Detection and Classification of Diabetic Retinopathy Using Modified Inception V3

Abini M. A.*, S. Sridevi Sathya Priya

Department of Electronics and Communication Engineering,
Karunya Institute of Technology and Sciences
Coimbatore, Tamilnadu, India
E-mails: abinima87@gmail.com, sridevi@karunya.edu

**Corresponding author*

Received: May 18, 2024 **Accepted: October 10, 2024**

Published: March 31, 2025

Abstract: *In the last decade, the prevalence of diabetic retinopathy (DR), a sight-threatening medical condition of diabetes mellitus, has markedly increased, impacting millions globally. The conventional method of diagnosing and classifying this condition was done through physical and detailed examination of fundus images by ophthalmologists, a process prone to human error and time-consuming. To overcome this challenge, artificial intelligence, particularly deep learning algorithms, has taken a position in automating the diagnosis of diabetic eye disease and categorization from fundus images. Various studies have confirmed the effectiveness of convolutional neural networks in this task, with Inception V3 emerging as a particularly successful architecture. In our current work, we adduce a novel approach for DR detection and categorizing it utilizing the Inception V3 architecture on fundus images. The learning rate is modified between le-3, le-4, le-5, and le-6 to investigate various optimization approaches. Our model, trained on the Asia Pacific Tele Ophthalmology Society datasets, accomplished an accuracy of 91.64% for a learning rate of le-5, which outperforms current approaches in diagnosing the five phases of DR.*

Keywords: *Deep learning, Convolutional neural networks, Retina images, Inception V3, Diabetic retinopathy.*

Introduction

Diabetic retinopathy (DR) is a prevalent condition of diabetes mellitus and is widely thought as a key cause of vision disability around the globe [14]. The prevalence of DR is anticipated to impact over 200 million individuals by 2040. It affects the photo-sensitive tissue at the back of the eye, the retina, and can aggravate to vision loss or blindness permanently if left untreated. DR may not solely be a vascular complication but involves neurodegenerative processes [17]. Increasing evidence indicates that DR may initially manifest as diabetic neuropathy, with neurodegenerative changes preceding visible vascular damage [19]. Retinal ganglion cells, being highly susceptible to metabolic changes, are thought to exhibit early functional impairments before the appearance of microangiopathy. DR progresses in stages, from mild Non-Proliferative Retinopathy (NPDR) distinguished by microaneurysms and hemorrhages, to severe stages including Proliferative Diabetic Retinopathy (PDR) with abnormal blood vessel growth [26, 27].

Fundus images of DR (Fig. 1) reveal distinct differences between early stage and advanced cases, including those with peripheral diabetic neuropathy. In early-stage retinopathy, images typically show microaneurysms, small retinal hemorrhages, and exudates, such as hard exudates and cotton-wool spots, without significant neovascularization. In contrast, advanced stages and peripheral diabetic neuropathy exhibit widespread retinal changes, including extensive microaneurysms and hemorrhages, increased density of exudates,

retinal edema (often in the macula), and neovascularization [10, 11]. Advanced cases may also present with vitreous hemorrhage. These differences in fundus images help differentiate between early and more severe stages of the disease.

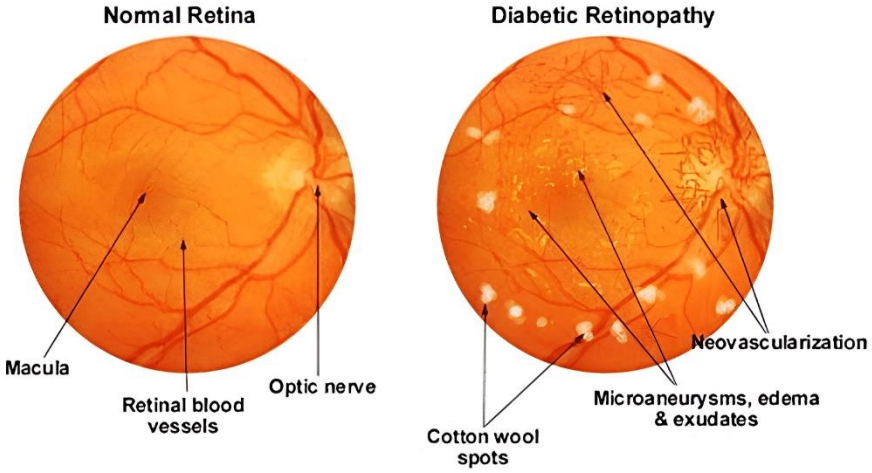


Fig. 1 The fundus image of normal retina and retina with diabetic retinopathy symptoms [29]

Early methods of DR screening involved physical examination of retinal images by trained ophthalmologists. Computerized systems supported with machine learning and deep learning (DL) techniques have elevated effects in auto DR screening [23]. These systems can analyze retinal images for signs of DR lesions, classify disease severity, and provide risk stratification for patients. Neural network studies offer advanced methods for early detection of diabetic mellitus [2, 3] by analyzing complex retinal and electrophysiological data. Using DL algorithms, these models can examine retinal images from optical coherence tomography and fundus photography to identify subtle structural changes that precede visible vascular damage. They can also analyze electrophysiological data from electroretinography to detect early functional alterations in retinal responses. By integrating multimodal data combining imaging, electrophysiological results, and clinical history, neural networks can predict and stratify risk levels for DR, enabling proactive management. Predictive analytics and personalized medicine approaches further enhance early intervention strategies by tailoring treatment based on individual risk profiles and early signs of disease progression.

DL has revolutionized medical image analysis by enabling models to learn hierarchical features directly from data, leading to improved diagnostic accuracy. Our research is motivated by the need for robust and accurate automated methods for DR detection and classification [7, 21]. By applying our neural network model to this subset of patients, we aim to uncover any latent retinal changes that could indicate early neurodegenerative processes, potentially refining the early diagnostic criteria for DR. We aim to leverage the most current DL architectures and innovative techniques, such as attention mechanisms, to enhance the performance of DR screening systems. By focusing on the Inception V3 architecture with an additional layer, we seek to improve the model's potential to identify subtle DR lesions and categorize disease severity levels accurately [28]. The classification of DR advancement within patients is categorized into five stages: class 0 no DR, class 1 mild, class 2 moderate, class 3 severe, and class 4 proliferative DR [6] (Fig. 2).

Harnessing DL technology offers the capability to leverage vast datasets efficiently, automatically discern data features, aid physicians in accurate and swift diagnoses, and enhance medical efficacy.

