Kelvin Open Science Publishers
Connect with Research Community

Research Article ***Volume 1 | Issue 2***

KOS Journal of Science and Engineering
https://kelvinpublishers.com/journals/science-and-engineering.php

Deep CNN models for Accurate Recognition of DR and Glaucoma from Fundus Retinal Images

Dr. BalKrishna Sharma*

Department of Computer Sciences and Applications, Mandsaur University, Mandsaur, India

***Corresponding author:** Dr. Bal Krishna Sharma, Department of Computer Sciences and Applications, Mandsaur University, Mandsaur, India, E-mail: bksharma7426@gmail.com

Received: July 02, 2025; **Accepted:** July 08, 2025; **Published:** July 10, 2025

Citation: BalKrishna S. (2025) Deep CNN models for Accurate Recognition of DR and Glaucoma from Fundus Retinal Images. *KOS J Sci and Eng.* 1(2): 1-6.

1. Abstract

Globally, diabetic retinopathy (DR) and glaucoma (GL) are major contributors to permanent vision loss hence, it is crucial to detect these conditions early and accurately in order to inhibit their progression. It is possible to misclassify fundus photos taken with varying field of view and resolution settings because key characteristics are lost. Specifically, this research suggests a multi-scale assisted DL model for precise DR and GL categorization. This research demonstrates automated diabetic retinopathy identification using a proprietary Convolutional Neural Network (CNN) model trained on the publicly accessible Messidor dataset. A foundation of deep learning supports the network. The proposed CNN architecture includes many convolutional layers with ReLU activation, max pooling, batch normalization, and dropout before a sigmoid-based output element for binary classification. Impressive performance is demonstrated by the experimental results, which include an F1-score of 99.42%, a recall of 99.90%, a precision of 98.86%, and an accuracy of 99.37%. To top it all off, the model's 99.98 AUC was far higher than that of ResNet50, KNN, and InceptionV3, which are considered benchmark models. The results show that the proposed CNN is strong and can be used in clinical settings to screen for diabetic retinopathy automatically and detect it early.

2. Keywords

Diabetic Retinopathy (DR), Glaucoma, Retinal Fundus Images, MESSIDOR dataset, Deep learning, Convolutional Neural Network (CNN)).

3. Introduction

Healthcare systems worldwide are increasingly challenged by the rising burden of chronic diseases, with diabetes mellitus standing out as a critical global health issue [1,2]. As the prevalence of diabetes continues to surge, so too does the incidence of its associated complications, which place considerable strain on healthcare infrastructure and patient quality of life [3]. Among these problems, ocular diseases are a major source of vision loss; the two most prevalent and serious eye disorders associated with diabetes are glaucoma and diabetic retinopathy (DR) [4]. A microvascular problem of diabetes, diabetic retinopathy, if left untreated, can gradually harm the retina, leading to gradual vision loss or perhaps blindness [5]. Mild non-proliferative abnormalities are the first stage, while severe proliferative retinopathy with macular oedema and retinal haemorrhage is the last stage [56,7]. In a similar vein, glaucoma, which is a cluster of ocular neuropathies defined by optic nerve injury, can coexist with DR and further endanger eyesight, especially as it frequently does not present symptoms until irreversible damage has already taken place [8]. The coexistence of these conditions, especially in diabetic patients, amplifies the complexity and urgency of timely and accurate diagnosis [9].

Retinal fundus imaging is very important for finding and

classifying diabetic retinopathy and glaucoma correctly, which is necessary for early treatment and keeping your sight [10,11]. To diagnose ocular diseases, fundus pictures are essential because they show the retinal components in great detail, including the macula, optic disc, and blood vessels. Particularly in large-scale screening programs, manual image evaluation is laborious, subjective, and prone to inconsistency owing to inter-observer variability [12,13]. The creation of automated diagnostic frameworks that are effective and dependable is in high demand due to these restrictions. DL and ML have recently [14,15] risen to prominence as a result of these difficulties, thanks to their capacity to accurately complete complicated picture classification tasks. An outstanding skill of CNN, a subtype of DL models, has been shown to automatically extract and learn hierarchical visual features from retinal pictures. This allows for reliable categorization of DR and Glaucoma, even under changeable imaging conditions [16-18]. Such AI-powered approaches provide scalable answers for mass screening while decreasing reliance on human experts' interpretations [19,20]. This study explores the application of a deep CNN technique for the simultaneous classification of retinal fundus pictures for glaucoma and diabetic retinopathy. With the aim of improving early detection and helping ophthalmologists provide data-driven eye care in a timely manner, the suggested approach tackles important problems such feature heterogeneity, class imbalance, and diagnostic accuracy.

3.1. Motivation with contribution

DR and glaucoma are leading causes of permanent blindness, and their prevalence has increased in tandem with the worldwide prevalence of diabetes. While it is crucial to discover these problems early on, the existing manual screening procedures are labour-intensive, resource-intensive, and vulnerable to human subjectivity, which can lead to a lack of consistency. Retinal fundus scans can show modest and overlapping signs of DR and glaucoma, making definitive diagnosis even more challenging. Because of this, scalable and automated diagnostic solutions are critically needed. This study aims to construct a deep CNN-based framework that can reliably classify both DR and Glaucoma from retinal pictures. The motivation for this effort comes from the limits of standard screening procedures as well as the potential of AI in medical imaging. In the long run, this technique can help minimize diabetic avoidable blindness by improving early diagnosis, decreasing diagnostic errors, and bolstering large-scale screening programs. The main contribution focuses on developing robust diabetic retinopathy and glaucoma detection systems for automated retinal screening applications:

- Modern pre-processing methods were used to increase the consistency and quality of the images in the Messi or dataset, making it more suitable for diabetic retinopathy identification.
- Performed EDA on the Messi dataset to visualize diabetic retinopathy severity levels, uncover class imbalance, and extract insights to inform pre-processing and model training decisions.
- Rotated, flipped, zoomed, and adjusted brightness and contrast among other data augmentation procedures to correct data imbalance.
- Creation of a new CNN architecture tailored for the detection of ocular diseases.
- Creating a comprehensive evaluation approach that includes recall, accuracy, precision, and F1-score four industry-standard categorization metrics to ensure robust model validation and practical application in clinical situations.

