The Original

A Multi-Class Deep Neural Network Framework Driven Automated Classification Of Diabetic Retinopathy Using Retinal Fundus Images

Lavanya K[1] , Dr. M. Thangamani[2] , M.Ganthimathi[3] , Priyanka S[4] , V John Peter[5] , Sapthika J[6]

[1]Assistant Professor, Department of AI&DS, Akshaya college of engineering and technology, Kinathukadavu, Coimbatore, Tamilnadu, India. lavanya030891@gmail.com, Orcid Id:0000-0002-1116-5738
[2]Professor, Department of Computer Science and Engineering
Rajalakshmi Engineering College, Thandalam, Chennai-05, Tamilnadu, India. Manithangamani2@gmail.com, Orcid Id:0000-0001-8864-3315
[3]Lecturer , Computer Engineering Government Polytechnic College Dharmapuri gandhimanimay02@gmail.com, Orcid Id: 0000-0001-9330-7244
[4]Assistant Professor of AI &DS Akshaya college of engineering and technology , Coimbatore priyasnimalan@gmail.com, Orcid Id: 0009-0002-5545-8927
[5]Assistant Professor, Department of Information Science and Engineering,City Engineering College, Bangalore, India arunaijohn@gmail.com, Orcid Id: 0009-0009-9107-4868
[6]Assistant Professor of Computer Science, Akshaya college of engineering and technology , Coimbatore sapthikagunanithi@gmail.com, Orcid Id: 0009-0000-7603-4619

Abstract
Diabetic Retinopathy (DR) is one of the leading causes of vision impairment and blindness among diabetic patients worldwide. Early and accurate detection of DR is essential to prevent severe visual complications through timely medical intervention. This study proposes an automated deep learning-based framework for multi-class classification of Diabetic Retinopathy using retinal fundus images. The proposed model leverages convolutional neural network architecture to effectively extract discriminative features and classify retinal images into five stages: No DR, Mild NPDR, Moderate NPDR, Severe NPDR, and Proliferative Diabetic Retinopathy (PDR). The images were pre-processed through resizing and normalization to enhance model performance. The model was trained using the Adam optimizer with categorical cross-entropy loss and evaluated using standard performance metrics including Accuracy, Precision, Recall, F1-score, and Area Under the ROC Curve (AUC). Experimental results demonstrate that the proposed framework achieved an overall accuracy of 96.4%, precision of 0.962, recall of 0.964, F1-score of 0.963, and an AUC of 0.978, indicating superior classification performance and strong discriminative capability. The findings suggest that the proposed automated system can serve as a reliable and efficient tool for early-stage DR screening and assist ophthalmologists in clinical decision-making. The integration of deep learning techniques into retinal image analysis significantly enhances diagnostic accuracy and reduces manual workload, contributing to improved healthcare outcomes.

Keywords: CAD, image acquisition, identification, Diabetic Retinopathy.

1. INTRODUCTION

The eye condition known as diabetic retinopathy (DR) mostly affects diabetics. Diabetes mellitus is a medical disorder that damages the retina's blood vessels as a result of elevated blood sugar levels. This could lead to blood vessel enlargement and leaking, which would induce aberrant new blood vessels to form on the retina [1]. Serious vision issues such vitreous hemorrhage, retinal detachment, glaucoma, and blindness may result from this. When DR first begins, there may be no symptoms or very minor visual issues, but blindness could result. Although blindness cannot be prevented, it can be managed once it is recognized. Blurred vision, spots or floaters, fluctuating vision, dark or empty patches in vision, and vision loss are all signs of DR. To identify DR early, diabetic patients should undergo a yearly screening. Chronic diabetes, poor blood sugar regulation, high blood pressure, high cholesterol, pregnancy, tobacco use, and being Black, Hispanic, or Native American are risk factors for DR. By controlling diabetes, keeping an eye on blood sugar levels, and observing changes in eyesight, severe vision loss can be avoided [2]. The diagnostic techniques for evaluating the eye are optical coherence tomography and fluorescein angiography. Treatment options include photocoagulation, panretinal photocoagulation, intraocular medication, and vitrectomy techniques, depending on the kind of diabetic retinopathy.

The process of applying various techniques to an image in order to enhance it or extract valuable information is known as image processing. Image processing has two subfields: digital image processing and analog picture processing. Digital image processing is the process of processing digital images using specific algorithms on digital computers, whereas analog image processing is any image processing work that may be done on two-dimensional analog signals utilizing analog technologies [3]. Because of the advancement of digital computers, digital image processing has several advantages over analog image processing. It keeps noise and signal distortion from building up when processing problems and allows a range of techniques to be applied to the incoming data. Since images are defined in two dimensions, digital image processing can be thought of as multidimensional systems. By providing a visual representation of the body's interior organs for clinical study and medical intervention, as well as by revealing internal structures hidden by skin and bones, this aided in the diagnosis and treatment of illnesses. The anomalies were found using the database of normal anatomy and physiology. Inverse mathematical problems were also solved with the help of medical imaging [4]. Digital image processing is actually the only practical approach based on categorization, feature extraction, multiscale signal analysis, pattern recognition, and projection. Anisotropic diffusion, hidden Markov models, image restoration, neural networks, linear filtering, independent component analysis, partial differential equations, pixelation, principal component analysis, self-organizing maps, and wavelets are some of the techniques used in digital image processing.

