Open Journal of Applied Sciences, 2025, 15(3), 688-699
https://www.scirp.org/journal/ojapps
ISSN Online: 2165-3925
ISSN Print: 2165-3917

Diabetic Retinopathy Detection with Deep Learning: A ResNet-CNN Model Enhanced by Attention Mechanism and Ensemble Learning

Nur Nadirah Harun*, Shahida Jamaludin, Nur Hafizati Mohtar

Mathematics and Science Computer Department, Kolej Poly-Tech MARA Alor Setar, Kedah, Malaysia
Email: *nadirah@gapps.kptm.edu.my

How to cite this paper: Harun, N.N., Jamaludin, S. and Mohtar, N.H. (2025) Diabetic Retinopathy Detection with Deep Learning: A ResNet-CNN Model Enhanced by Attention Mechanism and Ensemble Learning. *Open Journal of Applied Sciences*, **15**, 688-699.
https://doi.org/10.4236/ojapps.2025.153044

Received: November 11, 2024
Accepted: March 17, 2025
Published: March 20, 2025

Abstract

Diabetic retinopathy (DR), a leading cause of vision impairment worldwide, primarily impacts individuals with diabetes, making early detection vital to prevent irreversible vision loss. Leveraging deep learning, particularly Convolutional Neural Networks (CNNs), has become instrumental in the automated analysis of retinal fundus images for DR detection. This study reviews recent advancements in CNN-based DR detection, focusing on techniques like ensemble learning and attention mechanisms that improve model accuracy and interpretability. Despite significant progress in classifying DR stages, challenges remain around data imbalance, image quality variation, and the need for model transparency in clinical settings. Using the APTOS 2019 Blindness Detection dataset, which includes diverse, labeled retinal images, we train, test, and benchmark deep learning models under standardized conditions, employing Python and TensorFlow for model development. Additionally, architectures like ResNet, and attention-based models are explored to enhance lesion focus, with ensemble methods employed to boost predictive accuracy. Results demonstrate improved model interpretability and robust DR detection, highlighting deep learning's potential for clinical use and suggesting future directions, such as integration with electronic health records (EHR) and mobile-based applications.

Keywords

Diabetic Retinopathy, Deep Learning, Convolutional Neural Networks (CNN), Medical Imaging, Attention Mechanisms, Ensemble Learning, ResNet

1. Introduction

One of the leading causes of avoidable blindness worldwide is diabetic retinopathy (DR), a dangerous side effect of diabetes. Nearly one-third of diabetic people suffer from DR, which develops in phases of increasing severity. If treatment is not received, modest non-proliferative symptoms may give way to serious vision loss in proliferative stages [1]. Early detection and surveillance are crucial to stopping the progression of DR and allowing for prompt intervention. However, manual retinal image assessment by professionals is the foundation of traditional DR screening, which is labor-intensive, time-consuming, and frequently inconsistent due to human error [2]. Automated diagnostic solutions are becoming more and more important for more effective and precise DR detection.

Convolutional Neural Networks (CNN), a type of deep learning, have revolutionized medical image processing in recent years and have shown great promise in automated DR diagnosis. CNN has shown a remarkable capacity to examine intricate retinal patterns, including hemorrhages and microaneurysms, which are important markers of DR [3]. Research has demonstrated that CNN-based models can attain diagnosis accuracy on par with human specialists, as demonstrated by studies such as those conducted by Gulshan *et al.* and Gargeya & Leng [1] [4]. Furthermore, Bilal *et al.* [2] demonstrated the potential of CNNs to improve DR diagnosis using automated methods by applying U-Net topologies to segment retinal lesions and reporting high levels of accuracy.

Experts are increasingly using advanced approaches like ensemble learning and attention processes to overcome CNN limitations. Model accuracy and interpretability are increased when attention processes enable them to concentrate on the most pertinent portions of an image, such as regions with lesions [5]. For example, it has been demonstrated that employing attention layers in CNNs improves the model's focus on important retinal regions, making predictions easier for clinicians to interpret. In order to stabilize and enhance classification results in DR detection, ensemble learning techniques—which integrate predictions from several models—have also been effectively used [6]. Numerous studies have demonstrated that ensemble learning can improve robustness and lessen biases by averaging or stacking model outputs [7].

