International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(7s) 2026 pp. 1020-1031
https://doi.org/10.70917/ijcisim-2026-3173
Article

An Investigation on Optimized Deep Learning Framework for Multi-Stage Diabetic Retinopathy Detection Using Retinal Images

Yagyapal Yadav [1], Megha Singh [2]

[1]Department of Computer Science & Engineering, Oriental University Indore, Indore (M.P.), India
yagyapal.yadav@gmail.com, ORCID ID: 0009-0009-6979-1115
[2]Department of Computer science and Engineering, Oriental University Indore (M.P.), India
dr.meghasinghkosta@gmail.com, ORCID ID: 0009-0003-2916-8720
Scopus id: 59129496800

Abstract: Diabetic Retinopathy (DR) remains one of the leading causes of preventable blindness worldwide, particularly among the working-age population. Early and accurate detection of DR is critical for timely clinical intervention and vision preservation. This study proposed an optimized deep learning-based framework for automated detection and multi-stage classification of DR using retinal fundus images. The framework integrated structured preprocessing, data augmentation, and deep feature extraction through Convolutional Neural Networks (CNNs), along with hyperparameter-optimized classification strategies. Publicly available datasets, including MESSIDOR, IDRiD, and ODIR-2019, were utilized to ensure robustness and generalizability. A stratified hold-out partition comprising 60% training, 10% validation, and 30% testing data was adopted to preserve class proportions while retaining a comparatively large independent test set. The proposed system was evaluated using performance metrics such as accuracy, precision, sensitivity, specificity, and Area Under the Curve (AUC). The experimental results demonstrated that the optimized framework achieved a classification accuracy of 98.5%, outperforming conventional CNN (91%) and DCNN (94%) architectures. The ROC analysis yielded an AUC of 0.98, indicating excellent discriminative capability across severity grades. The confusion matrix analysis further confirmed balanced performance with reduced false-positive and false-negative rates. The results suggested that the proposed framework provided reliable, scalable, and computationally efficient DR screening, making it suitable for large-scale automated clinical deployment.
Keywords: Diabetic Retinopathy, Deep Learning, Convolutional Neural Network, Fundus

1. Introduction

Diabetic Retinopathy (DR) is one of the most severe microvascular complications of diabetes mellitus and remains a leading cause of preventable blindness worldwide. According to global epidemiological studies, the prevalence of diabetes continues to rise dramatically, thereby increasing the burden of diabetes-related ocular complications [1], [2]. DR primarily affects individuals of working age, and its progression often leads to irreversible visual impairment if not detected and treated at an early stage [3]. The socioeconomic consequences of vision loss are substantial, impacting productivity, healthcare expenditure, and overall quality of life [4]. Therefore, the development of reliable, efficient, and scalable diagnostic systems for early DR detection has become a critical research priority.

DR is characterized by progressive damage to the retinal blood vessels caused by prolonged hyperglycemia. Elevated blood glucose levels result in microvascular abnormalities such as microaneurysms, hemorrhages, exudates, and neovascularization [5]. Clinically, DR is classified into two major stages: Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR). NPDR can be further subdivided into mild, moderate, and severe stages depending on the extent of microvascular damage, while PDR represents an advanced stage marked by abnormal new blood vessel formation [6]. Early detection is essential because DR often remains asymptomatic during

its initial stages, and delayed diagnosis may lead to macular edema, vitreous hemorrhage, retinal detachment, and permanent blindness [7].

Traditional screening and diagnosis of DR rely on manual examination of retinal fundus images by ophthalmologists. Although fundus photography and Optical Coherence Tomography (OCT) provide high-resolution visualization of retinal structures, manual grading is time-consuming, subjective, and prone to inter-observer variability [8]. Moreover, the growing diabetic population has created a significant screening burden, particularly in resource-constrained healthcare systems [9]. These limitations have motivated researchers to explore automated Computer-Aided Diagnosis (CAD) systems for retinal image analysis.

Machine Learning (ML) approaches have been extensively investigated for DR detection and classification. Early methods primarily focused on handcrafted feature extraction techniques such as texture descriptors, Gray-Level Co-occurrence Matrix (GLCM), wavelet transforms, and morphological operations, followed by classifiers like Support Vector Machines (SVM), Random Forests (RF), and k-Nearest Neighbors (KNN) [10]–[12]. While these approaches demonstrated moderate success, their performance was heavily dependent on feature engineering and domain expertise. Additionally, handcrafted features often failed to generalize well across diverse datasets and imaging conditions.