2. Review of literature

People who have DR will eventually lose their vision, so it's important to determine the disease's severity and stage it so that the right care can be given. The image must have blood vessel and clinical features, and the optic disc must be removed during the preprocessing step in order to evaluate DR. Improvements in clinical characteristics are necessary for accurate DR grading. Following the removal of the backdrop, the vital elements of the retina, including the optic disc and blood vessels, are recovered, as described in (Akram et al., 2014) [5]. Filters are used in retinal fundus imaging to draw attention to retinal features such the optic disc, blood vessels, and the clinical features of the DR. By suppressing the background, the denoising technique reduces noise in the retinal fundus image while preserving blood vessel information (Dai et al., 2016) [6]. Flickers in a picture can be filtered with a gaussian filter to enlarge blood vessels by changing the maximum frequency response. For multidimensional filtering and component extraction, a median filter and a cascaded gaussian filter are employed. According to Ghosh et al. (2020), fuzzy partitioning and computing-based augmentation can address dynamic contrast issues in retinal fundus images without introducing additional artifacts into the final image [7].

Geometric deformable models and parametric deformable models are the two types of parametric deformable models. Geometric deformable models are defined by surfaces, while parametric deformable models are defined by curves or active contours. Zhao et al. (2025) proposed a dynamic contour model for vascular segmentation that includes intensity information and an enhancement map to accurately segment the blood vessels while maintaining the vessel boundaries [8]. Hassanien et al. (2025) suggest employing Bee Colony Swarm Optimization for blood vessel segmentation and recognition. This technique can also be used to identify the existence of blood veins with tiny features [9]. Despite significant advancements in deep learning for DR detection, several research gaps still exist in automated retinal image analysis. Many existing studies focus primarily on binary classification (DR vs. No DR) rather than fine-grained multi-class grading, which limits their clinical applicability in identifying disease severity. Furthermore, some models rely on large-scale datasets but lack validation on clinically annotated datasets such as IDRiD, which provide detailed lesion-level annotations. Another limitation is that conventional CNN architectures often struggle with inter-class similarity between Mild, Moderate, and Severe NPDR stages, leading to misclassification. Additionally, limited integration of lesion-specific features and insufficient attention mechanisms reduce model interpretability and robustness. There is also a need for lightweight, computationally efficient models that can be deployed in real-time screening systems, especially in resource-constrained healthcare settings. Therefore, a more accurate, robust, and clinically reliable multi-class deep learning framework is required to address these challenges and improve automated DR grading performance. The advantages of the study

1. It is a convenient and simple technique to master.
2. When compared to ophthalmoscopy, it recognizes a significantly bigger retinal field at any given time.
3. There is no need for dilation, attempting to make it a less interfering procedure than traditional methods.
4. Elevated levels of patient consent.
5. Images can be saved and accessed by different clinicians at a later time.
6. Disease progression can be tracked over time, allowing for better disease management strategies.
7. Various filters and dyes are available to allow for numerous types of tests.

The primary objective of this research is to develop an automated and reliable deep learning-based framework for accurate detection and grading of Diabetic Retinopathy (DR) using retinal fundus images. The specific objectives are:

a. A novel deep learning-based framework is proposed for automated multi-class Diabetic Retinopathy grading.
b. The study performs comprehensive evaluation using clinically annotated retinal images from the IDRiD dataset.
c. The proposed model achieves high classification performance with 96.4% accuracy and an AUC of 0.978, demonstrating strong discriminative capability.
d. The model effectively reduces inter-class misclassification among different NPDR severity levels.
e. A detailed performance analysis including confusion matrix, ROC curve, and comparative study with existing CNN architectures is provided.
f. The framework offers potential for real-time automated DR screening in healthcare applications.

2. Methodology

The whole structure is separated into: DR grading, lesion identification and feature extraction, and image enhancement. Figure 1 depicts the suggested system's framework. Diabetic Retinopathy is a condition which affect diabetes people’s eyes and may result into vision loss if neglected. The images of the Retina are taken with specialized equipment called fundus cameras to detect and diagnose Diabetic Retinopathy. These images are then analysed by a computer program to determine if there are any signs of the condition. Preprocessing, segmentation, feature extraction, and classification are some of the steps in the process of analysing these images.