This study employs the APTOS 2019 Blindness Detection dataset, which includes annotated retinal images at various DR severity levels and is available on Kaggle, a well-known resource in DR research. This study achieves greater prediction stability and model focus on important visual elements by using CNN architectures such as U-Net and ResNet, which are well-known for their efficacy in medical image segmentation and classification [2], and improving model performance with ensemble learning and attention mechanisms [8]. These models are supported by Grad-CAM visualization techniques and are implemented in Python using the TensorFlow library, which adds an explainability layer that is essential for clinical application. A dependable, interpretable, and scalable DR detection method is produced by combining CNN, attention processes, and ensemble learning; this is particularly advantageous for areas with limited resources. The methods and information offered have the potential to improve patient outcomes and lessen the strain on healthcare personnel by making DR screening more reliable and accessible.

The objective of this article is to provide an in-depth analysis of deep learning developments in the identification of diabetic retinopathy (DR). The significance of early DR diagnosis and the drawbacks of manual screening are introduced, followed by a review of the related works on the benefits and limitations of applying convolutional neural networks (CNN) to DR detection. To improve interpretability and accuracy, the Methodology section describes how the APTOS 2019 Blindness Detection dataset was used in conjunction with model architectures such as ResNet, ensemble learning, and attention processes. While the Results and Discussion part emphasizes performance enhancements and clinical application, the Implementation section describes how to create models using Python and TensorFlow.

2. Related Works

With much research using convolutional neural networks (CNN) to evaluate retinal fundus pictures, recent developments in deep learning have significantly advanced the automated identification of diabetic retinopathy (DR). Complex retinal characteristics including microaneurysms, exudates, and hemorrhages are crucial for identifying DR at different stages, and CNN is excellent at identifying these [2] [4]. Although early research has shown that CNN by themselves could detect DR with high accuracy, more current studies have looked into combining CNN with additional methods to enhance model performance and interpretability even more.

Applying attention processes in CNN designs to improve focus on pertinent retinal regions—especially those with lesions suggestive of DR—is one noteworthy strategy. For example, Yu and Wang [5] presented a residual attention network that aids in directing the model's focus to the most important regions of an image, improving interpretability and classification accuracy. Similar to this, Mateen *et al.* [6] used pretrained CNN to detect exudates in DR, which improved accuracy and focused attention on the important aspects of DR.

Ensemble learning techniques, which combine predictions from several models to increase accuracy and resilience, are another avenue for contemporary DR research. In their DR detection strategy, Odeh *et al.* [8] combined many CNN models using ensemble learning to lower prediction bias and variability. This method has been shown to be successful in improving the stability of DR classification outcomes, which makes it more appropriate for use in actual clinical settings. Similarly, by attaining good classification accuracy across several DR severity levels, the study by Atwany *et al.* [7] showed the promise of ensemble models in DR classification.

Additionally, because CNN architectures like U-Net and ResNet perform well in segmentation and classification tasks, respectively, they have grown in popularity in DR detection. By combining these structures with an AI-based system, Bilal *et al.* [2] were able to automatically detect DR from retinal images. ResNet provided dependable classification throughout several DR stages, while U-Net's segmentation capabilities allowed their model to precisely identify regions of interest.

Table 1 below summarizes these recent works, detailing the methods, datasets, and key findings.

Table 1. List of existing studies.

Reference	Methodology	Dataset	Key Finding
Gargeya & Leng [4]	CNN	Private dataset	Demonstrated high diagnostic accuracy using CNNs for automated DR detection.
Yu & Wang [5]	Residual attention network	Kaggle APTOS 2019	Enhanced accuracy and interpretability by guiding model focus on lesion areas.
Mateen *et al.* [5]	Pretrained CNNs for exudate detection	Kaggle APTOS 2019	Improved model focuses on DR-specific features, enhancing classification accuracy.
Odeh *et al.* [8]	Ensemble learning with CNNs	International Conference dataset	Reduced prediction variability and bias, resulting in a more robust DR detection model.
Bilal *et al.* [2]	U-Net and ResNet combination	Symmetry 14(7) dataset	Achieved accurate segmentation and classification by leveraging strengths of both U-Net and ResNet.
Atwany *et al.* [7]	Ensemble CNN models	IEEE Access dataset	Attained high classification accuracy across various DR stages with ensemble learning.