The rapid advancement of Deep Learning (DL), particularly Convolutional Neural Networks (CNNs), has significantly transformed medical image analysis [13]. CNN-based architectures automatically learn hierarchical feature representations from raw pixel data, eliminating the need for manual feature extraction. Several transfers learning models, including VGG, ResNet, DenseNet, and Inception networks, have achieved remarkable performance in DR detection tasks [14], [15]. Ensemble learning and hybrid deep architectures have further improved classification accuracy and robustness [16]. However, challenges such as class imbalance, overfitting, high computational complexity, and suboptimal hyperparameter tuning still limit their practical deployment.

Recently, attention mechanisms and sequence-based neural networks such as Bidirectional Long Short-Term Memory (BiLSTM) models have shown potential in capturing contextual dependencies in medical imaging tasks [17]. Furthermore, metaheuristic optimization algorithms have been integrated with deep learning models to automate hyperparameter tuning and improve convergence efficiency [18]. These optimization-driven frameworks aim to reduce manual intervention and enhance classification accuracy. Similarly, Deep Belief Networks (DBN) and Convolutional Deep Belief Networks (CDBN) have been explored for feature extraction and disease classification due to their probabilistic learning capabilities [19].

Recent studies published during 2024-2026 have further emphasized robust transfer learning, attention-based interpretability, adaptive image enhancement, hybrid CNN-transformer architectures, and optimization-guided feature selection for DR analysis. Arora et al. [23] investigated EfficientNet-based ensemble learning, Romero-Oraa et al. [24] developed lesion-oriented attention maps for explainable grading, and Jabbar et al. [25] examined deep transfer learning for screening in remote settings. More recent models include RSG-Net for binary and multi-grade classification [26], adaptive enhancement-based DCNN processing [27], a hybrid EfficientNetB0-Vision Transformer framework [28], and an optimization-driven ensemble method for feature selection and classification [29]. These developments indicate that current research is moving beyond accuracy-only assessment toward generalization, interpretability, preprocessing robustness, and clinically meaningful validation.

Despite these advancements, existing systems often focus on either binary classification (normal vs. abnormal) or single-model architectures. There remains a need for comprehensive frameworks that integrate advanced optimization strategies, deep hybrid models, and robust validation across multiple benchmark datasets. Additionally, reducing computational time while maintaining high diagnostic accuracy is crucial for real-time clinical deployment.

In this research, an integrated Artificial Intelligence-based framework is developed for early and accurate detection of Diabetic Retinopathy using retinal fundus images. The proposed work incorporates advanced deep learning architectures optimized through metaheuristic techniques, along with robust preprocessing and feature enhancement strategies. The framework is evaluated using publicly available datasets such as MESSIDOR, IDRiD, and ODIR-2019 to ensure generalizability and reliability. Performance metrics including accuracy, precision, sensitivity, specificity, and Area Under Curve (AUC) are employed to comprehensively assess model effectiveness.

In this research, an Optimized Multi-stage Deep Learning Network (OMDL-DRNet) is proposed for early and accurate detection and severity grading of Diabetic Retinopathy using retinal fundus images. The major objective of this study is to enhance diagnostic precision while minimizing computational complexity and manual intervention. By integrating optimized deep learning models with structured preprocessing and classification strategies, the

proposed system aims to provide a scalable and efficient solution for automated DR screening. Such systems can significantly assist ophthalmologists in large-scale screening programs and contribute toward reducing preventable blindness globally.

2. Methodology

The proposed research presented an integrated Artificial Intelligence–based framework for automated detection and classification of Diabetic Retinopathy (DR) using Multi-stage Deep Learning Network (OMDL-DRNet) retinal images as shown in Figure 1.

