaptive Histogram Equalization (CLAHE) preprocessing to enhance image quality. The MCNN is then trained on both CLAHE-processed and unprocessed versions of the same images. The research evaluates CLAHE preprocessing's effect on mild DR detection by analyzing the model's performance across both datasets, seeking to quantify any accuracy improvements. This approach leverages modern machine learning techniques to potentially improve early diagnosis of DR, addressing a critical need in ophthalmological practice and diabetic care.*

Keywords *- AI, CLAHE, Deep learning, Diabetic Retinopathy, MCNN.*

1. Introduction

Diabetes patients may experience diabetic retinopathy, an eye condition. The biggest global cause of vision loss in adults is diabetic retinopathy. Early detection is necessary to slow the disease's progression. One of the most widespread chronic diseases in developed nations, diabetic retinopathy is the main factor in middle-aged people losing their eyes. Small alterations in the retinal capillaries are the first signs of DR. Micro aneurysms, which are small disturbances of the retinal capillary, are the first identifiable abnormalities. Intraregional bleeding is caused by deformed micro aneurysms. This causes the initial stage of diabetic retinopathy, often known as mild non-proliferative diabetic retinopathy [1-3].

Fundus image is most suitable for screening purposes because the eye fundus responds to a list of vascular diseases. Indeed, the ability to extract fundus images, as well as methods of efficiently processing images in order to identify pathologies, have a clear link to the screening approach's outcome [4-6]. However, clinical diagnosis of DR might be challenging in low-resource settings because there are not enough ophthalmologists to treat every diabetes patient. Frequent fundus checks can prevent DR-related blindness. Manual diagnosis takes a lot of time and is not precise [7]. As a result, several computer vision approaches for automatically detecting diabetic retinopathy and its phases from retinal images have been developed. Deep Neural Networks (DNNs) are frequently referred to as brain-inspired systems of deep learning [8-10]. It may create a distributed depiction of data by becoming familiar with a variety of high-level attributes or attribute types that are directly extracted from the massive original data. Despite this, there are still issues with using neural networks in medical research. For starters, adequate real-world medical photographs, particularly for some specialized disorders, are difficult to come by. In addition, the accessibility of labelled medical data is usually constrained.

Secondly, because DR characteristics are so complicated, they are probably to interact with other lesions, and DR's minute lesions are difficult to detect if picture quality is inadequate. Fundus pictures are labelled by a human operating technique that is vulnerable to subjectivity, according to medical publications. Furthermore, utilizing a single model that is trained with a limited collection of medical imaging data and unavoidable image noise makes it difficult to successfully achieve high illness detection accuracy. Transfer learning is, hence, the primary method used in deep learning [11-13]. Furthermore, utilizing a single model that is trained

with a limited collection of medical imaging data and unavoidable image noise makes it difficult to successfully achieve high illness detection accuracy. Aside from deep learning approaches, several image-processing techniques have been created. Complex characteristics are manually detected in image processing algorithms [14, 15]. Diabetic retinopathy may be recognized early, which can significantly lower the chance of blindness. However, because of the variable morphology of diabetic retinopathy at different stages, automated diabetic retinopathy detection is a difficult process. However, to enhance the efficiency and effectiveness of feature extraction and detection, it is necessary to develop a novel solution.

The field of Diabetic Retinopathy (DR) detection using deep learning faces several critical challenges that hinder its effectiveness in clinical settings. Current preprocessing methods often compromise image quality by overemphasizing edges while eliminating crucial fine details, leading to suboptimal feature extraction. The arbitrary scaling of images necessitates increasingly complex models, raising concerns about overfitting and generalization errors.

Additionally, existing approaches struggle with accurate multi-class grading of DR severity, which is essential for timely intervention and personalized treatment. These limitations collectively point to a need for innovative solutions that can preserve important retinal details during preprocessing, balance model complexity with generalization capability, and improve the precision of DR severity classification. Such advancements would need to be clinically viable, integrating seamlessly with existing healthcare systems while offering improved accuracy and efficiency. Addressing these challenges could significantly enhance the early detection and management of DR, potentially reducing vision loss among diabetic patients worldwide. This objective of the study involves the following processes.

1. This study introduces a Minimal Convolutional Neural Network (MCNN) approach for analyzing retinal fundus images. The proposed model aims to detect and classify early signs of diabetic retinopathy, distinguishing between cases with mild symptoms and those without any indications of DR. By focusing on a streamlined neural network architecture, the research seeks to process ocular images for accurate diabetic retinopathy screening efficiently.
2. Utilize a mixture of Kaggle and Messidor publicly available datasets. Before passing the dataset to the proposed MCNN, we preprocessed the data using the CLAHE method.
3. The MCNN is trained using both CLAHE-processed and unprocessed datasets. We evaluate the model's performance by comparing accuracy, precision, recall, and F1 scores. To further assess the classification model's effectiveness, we also generate a ROC curve for analysis.

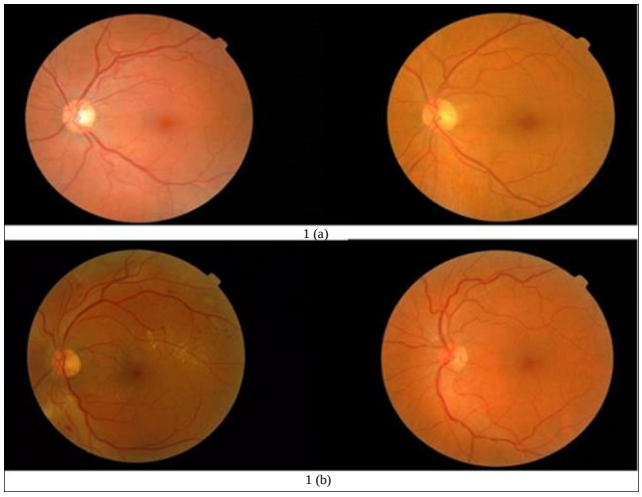


Fig. 1 a) Normal Fundus image, and b) Fundus image showing mild Diabetic Retinopathy (DR).