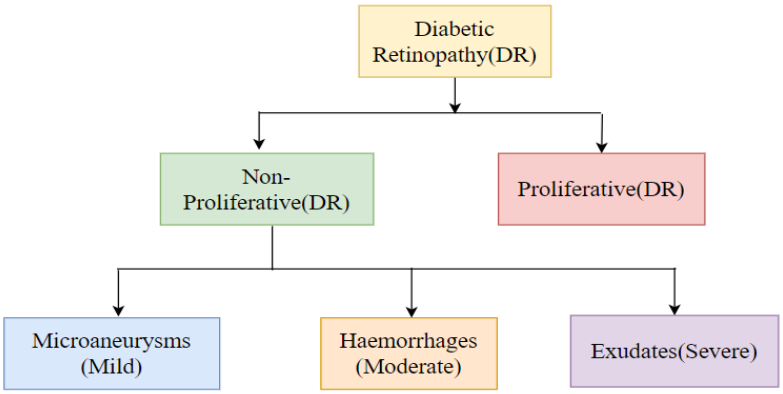


Fig. 2 Different lesions present in each stage of diabetic retinopathy [23]

To facilitate early detection and treatment, our study employed a deep learning framework utilizing Inception V3 transfer learning on fundus images, categorizing them into five distinct classes. Here we present a comprehensive study on diabetic retinopathy detection and classification using modified Inception V3. Our objectives include:

- Investigating the potential of the Inception V3 architecture in nabbing relevant features from retinal images.
- In order to mitigate the problem of class imbalance and enhance precision, the model under consideration employed image augmentation methods throughout the training process.
- Modifying the internal layers of the pre-trained model to enhance their performance in detecting DR.
- Evaluating the outputs of the proposed approach on benchmark datasets and comparing it with existing methods to showcase advancements in DR screening technology.

Related works

Prior studies utilized diverse methodologies to identify DR and extract relevant features. Although numerous systems based on image processing, machine learning, and deep learning have been developed by researchers, proving their usefulness and significance in the field of ophthalmology, better outcomes are yet achievable. InceptionResNet-V2 was previously taught using transfer learning, then a custom block of convolutional neural networks (CNN) layers was built over InceptionResNet-V2 to form the meld model, according to proposal of Gangwar and Ravi [6]. On the Messidor-1 and APTOS datasets, the model has evaluation correctness of 72.33% and 82.18%, respectively. Kobat et al. [12] highlighted DR as a frequent consequence of diabetes that impacts the blood vessels of the retina, which are essential for vision. To aid in the diagnosis of DR, they used a pre-trained DenseNet model to partition fundus pictures into horizontal and vertical patches. Their model achieved an impressive 84.90% accuracy rate using 10-fold cross-validation approaches, as shown by rigorous validation using several datasets. In addition, Pamadi et al. [20] performed binary and multinomial classification tasks on fundus images using the MobileNet-V2 architecture. The decision to use this architecture was based on its efficient training time and compatibility with mobile systems. Sanjana et al. [24] created a binary classification method for DR using five distinct learning models: Xception, InceptionResNet-V2, MobileNet-V2, DenseNet-121, and NASNet-Mobile. The best validation accuracies were obtained by these models is in the following order: 86.25%, 96.25%, 93.75%, 81.25%, and 80.00%, respectively. Qummar et al. [22] trained an ensemble of five deep CNN models, namely ResNet-50, Xception,

Inception V3, Dense-169, and Dense-121, utilizing a publicly accessible Kaggle data of retina images. Their ensemble model achieved an accuracy of 80.70%. Lahmar and Idri [13] conducted an empirical comparison of 28 deep combo architectures and seven end-to-end DL architectures for automatic binary classification of referable DR. They utilized seven DL feature extraction methods (DenseNet-201, VGG-16, VGG-19, MobileNet-V2, Inception V3, InceptionResNet-V2, and ResNet-50) along with four classifiers for hybrid architectures. Liu et al. [16] utilized transfer learning with Inception V3 and ResNet-50 models, incorporating CLAHE, grayscale transformations, and data augmentation techniques for DR lesion detection and classification. Comparisons between models employing similar data and training methods affirmed that with Inception V3 structure they achieved higher accuracy compared to those using ResNet-50. Deshpande and Pardhi [4] successfully developed a CNN model for diagnosing retinopathy and other retinal diseases, powering a pre-trained VGG-16 framework. Their model not only diagnoses the diseases but also provides information about their severity. They achieved a model accuracy of 74.58%, offering potential assistance to physicians for various purposes. Majumder and Kehtarnavaz [18] combined a regression model with Xception architecture to classify the five stages of DR. This integration resulted in accuracies of 86.00% and 82.00% for the APTOS 2019 and Kaggle EyePacs datasets, respectively. Table 1 summarizes the related literature reviews.

Table 1. Summary of related works

Reference	Method	Dataset	Accuracy
Gangwar and Ravi [6]	InceptionResNet-V2	Messidor-1 APTOS	72.33% 82.18%
Deshpande and Pardhi [4]	VGG-16	APTOS 2019	74.58%
Pamadi et al. [20]	MobileNet-V2	Kaggle APTOS	78.00%
Qummar et al. [22]	Ensemble model	Kaggle EyePacs	80.70%
Majumder and Kehtarnavaz [18]	Xception Model with regression	Kaggle EyePcas APTOS 2019	82.00% 86.00%
Liu et al. [16]	Inception V3	APTOS 2019	83.61%
Lahmar and Idri [13]	Ensemble Model	Messidor-2	84.10%
Kobat et al. [12]	Densenet-201	APTOS 2019	84.90%
Sanjana et al. [24]	Xception, InceptionResNet-V2, MobileNet-V2, DenseNet-121, NASNet-Mobile	Kaggle APTOS	86.10%

Materials and methods

Dataset

The present investigation employed the APTOS 2019 blindness detection dataset [28], comprising 3 662 retinal images obtained under varying illumination conditions. The Aravind Eye Hospital in India provided the images utilized in this study. The retinal images in the dataset were classified into five distinct categories, each representing a different degree of diabetic retinopathy: class 0 (which signified no DR), class 1 (which indicated mild DR), class 2 (which represented moderate DR), class 3 (which represented severe DR), and class 4 (which corresponded to proliferative DR) [9, 15, 16]. Fig. 3 provides a summary of the sample distribution according to severity level, with the quantity of images assigned to each

class denoted. To ensure precise model evaluation and training, the dataset was modified to include an equivalent number of instances of each severity level.

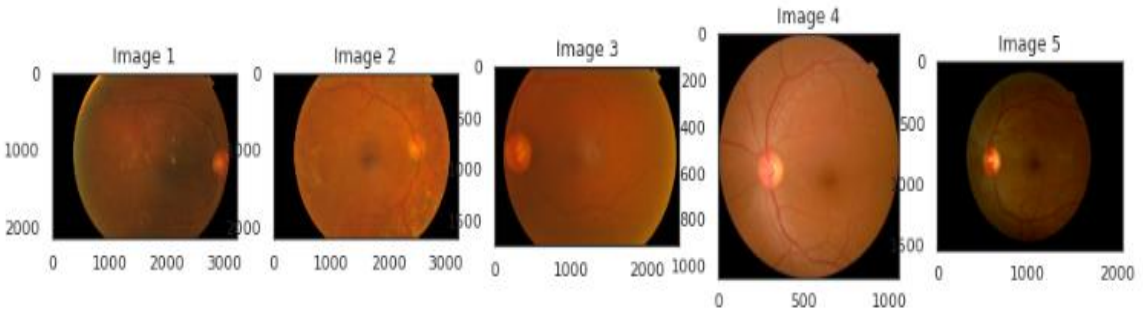


Fig. 3 Sample of images in the dataset [28]

Data preprocessing and augmentation

Preprocessing in DR involves several stages to enhance the quality of retinal images for analysis. This includes steps such as image cropping to remove irrelevant areas, color normalization to standardize appearance, an enhancement to improve visibility of structures, noise reduction to remove artefacts, normalization and rescaling for uniformity. A widely used data augmentation technique named SMOTE, overcomes the challenge in classification tasks like class imbalance. When dealing with imbalanced datasets, the scarcity of samples in the minority class can bias the classifier toward the larger class. By balancing the class division, SMOTE not only addresses this issue but also mitigates overfitting and enhances the classifier's generalization performance. It achieves this by diversifying the dataset, thereby reducing the risk of overfitting. By applying SMOTE to the preprocessed image categories, the capability of the dataset was expanded to a total of 9 025 images, with 1 805 images for each DR class.

