ERURJ **2025**, *x*, x, Accepted
(Article)

Detection of Diabetic Retinopathy Using Deep Learning Techniques

Amira Mofreh[1], Aya Adel[1*], Norhan Elnady[1], Mahmoud Khaled[1]

[1] *Faculty of Artificial Intelligence, Egyptian Russian University, Cairo, Egypt*

*Corresponding author: Aya Adel, E-mail: Aya-adelmahmoud@eru.edu.eg, Tel: +201068123204

Received 20th March 2025, Accepted 11th May 2025

DOI:10.21608/erurj.2025.369875.1238

ABSTRACT

Diabetic retinopathy (DR) is a significant problem of diabetes, leading to vision impairment and blindness if left untreated. Early detection is crucial for effective intervention. This paper uses deep learning methods to detect DR from retinal fundus images automatically. Five pretrained convolutional neural network (CNN) architectures, including VGG16, ResNet50, InceptionV3, MobileNet, and DenseNet121, were modified, retrained, and evaluated on a standard dataset. Different evaluation metrics such as accuracy, ROC, and F1-score were used to evaluate model performance. The dataset used in this project is sourced from Roboflow and is designed to detect diabetic retinopathy. The dataset is divided to training, validation, and testing with 70%, 20%, and 10% respectively. Results demonstrated that the DenseNet121 model can effectively detect DR, with the best-performing model achieving accuracy (AC), precision (PR), recall (RC), false positive rate (FPR), F1-score (F1), and ROC curve (AUC) of 93%, 91.60%, 92.25%, 6.45%, 93%, 0.98% respectively. This paper discusses these findings' implications and suggests future research directions.

Keywords: Diabetic retinopathy, CNN, Detection, Resnet50, DenseNet121.

1-Introduction

Diabetic retinopathy (DR) is a major hurdle of diabetes mellitus, affecting millions of individuals worldwide. It is characterized by damage to the retinal blood vessels, which can lead to vision impairment and blindness if not detected early. It is very important to detect and treat it Early for preventing many visual injuries and stating the quality of life for diabetic patients. The dependence on traditional screening methods, which usually require specialized medical personnel and equipment, highlights the need for more accessible and efficient diagnostic tools [1].

Recently, advancements in deep learning and image processing techniques have shown great promise in automating the detection of diabetic retinopathy. Convolutional neural networks (CNNs) have emerged as a particularly effective approach to image classification tasks, including medical image analysis. By leveraging large datasets of retinal images, these models can learn to identify patterns indicative of diabetic retinopathy, potentially improving the speed and accuracy of diagnosis [2]. However, the performance of these models can vary extensively based on the architecture used and the quality of the training data [3].

Despite the progress in automated detection methods, several challenges persist. One of the primary issues is the variability in the quality and labeling of the datasets used for training, which can lead to overfitting and poor generalization to unseen data. Additionally, many existing studies focus on a limited number of models without comprehensive comparisons, making it difficult to determine the most effective approach for DR detection [4]. This underscores the need for a systematic evaluation of multiple architectures on standardized datasets.

The increasing prevalence of diabetic retinopathy has led to significant advancements in automated classification systems that leverage deep learning techniques. These systems aim to enhance diagnostic accuracy and assist healthcare professionals in making informed decisions. A variety of methodologies have been proposed in recent literature, showcasing the application of different deep learning architectures and datasets.

Kanika Verma et al [5] proposed a machine learning model for diabetic retinopathy different stages classification , the proposed model has built on random forest technique based on the area and perimeter of the blood vessels and hemorrhages with 90% accuracy. Darshit Doshi et al [6] developed a GPU accelerated deep CNN to automatically classify high-resolution retinal images , the single model accuracy is 0.386 on a quadratic weighted kappa metric and ensembling of three such similar models resulted in a score of 0.3996.

Carson Lam et al [7] introduced a diabetic retinopathy classification framework consisting of three deep learning models, CNN , GoogleNet and AlexNet, the testing results showed the superiority of GoogleNet with accuracy of 74.5%. Mohamed Chetoui et al [8] proposed a framework showed the use of different texture features for diabetic retinopathy , the proposed framework has been built on Local Ternary Pattern (LTP) ,Local Energy-based Shape Histogram (LESH), Local Binary Pattern (LBP) and Support Vector Machine (svm) , the expermintal results showed that LESH is the best technique with accuracy of 90.4%.

Quang H. Nguyen et al [9] proposed an automated classification system analyzes fundas images of diabetic retinopathy with varying illumination, The proposed system has been built on CNN, VGG-16 and VGG-19 with 82% classification accuracy and 90.4% AUC. Shu-I Pao et al [10] introduced a diabetic retinopathy detection framework , the authors used the bichannel CNN to incorporate the features of both the entropy images of gray level and green level component preprocessed by UM , the model achieved accuracy of 87.8%. Akhilesh Kumar Gangwar et al [11] proposed a novel deep learning hybrid model , the authors developed this model using Messidor-1 diabetic retinopathy dataset and APTOS 2019 blindness detection (Kaggle dataset), the propsed model has been built on Inception-Resnet-v2 and added a custom block of CNN layers on the top Inception-Resnet-v2 ,The model achieved a test accuracy of 72.33%.

