ORIGINAL ARTICLE https://doi.org/10.52206/jsmc.2025.15.2.1149 OPEN ACCESS

Hybrid Deep Learning Model for Diabetic Retinopathy Severity Detection and Classification

Qammar Mehmood,[1] Abdullah Amer,[1] Moazam Ali,[1] Hafiz Ishfaq Ahmad,[1] Aafaque Ahmad[2]

1. Bahria University, Islamabad, Pakistan
2. Saidu Group of Teaching Hospitals, Swat, Pakistan

ABSTRACT

Background: A leading cause of vision loss and common complication in patients with diabetes is Diabetic retinopathy. The detection of these complications can be done in the early stages of the disease and can be managed through non-invasive retinal image analysis. However, traditional manual detection methods have low accuracy and high error rates, making it difficult to reliably identify key indicators like microaneurysms, hemorrhages, and exudates. This underscores the need for improved diagnostic methods.

Objectives: The objective of this research study is to propose a diagnostic system using deep learning models to improve the accuracy and reliability of diabetic retinopathy detection.

Materials and Methods: This study proposes an automated system for diabetic retinopathy detection using deep learning techniques. Two deep learning techniques, EfficientNet-B3 and ResNet18 CNN, were trained on retinal and non-retinal image datasets to identify indicators like microaneurysms, hemorrhages, and exudates.

Results: Our experiments demonstrated high accuracy in detecting and classifying diabetic retinopathy. The deep learning model achieved an accuracy of 98.18%, while the verification model reached 99%. These results highlight the system's ability to reliably identify critical features such as microaneurysms, hemorrhages, and exudates.

Conclusion: The automated system developed for diabetic retinopathy detection, powered by deep learnng models, significantly enhances diagnostic accuracy and efficiency. Its high performance in identifying key retinal indicators offers a practical solution for early detection and monitoring, especially in resource-limited areas. This approach shows potential for enhancing patient outcomes and alleviating the global impact of diabetic retinopathy by providing accessible and dependable diagnostic support.

Keywords: Diabetic retinopathy (DR), Deep learning models, EfficientNet-B, ResNet18, Medical image classification.

This article may be cited as: Mehmood Q, Amer A, Ali M, Ahmad HI, Ahmad A. Hybrid Deep Learning Model for Diabetic Retinopathy Severity Detection and Classification. *J Saidu Med Coll.* 2025;15(2):218-225. https://doi.org/10.52206/jsmc.2025.15.2.1149

INTRODUCTION

Diabetic Retinopathy (DR) is a severe diabetes-related complication that can cause substantial vision impairment and potentially lead to blindness. This disease damages the blood vessels of the retina, the light-sensitive layer at the back of the eye, leading to vision loss if left untreated. About 80% of patients with diabetes for over 10 years are affected by DR and causes blindness in these patients.[1] Two stages of DR in which the disease progresses are Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR). In the NPDR, blood vessels of the retina weaken and the starts leaking fluid and swelling appears in it.[1,2] In contrast, PDR is a more advanced stage of the disease in which the blood vessels become fragile leading to retinal detachment and vision loss.[1,3] It is essential to detect and monitor these retinal abnormalities as it may result in vision loss if untreated.

The timely diagnosis of Diabetic retinopathy (DR) is critical for the effective management of the disease. However, manual detection is prone to errors due to subtle retinal abnormalities and is resource-intensive, particularly in regions with limited access to skilled ophthalmologists. Existing automated methods have shown promise but often lack the accuracy or scalability required for clinical adoption. This study aims to address these gaps by investigating whether an integrated AI system combining advanced deep learning models can provide a more reliable and efficient solution for DR severity detection.