This distribution ensures an equal sample balance in each category. The number of images under each label before augmentation and after augmentation is shown in Fig. 4.

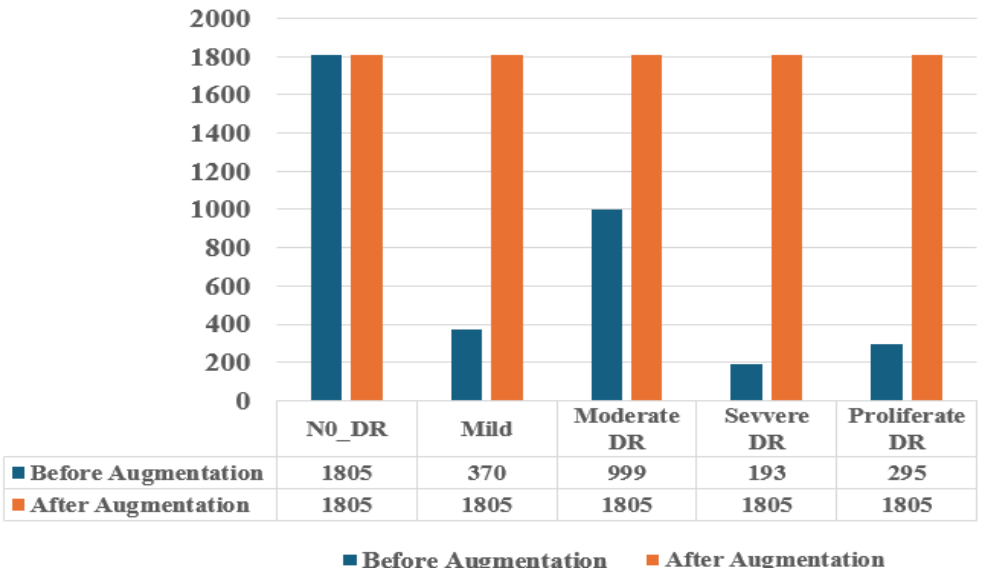


Fig. 4 The number of images under each label

Proposed methodology

The system architecture used in this work tries to measure the severity of DR using the dimensions. The neural network model designed for DR classification uses transfer learning using the Inception V3 architecture and is pre-trained on a large-scale picture dataset. By setting the trainable attribute of certain layers in the base model to false, the model freezes those layers during training, preventing their weights from being updated. This approach is beneficial when using pre-trained models like Inception V3, as it allows leveraging the learned features from a large dataset (like ImageNet) while customizing the final layers for a specific task such as DR detection. The model incorporates a series of layers to extract and learn complex features from retinal images. The base Inception V3 model serves as a feature extractor, combined by a layer with a 40% dropout rate to prevent overfitting. A flattening layer reshapes the features into a vector, which is then normalized using batch normalization to stabilize the training process. Subsequently, dense layers with 256, 128, and 32 neurons, initialized using the He uniform initializer, are introduced with batch normalization and ReLU activation functions to capture hierarchical patterns and introduce non-linearity. Dropout layers are again applied to prevent overfitting during training. The final layer will be with five neurons and a softmax activation function for multi-class classification into the stages of DR. This comprehensive architecture, integrating transfer learning, regularization techniques, normalization, and appropriate activation functions, is designed to achieve accurate and robust DR classification in medical imaging applications.

Let X represent the input retinal images. The Inception V3 model i, is pre-trained on the ImageNet dataset and certain layers L_i (where $i \in \{1, 2\ldots. k\}$) are frozen:

$$I(X) = F\ (X, \{Wi\}^k i = 1, \{Wj\}^n j = k + 1) \quad (1)$$

where $\{Wi\}_{i=1}$ are the frozen weights, and $\{Wj\}j = k + 1$ are the trainable weights.

A dropout layer with a rate of 0.4 is applied to the output of the Inception V3 model to prevent overfitting:

$$X_{drop} = Dropout\ (F\ (X, \{Wi\}^k i = 1, \{Wj\}^n jk + 1), 0.4) \quad (2)$$

A flattening layer reshapes the output features into a vector:

$$X_{flat} = Flatten(X_{drop}) \quad (3)$$

Batch normalization is applied to the flattened features to stabilize the training process:

$$X_{norm} = BatchNorm(X_{flat}) \quad (4)$$

Subsequently, dense layers with 256, 128, and 32 neurons are introduced, each initialized with the *He uniform initializer*, followed by ReLU activation functions and batch normalization:

$$H_1 = ReLU\ (BatchNorm(Dense(X_{norm}, 256, HeUniform))) \quad (5)$$

$$H_2 = ReLU\ (BatchNorm(Dense(H_1, 128, HeUniform))) \quad (6)$$

$$H_3 = ReLU\ (BatchNorm(Dense(H_2, 32, HeUniform))) \quad (7)$$

Dropout layers are applied to these dense layers to further prevent overfitting:

$$H_1^{drop} = Dropout(H_1, 0.4) \quad (8)$$

$$H_2^{drop} = Dropout(H_2, 0.4) \quad (9)$$

$$H_3^{drop} = Dropout(H_3, 0.4) \quad (10)$$

The final layer consists of five neurons and a softmax activation function for multi-class classification:

$$Y = SoftMax(Dense(H_3^{drop}, 5)) \quad (11)$$

The proposed comprehensive architecture integrates transfer learning, regularization techniques, normalization, and appropriate activation functions, aiming to achieve accurate and robust DR classification in medical imaging applications (Fig. 5).