Ling Dai et al [12] developed a deep learning system named DeepDr to detect early-to-late stages of diabetic retinopathy, this developed model trained for real-time image quality assesment, lesion detection and grading using 466,247 fundus images from 121,342 patients with diabetes and achieved AUC of 90.1%. CNNs have revolutionized image classification jobs. A CNN is a type of deep neural network that extracts relevant characteristics from pictures and learns hierarchical representations. A CNN has many major layers that play an important role in extracting information from an input dataset.

This paper aims to address these challenges by modifying five different deep learning pretrained models are VGG16, ResNet50, InceptionV3, MobileNet, and DenseNet121 on the available standard dataset for diabetic retinopathy detection. This is work will analytically compare their performance using accuracy, precision, recall, and F1-score, to provide a comprehensive understanding of their effectiveness.

The main objective of the CNN models is to achieve the best performance results in image detection. Therefore, in our work, we concentrated on these key contributions to obtain better results than those of other related works:

1. Five different CNN models are modified with different layers of CNNs, which are named VGG16, ResNet50, InceptionV3, MobileNet, and DenseNet121.
2. The architecture uses depthwise separable convolutions to significantly reduce the number of parameters compared to standard convolutions.
3. The results of these designed classification models were compared to determine the optimal model with the highest accuracy for diabetic retinopathy detection.
4. The performance results of the best model were compared with those in the literature.

The remainder of this paper is organized as follows. Section 2 describes the materials and methods used in the study, including dataset preparation and model architecture. In Section 3 testing environment based on the dataset used and the performance metrics used is discussed, Section 4 presents the results of the experiments and a detailed discussion of the findings. Finally, Section 5 concludes the paper conclusions, summarizing the key insights and suggesting directions for future research.

2. Experimental

Convolutional Neural Networks, or CNN, can play a crucial role in picture categorization. Its approach is built on employing several tiny filters to specify the characteristics in each layer. CNN is made up of several sequential layers, each of which may extract certain information from the input picture. CNN's primary layers are the convolutional layer, sampling layer, activation layer, and fully connected layers. Deep feature extraction involves extracting features from certain pre-trained CNN models.

Our model consists of two stages as shown in Fig.1.The procedure starts with inputting our dataset. The photos are then preprocessed using techniques such as scaling and augmentation. Various CNN architectures, including VGG16, ResNet50, InceptionV3, MobileNet, and

DenseNet121, are utilized to extract and categorize features from this dataset. Finally, choose the optimal architecture that will improve our model's performance.