Various advanced machine learning and artificial intelligence models can be used for the detection of DR in early stages. It requires high resolution images captured with specialized cameras called the Fundus images. These images help medical practitioners to see the internal structures of the eye, including the retina, optic disc, and macula. Common retinal diseases require the detection of various features in the context of DR. These includes microaneurysms, hemorrhages, and exudates.[4] Microaneurysms, the earliest sign of DR, are small swellings on blood vessels that may leak blood. Retinal hemorrhages involve bleeding into the retina and can impair vision if undetected. Exudates, fluid leaks from blood vessels, appear as yellowish patches and are a major cause of visual loss in non-proliferative DR.[1,4,5]

In recent years, deep learning-based methods have gained substantial momentum in DR diagnosis, aiming to improve both accuracy and automation. One notable study focuses on Deep Transfer Learning-based Computerized Diagnostic Systems (DTL-CDS) to classify the severity of DR into multiple stages using the IDRiD dataset.[6,7] Several deep inductive transfer learning models, such as Inception V3, ResNet34, EfficientNet B0, VGG16, and Xception, were evaluated in this study. The goal was to optimize feature extraction techniques and classification layers, leading

Correspondence: Dr. Hafiz Ishfaq Ahmad
Saidu Group of Teaching Hospitals, Swat, Pakistan
Email: hiahmad.buic@bahria.edu.pk

Received: November 12th, 2024 - 1st Revision: January 01st 2025
2nd Revision: February 27th 2025 - Accepted: March 15th, 2025

to better performance. To further improve the model, a global average pooling layer was integrated to reduce overfitting while capturing essential features. The final classification was achieved using a softmax layer, and the model reached an impressive accuracy of 99.36%, with excellent metrics in precision, recall, F1-score, and AUC, setting a new benchmark in DR diagnosis.[6]

In another study, a hybrid diagnostic method combines multi-headed Convolutional Neural Networks (CNNs) with vision transformers to enhance DR diagnosis. 8 They used CNN and vision transformers in their research study where the CNN mines the local spatial features of the retinal images. However, the vision transformer is implemented for the extraction of contextual features. This combination of the CNN and vision transformers allows the hybrid model to detect and classify the stages of DR accurately.[2,8]

On the other hand, random forest algorithms focuses on using segmentation and feature extraction for classification in the non-invasive retinal imaging.[9] This classification model achieved up to 90% accuracy in distinguishing non-proliferative DR (NPDR) against the proliferative DR (PDR). Results of the study presented enhancement in the early detection of DR and providing a novel technique for effective diagnosis of retinal diseases.[9]

In recent years, deep learning has seen significant advancements in the domain of diabetic retinopathy detection.[1,2,7] Transfer learning has been widely adopted, allowing pre-trained models to excel in specialized tasks with limited datasets. Ensemble methods have also emerged, combining the strengths of multiple architectures to achieve higher robustness and accuracy.[10,11] Furthermore, interpretability techniques have gained traction, enabling clinicians to understand the reasoning behind AI decisions. By leveraging these advancements, the aim of this study is to improve the accuracy, reduce errors and design a tool for automatically detecting DR at its earliest stages for improving patient quality of life.

MATERIALS AND METHODS

This section explains the dataset used for the research study, various preprocessing techniques used and the techniques for the training of our mode. To enhance the clarity and understanding of the research process, a methodology diagram has been included (see Figure 01). This visual representation provides a structured overview of the research framework, illustrating the key phases, techniques, and tools employed throughout the study. Ethical approval was not required for this study as it did not involve human or animal subjects. The section ends with the evaluation of the proposed model for the early detection and classification of diabetic retinopathy (DR).