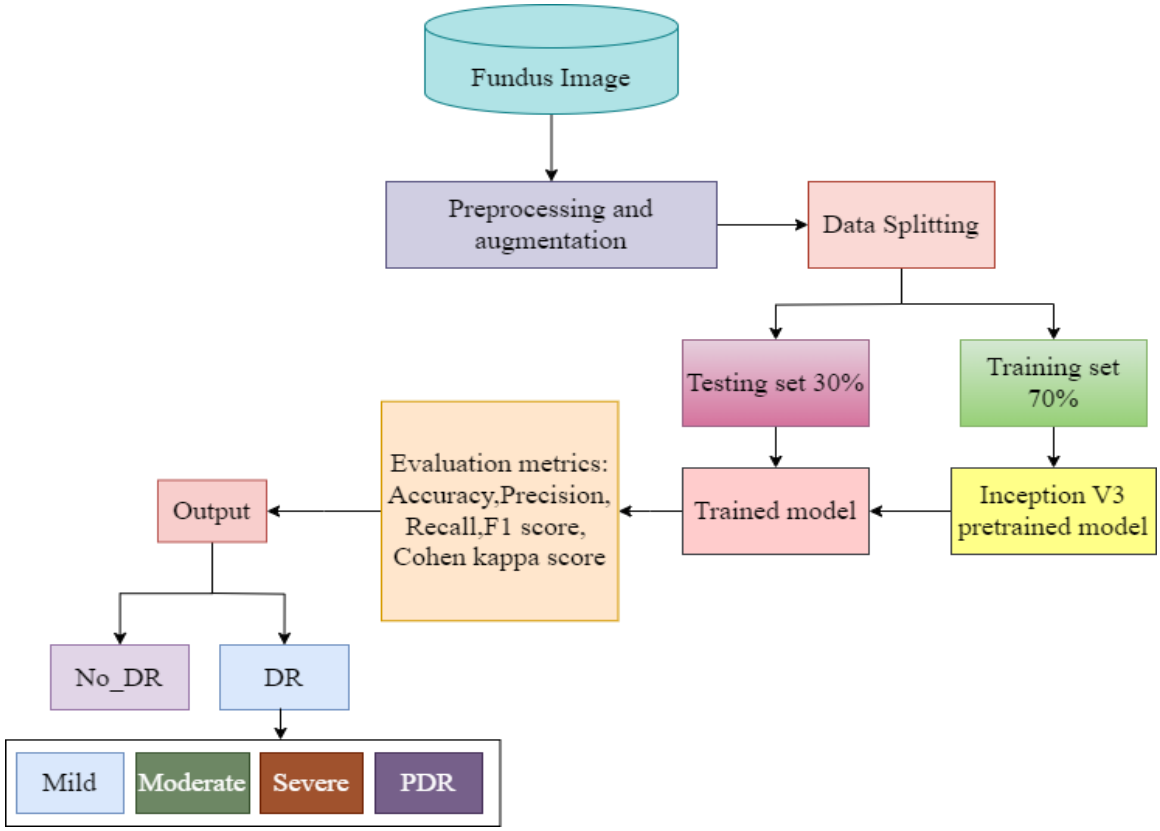


Fig. 5 Proposed block diagram of diabetic retinopathy grading using Inception V3

Inception V3

Inception V3 is a convolutional neural network architecture that leans on the successful GoogLeNet model, renowned for its exactitude in classifying biomedical data using transfer learning. Inception V3 introduces an inception module, inspired by GoogLeNet, which combines multi-sized convolutional filters into a single filter. This innovative approach minimises the learnable parameters, leading to lower computational complexity. The Inception V3 architecture incorporates three types of Inception modules: Type A, B, and C, as depicted in Fig. 6. These modules are meticulously crafted convolutional units for extracting distinctive features while minimizing the parameter count. Every Inception module has numerous pooling and

convolutional layers that operate simultaneously [25]. The number of parameters is significantly decreased by using tiny convolutional layers, such as 3×3, 1×3, 3×1, and 1×1 layers. Five Inception A modules, four Inception B modules, and two Inception C modules make up the sequence of stacked modules that makes up Inception V3. During training and testing, the pictures were down-sized to 299×299 pixels, which is the default input image dimensions of Inception V3. The convolutional layers and Inception modules cause the feature map dimensions to change to 8×8 with 2.048 channels overall. Even though the original Inception V3 network produced 1.000 classes, we changed the last layer's output channel count from 1.000 to 5 in order to meet our classification needs.

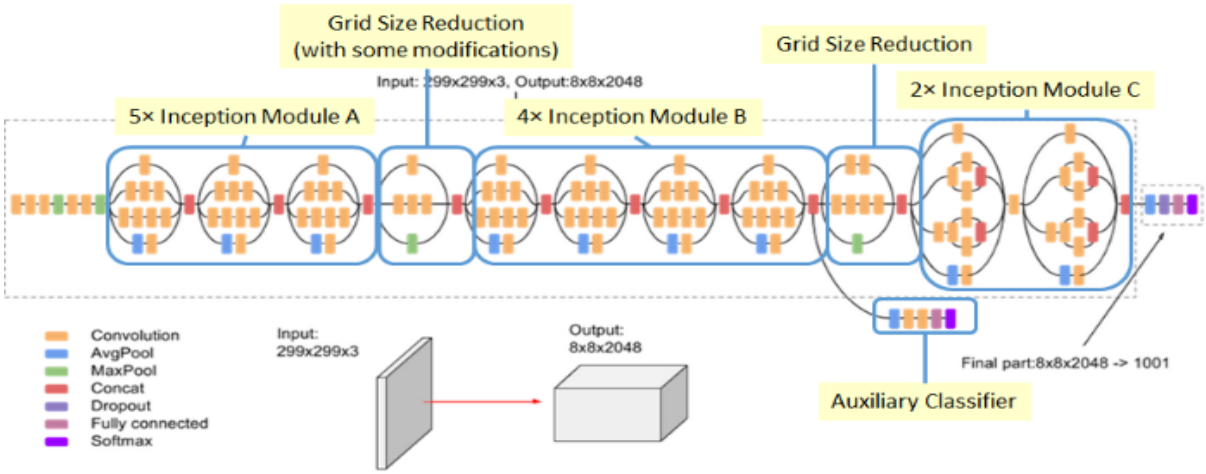


Fig. 6 Inception V3 architecture [25]

Performance evaluation measures

After completing the training phase, the models underwent testing using a separate test set comprising 2 708 fundus images. The evaluation of model performance was conducted using various metrics including overall classification accuracy, recall (sensitivity), precision, F1-score, and Cohen kappa score [1]. These metrics provide insights into different aspects of model performance:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{12}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{13}$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{14}$$

$$\text{F1-score} = \frac{2\times(\text{Recall}\times\text{Precision})}{(\text{Recall}\times\text{Precision})} \tag{15}$$

$$Cohen\ kappa\ score\ k = \frac{(Po - Pe)}{(1 - pe)} \tag{16}$$

where *TP*, *TN*, *FP*, and *FN* stand for true positive, true negative, false positive, and false negative samples, respectively.

Results and discussions

A system equipped with a GPU NVIDIA GTX1650, an Intel i5-1135G7 processor, and Windows 11 with Python libraries such as Keras and TensorFlow were utilized for various stages of the experiment, including preprocessing, augmentation, and implementing CNN.