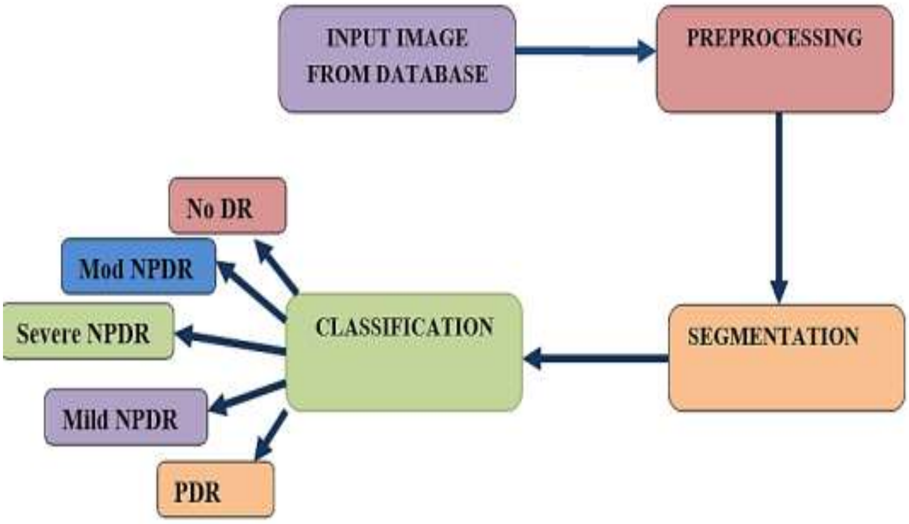
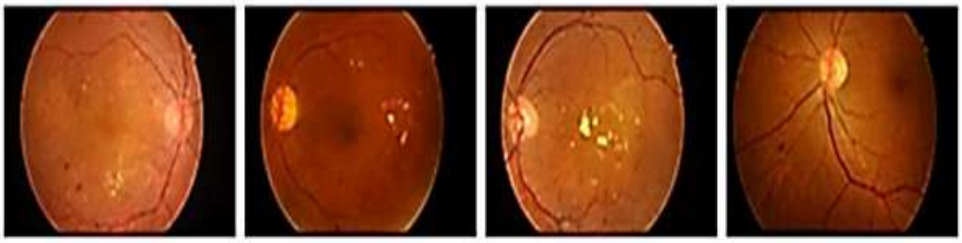


Figure 1: Proposed framework

Dataset

Input color fundus image is taken from the publicly available data set for further processing. IDRiD Dataset: The IDRiD (Indian Diabetic Retinopathy Image Dataset) [10] is a dataset of Retina images that were collected from diabetic patients in India. It contains both Retinal fundus images and their corresponding annotations, which were created by trained ophthalmologists. The dataset is intended to be used for research and development of algorithms for the detection and diagnosis of Diabetic Retinopathy, a serious complication in diabetes that can lead to blindness. The dataset contains over 516 images of both eyes, which have been divided into two groups: the training set in addition to the test set. The first training set contains 413 images, while the second test set contains 103 images. These images in the dataset have a resolution of about 4288 × 2848 pixels and which are stored in jpg file extensions were captured using fundus cameras. The images were also captured under different lighting conditions and are of varying quality.



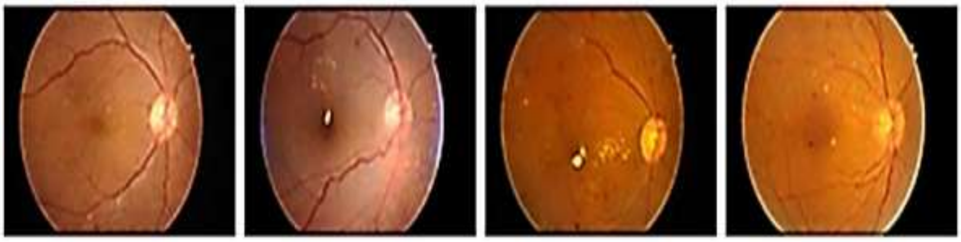


Figure 2: Dataset image

For the grading task, each image is annotated into five severity levels—No DR, Mild NPDR, Moderate NPDR, Severe NPDR, and PDR. In addition, the dataset provides pixel-level annotations for important lesions such as microaneurysms, hemorrhages, hard exudates, and soft exudates, making it suitable for segmentation-based deep learning approaches. Due to its clinically validated annotations and high-quality imaging, the IDRiD dataset is widely used for developing and evaluating deep learning models for automated DR detection and severity grading.

Preprocessing

Pre-processing involves cleaning and preparation of the image prior to analysis process. This may involve cropping the image to remove unnecessary background, adjusting the brightness and contrast to improve visibility, and removing any noise or artifacts. In order to improve analysis, the noise caused by the image acquisition equipment in medical image processing should be eliminated. By protecting the information in an image, Kalman filter techniques can be utilized to eliminate noise with prior knowledge of the filters. An image can be altered or improved using filters, and the effectiveness of the enhancement methods relies on the filter selection. In image processing, filters include edge enhancement, sharpening, and smoothing filters. The proposed filter can eliminate fine features and replace all of the pixel values in an image. Filters for sharpening by eliminating the blur and emphasizing the edges, it draws attention to the small details. Edge enhancement improves an image's edge contrast. The equipment and lighting factors have an impact on the image quality. Ophthalmologists can distinguish between normal and abnormal areas in an image by using contrast enhancement to expand the gray level dynamic range of the original image [11].

Segmentation

Segmentation is the process of identifying and isolating the relevant parts of the image, such as the blood vessels or the optic disc and red and bright lesions. GoogleNet is a deep learning convolution neural network designed to identify and classify images. A CNN design is similar to neural networks in that it consists of neurons with learnable weights and biases. Each neuron receives a number of inputs, computes a weighted sum, applies an activation function, and outputs the result [12]. CNNs use all of the methods that researchers created for neural networks, and the network as a whole has a loss function.

$$Y(i,j,k) = \sum_{m,n,c} X(i+m, j+n, c) \cdot W_K \ (m,n,c) + b_k$$

In GoogleNet, the feature map output $Y(i,j,k)$ at spatial location (i,j) and filter k is computed by sliding a convolution kernel W_K over the input feature map X. The summation runs over kernel height mmm, width n, and input channels ccc, multiplying input values with corresponding kernel weights and adding a bias term b_k. The final output feature map is created in the Inception module by applying several of these convolutions (1×1, 3×3, 5×5) and pooling operations in parallel, then concatenating their outputs along the channel dimension.