2. Literature Survey

Different deep learning models exist today to classify and predict Diabetic Macula Edema (DME) and Diabetic Retinopathy (DR). Various researchers are working on it, and fruitful results are coming to solve the DR patient's problems automatically. It also helps the clinical practitioners do their jobs easily. However, the limitations of dataset availability and credibility still exist. Deep learning techniques have recently enhanced the precision and speed of identifying diabetic retinopathy, marking a notable improvement in diagnostic capabilities.

Researchers have developed more sophisticated convolutional neural network architectures that can analyze retinal images with greater precision, often incorporating attention mechanisms to focus on subtle features indicative of the disease. There has been a shift towards multi-modal approaches, combining various imaging techniques like fundus photography and optical coherence tomography for a more comprehensive analysis. Explainable AI models have gained traction, providing transparency in the decision-making process and increasing trust among clinicians. Transfer learning and few-shot learning techniques have addressed the challenge of limited labeled data, while real-time screening systems have made deployment in resource-constrained settings more feasible. Integration with electronic health records has enabled more personalized risk assessments, and advancements in severity grading have improved treatment planning. These innovations collectively signify a major advancement in the timely identification and treatment of diabetic retinopathy, offering the potential to minimize vision impairment risks for a vast number of individuals across the globe.

S. Qummar et al. [16] introduced an ensemble model for detecting all DR stages using the Kaggle dataset. They resized 3888×2951 images to 786×512 and employed up-sampling

and down-sampling to balance the data. The data was divided into three sets: 64% for training, 20% for testing, and 16% for validation. The learning parameters were fine-tuned using Nesterov-accelerated Adaptive Moment Estimation. An NVIDIA Tesla k40 GPU was employed, resulting in 80.8% model accuracy. The highest recall (0.97) was observed in class 0, while class 1 had the lowest (0.54). With 113 samples, class 4 achieved the highest AUC of 0.97. Marginal improvements were noted when using a 0.0001 learning rate.

T. Li et al. [17] developed a Convolutional Attention Block (CAB) for area-wise feature detection in imbalanced DR data and a Global Attention Block (GAB) for small lesion information. These were combined to create CABNet for DR grading. The model comprises a backbone, GAB, CAB, and classifier, with various backbones tested. MobileNet 1.0 performed best, achieving 0.8569 accuracy and 0.8794 Kappa score. The study utilized Messidor, EyePACS, and DDR datasets, employing 512x512 input resolution and data augmentation to prevent overfitting. Adam optimizer trained each module over 70 epochs, with an initial 0.005 learning rate, 16 batch size, and cross-entropy loss function.

T. Araújo et al. [18] proposed the neovessel (NV) generation algorithm for augmenting Proliferative Diabetic Retinopathy (PDR) images. This algorithm generates DR-labeled datasets to enhance model training. NVs were placed near the Optic Disc (OD) with 0.15 probability, never at 0.40, and 0.45 for both locations. The Region of Interest (ROI) was set to 25% of the image size. Vessel segmentation used Otsu thresholding on the green channel, while OD segmentation employed the UOLO framework, validated across public datasets. Color assignment considered nearby surroundings. The DR Graduate model assessed the augmentation's effectiveness, showing improved no vessel detection capacity.

L. Qiao, Y. Zhu, and H. developed a framework to identify micro aneurysms in ocular fundus images [19]. They employed a semantic segmentation algorithm to categorize fundus images as normal or affected and detect micro aneurysms. The team focused on recognizing Non-Proliferative DR (NPDR) using deep CNN, with eye fundus images as input. They sourced their dataset from Ieee-Dataport.org. The initial preprocessing phase enhances dark lesion borders, facilitating background separation. Subsequently, an optimized wideband pass filter is applied to improve exudate contrast. Their lesion detection system comprises four key stages: vessel extraction, optical disc elimination, potential lesion identification, and pre/post-processing procedures. This approach enables comprehensive analysis of fundus images for early DR detection.

K. Shankar et al. [5] introduced a novel approach for identifying and categorizing diabetic retinopathy stages. Their study utilized the MESSIDOR dataset, employing CLAHE for contrast enhancement while avoiding unwanted noise amplification. Image segmentation was performed using a histogram-based model, followed by feature extraction with HPTI-v4. The experiment parameters included a 0.9 velocity, 500 epochs, and a 0.001 learning rate. Bayesian optimization was applied to determine the optimal hyperparameter combination. The HPTI-v4 method achieved impressive results: 99.49% accuracy, 98.83% sensitivity, and 99.68% specificity. This comprehensive approach demonstrates significant potential for improving diabetic retinopathy detection and classification in clinical settings.

N. Barhate et al. [20] developed a model combining an Auto Encoder (AE) with a VGG network (VGG AE) to address overfitting issues during architecture testing. Their approach utilizes the AE's dual-component structure: the encoder learns to compress input images into concise representations, while the decoder is trained to reconstruct these representations back into full images. This innovative design aims to enhance model robustness and generalization capabilities, potentially improving the accuracy and reliability of diabetic retinopathy detection systems in clinical applications. The encoder network is enjoying the VGG network. For training, EyePACS datasets are used, and the images are resized to 256×256 size in the preprocessing stage. Rotation and flipping methods were used to augment the data. Sigmoid, tanh, and ReLU activation functions are used, and the researchers found that the ReLU activation function gives more accuracy and the best result. The proposed architecture, VGG AE, shows the test accuracy at 76.27% and reduces the overfitting effectively.