The suggested model aims to detect DR in its early stages, using a dataset of 9 025 retinal pictures for training, testing, and validation. The datasets contained data for both training (70%) and testing (30%). The offered settings describe how to train the Modified Inception V3 model. ReLU activation functions are used throughout the model design with input pictures of 299×299 pixels and having three RGB channels each. Training is done with a batch size of 16 samples, and the Adam optimizer is used for parameter changes across 50 epochs. The learning rate is modified between le-3, le-4, le-5, and le-6 to investigate various optimization approaches. During training, the performance of the model is assessed using the categorical cross entropy loss function, which evaluates the discrepancy between predicted and actual labels. All parameters used for the study are given in Table 2.

Table 2. Hyper parameters used for the study.

Parameters	Value
Input image dimension	(299×299, 3)
Activation functions	ReLU
Batch size	16
Optimizer	Adam
Epochs	50
Learning rate	le-3, le-4, le-5, le-6
Loss function	Categorical cross entropy

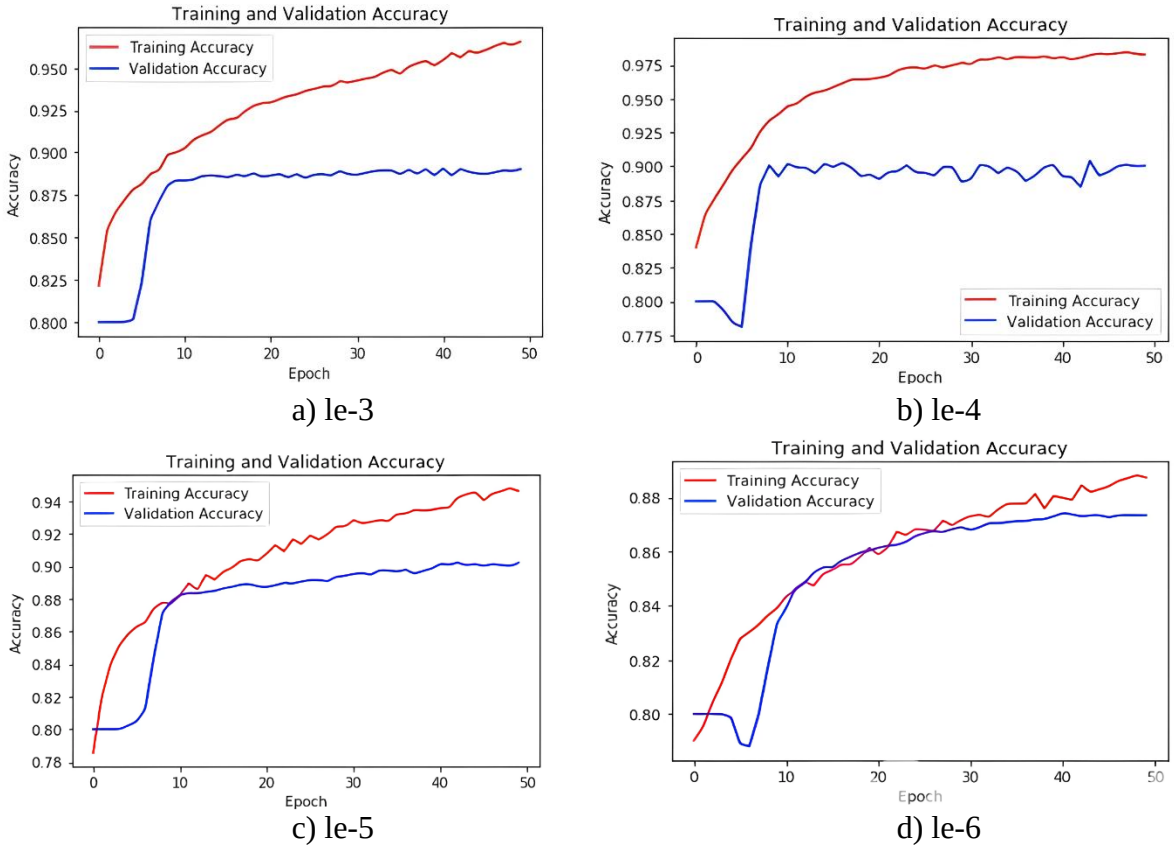
These configurations form the groundwork for training the model and allow for experimentation to achieve optimal performance in multi-class classification tasks. Adjusting parameters such as the learning rate and batch size may have a substantial influence on training efficiency and model accuracy. The presented table shows the performance characteristics of a modified Inception V3 model at various learning rates. As the learning rate increases from le-3 to le-6, so do precision, recall, F1-score, Cohen kappa score, testing accuracy, and loss. Notably, the model obtains the maximum precision, recall, F1-score, Cohen kappa score, testing accuracy, and lowest testing loss at a learning rate of le-5, suggesting optimal performance. In contrast, with a learning rate of le-6, there is a decline in precision, recall, F1-score, Cohen kappa score, testing accuracy, and an increase in testing loss compared to other learning rates, suggesting inferior performance. Furthermore, visualizing model training and validation accuracy curves, loss curves, and precision at different learning rates could provide deeper insights into the training dynamics and the impact of learning rate on model performance. Overall, selecting an optimal learning rate is crucial for achieving high model performance and generalization capability, as it directly influences the convergence and stability of the training process.

Table 3 shows the performance of the proposed model on the APTOS dataset. The increased performance of the model can be ascribed to the use of enhanced feature extraction and selection procedures. Furthermore, the higher kappa value highlights the ability of the proposed model to accurately categorize images with multiple labels.

Fig. 7 shows training and validation accuracy curves for different learning rates. Fig. 8 shows training and validation loss curves for different learning rates.

Table 3. Comparison of all the performance metrics for multistage classification

Model	Learning rate	Precision, (%)	Recall, (%)	F1-score	Cohen kappa score	Testing accuracy, (%)	Testing loss
Modified Inception V3	le-3	0.7534	0.7648	0.7104	0.5540	88.25	0.6582
	le-4	0.7442	0.7435	0.6639	0.5088	88.12	0.5518
	le-5	**0.9310**	**0.9310**	**0.9320**	**0.8705**	**91.64**	**0.3094**
	le-6	0.7324	0.7235	0.6327	0.4483	87.82	0.3438



a) le-3 b) le-4

c) le-5 d) le-6

Fig. 7 Training and validation accuracy curves for different learning rates

In the classification report provided, each class represents a different level of severity for diabetic retinopathy, a complication of diabetes that affects the eyes. Precision measures the accuracy of positive predictions, recall measures the proportion of actual positives correctly identified, and the F1-score combines both precision and recall into a single metric. Notably, for the "No DR" class, precision is high at 0.94, indicating a low rate of false positives, and recall is also high at 0.97, suggesting effective identification of actual cases. For "Mild DR", both precision and recall are slightly lower but still robust at 0.90 and 0.91, respectively. However, for "Moderate DR", while precision remains relatively high at 0.88, recall drops to 0.79, indicating a higher rate of false negatives, possibly implying challenges in accurately identifying this severity level. "Severe DR" demonstrates high precision and recall at 0.94 and 0.96, respectively, indicating strong performance in identification. Lastly, "Proliferative DR" shows good precision at 0.87 and recall at 0.90, suggesting effective identification with a slightly higher rate of false positives compared to other classes. Overall, this model demonstrates solid performance in distinguishing between various severity

levels of DR, with particular strengths in identifying the absence of DR and severe cases, while also indicating some room for improvement in accurately classifying moderate cases. Table 4 shows the Evaluation measures for the learning rate le-5.