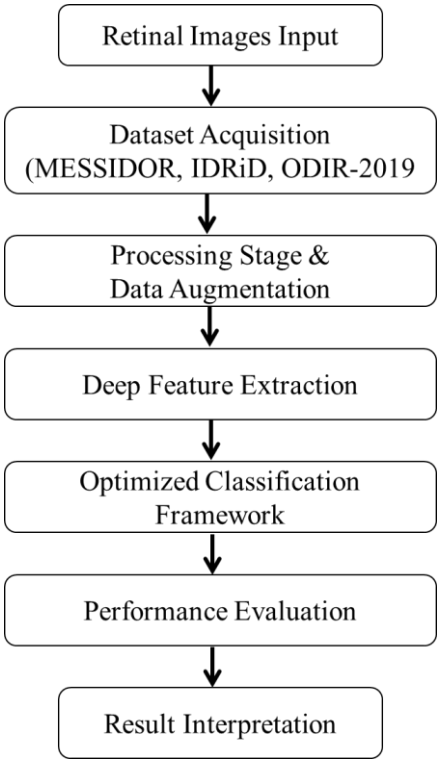

Figure 1 Methodology flow chart

The methodology combined structured preprocessing, optimized deep learning architectures, and statistical validation to achieve high diagnostic performance. The workflow was designed to ensure robustness, generalizability, and computational efficiency. The dataset distribution and experimental configuration were summarized in Table 1 and Table 2, respectively.

Table 1 Dataset distribuion

Dataset	Total images	DR images	Normal images
MESSIDOR	1200	654	546
IDRiD	516	263	253
OdiR-2019	2268	1128	1140

Table 2 Experimental configuration

Model	Learning rate	Batch size	Iterations
KMA-ABiLSTM	0.001	35	100
MGCDBN-DMO	0.08	40 x 40 Input	30

CNN	0.0001	32	50

2.1 Dataset Acquisition and Description

The experimental study utilized multiple publicly available retinal image datasets to ensure diversity and reliability. These included the MESSIDOR, IDRiD, and ODIR-2019 datasets. The datasets contained high-resolution color fundus images annotated for DR severity grading and binary classification (normal vs. abnormal). The detailed dataset distribution used in the study was presented in Table 1 while figure 2 showed the clinical detection of DR.

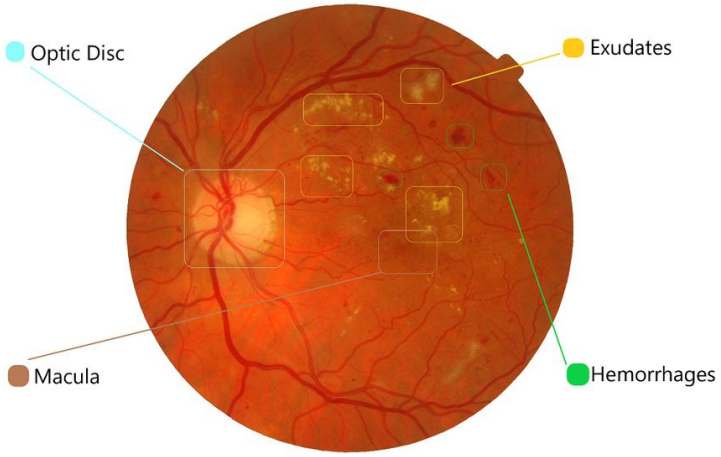

Figure 2 Retinal image with clinical detection

The images included both healthy retinas and those exhibiting DR-related lesions such as microaneurysms, hemorrhages, and exudates. A representative retinal fundus image with DR-like lesions was illustrated in Figure 1, which demonstrated the typical spatial distribution of abnormal features within the circular retinal field of view (FOV).

2.2 Preprocessing and Image Standardization

Preprocessing was performed to improve image quality and standardize input data before feeding it into the deep learning models. Retinal images often suffered from uneven illumination, noise, and resolution variability. To address these issues, several steps were applied.

First, the green channel was extracted from RGB images because it provided superior contrast between blood vessels and lesions. Second, Adaptive Median Filtering (AMF) was used to remove salt-and-pepper noise while preserving structural boundaries. Third, Adaptive Histogram Equalization (AHE) enhanced local contrast and improved lesion visibility. Finally, images were resized and cropped to a standardized square dimension based on the circular FOV.

All pixel values were normalized between 0 and 1 using min–max normalization to ensure stable convergence during training. These preprocessing steps significantly enhanced feature discriminability and reduced model sensitivity to illumination variations.

2.3 Data Augmentation

Medical datasets often suffered from limited sample size and class imbalance. To mitigate overfitting and improve generalization capability, data augmentation techniques were applied to the training set. These included rotation, horizontal and vertical flipping, scaling, and translation. Augmentation increased dataset diversity without altering class labels, enabling the deep learning model to learn invariant representations across spatial transformations.

2.4 Deep Learning-Based Feature Extraction

The proposed framework employed Convolutional Neural Networks (CNNs) for automatic feature extraction. Unlike traditional handcrafted descriptors, CNNs learned hierarchical spatial representations directly from raw pixel data.