Researchers developed a novel architecture called the Cross-disease Attention Network (CANet) to assess DR and DME severity while examining their interconnections [21]. This model incorporates specialized attention modules for each condition and their interdependencies, capturing unique characteristics and inter-disease relationships. These components are embedded within a deep neural framework for feature extraction. The system utilizes ResNet50 to generate feature maps, with the coarsest resolution map feeding into condition-specific attention modules. The disease-dependent module then analyzes intrinsic connections between disorders, generating features for DMR and DR. Two loss functions are applied for grading, with additional values to comprehend disease-related features. During training, images were resized to 350×350 and augmented. Adam optimizer was used with a 0.0003 initial learning rate, training for 1000 epochs with 40 batch sizes. Backbone model selection involved joint training for DR and DME grading, comparing ResNet34 (81.4%), Resnet50 (82.0%), and DenseNet161 (78.9%) on the Messidor dataset. CANet achieved 96.3% AUC for DR and 92.4% for DME, with accuracies of 92.6% (DR) and 91.2% (DME).

Z. Khan et al. [22] developed a multi-class classification model combining VGG16 and spatial pyramid pooling. The SPP layer reduces input images from 1024x1024 to 224x224,

creating a fixed-size output vector between the last convolution and the first connected layer. A Network-in-Network (NiN) tops the SPP layer to detect nonlinear data patterns. Transfer learning modifies VGG's fully connected layer while freezing convolution layers. Training used 8 batch sizes for 50 epochs, resulting in 52% fewer parameters and 95% micro AUC.

F. Saeed [23] proposed an intelligent DR grading system without image preprocessing. It employs a two-stage fine-tuning approach: first, a pre-trained CNN model embeds DR lesion structure using lesion ROIs. Second, fully connected layers are replaced by a Principal Component Analysis (PCA) layer to mitigate overfitting. The study utilized EyePACS and Messidor datasets for validation, demonstrating the model's effectiveness on challenging public datasets. The two data augmentation methods are employed in which they extract the ROIs around the lesion with distinct sizes and resize them to the same size, 64×64. Again, rotate each ROI in four different directions and flip horizontally. The ResNet152 with CONV1 and Gradient Boosting (GB) layer achieved the best result. The proposed method gave 99.73% accuracy. The Kappa scores are 98.45% and 96.67%, respectively.

C-H. Hua et al. [24] proposed a CNN having a two-fold feature augmentation capability, which provides more generalization at the feature level. The model is validated using a small-scale data set from Kyung Hee University Medical Center (KHUMC). The model uses ResNet residual blocks as the primary feature extractor. The two-fold feature augmentation contains weight sharing convolution kernels and reverses cross-stream. The RCA stream contains three components by itself. The first one is Self-Contest Aggregation (SCA), the second one is Pairwise Reverse Attention (PRA), and the final one is Multi-Level Fusion (MLF). All the images are rescaled to different sizes, like 224×224, 448×448 and 600×600. RCA is evaluated by the Messidor dataset, which achieves 94.8% accuracy.

According to Kaushik, Singh et al. [25], the dataset of eye images that is currently available has several color aberrations and unnecessary illuminations, which affect the model's accuracy. The study developed a stacking deep learning model and an image-processing framework for diagnosis. The gray world color constancy algorithm improves brightness normalization and image quality. The stack generalization model is prepared using three distinct CNNs. The dataset used is EyePACS, and to enrich the dataset, data augmentation is applied, including horizontal and vertical flips, width shift, height shift, fill mode, and zoom range. Test accuracy for binary classification using the stack model was 97.92%, and for multiclass classification, it was 87.45%.

Rajkumar et al. [26] employed transfer learning with ResNet-50 to classify diabetic retinopathy stages. They used a Kaggle-derived dataset, resized to 512x512 and labeled 0-4 by severity. The model, utilizing ReLU activation, ran on a DGX server with two Tesla V100 GPUs. It achieved 89.4% accuracy, 97% specificity, and 57% sensitivity. Comparisons with ResNet18, VGG19, Inception V3, AlexNet, and VGG19 showed ResNet 50's superior performance in all metrics.

Kaushik et al. [25] proposed a novel image processing schema and layered deep learning approach to eliminate unnecessary reflectance components in retinal images. They employed grayscale global color constancy for brightness normalization and image quality enhancement. The method's effectiveness was evaluated using the Peak Signal-to-Noise Ratio (PSNR) and Mean Squared Error (MSE) of normalized images. A layered generalization of Convolutional Neural Networks (CNN) was developed for a computer-aided diagnostic system. However, this approach is currently limited to binary classification of diabetic retinopathy.

Kalyani and colleagues introduced an innovative approach for identifying and classifying diabetic retinopathy [27]. Their method employs convolutional and primary capsule layers to extract key features from retinal images, while class capsule and softmax layers assess the probability of an image belonging to a particular diagnostic category. Utilized the Messidor dataset to confirm the efficacy of the proposed network change on four performance metrics. But this approach uses a lot of computer power, which makes it time- and energy-consuming. The authors developed a capsule network (CapsNet) for diabetic retinopathy detection, outperforming modified AlexNet with 97.98% accuracy.