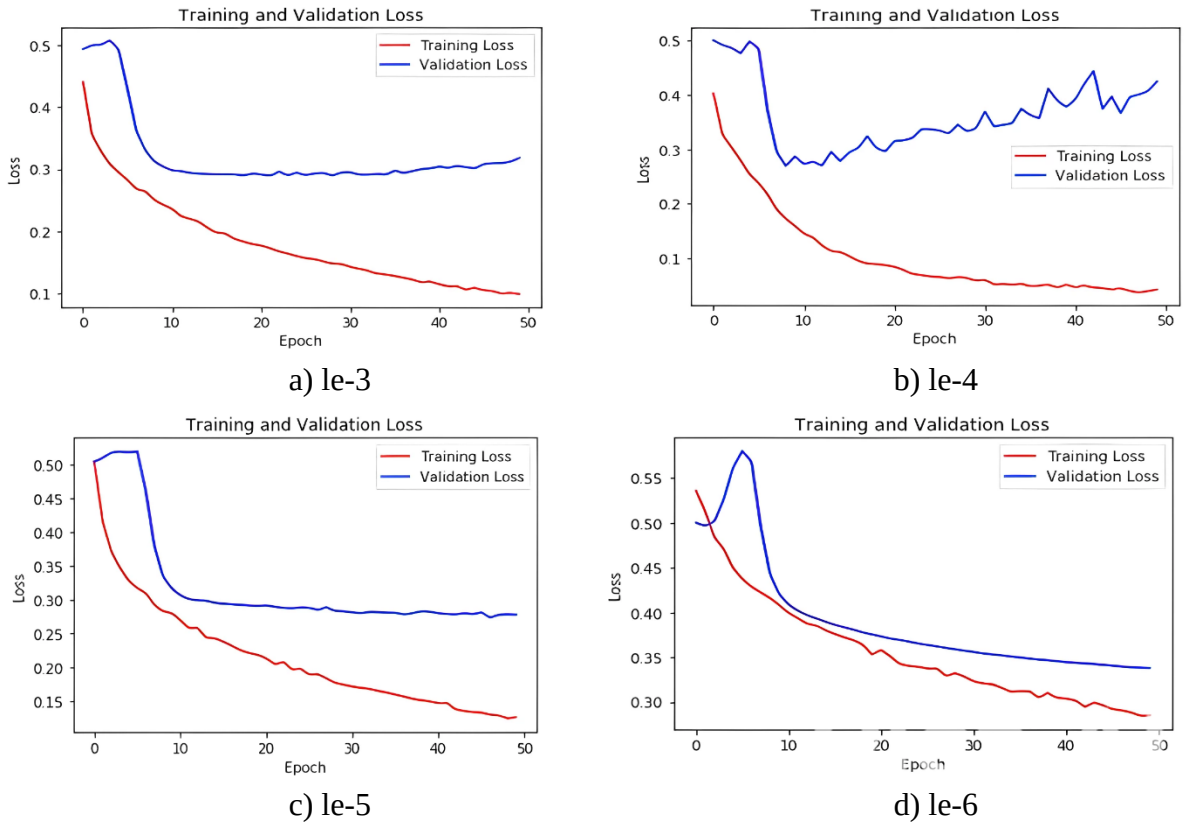


Fig. 8 Training and validation loss curves for different learning rates

Table 4. Performance metrics based on Inception V3 model features (le-5)

Classes	Precision, (%)	Recall, (%)	F1-score
No DR	0.94	0.97	0.95
Mild	0.90	0.91	0.91
Moderate	0.88	0.79	0.83
Severe	0.94	0.96	0.95
Proliferative DR	0.87	0.90	0.89

Table 5 shows a model's training and validation performance throughout numerous epochs. As the number of epochs rises, both training and validation accuracy progressively improve, demonstrating that the model is learning from the training data and generalizing well to previously unknown validation data. The training loss falls steadily across epochs, showing that the model's predictions are more closely aligned with the training labels. Similarly, the validation loss lowers across epochs, indicating that the model's performance on unseen validation data increases with training. Notably, towards the conclusion of training, the model achieves excellent accuracy on both the training and validation sets, showing that it can learn and generalize effectively. Furthermore, the reduced difference between training and validation accuracy in subsequent epochs indicates that the model is not overfitting and can generalize effectively to fresh data. Overall, the table shows that the model's performance gradually

improves across training epochs, with both accuracy metrics increasing and loss metrics dropping, suggesting effective model training.

Table 5. Comparison of results for various epochs for multistage classification

Epochs	Training accuracy, (%)	Validation accuracy, (%)	Training loss	Validation loss
10	0.8165	0.8014	0.1691	0.1885
20	0.8677	0.7868	0.1207	0.1556
30	0.9041	0.8264	0.0918	0.1420
40	0.9314	0.8432	0.0715	0.1327
50	0.9344	0.8774	0.0646	0.1214

The performance evaluation of the optimized CNN model is graphically demonstrated using the confusion matrix (Fig. 9) and receiver operating characteristic (ROC) curves (Fig. 10), which offer valuable insights into its classification accuracy. In Fig. 9, a normalized confusion matrix is provided to further analyses the model's accuracy in classifying each specific degree of DR. The confusion matrix was created by implementing the suggested approach on the test set. The suggested technique outperforms the baseline methods in all F1-score, accuracy, sensitivity, and specificity parameters, as shown in Table 4. Furthermore, to further analyses the connection between the rate of correctly identified positive cases and the rate of incorrectly identified negative cases, Fig. 10 presents the ROC curve. As the false positive rates grow, the ROC curve demonstrates the balance between the model's capacity to correctly identify unhealthy retinas and its tendency to incorrectly classify healthy retinas. This indicates that the model's confidence in recognizing non-healthy retinas diminishes.