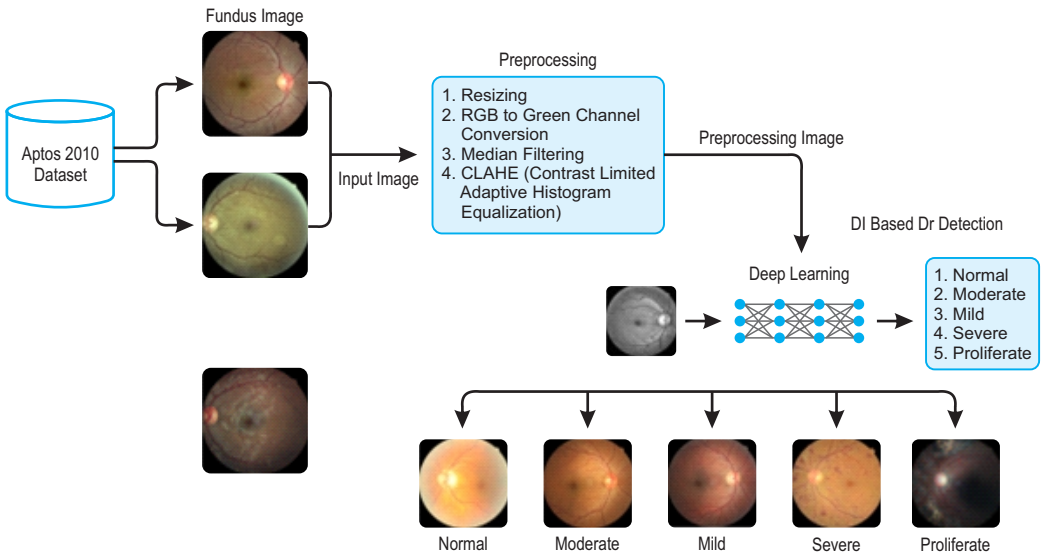


Figure 01. Research Methodology Framework illustrating the key phases, techniques, and tools employed in the study.

Datasets

The dataset used for the training and testing is available on Kaggle[12] namely APTOS 2019 Blindness Detection, that contains 35,108 retinal images. All the images in the dataset are labeled based on their severity. The DR severity scale is supplemented with each image and the scale is defined as: 0 for Normal, 1 for Mild, 2 for Moderate, 3 for Severe, and 4 for Proliferative Diabetic Retinopathy (PDR). However, to distinguish between retinal and non-retinal images a second dataset is downloaded from Kaggle open-source images[12] that consist images of cars, flowers, pets etc. The dataset was divided into training, validation, and test sets using an 80:10:10 split, ensuring class balance across all subsets to maintain consistency and fairness in model evaluation.

Training Models

The choice of EfficientNet-B3 and ResNet18 for this study was driven by their proven effectiveness in medical imaging tasks. EfficientNet-B3 offers a superior trade-off between computational efficiency and classification accuracy, while ResNet18 was selected for its simplicity and robustness in handling high-resolution images. Compared to other architectures, such as DenseNet and Inception, these models demonstrated better adaptability to the dataset while maintaining computational feasibility for large-scale analysis.[13]

EfficientNet-B3 is a state-of-the-art architecture of convolutional neural network, and it achieves improved performance while maintaining computational efficiency. The model is designed with a few convolutional layers including depth wise separable convolutions, squeeze-and-excitation blocks, and feature fusion layers. This enables the model to effectively detect features both local and global.

In the initialization phase of the model, pre-trained weights were used. These pre-trained weights were achieved on training the model on ImageNet dataset. The last layer of the EfficeintNet-B3 model is tailored for the prediction of the severity level of DR. A total of 1536 input features and 5 output units (severity scale from Normal to Proliferative) were connected. The model was trained using a batch size of 32, an initial learning rate of 0.001 with a step decay schedule, a dropout rate of 0.5, and a total of 50 epochs, selected based on preliminary experimentation for optimal performance. The initial learning rate was set to 0.001 and adjusted using a step decay schedule to optimize convergence during training.

System Architecture

The architecture of the proposed model integrates the front end (web-based interface) with the backend (deep learning pipeline). ResNet18 model is included in the study that performs a preliminary check. This check ensures that the images selected are suitable for analysis. The input of the model is the high-resolution retinal images which are then preprocessed for improving the image quality for the extraction of various features.

The architecture of EfficientNet-B3 is extended with a global average pooling layer, dense layers with 512 and 128 neurons and a dropout layer. Additional layers for the ResNet18 include the global average pooling layer, a dense layer with 512 neurons followed by a dropout layer. These architectures leverage the pre-trained feature extraction capabilities of ResNet50 and EfficientNetB4, making them well-suited for diabetic retinopathy severity detection due to their robust performance and efficiency.

Our system is integrated with advanced segmentation techniques, including the detection of both microaneurysms and hemorrhages, to capture a broader range of DR-related anomalies. The segmented regions undergo an expanded feature extraction process, analyzing edges, textures, shapes, and higher-level structures. These features are the inputs to the EfficientNet-B3 model, that categorizes the disease into stages (Normal, Mild, Moderate, Severe, and Proliferative DR). Cross-validation and ensemble learning techniques will be employed to fine-tune classification thresholds.

Data Flow Model

The data flow model for the retinal image analysis process begins with the collection of high-resolution retinal images. These factors often result in approximately 10% of the images containing artifacts that hinder human grading. To address these challenges, preprocessing becomes essential to ensure successful automated abnormality detection. Variations in camera settings, illumination, and pigmentation require normalization against a reference image. For images exhibiting low contrast, local contrast enhancement is applied as a second preprocessing step.