By employing an inception module as the foundational layer and stacking it on top of itself, GoogleNet aimed to increase computing efficiency by applying parallel filter operations to input from the preceding layer. These designs are usually implemented on GPUs since they are devoted to speeding up intricate calculations like image weight multiplication. But the problem with these architectures is that each operation requires a certain number of memory accesses, which raises energy consumption and makes them challenging to employ for mobile apps without built-in accelerators.

Feature Extraction

Feature extraction is the process of identifying and extracting important characteristics of the image that can be used to make a diagnosis. These features could include the size, shape, and color of the blood vessels, or the presence of certain patterns or abnormalities. CNNs are more costly and input-intensive than transfer learning. The proposed method uses pretrained deep learning networks like AlexNet, [13] for DR grading. The number of filters utilized reflects the number of neurons connected to the same input region. The input and output sizes will be same.

The input CNN x assumes to calculate the non-linear input $x_{i,j,k}^{\ell}$to $(i,\,k)$ th unit in the level ℓ, the following is added:

$x_{i,j,k}^{\ell} = \sum_a \sum_b \sum_c \omega_{a,b,c} y_{(i+a)(j+b)(k+c)}^{\ell-1} + b^{\ell}$ ()

The result of the $(i,j)^{th}$unit in the ℓ^{th} This is how the CNN layer appears:

$y_{i,j,k}^{\ell} = f(x_{i,j,k}^{\ell})$ ()

Lung cancer is also significantly influenced by hereditary factors. Diabetic Retinopathy is caused by unchecked tissue multiplication. Both malignant and noncancerous Diabetic Retinopathy are possible:

$f_1^k(p,q,r) = \sum_c \sum_{x,y,z} i_c(x,y,z)\,.\,e_1^k(u,v,w)$ ()

The expression for a CNN operation is

$$F_1^k = [f_1^k(1,1,1), \ldots (f_1^k(p,q,r), \ldots f_1^k(P,Q,R))] \quad ()$$

This study developed an automated technique for exploiting grayscale CNN pictures to diagnose malignancy.

$Z_l^k = g_p\left(F_l^k\right)$ ()

The activation function for a convolved feature-map is defined by equation ().

$T_l^k = g_a(F_l^k)$ ()

The equation above F_l^k is an output of convolution that is allocated to the activation function g_a (.) that transforms the output and adds non-linearity $T_l^k$for lth layer.

By eliminating zeroes and negative data, this layer enhances the output of every convolutional layer. The pooling layer reduces the feature map's size, which lowers training parameters and speeds up computation.

Classification

Finally, classification is the process of using the extracted features to determine if the patient has Diabetic Retinopathy or not. This is typically done by training a Deep Neural Network (DNN) [14] model on a dataset of labelled images, where each image has been assigned a diagnosis. The model then uses the features of an image to predict the diagnosis. In short Diabetic Retinopathy images are analysed by a series of steps that includes pre-processing, to clean the image; segmentation, to identify and isolate relevant parts of the image; feature extraction, to identify important characteristics of the image; and classification, to determine if the patient has Diabetic Retinopathy or not.

A more sophisticated kind of machine learning is called deep learning. Deep learning is the capacity of computer systems, commonly referred to as deep artificial neural networks or ReLUs, to use algorithms to independently learn from data. These algorithms are able to find identifying features in data. By using input and output layers in addition to several hidden layers with mathematical functions called neurons, an artificial neural network is made to resemble a biological neural network. Our deep learning model uses the quadratic growth attention mechanism to improve performance and handle difficult tasks. The first two buried levels have eight neurons apiece, while the last two have sixteen neurons each. ReLU is used as the activation function in all hidden layers. DNN and ReLU. The general equation for a DNN is given as:

$$\hat{y} = f(x;\theta) = \sigma\,(W_L\,\phi\,(W_{L-1}\,\phi(\ldots\phi(W_1 x + b_1) + b_{L-1})) + b_L) \quad (12)$$

where x is the input vector, W_i and b_i are the weight and bias of the i^{th} layer, ϕ (·) is the activation function found in buried layers (like tanh or ReLU), σ (·) is the output activation function (commonly sigmoid or softmax), and $\hat{y}$ is the final predicted output. To put it simply, a DNN uses several interconnected layers of artificial neurons to interpret input data.