The model uses convolutional and capsule layers for feature extraction and classification. Trained on the Messidor dataset, it identifies four stages of retinopathy. Future work involves expanding to five-stage classification and additional datasets for improved early diagnosis. The authors [28] developed a fast, accurate framework for detecting diabetic retinopathy levels from fundus images. Using CLAHE preprocessing and a lightweight parallel CNN, we extracted 120 key features. An ELM model then classified DR levels. Our approach outperformed state-of-the-art models on two datasets, offering quicker processing and earlier detection to prevent vision loss. A novel Deep ML-FEC system was created, utilizing a pre-trained convolutional neural network to autonomously recognize and categorize retinal abnormalities in fundus images [29]. This model differentiates five stages of diabetic retinopathy, including clinically significant macular edema. With 94.40% accuracy and 76.35% sensitivity, the system surpasses current alternatives, demonstrating the potential for both clinical implementation and widespread screening initiatives. The work developed by the researchers [30] DRCNNRB, a model using CNN and residual blocks to assess diabetic retinopathy severity automatically. It addresses performance degradation and vanishing gradient issues. Tested on the preprocessed DR 2015 dataset, it improves image clarity and accuracy.

3. Proposed Methodology

3.1. Dataset Preparation and Pre-processing

The study utilized publicly accessible fundus images from Kaggle and Messidor databases. A rigorous data cleansing process was implemented to ensure high-quality samples for analysis. To maintain image integrity, the collected dataset underwent standardization, being resized to 224x224x3 dimensions and converted to TIFF format for optimal preservation of visual characteristics. Table I shows the number of datasets used for the experiment from each category. The fundus images are preprocessed using the CLAHE method, and the output of the CLAHE preprocessed fundus is shown in Figure 2.

Table 1. Used dataset

Data Sets	Number of Images	Total Images
Messidor	337(No DR) 101(Mild DR)	438
Kaggle	1213(No DR) 1449(Mild DR)	2662

The contrast is increased using the contrast-constrained adaptive histogram equalization method. The noise is frequently amplified excessively by the AHE algorithm. To overcome this problem, the researchers used the CLAHE algorithm, which prevents overamplification by limiting the amplification. CLAHE considers the small regions of the images called tiles. Utilizing bilinear interpolation, adjacent tiles are blended to erase the false borders. The parameter clip limit is a threshold for contrast limiting, normally 40, and the tile grid size is 8×8. The total number of images is divided, like 1550 for mild DR and 1550 for no DR. Overall, 3100 quality images were selected for applying the CLAHE method. Image quality is one parameter to consider while training a model. The accuracy directly depends on the quality of the input images.

Contrast Limited Adaptive Histogram Equalization (CLAHE) improves upon traditional image enhancement methods by offering a more refined, context-sensitive approach. This technique divides an image into smaller sections, called tiles, and applies histogram equalization to each individually. By doing so, CLAHE can tailor its enhancement to the specific needs of different image areas, bringing out details in low-contrast regions while avoiding over-enhancement in areas that already have good visibility.

The CLAHE method calculates intensity histograms for each tile, applying a clip limit to prevent over-saturation. The clipped histograms are then rescaled, maintaining the original pixel count in each tile. CLAHE uses these adjusted histograms to transform pixel intensities, improving local contrast while preserving the overall intensity distribution. After processing, the enhanced tiles are smoothly blended back into the image, with interpolation techniques ensuring seamless transitions between tiles. The CLAHE algorithm follows a structured approach, beginning with image segmentation into tiles and computing histograms for each. It then applies a threshold-free mapping to determine clip limits, adjusting contrast accordingly. This technique encompasses multiple crucial phases: first, generating a histogram; second, determining and allocating surplus values; third, redistributing these excess quantities; and lastly, employing a cumulative distribution function to scale and map the results. Each step contributes to the overall enhancement of the image quality. This methodical approach allows CLAHE to enhance local image details while maintaining overall image integrity, making it particularly useful in fields like medical imaging, where preserving subtle features is crucial.

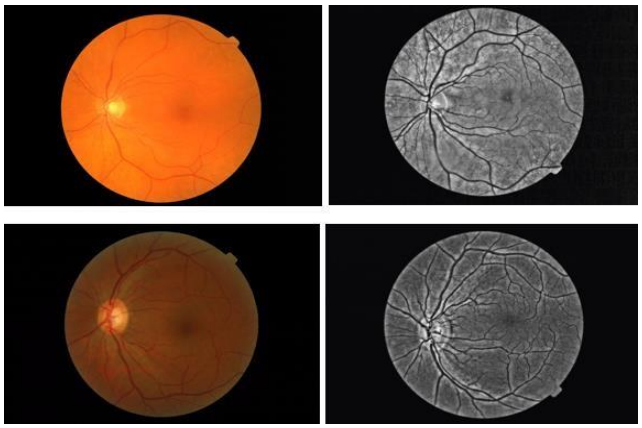


Fig. 2 Sample image before CLAHE (left) and after CLAHE (right) taken from Messidor

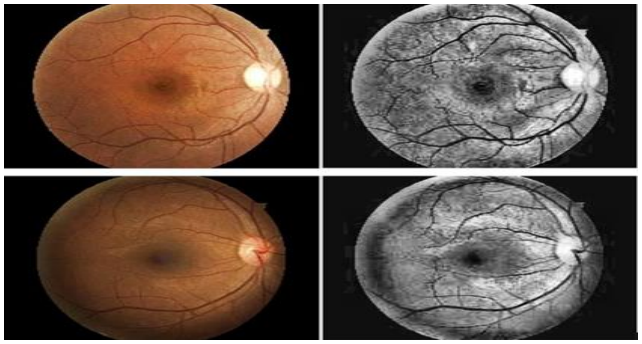


Fig. 3 Sample image before CLAHE (Left) and after CLAHE (Right) taken from Kaggle