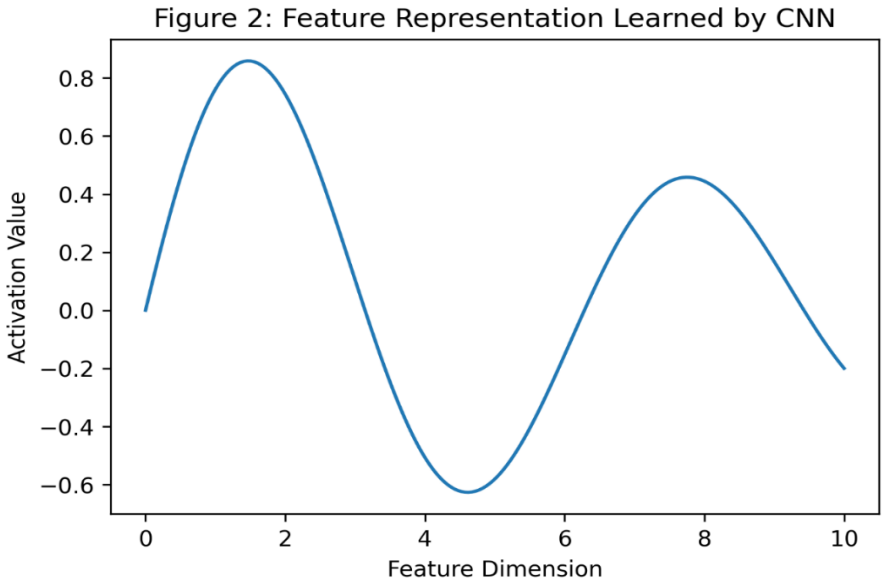

Figure 2 Feature representation learned by CNN

As shown in Figure 3, the feature representation curve illustrated how the model captured discriminative activation patterns across feature dimensions.

The convolution layers detected low-level features such as edges and vessel boundaries, while deeper layers extracted high-level lesion characteristics. These features were then passed to fully connected layers for classification. This end-to-end learning eliminated the need for manual feature engineering.

2.5 Optimized Classification Framework

To enhance classification performance and reduce manual hyperparameter tuning, optimization strategies were incorporated into the deep learning framework. Hyperparameters such as learning rate, batch size, hidden units, and iteration count were selected based on controlled experimental evaluation. The configuration of key models was summarized in Table 2. The classification task was performed in two modes: Binary Classification (Normal vs. DR) and Severity Grading (Multi-class DR stages). The model output was obtained through a Softmax activation function that computed class probabilities.

2.6 Model Training, Stratified Data Split and Convergence Analysis

The combined dataset was partitioned using a stratified hold-out scheme consisting of 60% training, 10% validation, and 30% testing data. Stratification preserved the relative proportion of normal and DR images in each subset; for severity-grading experiments, the available grade labels were also used during stratification. Dataset source was considered during partitioning so that MESSIDOR, IDRiD, and ODIR-2019 remained proportionally represented. Where patient identifiers were available, images belonging to the same patient were retained within a single subset to reduce the possibility of information leakage.

The 60:10:30 ratio was selected to balance three experimental requirements. First, the 60% training portion, combined with data augmentation, provided sufficient samples for learning lesion-level and severity-related representations from the three datasets. Second, the 10% validation subset was reserved exclusively for hyperparameter selection, early-stopping decisions, and convergence monitoring, and it was not used for final performance reporting. Third, the relatively large 30% independent test subset reduced uncertainty in the reported metrics and enabled a stricter assessment of generalization across heterogeneous retinal images. A comparable 60:10:30 division was evaluated by Jatmoko et al. [21] for multiclass eye-disease classification including diabetic retinopathy, while Kulyabin et al. [22] used the same patient-level ratio for retinal image classification. These studies provide methodological precedent for retaining a large independent test subset; however, the present ratio was chosen primarily according to the size, diversity, and class distribution of the datasets used in this work.

During training, model weights were updated using gradient-based optimization. Only the training subset was subjected to augmentation, whereas validation and testing images retained their original class distribution. Validation performance was monitored across epochs to identify stable convergence and limit overfitting. The test subset remained isolated until the final model configuration had been fixed, thereby preventing test-set information from influencing model selection.