The first preprocessing operation is the conversion of the color image to a grayscale representation, followed by the isolation of the green channel as shown in Figure 01. The green channel of the images is selected because it offers better visibility by offering optimal contrast for diagnosing retinal irregularities in comparison to the other two channels. Next, the selected channel is resized to a 1024x1024 aspect ratio to ensure consistency.

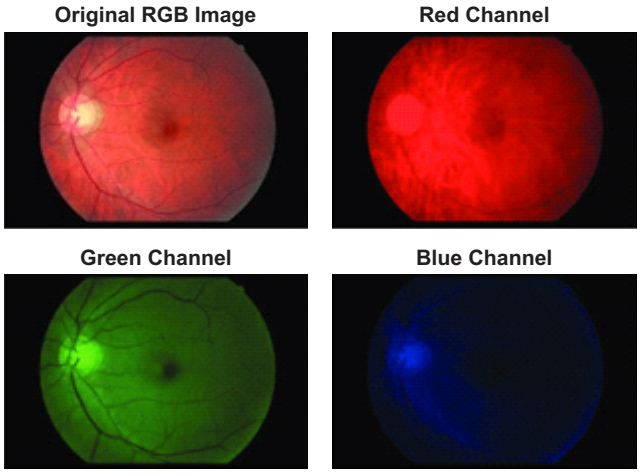


Figure 02. Three channel image

Resizing and normalization are essential preprocessing steps as resizing ensures consistent input dimensions for the model, while normalization scales pixel values to a standard range, enhancing the model's convergence and performance during training. Median filtering is used for removing the noise from the image. The process passes through the image pixel by pixel and changes the value of each pixel by the median

value of its neighboring pixel. This addresses the 'salt and pepper' noise from the image. Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied in the next step for enhancing the contrast and dynamic range of the image.

Segmentation and Classification

The output of the previous step is an image, and segmentation is applied for detection of vessels and later removing them. The blood vessels in the images are then separated from other structures using a top-hat filter. Thresholding applied on the resultant images for isolating bloods vessels, hemorrhages, and microaneurysms. The thresholding removes all the features that are below a defined threshold, leaving only the blood vessels.

The size of the blood vessels is then used to differentiate between the microaneurysms and hemorrhages. Structures below a defined area threshold are considered microaneurysms, while larger structures are identified as hemorrhages. Morphological opening is applied to remove thin blood vessel fragments and refine the segmentation of red lesions.

In the classification phase, distinct numerical features from the segmented image are used to categorize the data into different classes. The data augmentation pipeline includes random transformations such as rotation, horizontal flipping, zoom, and brightness adjustment to artificially increase the diversity of the training set and improve the model's generalization capabilities. For example, for brightness adjustment:

$$I_{new}(x,y) = aI(x,y) + \beta I_{new}(x,y)$$

where $I(x,y)$ is the pixel value at position (x,y), and a and B control the intensity and brightness adjustment. For a classification task, the categorical cross-entropy loss can be expressed as:[13]

$$\mathcal{L} = -\sum_{i=1}^{N} y_i \log(\mathcal{P}_i)$$

where y_i is the true label and p_i is the predicted probability for each class.

The process involves two main phases: training and testing. During training, the system identifies the key properties of image features and formulates a description for each classification category. In the testing phase, these defined features are used to classify new images based on the learned feature space. The implementation was carried out using Python, leveraging TensorFlow and Keras libraries, with training performed on a CPU for computational efficiency.

Statistical Analysis

Accuracy represents the ratio of correct predictions to the total cases analyzed. It serves as a simple measure of the model's overall predictive performance but doesn't account for the balance between classes, which is crucial for medical diagnosis systems where the cost of false negatives might be significantly higher than false positives.[14,15] For this study the accuracy was calculated as:

$$Accuracy = \frac{Number\ of\ correct\ predications}{Total\ number\ of\ predication}$$

In addition, quadratic weighted kappa (QWK) is used to evaluate the degree of agreement between two raters or classifications16. QWK is an extension of Cohen's kappa statistic, which measures inter-rater agreement for categorical items, but QWK is specifically designed to handle ordered categories and give more weight to larger disagreements. Diabetic retinopathy severity is typically categorized into ordinal levels, such as no retinopathy, mild, moderate, severe, and proliferative diabetic retinopathy. These levels have a natural order, making QWK an appropriate measure because it accounts for the ordered nature of these categories.[13,15]

Moreover, QWK's quadratic weighting means that misclassifying a case by a large margin (e.g., predicting “no retinopathy” when it is actually “severe retinopathy”) is penalized more heavily than a smaller error (e.g., predicting “mild” when it is “moderate”). This aligns well with clinical significance, where larger diagnostic errors have more serious implications.