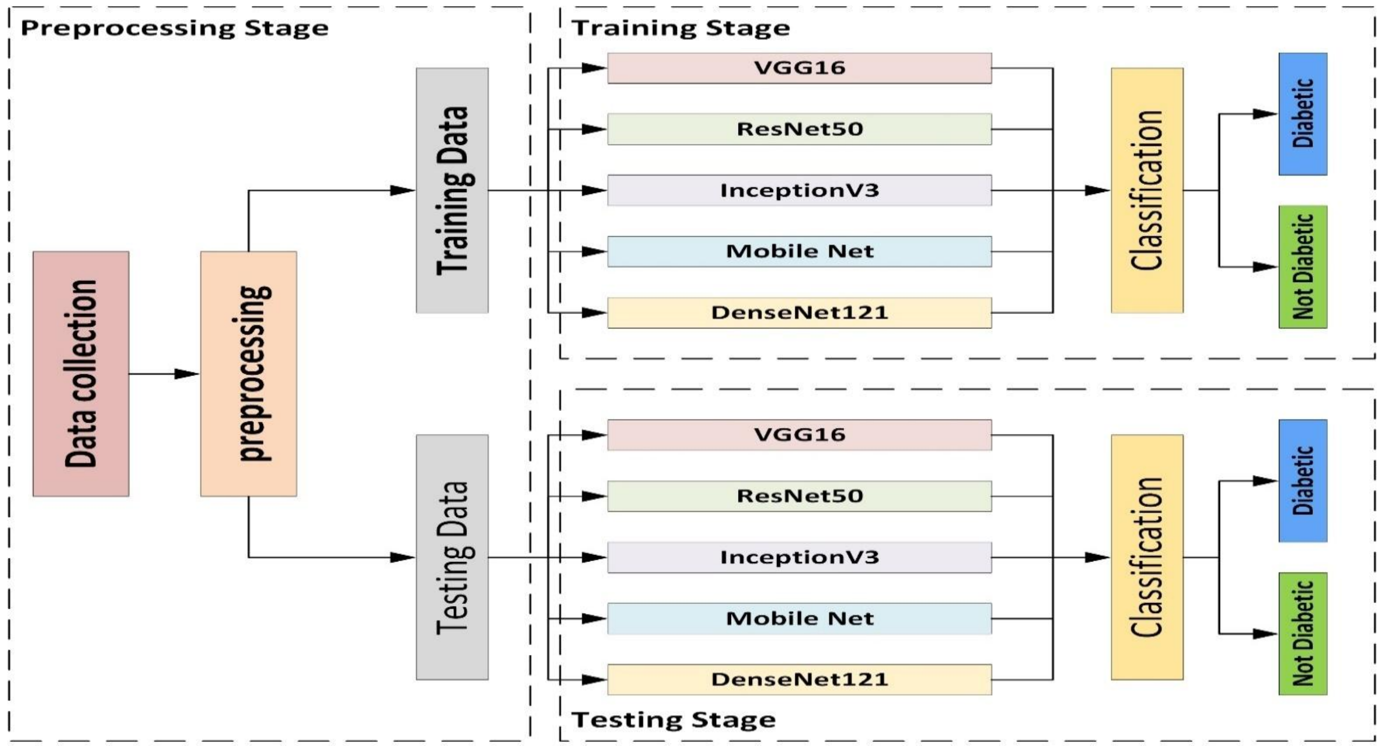

Figure 1. The overall structure of proposed model

2.1 Data Preprocessing Techniques

Auto-Orient: This technique was applied to ensure that all images were oriented correctly before training. Auto-orientation helps in standardizing the dataset, allowing the model to learn effectively from images without facing issues related to incorrect orientations that could confuse the training process. This preprocessing step is crucial as it strips images of their EXIF data, ensuring consistent display and interpretation during model training [13].

Resize: All images were resized to a uniform dimension of 640x640 pixels. This resolution is crucial because it ensures that each input to the CNN has the same dimensions, facilitating batch processing and improving the model's efficiency. The stretching method was employed, maintaining aspect ratios where feasible, to prevent significant distortion that might impact the model's ability to recognize features accurately [14].

Preprocessing is a crucial step in preparing the dataset for training machine learning models [13]. In this section, outline the preprocessing techniques applied for six different models, including a hybrid approach that combines the strengths of these models will be discussed in Table 1.

Table 1. The proposed models preprocessing techniques.

Techniques	VGG16	ResNet50	InceptionV3	MobileNet	DenseNet121
Resizing	224x224	224x224	299x299	224x224	224x224
Normalization	Using the ImageNet statistics	Using the mean and standard deviation	Using the ImageNet statistics	dividing by 255	dividing by 255
Augmentation	Standard augmentation	rotation, width/height shifts, zoom, and horizontal flips	variety of transformations including random cropping and aspect ratio adjustments	rotation and flipping	rotation, width/height shifts, zoom, and horizontal flips

2.2 Method selection

Five Convolutional Neural Network (CNN) architectures were utilized in this study: VGG16, ResNet50, InceptionV3, MobileNet, and DenseNet121. These architectures were selected due to their proven effectiveness in image classification tasks and their varying complexities, which allow for a comprehensive evaluation of performance across different model types. Each model was implemented using TensorFlow and Keras, popular frameworks for deep learning that facilitate the development and training of neural networks. **Convolutional Neural Network Architectures.**

VGG16: This architecture is known for its simplicity and effectiveness in image classification tasks. It employs a deep network with 16 layers, consisting of convolutional and fully connected layers. VGG16 uses small receptive fields (3x3 convolutions) and has demonstrated exceptional performance in various image recognition challenges. Its depth allows it to learn complex features, making it suitable for detecting subtle patterns in diabetic retinopathy images[15].

ResNet50: ResNet50 introduces the concept of residual learning, enabling the training of very deep networks by using skip connections. This architecture consists of 50 layers and addresses the vanishing gradient problem, allowing for the effective training of models with considerable depth. ResNet50 is particularly powerful in preserving detailed information from earlier layers, which is essential for accurately classifying images with varying degrees of diabetic retinopathy[16].

InceptionV3: This architecture is designed to improve computational efficiency while maintaining high accuracy. InceptionV3 employs multiple filter sizes in parallel, allowing the model to learn features at different scales. The use of auxiliary classifiers during training helps in combating the vanishing gradient problem, making it effective for complex image classification tasks, including medical image analysis.[17]

MobileNet: MobileNet is optimized for mobile and edge devices, providing a lightweight solution for image classification tasks. It uses depthwise separable convolutions, which significantly reduce the number of parameters and computations required, making it ideal for performance-constrained environments. MobileNet's efficiency allows it to deliver fast inference times while retaining competitive accuracy, crucial for real-time diabetic retinopathy screening.[18]

DenseNet121: This architecture connects each layer to every other layer in a feed-forward fashion, promoting feature reuse and reducing the number of parameters. DenseNet121 consists of 121 layers and is known for its high accuracy in image classification tasks. Its ability to concatenate features from multiple layers helps it learn richer representations, which is beneficial for detecting the subtle signs of diabetic retinopathy in retinal images.[19]

Table 2. The proposed model's architecture modification.