4. MCNN Architecture

The Minimum CNN (MCNN) architecture presents a streamlined approach to image classification, specifically tailored for distinguishing between No-DR and Mild-DR in retinal images. This model adopts a sequential structure, building complexity through a series of carefully designed layer stacks. Each stack in the MCNN contributes to the overall goal of efficient feature extraction and classification. At the foundation of the MCNN lies the first layer stack. This initial stage employs a convolution layer with 64 channels, utilizing the ReLU activation function to introduce non-

linearity. Following this, a max pooling layer with a 2x2 filter helps to reduce spatial dimensions while retaining important features. The inclusion of a dropout layer at this early stage serves as a preventive measure against overfitting, enhancing the model's ability to generalize. Building upon this foundation, the second layer stack increases the model's capacity for feature detection. It introduces a convolution layer with an expanded channel size of 128, again paired with ReLU activation. This layer is followed by another max pooling operation and a dropout layer, maintaining the pattern established in the first stack while allowing for more complex feature representation.

The third stack layer represents the deepest level of feature extraction in the MCNN. Here, a convolution layer with 256 channels is employed, significantly increasing the model's ability to capture intricate patterns. As with previous stacks, this layer is complemented by ReLU activation, max pooling, and dropout, ensuring consistency in the architecture's approach to feature learning and regularization. The final stages of the MCNN architecture focus on classification. A dense layer comprising 64 neurons processes the features extracted by the previous layers. This is then connected to a fully connected layer with two neurons, corresponding to the binary classification task of differentiating between No-DR and Mild-DR cases. The model employs categorical cross-entropy as its loss function, a choice well-suited to classification tasks, and utilizes the RMSprop optimizer for efficient training. This combination of architectural elements and training parameters aims to provide accurate DR detection while maintaining computational efficiency. The model parameters are shown in Figure 4.

```
(1, 128, 128)
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 64, 128, 128)      640
_________________________________________________________________
activation (Activation)      (None, 64, 128, 128)      0
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 32, 64, 128)       0
_________________________________________________________________
dropout (Dropout)            (None, 32, 64, 128)       0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 10, 21, 128)       147584
_________________________________________________________________
activation_1 (Activation)    (None, 10, 21, 128)       0
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 5, 11, 128)        0
_________________________________________________________________
dropout_1 (Dropout)          (None, 5, 11, 128)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 1, 3, 256)         295168
_________________________________________________________________
activation_2 (Activation)    (None, 1, 3, 256)         0
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 1, 2, 256)         0
_________________________________________________________________
dropout_2 (Dropout)          (None, 1, 2, 256)         0
_________________________________________________________________
flatten (Flatten)            (None, 512)               0
_________________________________________________________________
dense (Dense)                (None, 64)                32832
_________________________________________________________________
activation_3 (Activation)    (None, 64)                0
_________________________________________________________________
dropout_3 (Dropout)          (None, 64)                0
_________________________________________________________________
dense_1 (Dense)              (None, 2)                 130
_________________________________________________________________
activation_4 (Activation)    (None, 2)                 0
=================================================================
Total params: 476,354
Trainable params: 476,354
Non-trainable params: 0
```

Fig. 4 Implementation of the MCNN model

5. Results and Analysis

The MCNN model undergoes training using two distinct fundus image sets: one without CLAHE processing and another with CLAHE applied. Prior to input, images undergo preprocessing. Model evaluation employs various metrics, including confusion matrices, precision, recall, accuracy, and F1 scores. Additionally, the AUC-ROC curve graphically illustrates MCNN performance, plotting True Positive Rate (TPR) against False Positive Rate (FPR) to visualize classification effectiveness across different thresholds.

5.1. MCNN with Non-CLAHE Images

Optimal results were achieved when training MCNN on non-CLAHE images over 20 epochs, reaching 0.9760 training accuracy. Figure 5 illustrates training and validation accuracy trends. The model ran on a local PC featuring a 4GB Nvidia Geforce GTX 1650Ti GPU, using Jupyter Notebook and Anaconda for development. Performance metrics are shown in Table 2, where Class 1 represents mild DR and Class 2 indicates no DR. For mild DR, the model achieved 97% precision, 98% recall, and 97% F1 score. No DR classifications yielded 98% precision, 97% recall, and 98% F1 score. Mathematical formulas for calculating these parameters are provided below.

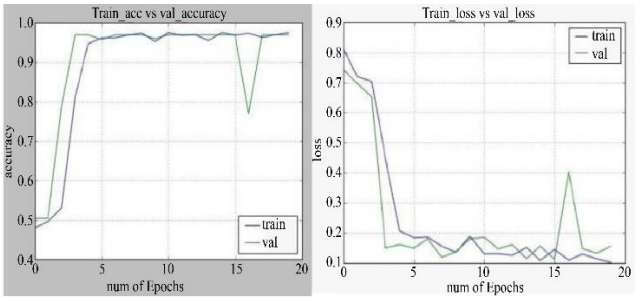


Fig. 5 Training and validation metrics for original (non-CLAHE) images: loss and accuracy across epochs

Table 2. Model performance measures

	Precision	Recall	F1-Score	Support
Class 1	0.97	0.98	0.97	439
Class 2	0.98	0.97	0.98	491
Micro Avg	0.97	0.97	0.97	930
Macro Avg	0.97	0.97	0.97	930
Weighted Avg	0.97	0.97	0.97	930

F1 Score = 2× (precision×recall) / (recall+ precision)

Accuracy = TN+ TP / (TP+ FP +TN+FN)

Recall (Sensitivity) = TP/ (FN+TP)

Precision = TP/ (FP+TP)

The performance of the MCNN with the Non-CLAHE model is checked using the confusion matrix. The confusion matrix can be drawn with the help of True Positive (TP) and True Negative (TN) values. A confusion matrix is the combination of predicted and actual values in a table format, which helps to find precision, recall, accuracy, and AUC-ROC. Figure 6 depicts the confusion matrix obtained and the normalised confusion matrix from the model MCNN with Non-CLAHE.