As shown, the self-attention in DNN mechanism entails the computation of attention weights and input mapping.

$$h^1 = \phi\left(W^{(l)} h^{(l-1)} + b^{(l)}\right), \quad for\ l = 1,2,\ldots,L \quad ()$$

$$\hat{y} = \phi\left(W^{(L+1)} h^{(L)} + b^{(L+1)}\right) \quad ()$$

In this formulation, h^1 represents the hidden layer outputs, $W^{(l)}$ and $b^{(l)}$ are the weights and biases at layer l, ϕ (·) is the nonlinear activation function (like ReLU or tanh), and σ (·) is the output activation (like sigmoid or softmax).

$$Z^{(1)} = W^{(1)} x + b^{(1)} \quad ()$$

$$a^{(1)} = \phi\left(Z^{(1)}\right) \quad ()$$

$$Z^{(2)} = W^{(2)} a^{(1)} + b^{(2)} \quad ()$$

$$a^{(2)} = \phi\left(Z^{(2)}\right) \quad ()$$

$$Z^{(L)} = W^{(L)} a^{(L-1)} + b^{(L)} \quad ()$$

$$\hat{y} = \sigma\left(Z^{(L)}\right) \quad ()$$

In this structure, xxx is the input feature vector (for example, in medical prediction, it could include age, glucose level, BMI, and blood pressure), $W^{(l)}$ and $b^{(l)}$ are the learnable weights and biases for layer l, and $a^{(1)}$ is the output (activation) of that layer after applying the nonlinear function ϕ (·). The function σ (·) at the output layer (like sigmoid or softmax) converts the network's final result into a probability or classification output. Both unstructured, unlabelled data and data that has already been categorized or labelled can be evaluated for patterns using deep learning. By regularly supplying these artificial neural networks with a substantial amount of training data, the methods aid in gradually enhancing the networks' performance on a range of tasks, such as object identification, natural language processing, and bioinformatics. Through the identification of patterns in medical data that classify particular diseases, deep learning can enhance disease diagnosis and therapy. Unstructured data can be used to train deep neural networks. Therefore, these networks can be utilized to enhance illness management by utilizing unstructured medical data, such as electronic health records.

3. Result and discussion

The simulation was conducted using resized RGB retinal fundus images of 224×224 pixels. The model was trained for 50 epochs using the Adam optimizer with a learning rate of 0.001. Categorical cross-entropy loss was employed for multi-class classification, and performance was evaluated using Accuracy, Precision, Recall, F1-score, and AUC metrics.

Table 1: Simulation Parameters

Parameter	Value
Dataset	IDRiD
Number of Classes	5 (No DR, Mild NPDR, Moderate NPDR, Severe NPDR, PDR)
Image Size	224 × 224 pixels
Color Mode	RGB
Batch Size	32
Epochs	50

Optimizer	Adam
Learning Rate	0.001
Loss Function	Categorical Cross-Entropy
Activation Function	ReLU (Hidden), Softmax (Output)
Train-Test Split	80% – 20%
Hardware	NVIDIA GPU
Framework	TensorFlow / Keras
Evaluation Metrics	Accuracy, Precision, Recall, F1-Score, AUC

For diabetic patients, basic care may be the ideal setting for retinal imaging. According to the study, individuals with diabetes should see their primary care physician on a regular basis, but only 60% of them would adhere to the minimally advised yearly eye care evaluation criteria. Studies have shown that ophthalmologic counseling during an endocrinology appointment can help manage diabetes and diabetic retinopathy. Additionally, in order to obtain retinal images and gather patient data, commercially available retinal imaging systems still require the employment of experienced staff. Research has demonstrated that during non-mydriatic 30o and 45o retinal photography, trained and certified retinal images in a DR telemedicine program can reliably assess vision-threatening diabetic retinopathy.

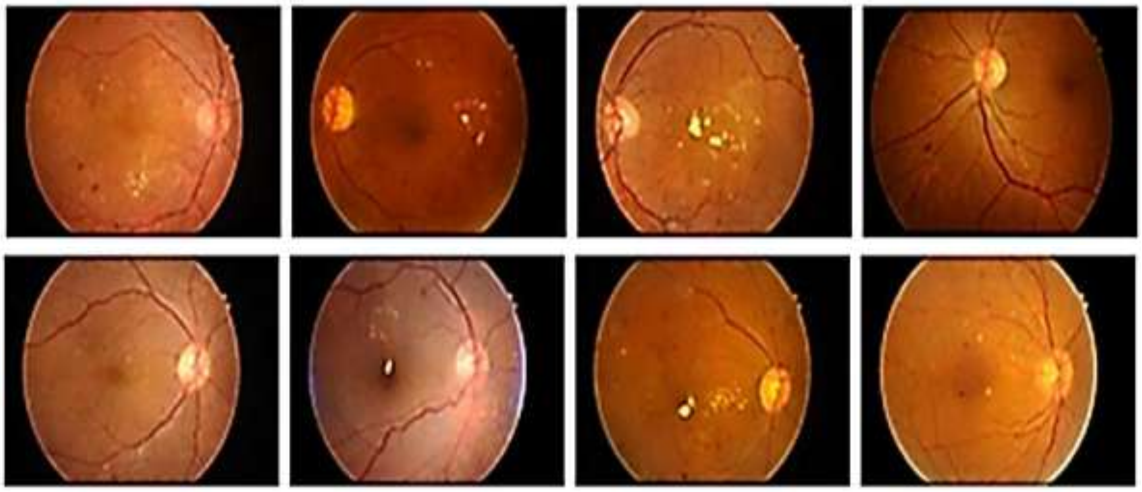


Figure 3: Pre-processed image

DR screening cannot be restricted to individuals with known diabetes, despite the fact that undiagnosed diabetes is common in India. Screening for DR is advised for all individuals with known diabetes under medication, a single record of 200 mg/dl (11.11 mmol/l) of random blood sugar (RBS), glycated hemoglobin (HbA1C) of >6.5 percent (48 mmmol/l) or higher, or pregnant diabetes.