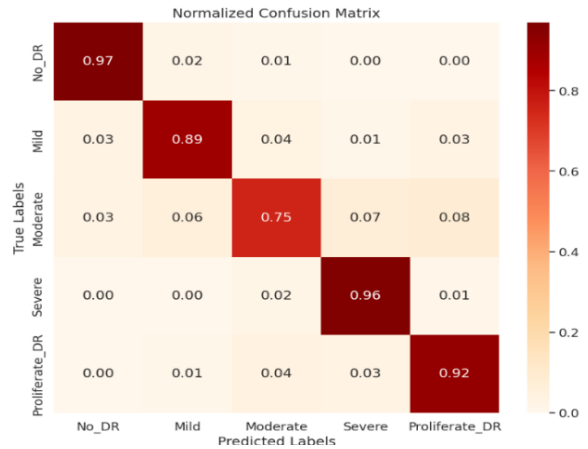


Fig. 9 Confusion matrix

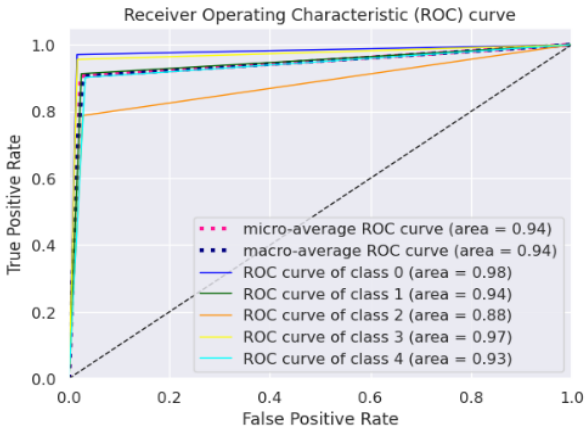


Fig. 10 ROC curve for proposed method

Comparing results with the state-of-the-art methods

The comparison of performance among different methods presented in Table 6 and Fig. 11 reflects their respective accuracies across various evaluation metrics. Dondeti et al. [5] approach, utilizing NASNet combined with γ-SVM, achieved an accuracy of 77.90%, demonstrating competitive performance. Majumder and Kehtarnavaz [18] proposed a modified DenseNet based squeeze excitation densely connected multitasking network (MSEDenseNet) with an accuracy of 81.00%, although showing a lower performance in recall and F1-score metrics. Kobat et al.'s [12] DenseNet-201 achieved the highest accuracy among the compared methods at 84.90%, showcasing strong performance across all evaluation metrics. Islam et al. [8] method utilizing supervised contrastive learning demonstrated an accuracy of 84.36%, with a slight drop in recall and F1-score metrics compared to accuracy.

Finally, the proposed method, utilizing a modified Inception V3 architecture, achieved the highest accuracy of 91.34%, demonstrating superior performance across all evaluation metrics, including precision, recall, and F1-score, indicating its effectiveness in classification tasks. Overall, the comparison highlights the varying performance levels of different approaches, with the proposed method exhibiting the highest accuracy and balanced performance across all metrics.

Table 6. Comparison of state-of-the-art methods with our proposed model for multi-class classification with APTOS 2019

Author	Models	Test accuracy, (%)	Precision, (%)	Sensitivity (Recall), (%)	F1-score, (%)
Dondeti et al. [5]	NASNet + γ-SVM	77.90	76.00	77.00	75.00
Majumder and Kehtarnavaz [18]	MSE Dense Net	81.00	67.00	59.00	61.00
Kobat et al. [12]	Densenet-201	84.90	82.40	83.10	82.70
Islam et al. [8]	Supervised contrastive learning	84.36	73.84	70.51	70.49
Proposed method	**Modified Inception V3**	**91.34**	**90.6**	**90.6**	**90.6**

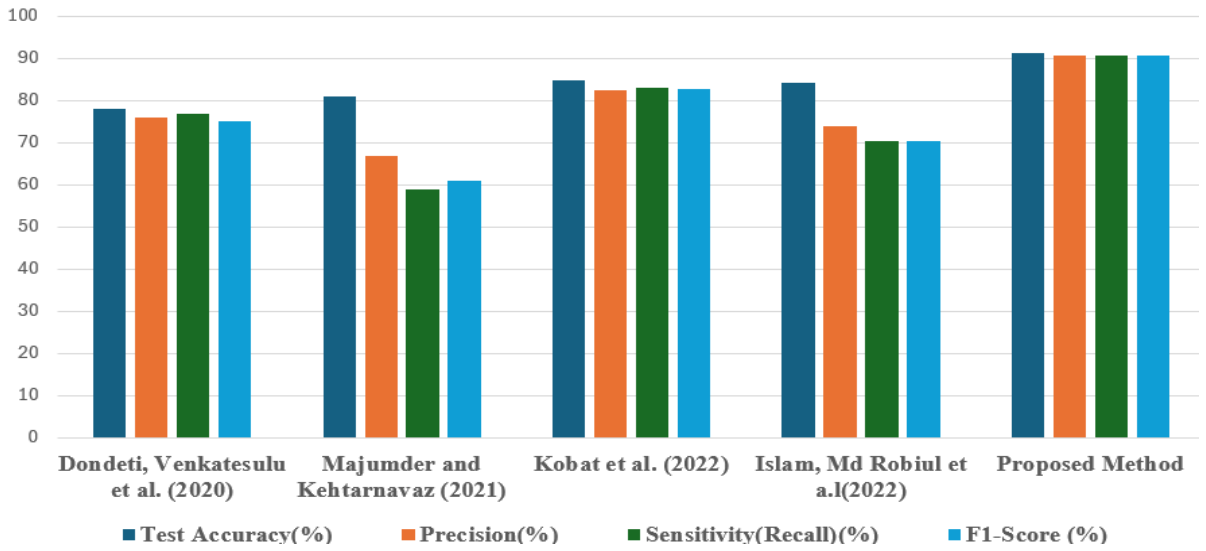


Fig. 11 Comparative analysis findings – APTOS

Conclusion and future scope

The global population of ophthalmologists may not be sufficient to fulfil the requirements for regulatory screenings due to the rapid increase in the projected number of diabetes mellitus patients in the future, thereby leading to a rise in the number of individuals affected by diabetic retinopathy. A safe and therapeutically relevant automatic detection technique has shown to be an excellent solution to the issue. Manual analysis of these photographs is a laborious and intricate process that necessitates the expertise of highly skilled professionals. We have effectively created a model using convolutional neural networks using a pretrained Inception V3 framework. This model is capable of detecting diabetic retinopathy and providing information about the disease's severity. The author of this work created a deep learning-based multiclass model to grade the severity of diabetic retinopathy. We have achieved a testing accuracy of 91.64% for multiclassification. This approach can aid clinicians in expediting the

diagnosis of this condition. Additional models can be created to diagnose various disorders, particularly those related to the eye. This might potentially be beneficial in early detection of certain disorders, hence preventing irreversible vision loss.