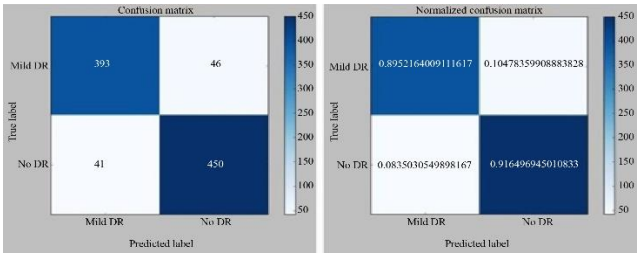


Fig. 6 Confusion matrix and normalized confusion matrix for Non-CLAHE images

5.2. MCNN with CLAHE Images

The MCNN is again trained with CLAHE images to compare the impact generated by the non-CLAHE images. The training with CLAHE images uses 20 epochs with the help of the same GPU configuration. The proposed model achieved a training accuracy of 0.9413. For mild DR classification, it demonstrated 89% precision, 90% recall, and an 89% F1 score. In the case of NO DR classification, the model showed 91% precision, 90% recall, and a 90% F1 score. Figure 7 shows the training validation loss and training validation accuracy.

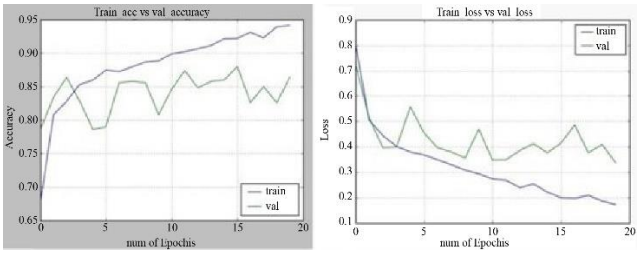


Fig. 7 Training and validation metrics for CLAHE-enhanced images: loss and accuracy over epochs

The following table 3 shows the precision, recall, and F1 score values obtained.

Table 3. Model performance measures

	Precision	Recall	F1-Score	Support
Class 1	0.89	0.90	0.89	439
Class 2	0.91	0.90	0.90	491
Micro Avg	0.90	0.90	0.90	930
Macro Avg	0.90	0.90	0.90	930
Weighted Avg	0.90	0.90	0.90	930

Figure 8 shows the confusion matrix and normalized confusion matrix that were derived from the MCNN model with CLAHE pictures. Out of 930 fundus images, 393 images are recognized as true positive and 450 images are recognized as true negative.

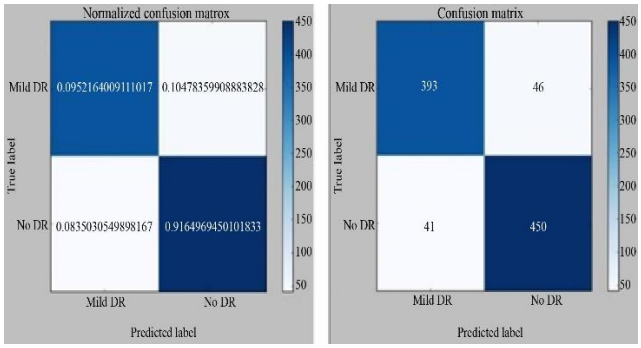


Fig. 8 Confusion matrix and normalized confusion matrix for CLAHE images

6. Discussions on CLAHE Impact

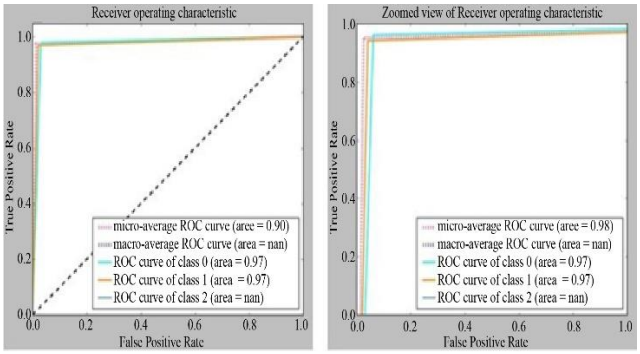
The application of Contrast Limited Adaptive Histogram Equalization (CLAHE) as a preprocessing step demonstrates a significant impact on mild diabetic retinopathy detection in this study. By enhancing the contrast and local details of fundus images, CLAHE appears to improve the visibility of subtle retinopathy features that may be challenging to detect in unprocessed images.

The comparative analysis likely reveals that the minimal CNN model achieves higher accuracy and improved feature extraction when trained on CLAHE-preprocessed images compared to unprocessed ones. This suggests that CLAHE effectively addresses some of the inherent limitations of fundus photography, such as non-uniform illumination and low contrast, which can obscure early signs of diabetic retinopathy.