To further evaluate the performance of the proposed model, additional metrics such as sensitivity, specificity, and F1-score were computed to complement the accuracy and QWK analysis. Sensitivity, also known as recall, highlights the model's ability to correctly identify samples for each class, while specificity measures the proportion of correctly identified negatives. F1-score, provides a balance between precision and recall. These metrics collectively offer a more comprehensive understanding of the model's predictive capabilities, particularly in handling imbalances in the dataset and ensuring robust classification across all diabetic retinopathy severity levels.

To evaluate the performance of the proposed models, we utilized a 10-fold cross-validation approach. This method involves splitting the dataset into k subsets, where the model is trained on 10-1 subsets and tested on the remaining one. This process is repeated 10 times, each time using a different subset for testing, ensuring that all data points are used for both training and validation. Cross-validation was chosen to minimize overfitting and provide a more reliable estimate of model performance across diverse data splits. This approach allows for a robust evaluation of the model's ability to generalize unseen data.

RESULTS

In evaluating the performance of machine learning models for Diabetic Retinopathy (DR) detection, accuracy and the confusion matrix are two pivotal

metrics that provide comprehensive insights into a model's efficiency in classifying DR stages and minimizing misclassifications.

Our model, which combines EfficientNet-B3 for classifying DR severity and ResNet18 for verifying retinal images, achieved an impressive accuracy of 98.18%. For comparison (see Table 01), other popular deep learning models also demonstrate strong performance, such as ResNet50 with an accuracy of 97.03%, Inception V3 at 87.88%, and DenseNet121 at 90.10%.[3,17] Notably, models like ResNet18 and VGG have been observed to perform similarly, reaching accuracies as high as 99%, underscoring their strong capabilities in image detection, particularly for tasks that involve complex medical images.[2,3,18]

Table 01. Comparison with other models using the same preprocessing techniques

Models	Accuracy
Efficentnet-B3	98.18%
ResNet50	97.03%
Inception V3	87.88%
DenseNet121	90.10%

The performance of various models in classifying Diabetic Retinopathy stages is summarized in the table below. Among the models evaluated, ResNet18 achieved the highest accuracy of 99%, demonstrating its superior ability to correctly classify retinal images into the appropriate severity stages (Table 02). In comparison, MobileNet achieved an accuracy of 98.31%, while VGG also performed well with an accuracy of 99%. ResNet7, on the other hand, showed a slightly lower accuracy of 97.06%.[10,19,20] These results highlight the effectiveness of ResNet18 in this task, surpassing the other models in terms of classification accuracy, and reinforcing its suitability for DR detection applications.

Table 02. Comparison accuracy for retinal and non-retinal images

Models	Accuracy
Resnet18	99.00%
Mobilenet	98.31%
VGG	99.00%
Resnet7	97.06%

Additionally, the use of a confusion matrix allows for a detailed analysis of model performance. The experiments showed that the model performed particularly well in classifying "No DR" and "Moderate" stages, evidenced by the higher numbers along the diagonal of the matrix as shown in Figure 02. However, there were some challenges in differentiating between the "Mild" and "PDR" (Proliferative Diabetic Retinopathy) stages, which were indicated by lighter matrix squares and fewer accurate predictions in those categories.