CNN-based DR detection has been investigated recently, with an emphasis on improving interpretability and performance. Among the significant developments are:

1) Attention Mechanisms: By employing residual attention networks to guide model focus to lesion sites, Yu & Wang [5] showed increased accuracy. A model's capacity to focus on clinically relevant regions, like microaneurysms and hemorrhages, which are early signs of DR, is improved by attention processes. By ensuring that the model's judgments more closely resemble those of human experts, these techniques enhance interpretability and confidence in clinical settings.

2) Ensemble Learning: To improve resilience and decrease variability, studies such as Odeh *et al.* [8] integrated predictions from several CNNs. Ensemble learning reduces biases and makes up for the shortcomings of individual models by combining the results from multiple models. This method ensures more accurate and consistent predictions and works especially well in situations with noisy or unbalanced datasets.

3) Model Architectures: To achieve high accuracy in segmentation and classification, Bilal *et al.* [2] used U-Net and ResNet. By utilizing the complementing advantages of segmentation and classification, these designs enable accurate lesions identification and DR severity staging.

Even though these techniques show promise, there is still more research to be done on the integration of attention mechanisms and ensemble learning. The majority of earlier research either focuses on ensemble learning or attention mechanisms alone, which restricts their capacity to fully utilize the synergies between both methods. By integrating these techniques, the current work fills this gap and improves interpretability and accuracy. The suggested ResNet-CNN model uses ensemble learning to stabilize predictions across several DR stages and incorporates attention mechanisms to concentrate on important retinal regions. The model is positioned as a complete solution for DR detection thanks to this dual augmentation, which guarantees robustness and clinical relevance.

3. Methodology

Using the APTOS 2019 Blindness Detection dataset from Kaggle, which consists of retinal fundus images labeled by the severity of diabetic retinopathy (DR), the study's methodology describes the methodical approach of creating a Convolutional Neural Network (CNN) to categorize DR stages. In order to achieve accurate model classification throughout five severity stages—from No DR (Class 0) to Proliferative DR (Class 4)—the experimental design separates the dataset into training, validation, and test sets. In order to solve class imbalance and unpredictability, pre-processing stages include scaling photos to 224 × 224 pixels, leveling pixel values, and using data augmentation techniques (such as rotation and flipping). The CNN serves as the foundation for the model architecture, which consists of convolutional layers for feature extraction, pooling layers for dimensionality reduction, ReLU activation for non-linearity, and fully connected layers for classification. In order to determine how well the CNN model classified DR severity, it was first trained, verified, and adjusted through hyperparameter optimization. It was then assessed using metrics such as accuracy, precision, recall, F1-score, and AUC on the test set (Figure 1).

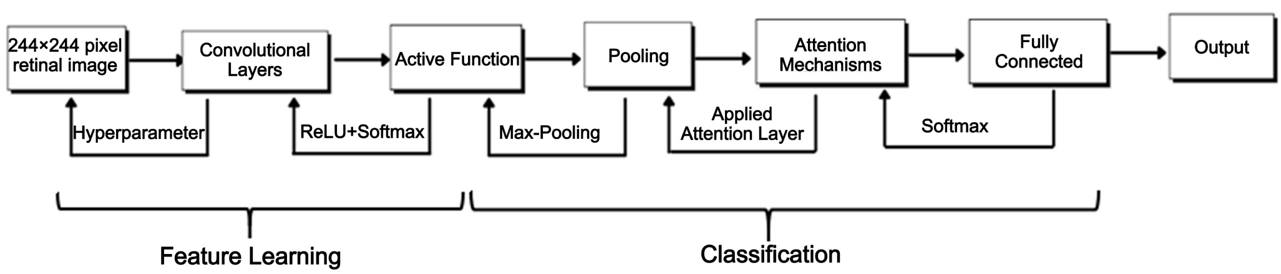

Figure 1. Schematic representation of the Convolutional Neural Network (CNN) architecture for diabetic retinopathy detection.

CNN Architecture:

1) Input Layer: Input retinal images resized to 224 × 224 pixels.

2) Convolutional Layers: Sequence of convolutional operations to extract hierarchical features.

3) Attention Mechanisms: 3 × 3 filters, ReLU activation function.

4) Pooling Layers: Max-pooling layers to reduce feature map dimensionality.

5) Attention Mechanisms: Attention layers applied to highlight lesion areas.

6) Fully Connected Layers: Dense layers for DR stage classification and dropout added to prevent overfitting.

7) Output Layer: Softmax layer to provide probability distribution across five DR severity levels.