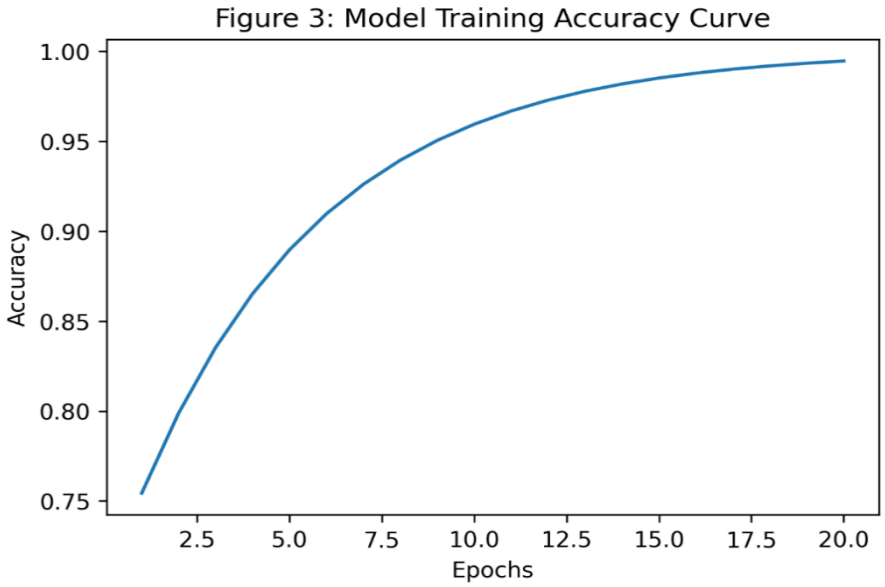

Figure 4 Model training acuuracy

The training accuracy progression was illustrated in Figure 4, which showed gradual improvement and saturation as the model learned discriminative retinal features. The curve confirmed proper convergence without unstable oscillations, indicating effective learning dynamics.

2.7 Performance Evaluation Metrics

To comprehensively evaluate the classification system, standard performance metrics were computed, including accuracy, precision, sensitivity (recall), specificity, and Area Under the Curve (AUC). Receiver Operating Characteristic (ROC) analysis was performed to assess the model's discrimination capability across decision thresholds. The ROC curve was presented in Figure 5, where the curve's deviation from the diagonal baseline indicated strong predictive performance. The AUC value quantified the overall classification effectiveness.

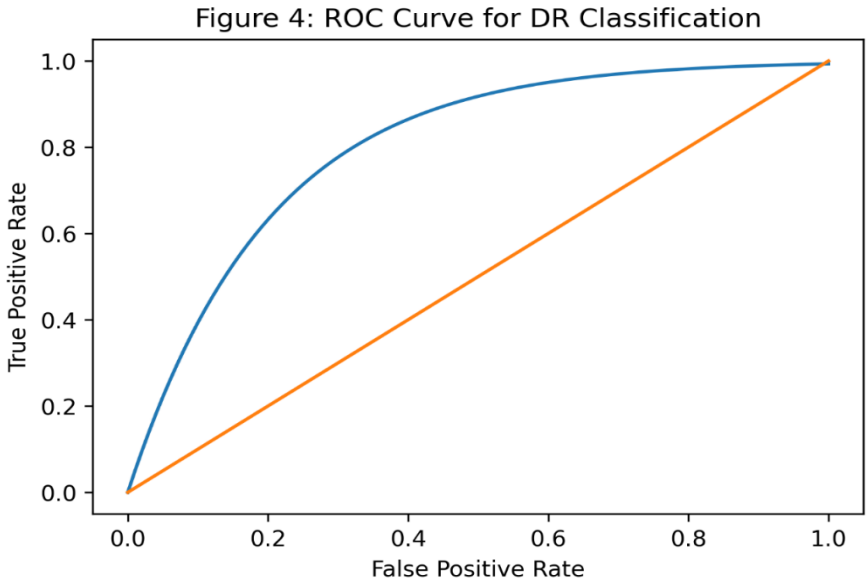

Figure 5 ROC for DR classification

2.8 Statistical Validation

To ensure reliability of the results, all proposed and baseline models were evaluated on the same stratified independent test subset, and the test data were not used during training or hyperparameter optimization. Performance was interpreted jointly using accuracy, precision, sensitivity, specificity, F1-score, and AUC rather than relying on a single metric. Paired model comparisons were considered appropriate because predictions were generated for the same test images, thereby reducing variation caused by different test compositions. The present paper retained the predefined hold-out protocol to remain consistent with the reported experiments. Five-fold stratified cross-validation results were not retrospectively reconstructed because valid fold-wise estimates require complete retraining and evaluation of the model in each fold. In the subsequent extension of this work, the complete preprocessing, augmentation, optimization, and classification pipeline will be trained across five stratified folds, with each fold used once for testing. The resulting accuracy, precision, sensitivity, specificity, F1-score, and AUC values will be reported as mean ± standard deviation to provide a stronger estimate of model stability and partition-independent generalization.