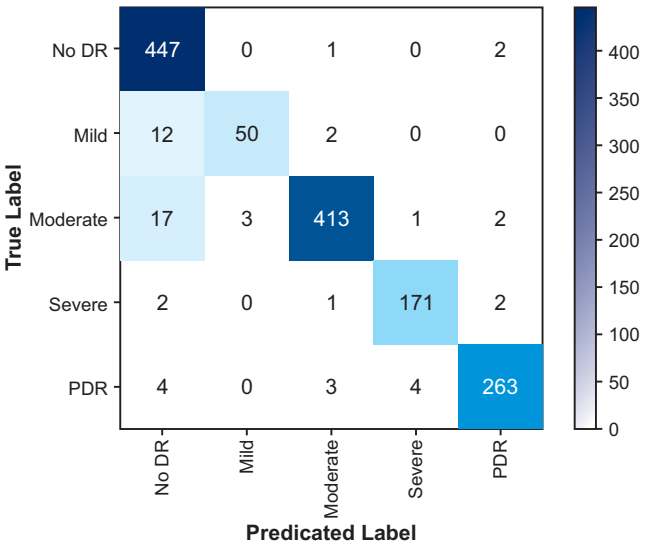


Figure 03. Results of the confusion matrix using the proposed model

The performance metrics, summarized in Table 03, demonstrate the robustness of the model in accurately classifying all classes. Notably, the model achieved the highest sensitivity of 99.33% for the No DR class, indicating its ability to effectively identify samples without diabetic retinopathy. The specificity was consistently high across all classes, with values exceeding 96%, reflecting the model's precision in distinguishing negative cases. Additionally, the F1-scores, which balance precision and recall, highlight the model's reliable performance, with values ranging from 85.47% (Mild) to 97.16% (Severe). These results underline the efficacy of the proposed model in handling imbalanced datasets and achieving robust classification for all severity levels.

Table 03. Performance Metrics for Diabetic Retinopathy Severity Levels

Class	Sensitivity	Specificity	Precision	F1-Score
No DR	99.33%	96.32%	92.74%	95.92%
Mild	78.13%	99.78%	94.34%	85.47%
Moderate	94.72%	99.27%	98.33%	96.50%
Severe	97.16%	99.59%	97.16%	97.16%
PDR	95.99%	99.47%	96.87%	96.87%

Throughout the training phase, the model showed a steady decline in loss and a consistent improvement in accuracy during the initial 20 epochs. However, after the 30th epoch, there was a slight plateau in both metrics, suggesting that the model was approaching its optimal performance. Despite this, the validation accuracy continued to fluctuate, indicating that the model was struggling with generalization at certain stages. This pattern of improvement followed by stabilization is typical in deep learning models as they converge to their best solution. No significant signs of overfitting were observed, but further tuning of hyperparameters could enhance performance in future iterations.

DISCUSSION

The proposed model using the EfficientNet-B3 in combination with ResNet18 achieved a remarkable accuracy of 98.18% on the detection of Diabetic Retinopathy (DR). This shows that the model developed is a reliable tool for classification of retinal images and accurate diagnosis of DR.

Our model outperformed most of the advanced models developed in literature review where they also have observed better performance. For example, an accuracy of 97.03% is achieved by the ResNet50 model that falls short of model's performance. The classification accuracy of 87.88% and 90.10% is achieved by the Inception V3 and DenseNet121, respectively.[13,17,21] This comparison of various deep learning models suggests that the proposed model using the EfficientNet-B3 architecture is a better solution for the classification of DR severity as it can capture complex patterns in images.

Overall, the better performance achieved by our hybrid model shows effectiveness in the diagnosis of DR and it stands out because both the image verification and DR severity classification is considered. In comparison, the ResNet50, Inception V3, and DenseNet121 also show strong results, however, our model ensures precision and reliability in image analysis due to its comprehensive.[2,20,22]

In addition, the high sensitivity values across most classes, particularly 99.33% for the No DR class, indicate the model's ability to correctly identify true positives, which is critical for early detection and timely intervention. Similarly, the consistently high specificity, exceeding 96% for all classes, demonstrates the model's reliability in minimizing false positives, ensuring accurate identification of non-affected cases. The strong F1-scores, such as 97.16% for the Severe class, reflect the balance between precision and recall, emphasizing the model's robustness in handling imbalanced datasets. These results, when considered alongside accuracy and QWK, confirm that the proposed model is well-suited for detecting varying levels of diabetic retinopathy severity. Moreover, the model's performance metrics align closely with or surpass those of state-of-the-art methods, underscoring its reliability and potential utility in real-world settings.