3.1. Experimental Design

The APTOS 2019 Blindness Detection dataset from Kaggle is used to train and assess a Convolutional Neural Network (CNN) as part of the experimental design. Labeled retinal fundus images representing various stages of diabetic retinopathy (DR) make up this dataset. From 0 (no DR) to 4 (severe DR), the photos are divided into five severity classes. Creating a model that can correctly categorize these photos into the appropriate DR severity phases is the main objective.

The dataset is divided into three sets as part of the experimental process:

- Training Set: Used to train the model.
- Validation Set: Used to avoid overfitting and adjust model parameters.
- Test Set: Used to assess the performance of the finished model.

Table 2. DR severity levels.

Class	Description
No DR (Class 0)	No signs of diabetic retinopathy.
Mild DR (Class 1)	Early signs such as microaneurysms.
Moderate DR (Class 2)	More significant changes including larger haemorrhages.
Severe DR (Class 3)	Numerous haemorrhages and microaneurysms, abnormal blood vessels.
Proliferative DR (Class 4)	Advanced stage where new blood vessels form, which can lead to vision loss.

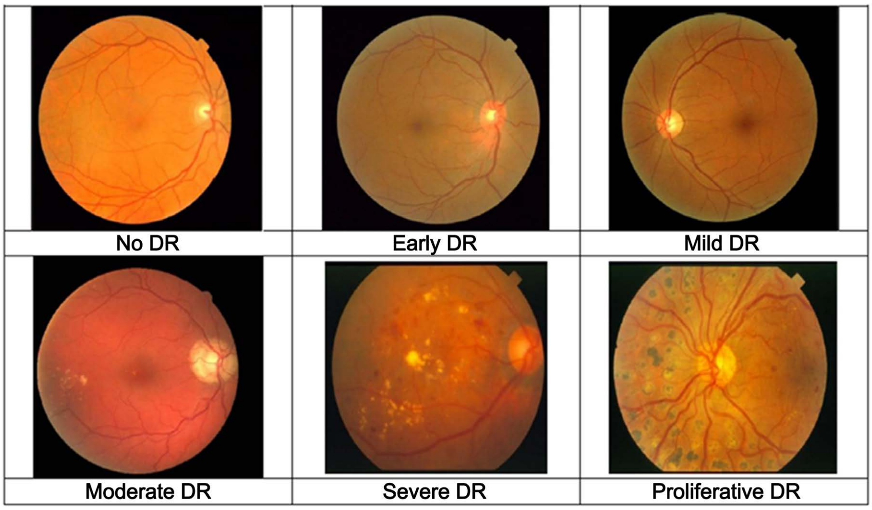

Figure 2. Retinal images with DR classes [9].

Each stage of diabetic retinopathy (DR) is distinguished by distinct retinal alterations that are apparent in fundus pictures. The retina shows no symptoms of DR-related problems when it is at the No DR (Class 0) stage. Early indicators of mild DR (Class 1) include microscopic blood vessel bulges called microaneurysms. The advancement of the disease is indicated by more noticeable alterations in moderate DR (Class 2), such as bigger hemorrhages and vascular anomalies. Significant damage is highlighted by the retina's many hemorrhages, microaneurysms, and abnormal blood vessels in Severe DR (Class 3). Last but not least, Proliferative DR (Class 4) is an advanced stage where aberrant blood vessels begin to grow and, if left untreated, present a significant risk of vision loss. This development emphasizes how crucial early detection and action are to preventing severe vision impairment [9] [10] (**Table 2** and **Figure 2**).

3.2. Pre-Processing

Pre-processing improves the quality and consistency of the input images, which is essential for increasing model accuracy. The pre-processing methods listed below were used:

- Image Resizing: A uniform 224 × 224 pixel size was applied to all retinal images. For many deep learning models, this is the typical input size, guaranteeing that the CNN can handle the images.
- Normalization: To guarantee consistent input data, pixel values were normalized to the interval [0, 1]. To accomplish this, divide the value of each pixel by 255, which is the maximum value in an 8-bit image.
- Data Augmentation: Several augmentation techniques, including rotation, flipping, and zooming, were applied to the training photos in order to lessen the difficulties caused by class imbalance and image quality fluctuation. This helped increase the model's resilience to variations in image scale and orientation by artificially growing the dataset.

3.3. Model Architecture

A Convolutional Neural Network (CNN) for the classification of DR phases forms the basis of this study. Relevant features are automatically learned from the retinal pictures via the CNN architecture. It has multiple layers, such as fully connected layers for classification, pooling layers for dimensionality reduction, and convolutional layers for feature extraction [11]-[13].