3.2. Justification and novelity

The proposed approach uses a custom CNN to ensure domain-specific feature extraction tailored for diabetic retinopathy detection, as opposed to using pre-trained models like ResNet50 or InceptionV3. Pre-trained models are designed for general image classification tasks and may not effectively capture subtle pathological patterns in fundus images. In contrast, the custom CNN is lightweight, built from scratch, and optimized through a specialized preprocessing pipeline, including grayscale conversion, Gaussian denoising, and histogram equalization, which significantly enhances retinal feature visibility. Targeted data augmentation further improves the model's sensitivity and generalization by addressing class imbalance. Because of this, the model performs better on all evaluation criteria, especially recall, showing that it can detect DR cases more correctly with fewer false negatives, which is crucial for clinical applications.

3.3. Structure of paper

The paper's outline is as follows: Results for diabetic retinopathy detection from the literature review are presented in Section II. In Section III, we explore the proposed CNN method in further detail. The experimental data and model comparisons are examined in Section IV. Lastly, this study concludes and recommends.

4. Literature Review

This section discusses the literature review on AI for accurate and efficient diabetic retinopathy and glaucoma detection. Table I provides a summary of the literature reviews discussed below:

Guzal Kangilbaeva (2025) Automated diabetic retinopathy diagnosis using DL algorithms in AI systems has progressed beyond visualizing and automatically segmenting the ocular fundus imaging DR dataset to diagnosing phases of diabetic retinopathy. In the medical sector, ophthalmology, screening for diabetic retinopathy, and the use of ML, AI, DL, and neural networks, have been able to diagnose this potentially blinding condition. AI programs have demonstrated sensitivity ranging from 82-99.1% AI programs undoubtedly help doctors diagnose vision-threatening diabetic retinopathy densely populated, low-income areas [21].

Karambelkar, et al. (2024) DL algorithms have dramatically improved the ability to screen, identify, segment, forecast, and categorize in numerous medical domains, such as those pertaining to the retina, heart, pathology, and abdominal regions. Among the most important causes of blindness in people of working age is diabetic retinopathy. For a good prognosis, early diagnosis of this illness is essential Researchers show how to recognize diabetic retinopathy staging using color fundus pictures and CNN [22].

Afshan, Chakraverti and Chhabra (2024) used DL methods, a standardization methodology was used to a dataset of 5,000 retinal photos to guarantee that each image is of the same size at 256×256 pixels, in order to automate the detection of glaucoma. To make datasets more diverse and models more

resilient, data augmentation methods including rescaling, rotation, and vertical and horizontal shifts are often employed. When training a model to identify glaucoma, the MobileNetV3 architecture makes use of two optimizers: Stochastic Gradient Descent (SGD) and Adam. Thanks to its high classification accuracy, the model is able to detect glaucoma effectively. This could lead to better automated glaucoma detection and, eventually, earlier diagnosis and treatment [23].

Hossain et al. (2024) Diagnose Diabetic Retinopathy from a patient's retinal imaging and determine if it is in its early, moderate, or advanced stages. Diabetic Retinopathy, Aptot2019 _blindness detection, and Messidor2 are the datasets utilized in the study. The research shows the comparison between CNN models, including Exception, Alex Net, VggNet-16,The ResNet-50 shows better performance than other models [24].

Kolte, et al. (2023) DL model, which can improve the model's accuracy while decreasing the amount of data needed for training. With a rate of 52.75%, SVM was the most accurate of the ML algorithms that were evaluated. Efficient-Net came out on top when compared to other DL approaches, boasting an impressive AUC of 0.91 and an accuracy rate of 91.8%. Efficient-Net and other DL algorithms excel at detecting diabetic retinopathy [25].

Butt, et al. (2022) Diabetic eye disease, also known as retinopathy, is a condition that gradually worsens in diabetics over time. Potential harm to the eyes could arise. In type 1 diabetes mellitus, elevated blood sugar levels do not go down. By integrating features retrieved from pre-trained CNN models using TL, a hybrid feature vector may be created, which improves the performance of DR detection for fundus images. The planned enhanced method achieved high levels of accuracy in both binary and multiclass classification, reaching 97.8% and 89.29%, respectively [26].

Mohamed, Elmohsen and Basha, 2021, Evaluate the DR characteristics' effects on each color channel, priorities the channels, and calculate their main components (PCA). Next, these channels are fed into the DL model, and the grading decision is made by applying a majority voting system to the model's output. Both local dataset (consisting of about 100 photos) and a publicly available dataset (containing about 80,000 color fundus images) were used to train and evaluate the proposed models. An increase of 85% in accuracy, 89% in sensitivity, and 96% in specificity was observed in DR multi-class classification results [27].

Existing studies show that DL and transfer learning methods can help find diabetic retinopathy and glaucoma, but there are still some areas that need more research. Many models depend on large, well-labelled datasets that aren't always very big or varied, which makes it hard to use these models in real-life clinical settings. In addition, the way things are done now often only uses one type of fundus picture and doesn't include multi-modal imaging data like OCT or visual field tests, which could help with accuracy in diagnosis. Furthermore, most studies put classification performance above interpretability, which is still a major obstacle to clinical use. Not much research has also been done on lightweight, resource-efficient designs that can be used in places with few resources, which is exactly where early diagnosis is most needed. Filling in these gaps could lead to AI-driven diagnostic systems that are more reliable, easy to understand, and flexible.

Table 1: Comparative analysis of recent studies on diabetic retinopathy and glaucoma detection using machine learning.