References

1. Alghamdi, H. S. (2022). Towards Explainable Deep Neural Networks for the Automatic Detection of Diabetic Retinopathy, Applied Sciences, 12(19), 9435.
2. Borré A., L. O. Seman, E. Camponogara, S. F. Stefenon, et al. (2023). Machine Fault Detection Using a Hybrid CNN-LSTM Attention-based Model, Sensors, 23(9), 4512.
3. Cizotto A. A. J., R. C. T. de Souza, V. C. Mariani, L. dos Santos Coelho (2023). Web Pages from Mockup Design based on Convolutional Neural Network and Class Activation Mapping, Multimedia Tools and Applications, 82(36), 38771–38797.
4. Deshpande A., J. Pardhi (2021). Automated Detection of Diabetic Retinopathy Using VGG-16 Architecture, International Research Journal of Engineering and Technology, 8(3), 2936-2940.
5. Dondeti V., J. D. Bodapati, S. N. Shareef, N. Veeranjaneyulu (2020). Deep Convolution Features in Non-linear Embedding Space for Fundus Image Classification, Revue d'Intelligence Artificielle, 34(3), 307-313.
6. Gangwar A. K, V. Ravi (2020). Diabetic Retinopathy Detection Using Transfer Learning and Deep Learning, Frontiers in Intelligent Computing: Theory and Applications, 1, 679-689.
7. Gupta S., S. Thakur, A. Gupta (2022). Optimized Hybrid Machine Learning Approach for Smartphone-based Diabetic Retinopathy Detection, Multimedia Tools and Applications, 81(10), 14475-14501.
8. Islam M. R., L. F. Abdulrazak, M. Nahiduzzaman, M. O. F. Goni, et al. (2022). Applying Supervised Contrastive Learning for the Detection of Diabetic Retinopathy and Its Severity Levels from Fundus Images, Computers in Biology and Medicine, 146, 105602.
9. Kale Y., S. Sharma (2022). Detection of Five Severity Levels of Diabetic Retinopathy Using Ensemble Deep Learning Model, Multimedia Tools and Applications, 82(12), 19005-19020.
10. Karki S. S., P. Kulkarni (2021). Diabetic Retinopathy Classification Using a Combination of EfficientNets, 2021 International Conference on Emerging Smart Computing and Informatics, 68-72.
11. Kassani S. H., P. H. Kassani, R. Khazaeinezhad, M. J. Wesolowski, et al. (2019). Diabetic Retinopathy Classification Using a Modified Xception Architecture, IEEE International Symposium on Signal Processing and Information Technology, 1-6.
12. Kobat S. G., N. Baygin, E. Yusufoglu, M. Baygin, et al. (2022). Automated Diabetic Retinopathy Detection Using Horizontal and Vertical Patch Division-based Pre-trained DenseNET with Digital Fundus Images, Diagnostics, 12(8), 1975.
13. Lahmar C., A. Idri (2022). Deep Hybrid Architectures for Diabetic Retinopathy Classification, Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 11(2), 166-184.
14. Lin X., Y. Xu, X. Pan, J. Xu, et al. (2020). Global, Regional, and National Burden and Trend of Diabetes in 195 Countries and Territories: An Analysis from 1990 to 2025, Scientific Reports, 10(1), 1-11.
15. Liu H., K. Yue, S. Cheng, C. Pan, et al. (2020). Hybrid Model Structure for Diabetic Retinopathy Classification, Journal of Healthcare Engineering, 2020(1), 8840174.
16. Liu K., T. Si, C. Huang, Y. Wang, et al. (2024). Diagnosis and Detection of Diabetic Retinopathy Based on Transfer Learning, Multimedia Tools and Applications, 83(35), 82945-82961.

17. Luo X., W. Wang, Y. Xu, Z. Lai, et al. (2024). A Deep Convolutional Neural Network for Diabetic Retinopathy Detection via Mining Local and Long-range Dependence, CAAI Transactions on Intelligence Technology, 9(1), 153–166.
18. Majumder S., N. Kehtarnavaz (2021). Multitasking Deep Learning Model for Detection of Five Stages of Diabetic Retinopathy, IEEE Access, 9, 123220-123230.
19. Ohri K., M. Kumar (2024). Supervised Fine-tuned Approach for Automated Detection of Diabetic Retinopathy, Multimedia Tools and Applications, 83(5), 14259-14280.
20. Pamadi A. M., A. Ravishankar, P. A. Nithya, G. Jahnavi, et al. (2022). Diabetic Retinopathy Detection Using MobileNetV2 Architecture, 2022 International Conference on Smart Technologies and Systems for Next Generation Computing, 1-5.
21. Paul B., S. Phadikar (2024). A Hybrid Feature-extracted Deep CNN with Reduced Parameters Substitutes an End-to-end CNN for the Recognition of Spoken Bengali Digits, Multimedia Tools and Applications, 83(1), 1669-1692.
22. Qummar S., F. G. Khan, S. Shah, A. Khan, et al. (2019). A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection, IEEE Access, 7, 150530-150539.
23. Ren C., W. Liu, J. Li, Y. Cao, et al. (2019). Physical Activity and Risk of Diabetic Retinopathy: A Systematic Review and Meta-analysis, Acta Diabetologica, 56, 823-837.
24. Sanjana S., N. S. Shadin, M. Farzana (2021). Automated Diabetic Retinopathy Detection Using Transfer Learning Models, 5th International Conference on Electrical Engineering and Information Communication Technology, 1-6.
25. Tsang S. (2015). Review: Inception-v3-1st Runner Up (Image Classification) in ILSVRC 2015. [Accessed: 26-Nov-2018].
26. Vyawahare H., P. Shinde (2022). Age-related Macular Degeneration: Epidemiology, Pathophysiology, Diagnosis, and Treatment, Cureus, 14(9), e29583.
27. Wilkinson C. P., F. L. Ferris III, R. E. Klein, P. P. Lee, C. D. Agardh, M. Davis, et al. & Global Diabetic Retinopathy Project Group (2003). Proposed International Clinical Diabetic Retinopathy and Diabetic Macular Edema Disease Severity Scales, Ophthalmology, 110(9), 1677-1682.
28. https://www.kaggle.com/c/aptos2019-blindness-detection/data /Diabetic Retinopathy Dataset APTOS obtained from Kaggle. (access date 04 March 2025).
29. Pittu V. P., S. R. Avanapu, J. Sharma (2013). Diabetic Retinopathy – Can Lead to Complete Blindness, International Journal of Science Inventions Today, 2(4), 254-265.

Abini M. A., Ph.D. Student
E-mail: abinima87@gmail.com

Abini M. A. received her M.Tech. degree in Very-large-scale integration (VLSI) and Embedded System from TocH Institute of Science and Technology, Arakkunnam, India, in 2011 and completed her B.Tech. degree in Electronics and Communication from Kerala Muslim Educational Association (KMEA) Engineering College, Aluva, India, in 2009. She is a Research Scholar at Karunya Institute of Technology and Sciences, Coimbatore, India. She is currently working as an Assistant Professor in Department of Electronics and Communication Engineering (ECE) at KMEA Engineering College, Aluva, India. Her research interests include deep learning, image processing and artificial intelligence. She is a life member of Institution of Electronics and Telecommunication Engineers (IETE), India and an associate member of Institution of Engineers (IE), India.

S. Sridevi Sathya Priya, Ph.D.
E-mail: sridevi@karunya.edu.in

Dr. S. Sridevi Sathya Priya received a B.Sc. degree of Engineering in Electronics and Communication Engineering in 2001 from Madras University, Chennai, India and M.Sc. of Engineering in Very-large-scale integration (VLSI) Design from Karunya Institute of Technology and Sciences, Coimbatore, India, in 2006. She completed her Ph.D. degree in Electronics and Communication Engineering under Karunya Institute of Technology and Sciences, Coimbatore, India, in 2017. She is a member of Institute of Electrical and Electronics Engineers (IEEE). She is currently working as an Assistant Professor in Electronics and Communication Engineering Department, Karunya Institute of Technology and Sciences, Coimbatore, India. She is a member of International Association of Engineers (IAENG). Her research work includes hardware implementation of security algorithms for various applications. She has more than 10 international publications and more than 25 publications in various conference proceedings.