The model's lower performance in certain classes, such as Mild and PDR (Proliferative Diabetic Retinopathy), could be attributed to several factors. One of the primary reasons is class imbalance, where images from less frequent stages, particularly the Proliferative Diabetic Retinopathy class, are underrepresented in the dataset, making it challenging for the model to accurately predict these stages. Additionally, the subtle visual differences between stages like Mild and Moderate DR may lead to misclassifications, as these stages often exhibit overlapping characteristics in terms of lesion appearance. The model may also struggle with limited data for certain classes, which hinders its ability to generalize and capture the fine details necessary for accurate classification. These challenges could be addressed by utilizing data augmentation techniques or applying advanced transfer learning strategies to improve model robustness and performance on these specific stages.

While AI models like ResNet18 and EfficientNet-B3 offer transformative potential in DR detection, their adoption must be accompanied by careful consideration of ethical concerns. To address the model's generalizability, it is important to note that the results are based on a specific dataset, and while the high accuracy suggests the model performs well within this context, further validation on diverse, unseen datasets is necessary to assess its real-world applicability. Cross-validation and testing on external datasets can provide a clearer understanding of its performance in varied scenarios.

CONCLUSION

The proposed hybrid model in this study combines two efficient models, ResNet18 and EfficientNet-B3, and shows remarkable achievement in the diagnosis of Diabetic Retinopathy (DR). The results demonstrate improvements in both accuracy and efficiency in detecting DR. The ResNet18 model performs the initial verification and screening of relevant images, while EfficientNet-B3 effectively classifies DR severity. This integration of AI streamlines the diagnostic process by reducing manual effort and mitigating risks, addressing time constraints, and enhancing overall efficiency. However, the model's performance is still limited by factors such as variations in image quality, and the inability to generalize perfectly to all real-world scenarios, such as diverse patient demographics and varying imaging conditions. Future studies incorporating larger and more diverse datasets, along with advancements in AI technologies like augmented reality (AR) and telemedicine, may further enhance DR care by offering personalized treatments, risk assessment, and ongoing monitoring, while addressing these current limitations.

Acknowledgements
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Ethical Consideration
This study was approved by Institutional Review Board of Bahria University, Islamabad, Pakistan.

Funding
There is no role of any funding agency in this study.

Conflict of Interest
Authors declared no conflict of interest.