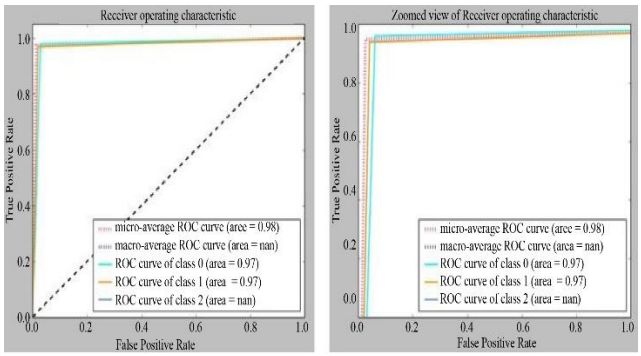
The enhanced performance observed with CLAHE preprocessing underscores the importance of image enhancement techniques in medical image analysis, particularly for conditions like mild diabetic retinopathy, where early detection is crucial for timely intervention and management.

The result shows that the training accuracy of MCNN with non-CLAHE images is 0.9760, and with CLAHE images is 0.9413. Both trainings took 20 epochs. MCNN with non-CLAHE shows better accuracy than MCNN with CLAHE. The generated feature map is shown in the following figure 10. It is the feature detected by Minimum CNN. A feature map can be generated by applying the filters on the input images, and the output of each layer is a feature map.

The following figure shows the ROC curve obtained when training with non-CLAHE images and CLAHE images.



(a) ROC of Non-CLAHE images



(b) ROC of CLAHE images

Fig. 9 ROC curve and its zoomed view of the MCNN model with CLAHE and Non-CLAHE images

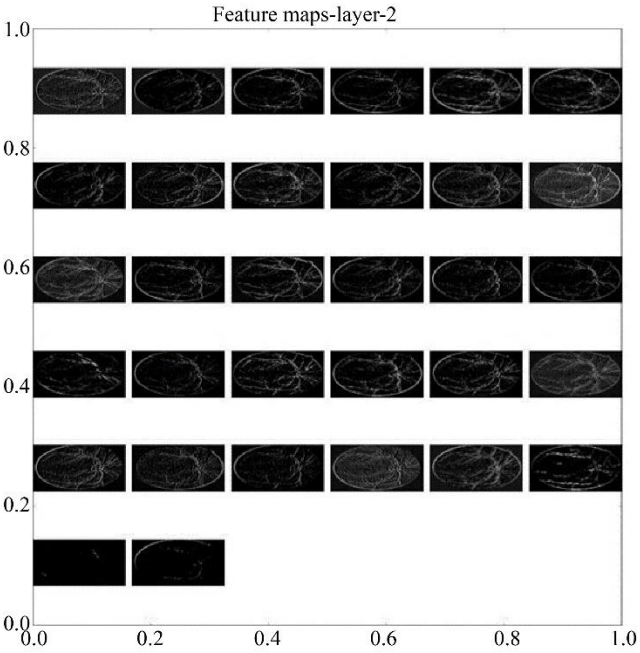


Fig. 10 Features detected by MCNN

7. Limitations and Future Work

A notable constraint in this research is the reliance on a single image enhancement method (CLAHE) and a basic CNN architecture. This narrow focus, while enabling direct comparisons, may not fully exploit the potential of advanced preprocessing and deep learning for diabetic retinopathy screening.

Subsequent studies could investigate diverse image enhancement techniques, including alternative contrast adjustment algorithms, noise reduction methods, or blood vessel extraction processes, to potentially uncover more effective strategies for emphasizing subtle retinal abnormalities associated with diabetic retinopathy. The study presumably focused on mild diabetic retinopathy detection, which is valuable but may not provide a complete picture of the disease's progression. A limitation here is the potential lack of analysis on moderate, severe, and proliferative diabetic retinopathy stages. Future research could extend the comparison to include all stages of diabetic retinopathy, providing a more comprehensive evaluation of the CLAHE preprocessing and minimal CNN model approach across the full spectrum of the condition. Another potential limitation is the size and diversity of the dataset used. Depending on the study's scope, the dataset may not fully represent the variety of fundus images.

Future research could address these areas by incorporating explainable AI techniques to provide insight into the model's decision-making process, optimizing the model for deployment on resource-constrained devices, and conducting clinical trials to assess the practical impact of the proposed approach on diabetic retinopathy screening and diagnosis in real-world healthcare settings encountered in clinical practice, including those from different ethnic groups, age ranges, or imaging devices. Future work should aim to validate the findings on larger, more diverse datasets and potentially explore the model's generalizability across different populations and imaging equipment.

8. Conclusion

The study proposes a simplified convolutional neural network for detecting and categorizing early diabetic retinopathy indicators, utilizing combined datasets from Kaggle and Messidor. The MCNN was trained on two image sets: one preprocessed with CLAHE and another without CLAHE. The preprocessing involves applying CLAHE to enhance image quality.

Performance analysis revealed that the model achieved 0.9413 accuracy with CLAHE-processed images, while non-CLAHE images yielded a higher accuracy of 0.9760. For photographs that are not CLAHE, the MCNN model performs better than CLAHE images. Researchers are interested in carrying out the experiment using the current mobile net architecture with an SVM linear combination in the next investigations. Owing to the scarcity of datasets, the author intends to work with medical professionals to establish a dataset so that the inquiry can be carried out with greater testing accuracy. The work can be improved in future by comparing different works.