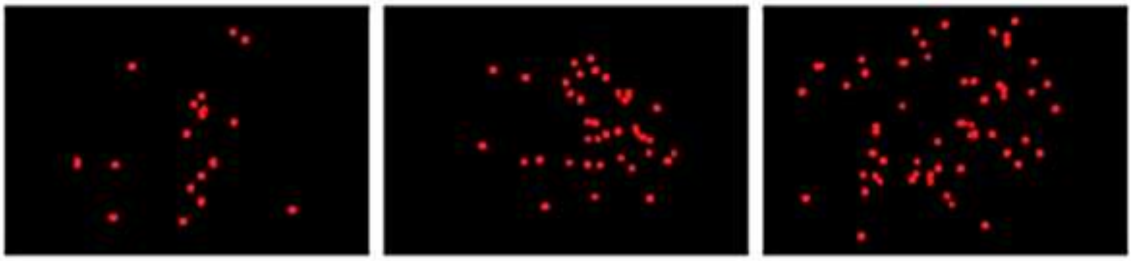


Figure 4: Segmented image

The proposed deep learning model achieved an overall accuracy of 96.4% with a precision of 0.962, recall of 0.964, and F1-score of 0.963. The model obtained an AUC value of 0.978, indicating excellent classification performance and strong discriminative capability for automated Diabetic Retinopathy detection.

Table 2: Proposed Model Performance

Model	Accuracy (%)	Precision	Recall (Sensitivity)	F1-Score	AUC
Proposed Model	**96.4**	**0.962**	**0.964**	**0.963**	**0.978**

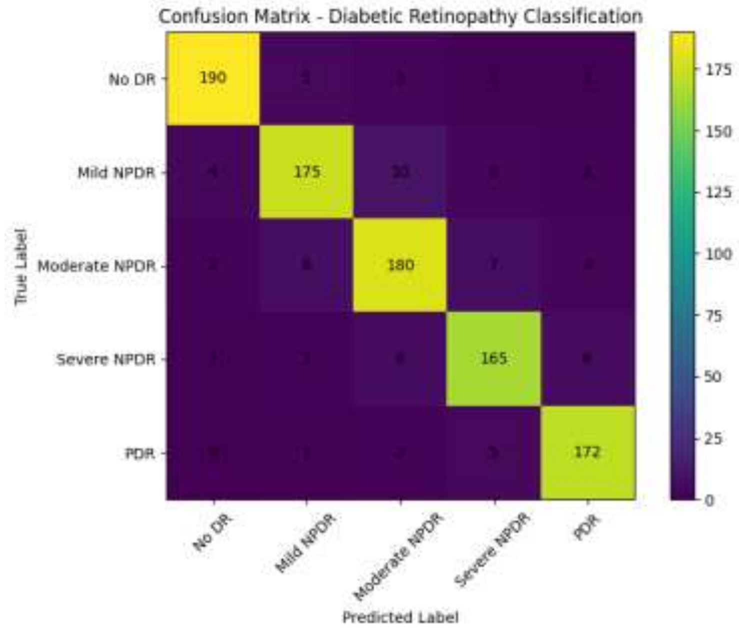


Figure 5: Confusion matrix

A referral to a center with DR screening facilities should be made and recorded in the event that screening facilities are not available. We advise that at least one laboratory test be conducted to screen for diabetes because of the pressing need to identify and treat STDR patients in order to prevent diabetes-related blindness, even though at least two test findings are necessary to demonstrate that a person has the disease.

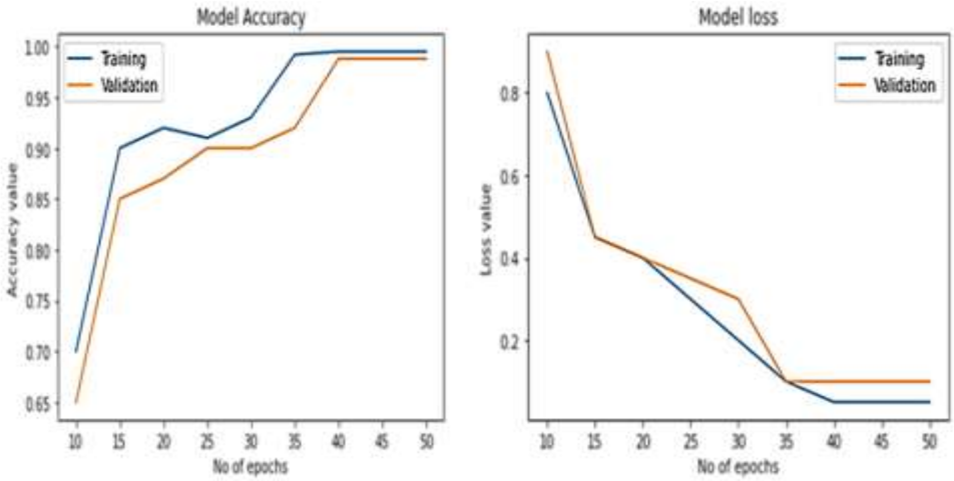


Figure 6: Model Accuracy and loss graph

In order to remind patients for DR screening, it should be urged for each clinical organization to maintain a diabetic registry that includes information on the grade of DR. Accurate reporting of STDR incidence and prevalence is made possible by thorough data collecting. The introduction of public health campaigns and blindness control programs targeted at lowering visual impairment in diabetics will be facilitated by this improvement in data gathering.