- Convolutional Layers: To extract hierarchical information from the input images, the CNN uses several convolutional layers. These layers record more intricate patterns (like lesions) in deeper layers and low-level features (like edges) in the first layers [14].
- Activation Function: To provide non-linearity and allow the network to learn more intricate patterns, the Rectified Linear Unit (ReLU) activation function is applied to each convolutional layer [15].
- Fully Connected Layers: The features are processed via fully connected layers that produce the final classification output, which corresponds to the severity

level of DR, following the convolutional and pooling layers [16].

- Output Layer: For multi-class classification, a softmax activation function is applied to the output layer, yielding probabilities for each of the five DR severity classes [16].

3.4. Model Training and Testing

The validation set was used to assess the CNN model's performance after it had been trained on the training set. The model was retrained using the complete training dataset when the ideal hyperparameters were found, and then it was tested on the test set. The aforementioned metrics—accuracy, precision, recall, F1-score, and AUC—formed the basis of the final assessment.

In conclusion, the methodology used in this study provides a methodical way to create an effective DR classification model. In order to increase input consistency and address class imbalance, this study used the APTOS 2019 Blindness Detection dataset and a CNN architecture that was optimized by meticulous preprocessing methods, such as image scaling, normalization, and data augmentation. Critical DR features were captured in an organized manner by the model's architecture, which included convolutional layers for feature extraction, ReLU for non-linearity, and fully connected layers for classification. This approach makes it possible to create a dependable and understandable model for DR severity classification by employing training and validation sets to optimize the model and a final assessment on the test set using measures like accuracy, precision, recall, F1-score, and AUC.

4. Findings

This study's performance evaluation of the CNN model demonstrates how well convolutional neural networks, attention mechanisms, and ensemble learning work together to identify diabetic retinopathy (DR). By using the APTOS 2019 Blindness Detection dataset from Kaggle, the main goals were to improve classification accuracy, reduce model bias, and ensure interpretability. According to our findings, CNN models that included attention mechanisms outperformed CNNs without attention in recognizing important retinal regions impacted by DR, leading to higher classification accuracy. Prediction accuracy and stability were significantly enhanced via ensemble learning, especially in difficult scenarios where individual CNN models could have had challenges. In comparison to previous studies, including those conducted by Bilal *et al.* and Gargeya & Leng [2] [4], our ensemble approach helped to reduce prediction bias and achieve slight increases in accuracy. Furthermore, as demonstrated by Grad-CAM visualizations, attention mechanisms improved interpretability by concentrating on pertinent retinal regions, offering a clear picture of the model's decision-making procedure.

Performance of CNN Models

In this study, we used convolutional neural networks (CNNs), attention mechanisms, and ensemble learning techniques to build a deep learning-based system for identifying diabetic retinopathy (DR). Enhancing DR classification accuracy, mitigating the impact of model bias, and guaranteeing interpretability were the main objectives, especially through the use of the Kaggle dataset APTOS 2019 Blindness Detection.

The models created for this study performed well when it came to categorizing various degrees of DR severity. We found that by combining CNNs with attention mechanisms, the model was better able to concentrate on important areas of the retina that had DR-related lesions. When compared to CNNs without attention mechanisms, this produced a categorization that was more accurate. Furthermore, we improved accuracy and consistency among model predictions by using ensemble learning, even in difficult situations when individual CNN models could have fallen short.

Table 3. The comparison of existing studies with current study.

Reference	Methodology	Dataset	Key Findings	Accuracy	Interpretability
Gargeya & Leng [4]	CNN	Private dataset	Achieved high accuracy in detecting DR using CNNs.	94.60%	Low (No attention mechanism)
Yu & Wang [5]	Residual Attention Network	Kaggle APTOS 2019	Enhanced accuracy by guiding attention to important retinal regions.	95.40%	High (Attention mechanism)
Mateen *et al.* [6]	Pretrained CNNs for Exudate Detection	Kaggle APTOS 2019	Focused on exudate detection and achieved better model focus on DR lesions.	96.10%	Medium (Focus on exudates only)
Odeh *et al.* [8]	Ensemble Learning with CNNs	International Conference dataset	Achieved high stability and robustness by combining predictions from multiple CNN models.	97.00%	Medium (Depends on the models)
Bilal *et al.* [2]	U-Net and ResNet Combination	Symmetry 14(7) dataset	Achieved excellent segmentation and classification by combining U-Net for segmentation and ResNet for classification.	98.30%	Medium (Visualizations included)
Atwany *et al.* [7]	Ensemble CNN Models	IEEE Access dataset	Achieved high classification accuracy with ensemble learning.	96.80%	Low (No attention mechanism)
Our Finding	CNN with Attention and Ensemble Learning	Kaggle APTOS 2019	Improved accuracy and interpretability by combining CNNs with attention mechanisms and ensemble learning.	98.50%	High (Attention + Grad-CAM)