Author	Dataset	Methodology	Key Findings	Advantages	Limitations	Future Work
Guzal Kangilbaeva [21]	Public DR fundus dataset	AI-based DR diagnosis using visualization & segmentation of fundus images with DL	AI programs show 82–99.1% sensitivity for vision-threatening DR diagnosis in low-income areas	High sensitivity; enables screening in underserved regions	Dataset details not provided; general discussion of AI programs	Validate on large, diverse datasets; integrate with clinical workflow
Karambelkar, et al. [22]	Color fundus images	CNN-based staging of DR using deep learning on fundus images	CNN effectively detects and stages DR, aiding early diagnosis	Supports early detection of DR in working-age adults	Specific accuracy/metrics not reported	Expand to multiclass severity grading and real-world testing
Afshan, Chakraverti & Chhabra, [23]	5000 retinal images	MobileNetV3 with preprocessing (resizing, greyscale, Gaussian noise), data augmentation	MobileNetV3 model achieves excellent classification accuracy for glaucoma detection	Standardized preprocessing improves robustness; efficient architecture	Only glaucoma addressed; no multi-disease detection	Extend to multi-disease diagnosis and real-time deployment
Hossain, et al. [24]	Diabetic Retinopathy, APTOS 2019, Messidor2	Comparison of CNN architectures: Xception, AlexNet, VGG-16,	ResNet-50 outperformed others in DR stage detection	Comprehensive comparison of CNNs; validated on multiple datasets	Focused on CNNs only; no hybrid/ensemble models explored	Explore hybrid or ensemble approaches for further improvement

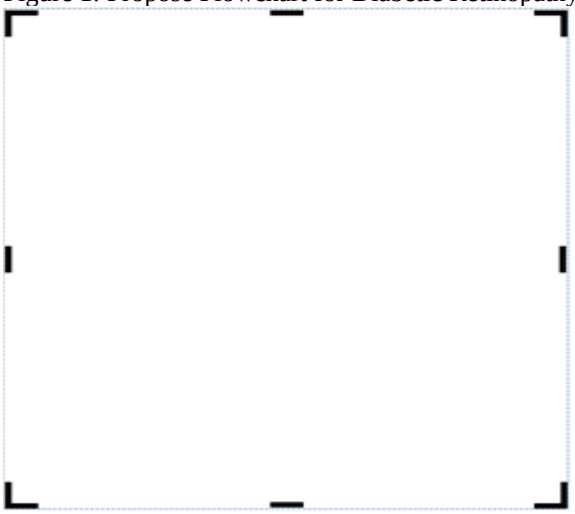
		ResNet-50				
Kolte, et al. [25]	Kaggle DR dataset	Image resizing + Haralick & Hu features + ML (SVM) vs. EfficientNet DL	EfficientNet achieved highest accuracy (91.8%) and AUC (0.91), outperforming SVM	Deep learning superior to classical ML; EfficientNet highly effective	SVM and classical ML underperform significantly	Investigate lightweight models and interpretability of EfficientNet
Butt, et al. [26]	Fundus images	Transfer learning on pre-trained CNNs with hybrid feature vector	Achieved 97.8% accuracy (binary) and 89.29% (multiclass) in DR detection	The hybrid TL model improves detection performance significantly	Dataset size and diversity not mentioned	Test on larger datasets; optimize for multiclass DR detection
Mohamed, Elmohsen & Basha, [27]	Public dataset (~80K images) + local test set (~100 images)	PCA-based channel selection + deep learning + majority voting	Increased DR multi-class accuracy: 85%, sensitivity: 89%, specificity: 96%	Channel-wise analysis boosts multi-class performance	Limited testing on local dataset (small size)	Expand evaluation on larger, diverse test sets

.

5. Methodology

This study's research technique is based on a complete pipeline that starts with using the publicly available Messidor dataset. Image cropping, greyscale to black and white conversion with thresholding, denoising with Gaussian blurring, and contrast enhancement with histogram equalization in YUV color space are all part of the initial stage of preprocessing. Figure 1 illustrates the Flowchart for diabetic retinopathy illustrates this process. The images are shrunk to a size of 200×200×3 pixels. Data augmentation approaches to fix class imbalance include rotating, flipping, zooming, and adjusting brightness and contrast. Training, validation, and test sets make up the dataset in a proportion of 70:10:20. CNNs are trained to classify data into two categories using a series of operations including batch normalization, max pooling, flattening, dropout, and ReLU activation. Recall, accuracy, precision, F1-score, ROC-AUC, and confusion matrix analysis are some of the approaches used to evaluate the model's performance.

Figure 1: Propose Flowchart for Diabetic Retinopathy.

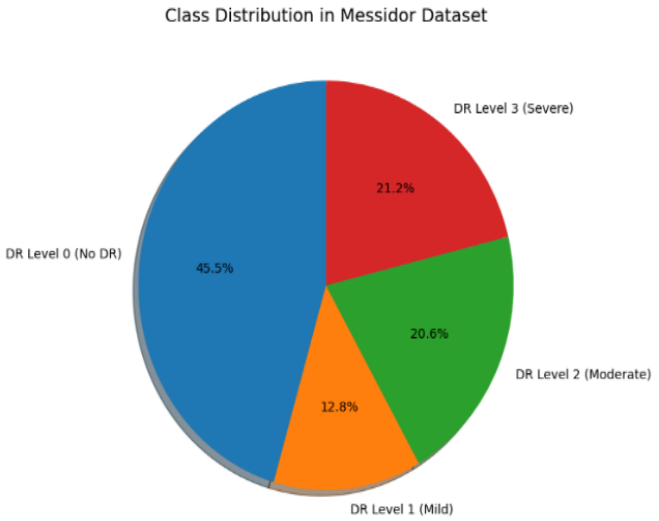


The detail explanation of the propose methodology steps shown in the flowchart is provided in the next section.

5.1. Data collection

At ADCIS39, you can find the Messidor benchmark dataset, which is open to the public. The skilled ophthalmologists have painstakingly annotated around 1200 color fundus photos. Pictures taken by three different ophthalmology clinics make up the dataset. In all, about 800 pictures were captured with and 400 without pupil dilatation. They use four categories to identify these photographs based on the severity of the retinopathy. As seen in **Figure 2**, the image data is divided into four categories.