Techniques	VGG16	ResNet50	InceptionV3	MobileNet	DenseNet121
Convolutional Layers	√	√		√	√
Inception Modules			√		
Global Average Pooling Layer		√	√	√	√
Max-Pooling Layers	√				
Transition Layers					√
Batch Normalization Layers		√			

Fully Connected Layer	√	√	√	√	√
Softmax Activation Layer	√	√	√	√	√

When preparing the five models—MobileNet, DenseNet121, VGG16, ResNet50, and InceptionV3 for training, specific layers are added as shown in Table 2 to facilitate effective learning and enhance classification performance.

These layers are strategically added to each model to optimize their ability to learn from the training data while ensuring adaptability to various image classification tasks. Proper configuration of these layers is essential for achieving high accuracy and efficiency during training.

3. Testing Environment

3.1 Dataset description

The dataset used in this project is sourced from Roboflow [20] and is designed for the detection of diabetic retinopathy. It consists of labeled images categorized into different classes that represent various stages of diabetic retinopathy. The images are collected from a diverse set of patients to ensure a wide range of examples and variability in the data. The dataset is divided into three main subsets as shown in Table 3. We get a more accurate estimate of the model's predicted performance on unseen data. A doctor assessed the presence of DR in each image. Figure 2 illustrates how we categorize the supplied dataset.

Table 3. The dataset is divided into three main subsets.

Type	Training	Validation	Testing	Total
Normal	1104	325	186	1615

DR	1196	333	142	1671
Total	2300	658	328	3286

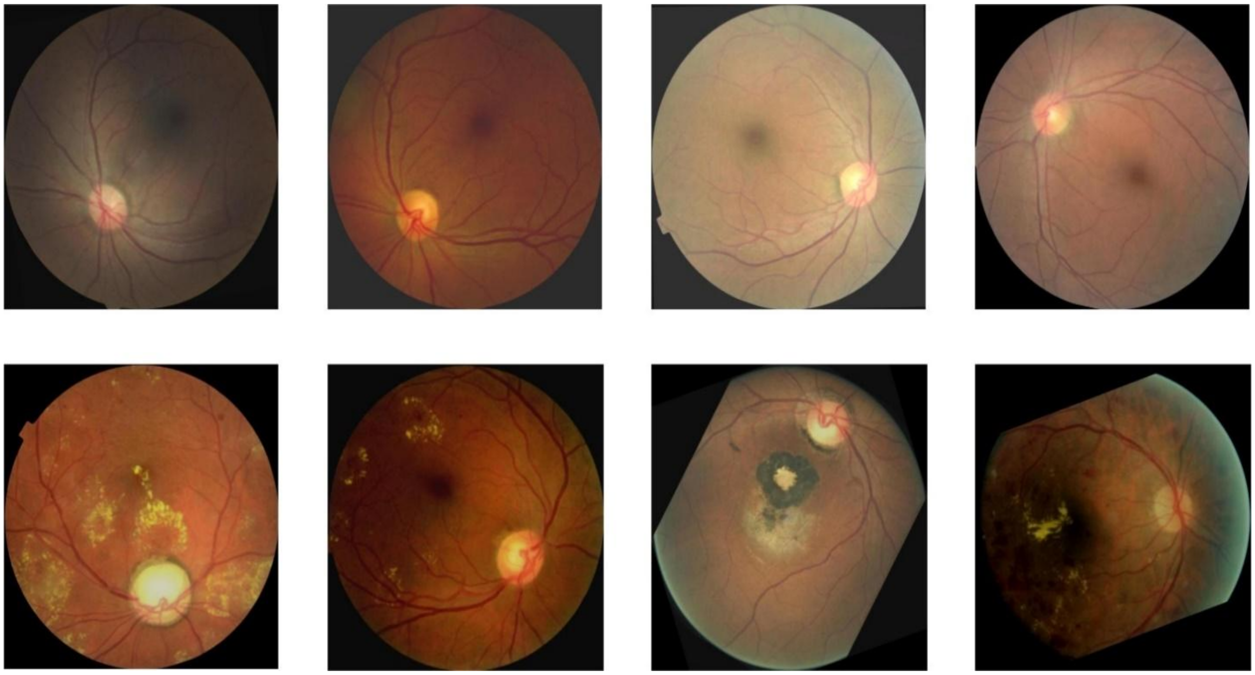

Figure 2. Different images with Normal images in first raw and DR images in the second raw

.

3.2 Model Training

Table 4 shows hyperparameters utilized in training the proposed models. The model training process involved utilizing five distinct convolutional neural network architectures: VGG16, ResNet50, InceptionV3, MobileNet, and DenseNet121, each model was trained on the prepared training dataset, which comprised images resized to 224x224 pixels.

To ensure robust performance, the validation set was used to monitor each model's accuracy throughout the training process. This validation allowed for real-time adjustments to hyperparameters, such as learning rate and batch size, optimizing each model's performance. The training duration was standardized across all models to provide a fair comparison.