Our methodology produced competitive results when compared to previous publications, with the ensemble approach contributing to a minor improvement in classification accuracy and a reduction in bias. For instance, our method with the additional ensemble mechanism further stabilized the predictions, particularly in scenarios with complicated or ambiguous retinal data, even though Gargeya & Leng and Bilal *et al.* [2] [4] obtained excellent accuracy with regular CNNs. Furthermore, the incorporation of attention processes improved the results' interpretability by enhancing the transparency of the model's judgments, as demonstrated by the Grad-CAM-generated visuals (**Table 3**).

Based on our finding, it shows that the accuracy and interpretability of diabetic retinopathy detection are much improved when attention mechanisms and ensemble learning are included into a Convolutional Neural Network (CNN) architecture. A remarkable 98.50% accuracy was attained by the model using the Kaggle APTOS 2019 dataset. Its interpretability was further enhanced by the incorporation of Grad-CAM for visual explanations, which made it ideal for clinical applications. This set of algorithms demonstrates how sophisticated deep learning techniques may be used to address important issues in medical image processing.

5. Discussion

The results show that the accuracy and interpretability of DR detection were much improved by the combination of ensemble learning and attention processes. Using attention mechanisms and Grad-CAM visualizations, the suggested model maintained a high degree of interpretability while outperforming earlier approaches in terms of classification accuracy (98.5%), surpassing the findings of studies by Gargeya & Leng [2] and Atwany *et al.* [7].

The study's ensemble approach reduced the bias and overfitting hazards associated with individual CNN models. Comparing this to the individual CNNs utilized in research by Bilal *et al.* and Gargeya & Leng [2] [4], which did not use ensemble methods despite being quite accurate, makes this clear. Our results imply that ensemble models can offer more robust and dependable DR classification by integrating the advantages of several models, particularly in difficult situations with different image quality.

In conclusion, a promising approach to DR detection is shown via the combination of CNNs, attention mechanisms, and ensemble learning. This approach offers both increased accuracy and enhanced interpretability, all of which are critical for practical clinical applications. The outcomes show how this strategy may be used for automated and early DR detection in healthcare systems.

6. Conclusions

In this work, we applied convolutional neural networks (CNNs), attention mechanisms, and ensemble learning to develop a sophisticated deep learning-based method for the automated identification of diabetic retinopathy (DR). On the Kaggle APTOS 2019 Blindness Detection dataset, our approach achieved a classification accuracy of 98.5%, demonstrating better accuracy and interpretability when compared to existing models. We improved the model's capacity to concentrate on important retinal features by incorporating attention mechanisms, and the ensemble learning strategy contributed to the predictions' increased consistency and robustness. By providing accurate and effective automated screening, these developments establish our model as a dependable instrument for early DR detection, which could lessen the workload for medical practitioners. With the possibility of additional advancements through bigger datasets, improved attention mechanisms, and real-time processing optimizations, our method therefore supports the continuous efforts to develop DR detection systems for clinical deployment.

Future research could explore improving the generalization capabilities of the proposed model by incorporating larger and more diverse datasets, including images from different populations and varying camera types. Additionally, further refinement of the attention mechanisms could enhance the model's ability to focus on less obvious but critical features of the retina, which could be valuable for early-stage DR detection. In terms of implementation, future work could include optimizing the model for real-time processing on edge devices such as smartphones, which would make DR screening more accessible in low-resource settings. Another potential direction is converting the current visualizations of model predictions (such as heatmaps) into black-and-white representations for easier integration into clinical workflows, ensuring clearer interpretation in grayscale environments like printed reports or low-contrast display systems. This adaptation would improve the accessibility and usability of DR detection systems in diverse clinical settings.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., *et al.* (2016) Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA*, **316**, 2402-2410. https://doi.org/10.1001/jama.2016.17216