Figure 2: Pie Chart for Distributions of Classes.



Retinopathy severity levels in diabetic eye disease are shown in **Figure 2**, which shows the class distribution of the Messidor dataset. At 45.5%, DR Level 0 (No DR) is the most common, followed by DR Levels 2, 3, and 1 with

corresponding shares of 21.2%, 20.6%, and 12.8%. There is a clear class imbalance in the sample across severity ratings, as seen by the distribution.

Figure 3: Enhancement of Sample Images from Messidor Dataset.

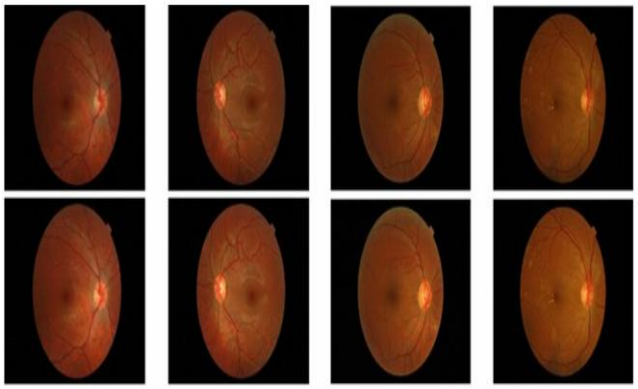


Figure 3 displays the enhanced fundus pictures obtained from the Messidor dataset. There are two columns of four photos each depicting the retina in the figure. DR research makes use of images that demonstrate the enhancement or preprocessing of ophthalmological data to better expose retinal characteristics such the macula, optic disc, and blood vessels.

5.2. Image preprocessing

Improving the quality to enable accurate and dependable model predictions is the primary goal of data preprocessing, an essential step in guaranteeing the success of DL procedures. The goal is to fix problems including data augmentation, noise, greyscale conversion, and picture cropping. Images are preprocessed in this research using a number of procedures that are detailed below.

Image cropping: The dark pixels surrounding the retina were clipped out of the images. Because of this, the annotation files no longer reflect the original lesion locations with bounding boxes [28]. There is an unflattering black backdrop in all of the photographs in the collection: A dark background is typically used to depict the fundus in images. The blood vessels, optic nerve, and retina are all located in this region. The Bounding Box Method for Cropping.

Gray scale conversion: Greyscale was applied to the images. Then, cv2.threshold() was used to apply the threshold, which ranged from 10 to 255. This process creates a binary image by separating the image's black backdrop from the remainder of the picture.

5.3. Image denoising

Noise reduces the clarity and interpretability of images by obfuscating details and distorting features, and can enhance the image's visual quality by eliminating noise. Using cv2.GaussianBlur() to apply a Gaussian blur to dataset in order to eliminate noise. As shown in Equation (1), a popular technique for image smoothing is the Gaussian blur, which uses a weighted average of the values of nearby pixels to replace each pixel's value.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{-x^2+y^2}{2\sigma^2}} \quad (1)$$

The standard deviation is denoted by σ, the distances from the centre of the kernel are denoted by x and y, and the Gaussian kernel is represented by G(x, y). The amount of image smoothing is affected by the size of the Gaussian kernel. From 1 × 1 to 7 × 7, tried out different kernel sizes. The loss of important features in the retinal images was substantial when using larger kernels, despite the fact that they provided more smoothing. An efficient method for balancing noise reduction was the 3 × 3 kernel.

5.4. Image enhancement

The images are given Histogram Equalisation (HE). To improve the contrast of a picture, histogram equalisation moves the intensities of pixels from the RGB colour space to the YUV colour space using cv2.cvtColor() [29]. Histogram Equalisation is a way to change the way the intensity values are spread out in a picture so that they are spread out evenly in the output image. The method for equalising a histogram is shown below in Equation (2):

$$C(k) = \sum_{i=0}^{k} H(i)$$

This is how you find the equalised intensity I′(x, y) of a pixel with intensity I(x, y), the Equation (3) defined as below:

$$I'(x,y) = \frac{(C(I(x,y)) - C_{min}}{M \times N - C_{min}} \times 255 \quad (3)$$

In this equation, M × N is the total number of pixels in the picture, and C_{min}Is the cumulative histogram's lowest nonzero value.

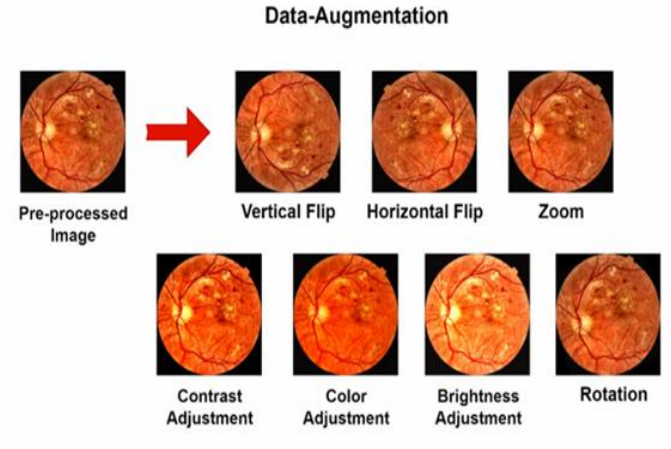
5.5. Image resizing

All the images are adjusted so they are the same size, which is 200 × 200 × 3 pixels. Here, pixel dimensions are given by the first two numbers, and the presence of the RGB (red, green, and blue) colour channels is indicated by the third number. Images are resized to a consistent resolution to guarantee input data format uniformity.

5.6. Data augmentation

Fewer samples of severe DR compared to normal or mild cases might be an issue with retinal imaging datasets like MESSIDOR, which are frequently small and may have an imbalanced class distribution. The training dataset is artificially expanded using data augmentation techniques (as shown in **Figure 4**). These techniques generate plausible modifications of existing photos without changing their diagnostic value.