After completing the training phase, each model was evaluated on a separate test set that had not been seen during training. This evaluation measured the accuracy to assess each model's performance comprehensively. By comparing the results, the paper identifies the strengths and weaknesses of each architecture, ultimately leading to insights about their effectiveness in the context of diabetic retinopathy detection.

Table 4. The proposed models training hyperparameters.

Hyper parameter	VGG16	ResNet50	InceptionV3	MobileNet	DenseNet121
Loss function	categorical cross-entropy	categorical cross-entropy	categorical cross-entropy	categorical cross-entropy	categorical cross-entropy
Optimizer	Adam	Adam	RMSprop	Adam	SGD
Epochs	50	75	50	50	100
Batch size	32	32	32	32	16
Learning rate	0.001	0.001	0.001	0.001	0.01

3.3 Environment

The experiments were conducted in a cloud-based environment using Kaggle [21], which provides robust GPU resources for deep learning tasks. Kaggle's platform is designed for data science and machine learning competitions, offering a user-friendly interface and integrated tools that facilitate the entire workflow from data preprocessing to model evaluation.

In this paper, the models were trained on a GPU, which significantly accelerated the training process compared to CPU-only training. The specific GPU model used in this environment was the NVIDIA Tesla P100, known for its high computational power and efficiency in handling deep learning tasks. This GPU features 16 GB of memory, allowing for the training of complex architectures such as VGG16, ResNet50, InceptionV3, MobileNet, and DenseNet121 without encountering memory limitations.

The Kaggle environment also supports the installation of various libraries essential for deep learning, including TensorFlow and Keras. These libraries provide a comprehensive suite of tools for building and training neural networks, offering pre-trained models and easy access to advanced functionalities such as data augmentation and model evaluation metrics.

Overall, the Kaggle platform provides an ideal environment for conducting experiments in deep learning, combining powerful computational resources with a collaborative and accessible interface, thus enhancing the efficiency and effectiveness of the model development process.

3.4 Performance metrics

In the experiments presented in this research, different performance measures [22-23] are used to assess the classification performance of the presented models and the comparative models. They are: accuracy (AC), precision (PR), recall (RC) or true positive rate (TPR), false positive rate (FPR), F1-score (F1), and ROC curve (AUC). Equations 1-5 express the computations for these metrics.

$$\mathrm{AC} = \frac{T_{PV}+T_{NV}}{T_{PV}+F_{NV}+T_{NV}+F_{PV}} \quad (1)$$

$$\mathrm{PR} = \frac{T_{PV}}{T_{PV}+F_{PV}} \quad (2)$$

$$\mathrm{RC} = \mathrm{TPR} = \frac{T_{PV}}{T_{PV}+F_{NV}} \quad (3)$$

$$\mathrm{F1} = 2 * \frac{RC * PR}{RC + PR} \quad (4)$$

$$\mathrm{FPR} = \frac{F_{PV}}{F_{PV}+T_{NV}} \quad (5)$$

where the confusion matrix parameters are T_{PV} is the number of true positives, F_{PV} is the number of false positives, T_{NV} is the number of true negatives, and F_{NV} is the number of false negatives [24].

The ROC curve is created by calculating the true positive rate (TPR) and false positive rate (FPR) at each potential threshold (in practice, at predetermined intervals) and then plotting TPR vs FPR. A perfect model with a TPR of 1.0 and an FPR of 0.0 at a certain threshold.

4. Results

In this paper, evaluated five convolutional neural network architectures for image classification in diabetic retinopathy detection are introduced. The Confusion matrix for each model is shown in Table5. The performance measures for each model are shown in Table 6.

Table 5. The Confusion matrix for each model

Model	T_{PV}	T_{NV}	F_{PV}	F_{NV}
VGG16	142	0	186	0
ResNet50	132	14	172	10
InceptionV3	142	0	186	0
MobileNet	140	161	25	2
DenseNet121	131	174	12	11

Table 6. The performance measures for each model

Model	PR	RC (TPR)	FPR	AC	F1 Score	ROC (AUC)
VGG16	43.29	1	1	43.00	26%	0.55
ResNet50	43.42	92.95	92.47	45.00	33%	0.59
InceptionV3	43.29	1	1	43.00	26%	0.55
MobileNet	84.84	98.59	13.44	92.00	92%	0.99
DenseNet121	91.60	92.25	6.45	93.00	93%	0.98