REFERENCES

1. Bora A, Balasubramanian S, Babenko B, Virmani S, Venugopalan S, Mitani A, et al. Predicting the risk of developing diabetic retinopathy using deep learning. *The Lancet Digital Health*. 2021;3(1):e10-e9. https://doi.org/10.1016/S2589-7500(20)30250-8
2. Nadeem MW, Goh HG, Hussain M, Liew S-Y, Andonovic I, Khan MA. Deep learning for diabetic retinopathy analysis: a review, research challenges, and future directions. *Sensors*. 2022;22(18):6780. https://doi.org/10.3390/s22186780
3. Atwany MZ, Sahyoun A, Yaqub M. Deep Learning Techniques for Diabetic Retinopathy Classification: A Survey. *IEEE Access*. 2022;PP:1- https://doi.org/10.1109/ACCESS.2022.3157632
4. Chaturvedi K, Bhandari V, Shrivastava R, Tiwari R. A Comparative Study on Diabetic Retinopathy Datasets for Data Accuracy Detection. *2024 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*. 2024:1-6. https://doi.org/10.1109/SCEECS61402.2024.10481879
5. Clinic M. Diabetic retinopathy - symptoms & causes 2023. https://www.mayoclinic.org/diseases-conditions/diabetic-retinopathy/symptoms-causes/syc-20371611
6. Vij R, Arora S. A novel deep transfer learning based computerized diagnostic Systems for Multi-class imbalanced diabetic retinopathy severity classification. *Multimedia Tools and Applications*. 2023:1-38. https://doi.org/10.1007/s11042-023-14963-4
7. Das D, Biswas SK, Bandyopadhyay S. A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning. *Multimedia Tools and Applications*. 2022;81(18):25613-55. https://doi.org/10.1007/s11042-022-12642-4
8. Saini NK, Ram D, Gyanchandani M. Multi-Headed CNN and Vision Transformer-Based Diabetic Retinopathy Classification. *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. 2023:1-6. https://doi.org/10.1109/ICCCNT56998.2023.10306806
9. Deepak NA, Savitha G, Deepak DJ, Supraj PK. Retinopathy Based Multistage Classification of Diabeties. Proceedings of the 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021)2021. Available from: https://api.semanticscholar.org/CorpusID:237953571
10. Ruamviboonsuk P, Tiwari R, Sayres R, Nganthavee V, Hemarat K, Kongprayoon A, et al. Real-time diabetic retinopathy screening by deep learning in a multisite national screening programme: a prospective interventional cohort study. *The Lancet Digital Health*. 2022;4(4):e235-e44. https://doi.org/10.1016/S2589-7500(22)00017-6
11. Abushawish IY, Modak S, Abdel-Raheem E, Mahmoud SA, Hussain AJ. Deep learning in automatic diabetic retinopathy detection and grading systems: a comprehensive survey and comparison of methods. *IEEE Access*. 2024; 12:84785-802. https://doi.org/10.1109/ACCESS.2024.3415617
12. Aptos 2019 blindness detection. 2019.
13. Dai L, Sheng B, Chen T, Wu Q, Liu R, Cai C, et al. A deep learning system for predicting time to progression of diabetic retinopathy. *Nature Medicine*. 2024;30(2):584-94. https://doi.org/10.1038/s41591-023-02702-z
14. Li C, Chua J, Schwarzhans F, Husain R, Girard MJ, Majithia S, et al. Assessing the external validity of machine learning-based detection of glaucoma. *Scientific reports*. 2023;13(1):558. https://doi.org/10.1038/s41598-023-27783-1
15. Muthusamy D, Palani P. Deep learning model using classification for diabetic retinopathy detection: an overview. *Artificial Intelligence Review*. 2024;57(7):185. https://doi.org/10.1007/s10462-024-10806-2
16. Sugihara S, Kajiwara T, Ninomiya T, Wakamiya S, Aramaki E. Semi-automatic Construction of a Word Complexity Lexicon for Japanese Medical Terminology. Proceedings of the 6th Clinical Natural Language Processing Workshop; 2024. https://doi.org/10.18653/v1/2024.clinicalnlp-1.29
17. Abdel Hussein Abdel Amir H, Majeed Hilal Almiahi O. Enhanced model to detection of Diabetic Retinopathy using Deep Learning techniques. *Journal of Al-Qadisiyah for Computer Science and Mathematics*. 2024. https://doi.org/10.29304/jqcsm.2023.15.31344
18. Ting DSW, Cheung CY-L, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *Jama*. 2017;318(22):2211-23. https://doi.org/10.1001/jama.2017.18152
19. Arora M, Pandey M, editors. Deep neural network for diabetic retinopathy detection. 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon); 2019: IEEE; https://doi.org/10.1109/COMITCon.2019.8862217

20. Dai L, Wu L, Li H, Cai C, Wu Q, Kong H, et al. A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nature Communications*. 2021;12. https://doi.org/10.1038/s41467-021-23458-5
21. Sikder N, Masud M, Bairagi AK, Arif ASM, Nahid A-A, Alhumyani HA. Severity classification of diabetic retinopathy using an ensemble learning algorithm through analyzing retinal images. *Symmetry*. 2021;13(4):670. https://doi.org/10.3390/sym13040670
22. Ríos H, Rodríguez FJ, Perdomo OJ, González FA, editors. A deep learning model for classification of diabetic retinopathy in eye fundus images based on retinal lesion detection. 17th International symposium on medical information processing and analysis; 2021: SPIE; https://doi.org/10.1117/12.2606319

DATA SHARING STATEMENT: The data that support the findings of this study are available on request from the corresponding author. The data is not publicly available due to privacy or ethical restrictions.

AUTHORS' CONTRIBUTION

The following authors fulfill authorship criteria as per ICMJE guidelines;

Mehmood Q, Amer A: Substantial contribution to conception and design, drafting the manuscript, final approval, agreed to be accountable for all the work.

Ali M: Substantial contribution to conception and design, data acquisition, drafting the manuscript, final approval, agreed to be accountable for all the work.

Ahmad HI: Analysis and interpretation of data, drafting the manuscript, final approval, agreed to be accountable for all the work.

Ahmad A: Analysis and interpretation of data, critical revision, final approval, agreed to be accountable for all the work.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License, which permits unrestricted use, distribution & reproduction in any medium provided that original work is cited properly.