References

[1] Sam E. Mansour et al., "The Evolving Treatment of Diabetic Retinopathy," *Clinical Ophthalmology*, vol. 14, pp. 653-678, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[2] Skylar Stolte, and Ruogu Fang, "A Survey on Medical Image Analysis in Diabetic Retinopathy," *Medical Image Analysis*, vol. 64, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] David A. Antonetti, Paolo S. Silva, and Alan W. Stitt, "Current Understanding of the Molecular and Cellular Pathology of Diabetic Retinopathy," *Nature Reviews Endocrinology*, vol. 17, pp. 195-206, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[4] Maximilian W.M. Wintergerst et al., "Diabetic Retinopathy Screening using Smartphone-Based Fundus Imaging in India," *Ophthalmology*, vol. 127, no. 11, pp. 1529-1538, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] K. Shankar et al., "Hyperparameter Tuning Deep Learning for Diabetic Retinopathy Fundus Image Classification," *IEEE Access*, vol. 8, pp. 118164-118173, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Gilbert Lim et al., "Different Fundus Imaging Modalities and Technical Factors in AI Screening for Diabetic Retinopathy: A Review," *Eye and Vision*, vol. 7, pp. 1-13, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Pelin Görgel, "Greedy Segmentation Based Diabetic Retinopathy Identification using Curvelet Transform and Scale Invariant Features," *Journal of Engineering Research*, vol. 9, no. 1, pp. 134-150, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] Wejdan L. Alyoubi, Wafaa M. Shalash, and Maysoon F. Abulkhair, "Diabetic Retinopathy Detection through Deep Learning Techniques: A Review," *Informatics in Medicine Unlocked*, vol. 20, pp. 1-11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] K. Shankar et al., "Automated Detection and Classification of Fundus Diabetic Retinopathy Images Using Synergic Deep Learning Model," *Pattern Recognition Letter*, vol. 133, pp. 210-216, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Morgan Heisler et al., "Ensemble Deep Learning for Diabetic Retinopathy Detection Using Optical Coherence Tomography Angiography," *Translational Vision Science & Technology*, vol. 9, no. 2, pp. 1-11, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Ibrahem Kandel, and Mauro Castelli, "Transfer Learning with Convolutional Neural Networks for Diabetic Retinopathy Image Classification. A Review," *Applied Sciences*, vol. 10, no. 6, pp. 1-24, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Rupa Patel, and Anita Chaware, "Transfer Learning with Fine-Tuned MobileNetV2 for Diabetic Retinopathy," *2020 International Conference for Emerging Technology (INCET)*, Belgaum, India, pp. 1-4, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[13] David Le et al., "Transfer Learning for Automated OCTA Detection of Diabetic Retinopathy," *Translational Vision Science & Technology*, vol. 9, no. 2, pp. 1-9, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] Nikki Gaikwad et al., "Automated System for Detection for Diabetic Retinopathy using Image Processing," *International Research Journal of Engineering and Technology*, vol. 7, no. 8, pp. 4980-4985, 2020. [Google Scholar] [Publisher Link]

[15] Sarni Suhaila Rahim, Vasile Palade, and Andreas Holzinger, *Image Processing and Machine Learning Techniques for Diabetic Retinopathy Detection: A Review*, Artificial Intelligence and Machine Learning for Digital Pathology, pp. 136-154, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[16] Sehrish Qummar et al., "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," *IEEE Access*, vol. 7, pp. 150530-150539, 2019. [CrossRef] [Google Scholar] [Publisher Link]

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

[18] Teresa Araújo et al., "Data Augmentation for Improving Proliferative Diabetic Retinopathy Detection in Eye Fundus Images," *IEEE Access*, vol. 8, pp. 182462-182474, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Lifeng Qiao, Ying Zhu, and Hui Zhou, "Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms," *IEEE Access*, vol. 8, pp. 104292-104302, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[20] Nikhil Barhate et al., "Reducing Overfitting in Diabetic Retinopathy Detection using Transfer Learning," *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India, pp. 298-301, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] Xiaomeng Li et al., "CANet: Cross-Disease Attention Network for Joint Diabetic Retinopathy and Diabetic Macular Edema Grading," *IEEE Transactions on Medical Imaging*, vol. 39, no. 5, pp. 1483-1493, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[22] Zubair Khan et al., "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture," *IEEE Access*, vol. 9, pp. 61408-61416, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[23] Fahman Saeed, Muhammad Hussain, and Hatim A. Aboalsamh, "Automatic Diabetic Retinopathy Diagnosis Using Adaptive Fine-Tuned Convolutional Neural Network," *IEEE Access*, vol. 9, pp. 41344-41359, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[24] Cam-Hao Hua et al., "Convolutional Network with Twofold Feature Augmentation for Diabetic Retinopathy Recognition from Multi-Modal Images," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 7, pp. 2686-2697, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[25] Harshit Kaushik et al., "Diabetic Retinopathy Diagnosis From Fundus Images Using Stacked Generalization of Deep Models," *IEEE Access*, vol. 9, pp. 108276-108292, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[26] R.S. Rajkumar et al., "Transfer Learning Approach for Diabetic Retinopathy Detection using Residual Network," *2021 6th International Conference on Inventive Computation Technologies (ICICT)*, Coimbatore, India, pp. 1189-1193, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[27] G. Kalyani et al., "Diabetic Retinopathy Detection and Classification using Capsule Networks," *Complex & Intelligent Systems*, vol. 9, pp. 2651-2664, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[28] Md. Nahiduzzaman et al., "Diabetic Retinopathy Identification using Parallel Convolutional Neural Network based Feature Extractor and ELM Classifier," *Expert Systems with Applications*, vol. 217, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[29] Tiwalade Modupe Usman et al., "Diabetic Retinopathy Detection using Principal Component Analysis Multi-Label Feature Extraction and Classification," *International Journal of Cognitive Computing in Engineering*, vol. 4, pp. 78-88, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[30] Rajasekhar Kommaraju, and M.S. Anbarasi, "Diabetic Retinopathy Detection Using Convolutional Neural Network with Residual Blocks," *Biomedical Signal Processing and Control*, vol. 87, no. A, 2024. [CrossRef] [Google Scholar] [Publisher Link]