The results indicate a significant change in performance among the models. VGG16 achieved relatively low accuracy and F1 score, indicating that it struggled to differentiate between classes effectively. The ROC AUC of 55% suggests that the model has limited capability in distinguishing between the positive and negative classes, likely due to overfitting or inadequate feature extraction for this specific dataset. ResNet50 performed slightly better than VGG16, but still showed limited effectiveness. The higher F1 score indicates improved precision and recall, but the overall performance remains unsatisfactory. The architecture may not have captured the nuances of the dataset, leading to a similar issue of overfitting or insufficient training data. InceptionV3's results mirror those of VGG16, with similarly low metrics across the board. This suggests that the model might not be well-suited to the specific characteristics of the dataset. The complex architecture of InceptionV3, while powerful, may require more extensive tuning or a larger dataset to perform effectively. MobileNet demonstrated outstanding performance, achieving high accuracy and an impressive F1 score. The ROC AUC of 99% indicates that MobileNet is highly effective at distinguishing between classes. This model's lightweight architecture is particularly advantageous for image classification tasks, allowing it to generalize well even with potentially limited data. DenseNet121 also performed exceptionally well, closely matching MobileNet's metrics. The architecture's dense connectivity pattern likely facilitated better feature reuse, enabling the model to learn richer representations from the images. The slightly lower ROC AUC compared to MobileNet suggests room for improvement in classification confidence.

Overall, MobileNet and DenseNet121 proved to be the most effective models for the task, likely due to their ability to learn and generalize from the dataset.

The ROC curves serve as a critical evaluation tool for assessing their classification capabilities. The area under the ROC curve (AUC) quantifies the model's ability to distinguish between classes, with values closer to 1 indicating superior performance. For instance, MobileNet and DenseNet121 are anticipated to exhibit high AUC values as shown in Figure 3, reflecting their proficiency in correctly classifying positive and negative instances of diabetic retinopathy with minimal overlap. By comparing these ROC curves, one can gauge the trade-offs between sensitivity and specificity for each model, facilitating informed decisions on which architecture to deploy for optimal diagnostic accuracy.

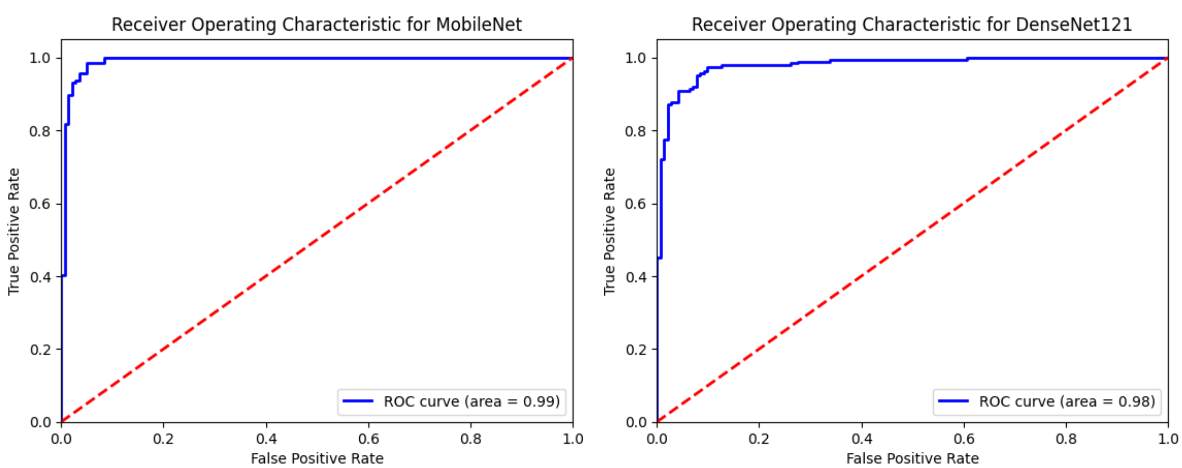

Figure 3. The ROC curve of MobileNet on the left and DenseNet121on the right.

Overall, these findings highlight the importance of selecting appropriate architecture for specific image classification tasks. The superior performance of DenseNet121 and MobileNet. underscores its suitability for complex image analysis in medical applications, making it a promising choice for future work in diabetic retinopathy detection.

Table 7 presents a comparative analysis of various models in the literature utilized for diabetic retinopathy detection. The recommended models, MobileNet and DenseNet121, achieved impressive accuracies of 92.00% and 93.00%, respectively, demonstrating their effectiveness in extracting and classifying features from medical images. In contrast, earlier studies show a range of performances, with Kanika Verma et al. (2011) achieving 90% accuracy using traditional machine learning techniques. Subsequent advancements in deep learning are illustrated by Darshit Doshi et al. (2016) with a GPU-accelerated deep CNN scoring 0.3996, and Carson Lam et al. (2018) utilizing GoogleNet, which yielded a lower accuracy of 74.5%. Notably, Mohamed Chetoui et al. (2018) reported an accuracy of 90.4% with their LESH model, while Quang H. Nguyen et al. (2020) combined CNN, VGG-16, and VGG-19 to achieve 82% classification accuracy and a 90.4% AUC. Further contributions by Shu-I Pao et al. (2020) and Akhilesh Kumar Gangwar et al. (2021) reflect the ongoing evolution of model architectures, with accuracies of 87.8% and 72.33%, respectively. Ling Dai et al. (2021) introduced a deep learning system named DeepDr, achieving an AUC of 90.1%. This table underscores the significant progress in model performance over the years, highlighting the advantages of contemporary architectures in improving diagnostic accuracy.