Figure 4: Sample Image of Data Augmentation Techniques.



A sample of each augmentation approach is shown in Figure 4. Pre-processed fundus pictures of the retina used to train DL algorithms to identify diabetic retinopathy. Starting from a single pre-processed image, multiple augmented versions are generated using geometric and photometric transformations. These include vertical and horizontal flips to introduce orientation variability, zoom to simulate scale changes, and rotation to make the model invariant to angular differences. Additionally, image quality is varied through contrast adjustment, color adjustment, and brightness adjustment to reflect different imaging conditions.

5.6. Data splitting

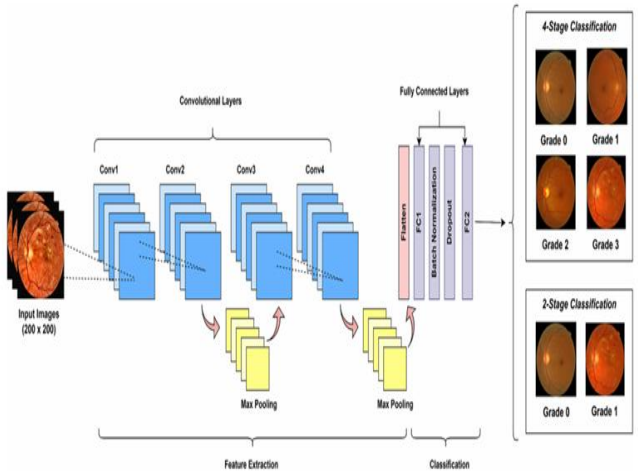
The augmented dataset consists of three parts: training, validation, and testing. Seventy percent of the dataset is used for training, ten percent for validation, and twenty percent for testing in this study. You use the training set to train your model, and the testing set to see how well it performs.

5.7. Proposed CNN model

One type of DL model commonly used for visual data analysis is a CNN. CNN uses a series of convolutional layers to train automatically and extract valuable features from incoming data [30,31]. The four convolutional layers consist of two max-pooling layers, one flattening layer, one fully-connected layer, one batch-normalization layer, and one dropout layer. **Figure 5** shows that the CNN has been partitioned into three equal parts.

The initial block is made up of two convolutional layers, with a 3 × 3 kernel size and 32 filters in each layer.

Figure 5: CNN architecture in DR Image.



This layer's job is to automatically activate ReLU to extract features from input retinal pictures [32]. It has a 64-filter convolutional layer and a 128-filter convolutional layer. Max pooling is required following two convolutional layers, each of which has 32 filters and ReLU activations, according to Equations (4-6).

$$F_1 = ReLU(W_1 * X + b_1) \quad (4)$$

$$F_2 = ReLU(W_2 * F_1 + b_1) \quad (5)$$

$$P_1 = MaxPool(F_2) \quad (6)$$

Where $*$ stands for convolution, W_1 and W_2 are 3x33 convolution kernels with 32 filters each, and P_1 is the feature map that was pooled. Used a smaller filter value in the previous convolutional layers of Block 1 to find low-level features in the input pictures, such as edges, textures, and basic shapes. Used bigger filters (64, 128) to pick up more general features like patterns, structures, or items as went deeper into the CNN network. There are two lower convolutional layers with ReLU activations and 64 and 128 filters. Then there is max pooling in Equations (7)-(9).

$$F_3 = ReLU(W_3 * p_1 + b_3) \quad (7)$$

$$F_4 = ReLU(W_4 * F_3 + b_4) \quad (8)$$

$$P_2 = MaxPool(F_4) \quad (9)$$

W_3 contains 64 filters, while W_4 Contains 128 filters. In order to make the feature maps one-dimensional arrays, the Flatten layer converts them. One must first "flatten" the data before it can be input into fully connected layers. They achieve this by converting the 2D feature maps generated by the convolutional layers into a 1D array. In order to transform the combined feature maps into a one-dimensional vector, they employ Equation (10).

$$Z = Flatten(P_2) \quad (10)$$

This transformation is crucial because in order for fully linked layers to do advanced reasoning and classification, they require a flattened input. The sigmoid activation function was utilised for 2-stage categorisation. A dense layer implemented subsequent to batch normalisation and dropout regularisation in Equation 11.

$$\hat{h} = Dropout\left(BatchNorm\left(W_f z + b_f\right)\right) \quad (11)$$

where areW_f weights and biases$b_f$of the fully connected layer in Equation (12).

$$\hat{y} = \sigma\left(Dropout\left(BatchNorm\left(w_f \cdot Flatten(P_2) + b_f\right)\right)\right) \quad (12)$$

where $\hat{y}$Is final output in Equation (13), P_2 results from two blocks of convolution and $\hat{y}$ This can be used to determine the chance of each input being either the presence or absence of retinopathy, and returns a probability between 0 and 1 for each input.

5.8. Performance matrix

Verifying a classification model's precision is possible with the use of a confusion matrix, a common table in ML. This matrix summarizes the expected and actual class labels for a particular set of test data [33]. The confusion matrix is available at this location In this context, TP, FP, TN, and FN represent the numbers of correct predictions made by the negative class, incorrect predictions made by the negative class, false positives made by the positive class, and true negatives made by the negative class, in that order. A confusion matrix can be used to compute several assessment metrics. This includes F1 score, recall, accuracy, and precision.