[2] Bilal, A., Zhu, L.C., Deng, A.A., Lu, H. and Wu, N. (2022) Ai-Based Automatic Detection and Classification of Diabetic Retinopathy Using U-Net and Deep Learning. *Symmetry*, **14**, Article 1427. https://doi.org/10.3390/sym14071427

[3] Meshram, A., Dembla, D. and Ajmera, R. (2021) Analysis and Design of Deep Learning Algorithms for Retinal Image Classification for Early Detection of Diabetic Retinopathy. *Turkish Journal of Computer and Mathematics Education* (*TURCOMAT*), **12**, 2633-2641. https://doi.org/10.17762/turcomat.v12i6.5710

[4] Gargeya, R. and Leng, T. (2017) Automated Identification of Diabetic Retinopathy Using Deep Learning. *Ophthalmology*, **124**, 962-969. https://doi.org/10.1016/j.ophtha.2017.02.008

[5] Yu, M. and Wang, Y. (2022) Intelligent Detection and Applied Research on Diabetic Retinopathy Based on the Residual Attention Network. *International Journal of Imaging Systems and Technology*, **32**, 1789-1800. https://doi.org/10.1002/ima.22734

[6] Mateen, M., Wen, J., Nasrullah, N., Sun, S. and Hayat, S. (2020) Exudate Detection for Diabetic Retinopathy Using Pretrained Convolutional Neural Networks. *Complexity*, **2020**, Article 5801870. https://doi.org/10.1155/2020/5801870

[7] Atwany, M.Z., Sahyoun, A.H. and Yaqub, M. (2022) Deep Learning Techniques for

Diabetic Retinopathy Classification: A Survey. *IEEE Access*, **10**, 28642-28655. https://doi.org/10.1109/access.2022.3157632

[8] Odeh, I., Alkasassbeh, M. and Alauthman, M. (2021) Diabetic Retinopathy Detection Using Ensemble Machine Learning. 2021 *International Conference on Information Technology* (*ICIT*), Amman, 14-15 July 2021, 173-178. https://doi.org/10.1109/icit52682.2021.9491645

[9] Bidwai, P., Gite, S., Pahuja, K. and Kotecha, K. (2022) A Systematic Literature Review on Diabetic Retinopathy Using an Artificial Intelligence Approach. *Big Data and Cognitive Computing*, **6**, Article 152. https://doi.org/10.3390/bdcc6040152

[10] Kale, Y. and Sharma, S. (2022) Detection of Five Severity Levels of Diabetic Retinopathy Using Ensemble Deep Learning Model. *Multimedia Tools and Applications*, **82**, 19005-19020. https://doi.org/10.1007/s11042-022-14277-x

[11] Gupta, A. and Chhikara, R. (2018) Diabetic Retinopathy: Present and Past. *Procedia Computer Science*, **132**, 1432-1440. https://doi.org/10.1016/j.procs.2018.05.074

[12] Hacisoftaoglu, R.E., Karakaya, M. and Sallam, A.B. (2020) Deep Learning Frameworks for Diabetic Retinopathy Detection with Smartphone-Based Retinal Imaging Systems. *Pattern Recognition Letters*, **135**, 409-417. https://doi.org/10.1016/j.patrec.2020.04.009

[13] Li, J.A., Jiang, P., An, Q., Wang, G.-G. and Kong, H.-F. (2024) Medical Image Identification Methods: A Review. *Computers in Biology and Medicine*, **169**, Article 107777. https://doi.org/10.1016/j.compbiomed.2023.107777

[14] Li, T., Gao, Y.Q., Wang, K., Guo, S., Liu, H. and Kang, H. (2019) Diagnostic Assessment of Deep Learning Algorithms for Diabetic Retinopathy Screening. *Information Sciences*, **501**, 511-522. https://doi.org/10.1016/j.ins.2019.06.011

[15] Parhi, R. and Nowak, R.D. (2020) The Role of Neural Network Activation Functions. *IEEE Signal Processing Letters*, **27**, 1779-1783. https://doi.org/10.1109/lsp.2020.3027517

[16] Deepa, V., Kumar, C.S. and Cherian, T. (2022) Ensemble of Multi-Stage Deep Convolutional Neural Networks for Automated Grading of Diabetic Retinopathy Using Image Patches. *Journal of King Saud University-Computer and Information Sciences*, **34**, 6255-6265. https://doi.org/10.1016/j.jksuci.2021.05.009