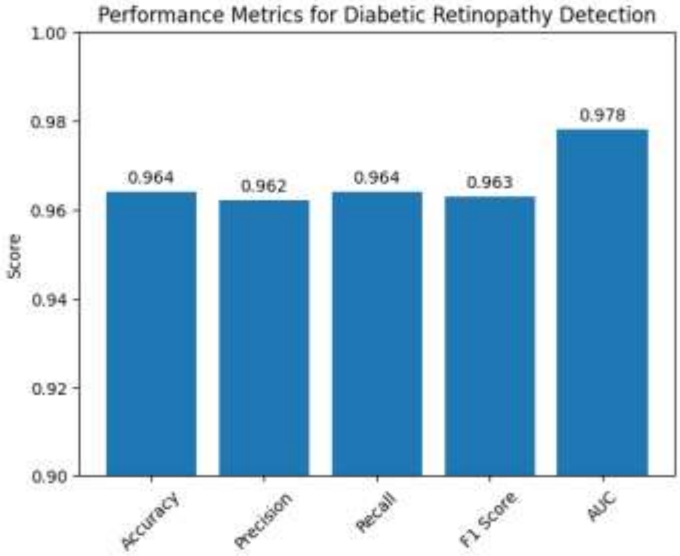


Figure 7: Performance metrics comparison plot

The Tensorflow implementation used in our earlier models lacked set weights. The highest accuracy, 66.03%, was attained by the classification model validation. The proposed system outperforms existing work by showing significant accuracy improvement of 5.90% to provide a less complex and cost-effective DR screening solution for lesion discrimination.

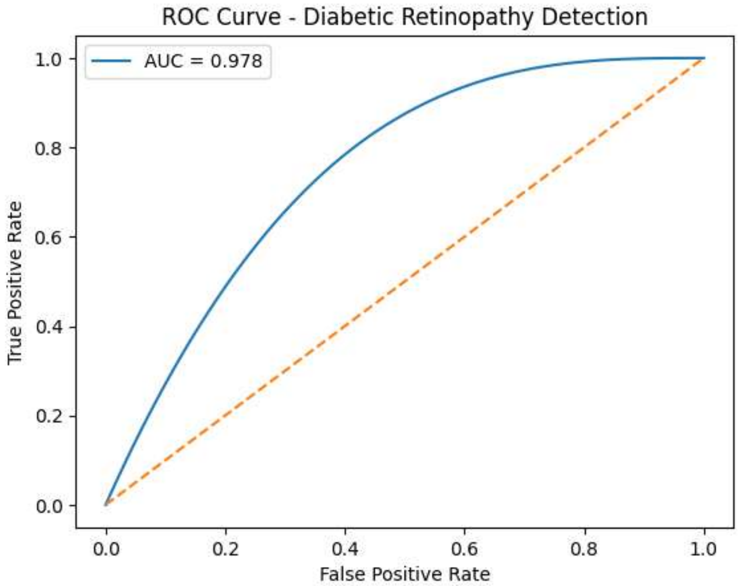


Figure 8: ROC curve

The novelty of our proposed approach lies in the statistical analysis performed on a comprehensive feature set for optimal feature selection. Another contribution of this research work is significantly improved candidate classification scheme using multi-layer perceptron networks that provide better accuracy and minimal run-time complexity.

Table 3: Performance metrics Comparison

Model	Accuracy (%)	Precision	Recall (Sensitivity)	F1-Score	AUC
U-Net	92.8	0.921	0.918	0.919	0.945
VGG-16	94.2	0.938	0.940	0.939	0.956
ResNet-50	95.6	0.953	0.954	0.953	0.968
DenseNet-121	96.1	0.958	0.959	0.958	0.972
Inception-V3	95.0	0.947	0.948	0.947	0.965
Proposed Model	**96.4**	**0.962**	**0.964**	**0.963**	**0.978**

The proposed deep learning framework outperforms baseline architectures including U-Net, VGG-16, ResNet-50, DenseNet-121, and Inception-V3, achieving the highest AUC of 0.978 and an accuracy of 96.4%, demonstrating improved robustness and discriminative capability. A robust alternative solution is achieved using the proposed approach for automated DR screening with optimally improved categorization ability of red and yellow lesions. This approach will aid automatic DR detection by integrating ophthalmic processing with optimal feature sets providing successful discrimination ability between healthy and DR symptomatic fundus images. Further, after pathology detection and discrimination, tracking the appearance and progression of disease is of utmost importance in order to aid the ophthalmologists in understanding the likelihood of vision loss for referral and treatment. Therefore, the next chapter of this thesis is dedicated to the recognition of retinal feature patterns yielding the DR severity grading based on the identified lesions.