5.8.1. Accuracy: This metric counts how many pixels in the dataset depict blood vessels in relation to how many pixels in the segmented image that were correctly allocated **[34]**. Then, in Equation (13).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad (13)$$

5.8.2. Precision: The ratio of correctly predicted positive

instances to total positive examples is one way to describe a model's accuracy rate. Equation (14) describes the model's capacity to prevent false positives:

$$Precision = \frac{TP}{TP+FR} \times 100 \quad (14)$$

5.8.3. **Recall:** The recall, sensitivity, or true positive rate is a measure of how many out of all positive cases were accurately predicted. Equation (15) describes the model's capacity to prevent false negatives:

$$Recall = \frac{TP}{TP+FN} \times 100 \quad (15)$$

5.8.4. **F1 Score:** The F-score quantifies the precision of a test. A dataset is considered imbalanced if its positive outcomes are less than its true positive findings, as shown in Equation (16):

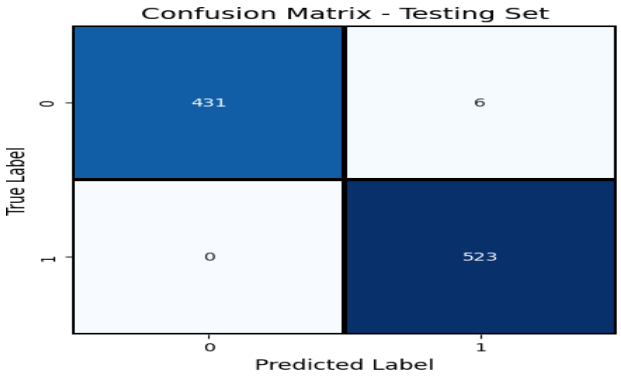
$$F1 - score = \frac{2 \times recall \times precision}{recall + precision} \quad (16)$$

5.8.5. **ROC Curve:** The true-positive rate and false-positive rate are used to plot the ROC curve, which is a graphical representation of the classifier's performance. A higher value for analysis of variance (AUC) [35] indicates a better classification rate; it is a representation of the performance model.

6. Results and Discussion

Using ML algorithms on the MESSIDOR dataset, this part shows the experimental results for diabetic retinopathy identification. The primary metrics employed to evaluate the proposed model for binary classification tasks include accuracy, precision, recall, and F1-score. Using matplotlib and other necessary Python libraries for medical image processing and analysis, the implementation was carried out in a Jupyter Notebook environment on Google Colab. The studies were carried out using a workstation that had the necessary hardware to handle the computational needs of DL models on high-resolution retinal images. It came with 64 GB RAM, an NVIDIA RTX 3090 GPU, and an Intel Core i9 CPU. To back up the efficacy of the suggested CNN method for automated screening and diagnosis in clinical ophthalmology settings, the following outputs give comprehensive insights into the findings of DR identification.

Figure 6: Confusion Matrix of CNN for Binary Classification.

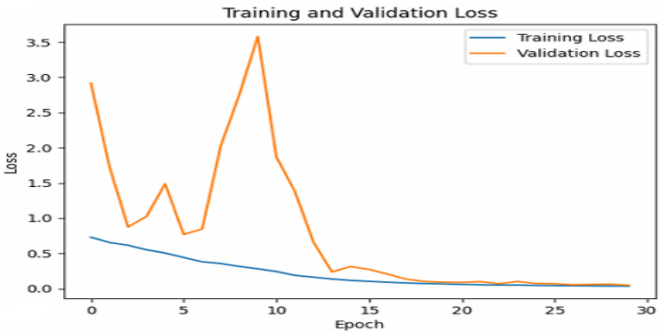


This testing set's binary diabetic retinopathy categorization using CNN is shown in Figure 6, which also contains the confusion matrix for that set. The matrix shows that the model can distinguish between cases of diabetic retinopathy and glaucoma; it has 431 class 0 true negatives, 523 class 1 true positives, 6 erroneous positives, and 0 false negatives. The results show that the categorization accuracy is high.

Figure 7: Accuracy Graph for Training and Validation of CNN.

The accuracy curves for training and validation of CNN-based binary classification are displayed in Figure 7. Over the course of 30 epochs, the blue line represents the training accuracy, which rose from 0.58 to almost 1.0. After some early fluctuations including a spike around epoch 5 the orange validation accuracy line settles at 1.0 after epoch 15, indicating that the model has successfully converged.

Figure 8: Loss for Training and Validation using CNN Model.



The CNN model's training and validation loss curves over 30 epochs are shown in Figure 8. The training loss is going down, which means the model is learning well. Validation loss, on the other hand, shows substantial initial fluctuation before levelling off following a steep decline beginning around the eighth epoch. This pattern suggests that while the model initially struggled to generalize, it achieved improved stability and convergence in later epochs, ultimately indicating reduced overfitting and better generalization performance.

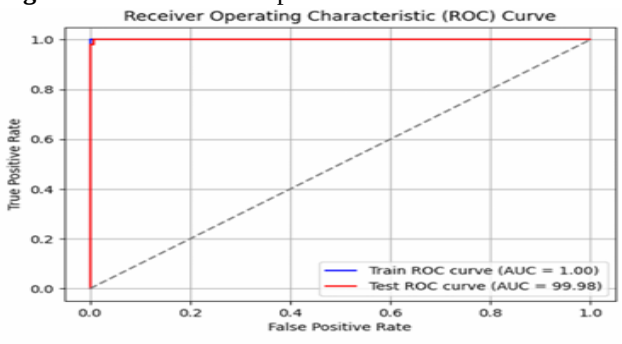
Table 2: Proposed Model Performance for Diabetic Retinopathy on the Messidor Dataset.

Evaluation Measure	Convolutional Neural Network (CNN)
Accuracy	99.37
Precision	98.86
Recall	99.90
F1-score	99.42

Using the MESSIDOR dataset, Table II presents the results of the performance evaluation of the proposed CNN model for diabetic retinopathy classification. The model demonstrates exceptional diagnostic performance with comprehensive metrics, including an accuracy of 99.37%. The high recall value indicates superior sensitivity in detecting diabetic retinopathy cases with minimal false negatives, while the balanced precision ensures low false positive rates. These outstanding results validate the CNN architecture's effectiveness for automated retinal disease

screening, demonstrating its clinical viability for early diabetic retinopathy detection and supporting ophthalmologists in diagnostic decision-making processes.

Figure 9: ROC-AUC Graph for CNN Model.