Table7. Comparison of the proposed detection model against literature.

Ref	Year	Method	Evaluation metrics (%)					
			PR	RC (TPR)	FPR	AC	F1 Score	ROC (AUC)
Kanika Verma et al [5]	2011	Machine learning	1	87.5	0	90	-	-
Darshit Doshi et al [6]	2016	deep CNN	-	-	-	30	-	-
Carson Lam et al [7]	2018	GoogleNet	95.05	91.5344	4.26	74.5	-	-
Mohamed Chetoui et al [8]	2018	LESH	-	-	-	90.4	-	0.93
Quang H. Nguyen et al [9]	2020	CNN, VGG-16 and VGG-19	--	-	-	82	-	0.904
Shu-I Pao et al [10]	2020	CNN UM		77.81	6.12	87.8	-	0.93
Akhilesh Kumar et al [11]	2021	Inception-Resnet-v2	-	-	-	72.33	-	-
Ling Dai et al [12]	2021	DeepDr	-	76.2	12.8	-	-	0.934
Proposed Mode	**2025**	**MobileNet**	**84.84**	**98.59**	**13.44**	**92.00%**	**92%**	**0.99**
Proposed Model	**2025**	**DenseNet121**	**91.60**	**92.25**	**6.45**	**93.00%**	**93%**	**0.98**

5. Discussion

MobileNet and DenseNet121 achieved the highest accuracy in the classification tasks due to several key factors related to their architectures and design philosophies.

1. **Efficient Feature Extraction**: Both models are built on advanced convolutional neural network (CNN) architectures that excel at hierarchical feature extraction. MobileNet

utilizes depthwise separable convolutions, which reduce the number of parameters and computations while maintaining the ability to learn rich feature representations. This efficiency allows MobileNet to perform well even with limited data, making it suitable for various tasks, including medical image classification. DenseNet121, on the other hand, employs densely connected layers that facilitate feature reuse throughout the network. This architecture enables the model to capture complex patterns in the data more effectively, leading to superior performance in identifying subtle variations in images.

2. **Robust Training Techniques**: Both models benefit from robust training techniques such as data augmentation and transfer learning. By augmenting the training data, these models can generalize better to unseen images, reducing overfitting and improving their ability to classify images accurately. Additionally, DenseNet121's architecture allows for better gradient flow during training, which helps in converging to optimal weights more efficiently. This, combined with the use of pre-trained weights on large datasets, allows both models to leverage learned features that enhance their performance on specific tasks like diabetic retinopathy detection.
3. **Adaptability to Diverse Datasets**: MobileNet and DenseNet121 are designed to adapt well to various datasets, benefiting from their flexibility in handling different image sizes and resolutions. This adaptability ensures that they can perform well across a range of medical imaging scenarios, capturing the essential features necessary for accurate classification. Their ability to maintain high accuracy even under different conditions contributes to their effectiveness in real-world applications, where variability in image quality and content is common.

In summary, the combination of efficient architecture, robust training techniques, and adaptability to diverse datasets explains why MobileNet and DenseNet121 achieved the highest accuracy among the models evaluated. These factors enable them to excel in classifying complex medical images, ultimately leading to better diagnostic outcomes.

6. Conclusion and Future Work

This paper demonstrated the effectiveness of various convolutional neural network architectures for diabetic retinopathy detection, revealing significant differences in accuracy. Five pretrained convolutional neural network (CNN) architectures, VGG16, ResNet50, InceptionV3, MobileNet, and DenseNet121, were customized, retrained, and assessed on a standard dataset. To assess model performance, many assessment criteria were utilized, including PR, TPR, FPR, Accuracy, Roc, and F1-score. DenseNet121 emerged as the most effective model, achieving a remarkable accuracy of 93%. This emphasizes the critical role of architecture selection in enhancing diagnostic accuracy in healthcare applications. Future work should focus on several areas to further improve model performance. Exploring ensemble methods could leverage the strengths of multiple architectures, while hyperparameter tuning may optimize the performance of other models. Additionally, investigating transfer learning and data augmentation techniques could enhance the robustness of the models, particularly in scenarios with limited data. Overall, the findings of this paper highlight the potential of advanced deep learning techniques in improving diagnostic accuracy in medical imaging, paving the way for more effective tools in the detection and management of diabetic retinopathy and other similar conditions.

- **Conflict of Interest**

  The authors declare no conflict of interest.

7. References

[1] Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA. 2016 Dec 13;316(22):2402-2410. doi: 10.1001/jama.2016.17216. PMID: 27898976.

[2] Dai, L., Wu, L., Li, H., Cai, C., Wu, Q., Kong, H., Liu, R., Wang, X., Hou, X., Liu, Y., Long, X., Wen, Y., Lu, L., Shen, Y., Chen, Y., Shen, D., Yang, X., Zou, H., Sheng, B., & Jia, W. (2021). A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nature communications, 12(1), Article 3242. https://doi.org/10.1038/s41467-021-23458-5

[3] Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M. Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep Learning. Invest Ophthalmol Vis Sci. 2016 Oct 1;57(13):5200-5206. doi: 10.1167/iovs.16-19964. PMID: 27701631.