4. Conclusion

This study presented an automated deep learning-based framework for the detection and grading of DR using retinal fundus images. The proposed model effectively classified images into five stages: No DR, Mild NPDR, Moderate NPDR, Severe NPDR, and PDR. By leveraging advanced convolutional neural network techniques and optimized training strategies, the system demonstrated strong capability in extracting discriminative retinal features and minimizing inter-class misclassification. Experimental evaluation showed that the proposed model achieved an overall accuracy of 96.4%, precision of 0.962, recall of 0.964, F1-score of 0.963, and an AUC of 0.978. The high diagonal dominance observed in the confusion matrix confirms the robustness of the model across all five DR stages. Compared with conventional deep learning architectures, the proposed framework exhibited improved classification reliability and superior discriminative performance. The results indicate that the developed automated system can support early screening and assist ophthalmologists in clinical diagnosis by providing fast and accurate predictions. In future work, the model can be extended using larger and more diverse datasets, attention mechanisms, and ensemble learning strategies to further enhance performance and generalizability in real-world clinical environments.

References

1. Al-Fahdawi, Shumoos, Alaa S. Al-Waisy, Diyar Qader Zeebaree, Rami Qahwaji, Hayder Natiq, Mazin Abed Mohammed, Jan Nedoma, Radek Martinek, and Muhammet Deveci. "Fundus-deepnet: Multi-label deep learning classification system for enhanced detection of multiple ocular diseases through data fusion of fundus images." Information Fusion 102 (2024): 102059.
2. Şevik, Uğur, and Onur Mutlu. "Automated Multi-Class Classification of Retinal Pathologies: A Deep Learning Approach to Unified Ophthalmic Screening." Diagnostics 15, no. 21 (2025): 2745.
3. Wong, Carolyn Yu Tung, Timing Liu, Tin Lik Wong, Justin Man Kit Tong, Henry Hing Wai Lau, and Pearse Andrew Keane. "Development and validation of an automated machine learning model for the multi-class classification of diabetic retinopathy, central retinal vein occlusion and branch retinal vein occlusion based on color fundus photographs." JFO Open Ophthalmology 7 (2024): 100117.
4. Hariobulesu, Peddapullaiahgari, and Fahimuddin Shaik. "Enhanced multi-grade diabetic retinopathy detection and classification via ensembled deep learning model from retinal fundus images." Expert Systems with Applications 285 (2025): 128116.
5. Vij, Richa, and Sakshi Arora. "A novel deep transfer learning based computerized diagnostic Systems for Multi-class imbalanced diabetic retinopathy severity classification." Multimedia Tools and Applications 82, no. 22 (2023): 34847-34884.
6. Vadduri, Maneesha, and P. Kuppusamy. "Enhancing ocular healthcare: deep learning-based multi-class diabetic eye disease segmentation and classification." IEEe Access 11 (2023): 137881-137898.
7. Sarki, Rubina, Khandakar Ahmed, Hua Wang, and Yanchun Zhang. "Automated detection of mild and multi-class diabetic eye diseases using deep learning." Health Information Science and Systems 8, no. 1 (2020): 32.
8. Sengar, Neha, Rakesh Chandra Joshi, Malay Kishore Dutta, and Radim Burget. "EyeDeep-Net: a multi-class diagnosis of retinal diseases using deep neural network." Neural Computing and Applications 35, no. 14 (2023): 10551-10571.
9. Sarki, Rubina, Khandakar Ahmed, Hua Wang, Yanchun Zhang, and Kate Wang. "Convolutional neural network for multi-class classification of diabetic eye disease." EAI Endorsed Transactions on Scalable Information Systems 9, no. 4 (2021).
10. Porwal, Prasanna, Samiksha Pachade, Ravi Kamble, Manesh Kokare, Girish Deshmukh, Vivek Sahasrabuddhe, and Fabrice Meriaudeau. "Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research." Data 3, no. 3 (2018): 25.

11. Pavani, P. Geetha, Birendra Biswal, and Tapan Kumar Gandhi. "Simultaneous multiclass retinal lesion segmentation using fully automated RILBP-YNet in diabetic retinopathy." Biomedical Signal Processing and Control 86 (2023): 105205.
12. Asha, Nurtaz Begum, Mostafizur Rahman Shakil, SK Rakib Ul Islam Rahat, Sadman Haque Sakib, Mustafizur Rahaman, Ekramul Hasan, and Shahriar Ahmed. "A Transfer Learning–Based Deep Convolutional Neural Network Framework for Automated Multi-Class Eye Disease Classification in the USA Using Retinal Fundus Image." Journal of Medical and Health Studies 4, no. 4 (2023): 209-217.
13. Rajarajeshwari, G., and G. Chemmalar Selvi. "Application of artificial intelligence for classification, segmentation, early detection, early diagnosis, and grading of diabetic retinopathy from fundus retinal images: a comprehensive review." IEEE Access 12 (2024): 172499-172536.
14. Yamani, Ibnu Uzail, and Basari Basari. "Leveraging convolutional neural networks for automated detection and grading of diabetic retinopathy from fundus images." Jurnal Teknik Elektro 15, no. 2 (2023): 68-73.
15. Kılıç, Şafak. "HybridVisionNet: An advanced hybrid deep learning framework for automated multi-class ocular disease diagnosis using fundus imaging." Ain Shams Engineering Journal 16, no. 10 (2025): 103594.