The ROC curve for the binary classification of DR and glaucoma in retinal pictures using CNN is shown in Figure 9. The red training ROC curve achieves perfect classification with AUC = 1.00, while the blue test ROC curve demonstrates excellent performance with AUC = 99.98. Both curves significantly outperform the diagonal reference line (random classifier), indicating superior model discrimination capability between diabetic retinopathy and glaucoma cases.

6.1. Comparison and discussion

The MESSIDOR dataset was used to compare diabetic retinopathy detection models in detail (Table III). Compared to its rivals, the suggested CNN model performs better across the board, with a 99.37% accuracy rate, 98.86% precision, 99.90% recall, and 99.42% F1-score. From the baseline models, there is a noticeable drop in performance When looking into ResNet50 [36] vs. KNN [37], they can see that ResNet50 obtains a higher accuracy (97.08%), precision (94.04%), recall (94.04%), and F1-score (94.04%), but KNN only manages a recall (90.62%), precision (90.71%), and accuracy (90.62%). Compared to other metrics, InceptionV3's 97% accuracy, precision, recall, and F1-score [38] are similar. When compared to the control models (ResNet50, KNN, and InceptionV3), the proposed CNN architecture achieves superior performance in the automated detection of diabetic retinopathy. It has a very high recall rate of 99.90%, which means it can identify positive cases very well while minimizing false negatives, which is important for medical diagnosis applications.

Table 3: Comparison Between Proposed and Existing Models' Performance for Diabetic Retinopathy Classification.

Matrix	Resnet50	KNN	InceptionV3	CNN
Accuracy	90.62	97.08	97	99.37
Precision	90.71	94	97	98.86
Recall	90.62	94	97	99.90
F1-score	90.52	94	97	99.42

The proposed CNN model beats baseline methods for detecting DR from fundus images by a large margin. Medical image analysis was its primary inspiration for its unique architecture, which improves classification accuracy by efficiently extracting both low- and high-level data. Unlike generic models, the CNN shows strong generalization with stable training and validation performance, reduced overfitting, and robust convergence. A further important feature for clinical diagnosis is its high sensitivity in detecting positive instances, which helps to reduce the likelihood of false negatives. The model's exceptional performance on all important evaluation criteria, along with its remarkable ability to distinguish between diabetic retinopathy and non-DR cases, proves that it is effective for automated screening of retinal diseases and might be integrated into clinical decision-support systems.

7. Conclusion and Future Work

DR ranks as the leading cause of eye problems and blindness among people with diabetes globally. Early identification and accurate classification of DR will help to considerably slow down the development of DR and thereby determine the timeliness of intervention and treatment. This research presents an effective DL-based framework for automated diabetic retinopathy detection using a custom-designed CNN. Through the integration of a robust image preprocessing pipeline and a lightweight, domain-specific CNN architecture, the proposed model achieves superior diagnostic performance, recording an accuracy of 99.37% and an AUC of 99.98%. These results significantly outperform conventional architectures such as ResNet50, InceptionV3, and KNN, validating the clinical relevance of the proposed model for early DR detection and decision support in ophthalmology. Despite its high performance, the study is subject to certain limitations. The binary classification formulation restricts the model's applicability in scenarios requiring fine-grained grading of DR severity. Additionally, evaluation was limited to a single dataset, potentially impacting generalizability across diverse populations and imaging conditions. Future research will focus on extending the model to multi-class classification to enable detailed severity grading.

8. References

1. V Singh, DK Pathak, P Gupta. (2023) Integrating Artificial Intelligence and Machine Learning into Healthcare ERP Systems: A Framework for Oracle Cloud and Beyond. ESP J. Eng. Technol. Adv. 3(2): 171-178.
2. A Mishra, A Majumder, D Kommineni, (2025) et al. Role of Generative Artificial Intelligence in Personalized Medicine: A Systematic Review. Cureus. 17(4): e82310.
3. N Patel. (2022) Quantum Cryptography In Healthcare Information Systems: Enhancing Security In Medical Data Storage And Communication. J. Emerg. Technol. Innov. Res. 9(8): 2022.
4. SA Pahune. (2023) How does AI help in Rural Development in Healthcare Domain: A Short Survey. IJRASET. 11(6): 4184-4191.
5. A Bajwa, N Nosheen, KI Talpur, et al. (2023) A Prospective Study on Diabetic Retinopathy Detection Based on Modify Convolutional Neural Network Using Fundus Images at Sindh Institute of Ophthalmology & Visual Sciences. Diagnostics. 13(3): 393.
6. A-O Asia, Cheng-Zhang Zhu, Sara A Althubiti, et al. (2022) Detection of Diabetic Retinopathy in Retinal Fundus Images Using CNN Classification Models. Electronics. 11(17): 2740.
7. S Pandya. (2023) Predicting Diabetes Mellitus in Healthcare: A Comparative Analysis of Machine Learning Algorithms," Int. J. Curr. Eng. Technol., vol. 13(6): 536-544.
8. Y Li, W Mitchell, T Elze, et al. (2021) Association Between Diabetes, Diabetic Retinopathy, and Glaucoma. Curr. Diab. Rep. 21(10): 38.
9. SR Sagili, C Goswami, VC Bharathi, et al. (2024) Identification of Diabetic Retinopathy by Transfer