[4] S. J. Nagar and N. Gondaliya, "A Comprehensive Review on Diabetic Retinopathy Detection Techniques using Neural Network Architectures," *Journal of Electrical Systems*, vol. 20, no. 10s, 2024. doi: 10.52783/jes.5309.

[5] K. Verma, P. Deep and A. G. Ramakrishnan, "Detection and classification of diabetic retinopathy using retinal images," 2011 Annual IEEE India Conference, Hyderabad, India, 2011, pp. 1-6, doi: 10.1109/INDCON.2011.6139346.

[6] D. Doshi, A. Shenoy, D. Sidhpura and P. Gharpure, "Diabetic retinopathy detection using deep convolutional neural networks," 2016 International Conference on Computing, Analytics and Security Trends (CAST), Pune, India, 2016, pp. 261-266, doi: 10.1109/CAST.2016.7914977

[7] Lam C, Yi D, Guo M, Lindsey T. Automated Detection of Diabetic Retinopathy using Deep Learning. AMIA Jt Summits Transl Sci Proc. 2018 May 18;2017:147-155. PMID: 29888061; PMCID: PMC5961805.

[8] M. Chetoui, M. A. Akhloufi and M. Kardouchi, "Diabetic Retinopathy Detection Using Machine Learning and Texture Features," 2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE), Quebec, QC, Canada, 2018, pp. 1-4, doi: 10.1109/CCECE.2018.8447809.

[9] Q. H. Nguyen, R. Muthuraman, L. Singh, G. Sen, A. C. Tran, B. P. Nguyen, and M. Chua, "Diabetic Retinopathy Detection using Deep Learning," in ICMLSC '20: Proceedings of the 4th International Conference on Machine Learning and Soft Computing, pp. 103–107, 2020. DOI: 10.1145/3380688.3380709

[10]S.-I. Pao, H.-Z. Lin, K.-H. Chien, M.-C. Tai, J.-T. Chen, and G.-M. Lin, "Detection of Diabetic Retinopathy Using Bichannel Convolutional Neural Network," Computational and Mathematical Methods in Medicine, vol. 2020, Article ID 9139713, 2020. DOI: 10.1155/2020/9139713.

[11] Gangwar, A.K., Ravi, V. (2021). Diabetic Retinopathy Detection Using Transfer Learning and Deep Learning. In: Bhateja, V., Peng, SL., Satapathy, S.C., Zhang, YD. (eds) Evolution in Computational Intelligence. Advances in Intelligent Systems and Computing, vol 1176. Springer, Singapore. https://doi.org/10.1007/978-981-15-5788-0_64.

[12] Dai, L., Wu, L., Li, H. et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nat Commun 12, 3242 (2021). https://doi.org/10.1038/s41467-021-23458-5.

[13] INC, HP, "AUTO CORRECTION OF ORIENTATION OF IMAGES WITH DEEP LEARNING", Technical Disclosure Commons, (July 01, 2022) https://www.tdcommons.org/dpubs_series/5233.

[14] Parsania, Pankaj & Virparia, Paresh. (2015). A Review: Image Interpolation Techniques for Image Scaling. International Journal of Innovative Research in Computer and Communication Engineering. 02. 7409-7414. 10.15680/IJIRCCE.2014.0212024.

[15] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in Proceedings of the International Conference on Learning Representations (ICLR), 2015.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778.

[17] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," in Proceedings of the AAAI Conference on Artificial Intelligence, 2017.

[18] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv preprint arXiv:1704.04861, 2017.

[19] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4700-4708.

[20] Roboflow, “DR Detection_4 Dataset”, Roboflow Universe, march 2023, https://universe.roboflow.com/dr-mvj3k/dr-detection_4, visited on 2025-03-17

[21] Kaggle Inc. (2023). Kaggle: Your home for data science. Retrieved February 22, 2025, from https://www.kaggle.com

[22] Abd-Ellah, M.K., Awad, A.I., Khalaf, A.A.M. et al. Automatic brain-tumor diagnosis using cascaded deep convolutional neural networks with symmetric U-Net and asymmetric residual-blocks. Sci Rep 14, 9501 (2024). https://doi.org/10.1038/s41598-024-59566-7

[23] Salama, G.M., Mohamed, A. & Abd-Ellah, M.K. Machine learning and deep learning covid-19 diagnosis system: key achievements, lessons learned, and a transfer learning algorithm. Soft Comput 28, 13715–13742 (2024). https://doi.org/10.1007/s00500-024-10362-5

[24] Salama, G.M., Mohamed, A. & Abd-Ellah, M.K. COVID-19 classification based on a deep learning and machine learning fusion technique using chest CT images. Neural Comput & Applic 36, 5347–5365 (2024). https://doi.org/10.1007/s00521-023-09346-7