Learning Based Retinal Images. In: 2024 9th International Conference on Communication and Electronics Systems (ICCES). IEEE. 1149-1154.
10. MZ Chauhan, Abdelrahman M Elhusseiny, Krishna S Kishor, et al. (2024) Association of Primary Open-Angle Glaucoma with Diabetic Retinopathy among Patients with Type 1 and Type 2 Diabetes: A Large Global Database Study. Ophthalmology. 131(7): 827-835.
11. RP Mahajan. (2025) Improvised Diabetic Retinopathy Detection Accuracy in Retinal Images Using Machine Learning Algorithms. Int. Res. J. 12(3): 155-161.
12. N Tamim, M Elshrkawey, H Nassar. (2021) Accurate Diagnosis of Diabetic Retinopathy and Glaucoma Using Retinal Fundus Images Based on Hybrid Features and Genetic Algorithm. Appl. Sci. 11(13): 6178.
13. BV Baiju, VC Sharmila, MB, et al. (2021) Glaucoma and diabetic retinopathy diagnosis using image mining. In: Proceedings of the First International Conference on Computing, Communication and Control System. I3CAC 2021, 7-8 June 2021, Bharath University, Chennai, India, EAI, 2021.
14. MA Mostafiz. (2025) Machine Learning for Early Cancer Detection and Classification : AI-Based Medical Imaging Analysis in Healthcare. Int. J. Curr. Eng. Technol. 15(3): 251-260.
15. RQ Majumder. (2025) Machine Learning for Predictive Analytics : Trends and Future Directions. Int. J. Innov. Sci. Res. Technol. 10(4).
16. J Liu Ella Gibson, Shawn Ramchal, et al. (2021) Diabetic Retinopathy Screening with Automated Retinal Image Analysis in a Primary Care Setting Improves Adherence to Ophthalmic Care. Ophthalmol. Retin. 5(1): 71-77.
17. RS Latha, GR Sreekanth, RC Suganthe, et al. (2022) Diagnosis of Diabetic Retinopathy and Glaucoma from Retinal Images Using Deep Convolution Neural Network. In: 2022 International Conference on Computer Communication and Informatics (ICCCI). 1-6.
18. RP Mahajan. (2025) Transfer Learning for MRI image reconstruction: Enhancing model performance with pretrained networks. Int. J. Sci. Res. Arch. 15(1): 298-309.
19. N Prajapati. (2025) The Role of Machine Learning in Big Data Analytics: Tools, Techniques, and Applications. ESP J. Eng. Technol. Adv. 5(2).
20. S Pandya. (2024) Integrating Smart IoT and AI-Enhanced Systems for Predictive Diagnostics Disease in Healthcare. Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol. 10(6): 2093-2105.
21. GE Kangilbaeva, F Bakhritdinova, A Jurabekova. (2025) Diagnosis of Diabetic Retinopathy Using Artificial Intelligence. Prerpints.org.
22. VH Karambelkar, G Thorat, AR Chopade, et al. (204) Detection of Medical Diabetic Retinopathy through Application Using Deep Learning Models. In: 2024 International Conference on Healthcare Innovations, Software and Engineering Technologies (HISET). 207-210.
23. R Afshan, AK Chakraverti, M Chhabra. (2024) Glaucoma Detection Approach based on Colour Fundus Pictures. In: 2024 11th International Conference on Computing for Sustainable Global Development (INDIACom). 87-92.
24. S Hossain, Md Tanvir Chowdhury, Md Monira Hosen Akanda, et al. (2024) Enhancing Early Diagnosis and Precise Classification of Type 2 Diabetes with Diabetic Retinopathy Using Machine Learning and Image Processing. In: 2024 7th International Conference on Information and Communications Technology (ICOIACT). 217-221.
25. AD Kolte, J Sharma, U Rai, et al. (2023) Advancing Diabetic Retinopathy Detection: Leveraging Deep Learning for Accurate Classification and Early Diagnosis. In: Proceedings of the 8th International Conference on Communication and Electronics Systems, ICCES.
26. MM Butt, DNFA Iskandar, SE Abdelhamid, et al. Diabetic Retinopathy Detection from Fundus Images of the Eye Using Hybrid Deep Learning Features. Diagnostics. 12(7): 1607.
27. E Mohamed, MA Elmohsen, T Basha. (2021) Improved Automatic Grading of Diabetic Retinopathy Using Deep Learning and Principal Component Analysis. In: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). 3898-3901.
28. WL Alyoubi, MF Abulkhair, WM Shalash. (2021) Diabetic Retinopathy Fundus Image Classification and Lesions Localization System Using Deep Learning. Sensors. 21(11): 3704.
29. G Ali, A Dastgir, MW Iqbal, et al. (2023) A Hybrid Convolutional Neural Network Model for Automatic Diabetic Retinopathy Classification From Fundus Images. IEEE J. Transl. Eng. Heal. Med. 11: 341-350.
30. MS Santos, CT Valadao, CZ Resende, et al. (2024) Predicting diabetic retinopathy stage using Siamese Convolutional Neural Network. Comput. Methods Biomech. Biomed. Eng. Imaging Vis. 12(1).
31. S Mathur, S Gupta. (2023) Classification and Detection of Automated Facial Mask to COVID-19 based on Deep CNN Model. In: 2023 IEEE 7th Conference on Information and Communication Technology, CICT.
32. S Nokhwal, P Chilakalapudi, P Donekal, et al. (2024) Accelerating Neural Network Training: A Brief Review," ACM Int. Conf. Proceeding Ser. 31-35.
33. M Youldash, Atta Rahman, Manar Alsayed, et al. (2024) Early Detection and Classification of Diabetic Retinopathy: A Deep Learning Approach. AI. 5(4): 2586-2617.
34. P Bidwai, S Gite, K Pahuja, et al. (2022) A Systematic Literature Review on Diabetic Retinopathy Using an Artificial Intelligence Approach. Big Data Cogn. Comput. 6(4).
35. PK Jena, B Khuntia, C Palai, et al. (2023) A Novel Approach for Diabetic Retinopathy Screening Using Asymmetric Deep Learning Features. Big Data Cogn. Comput. 7(1).
36. M Feng, Y Cai, S Yan. (2025) Enhanced ResNet50 for Diabetic Retinopathy Classification: External Attention and Modified Residual Branch. Mathematics. 13(10): 1557.
37. N Shaukat, J Amin, M Sharif, et al. (2022) Three-Dimensional Semantic Segmentation of Diabetic Retinopathy Lesions and Grading Using Transfer Learning. J. Pers. Med. 12(9): 1454.
38. MR Shoaib, Heba M Emara, Jun Zhao. et al. (2024) Deep learning innovations in diagnosing diabetic retinopathy: The potential of transfer learning and the DiaCNN model," Comput. Biol. Med. 169: 1-51.