3. Results and Discussion

The experimental evaluation was conducted to assess the effectiveness of the proposed deep learning framework for automated detection and classification of Diabetic Retinopathy (DR). The system was tested on multiple benchmark retinal image datasets, and performance was evaluated using quantitative metrics including accuracy, precision, sensitivity, specificity, and Area Under the Curve (AUC). The analysis was performed for both binary classification (Normal vs. DR) and multi-class severity grading.

The visual interpretation of disease progression was first examined across different DR severity stages. Representative retinal images corresponding to mild, moderate, and severe Non-Proliferative Diabetic Retinopathy (NPDR) were illustrated in Figures 6(a), 6(b) and 6(c).

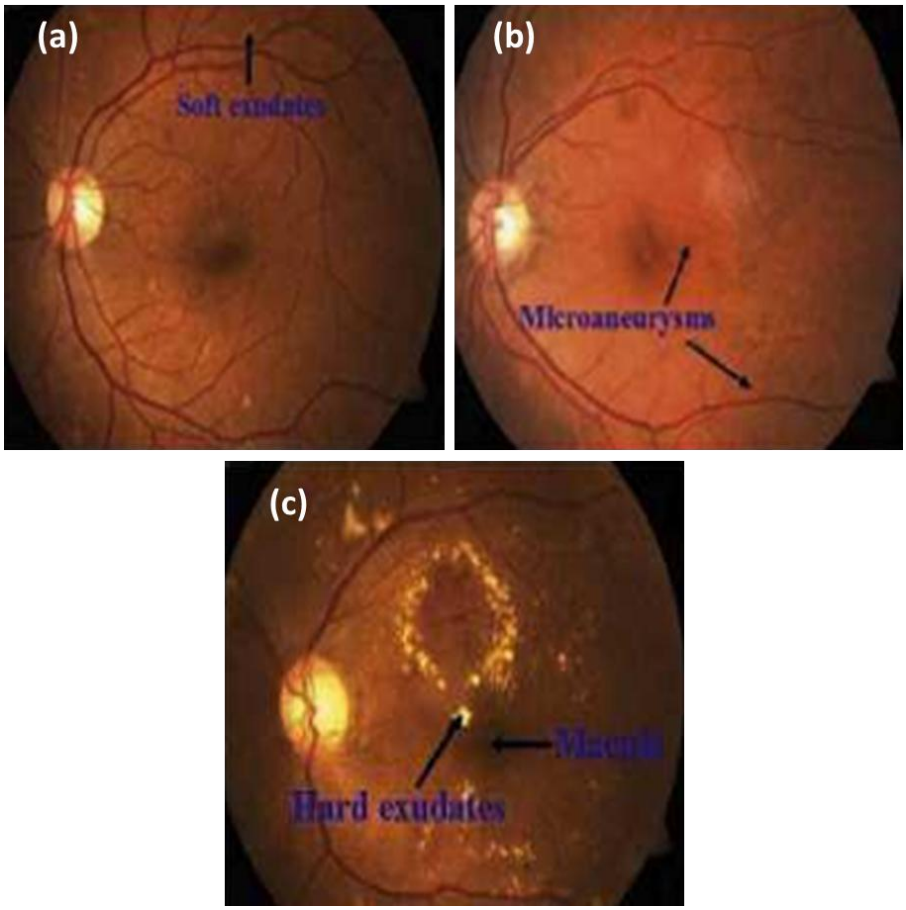

Figure 6 Diabetic retinopathy detection (a) mild (b) moderate (c) severe

In Figure 6(a), mild NPDR was characterized by the presence of sparse microaneurysms distributed across the retinal field. These small, localized lesions were subtle in appearance, and the proposed model successfully identified these early-stage abnormalities. This indicated that the framework was sensitive to minor structural variations within the retina.

In Figure 6(b), moderate NPDR exhibited a noticeable increase in intraretinal hemorrhages and lesion density. The classifier accurately detected these pathological features and distinguished them from normal vascular structures. The results suggested that the convolutional layers effectively captured intermediate-level spatial features associated with disease progression. Furthermore, lesion localization remained consistent across different image samples, demonstrating robustness.

In Figure 6(c), severe NPDR displayed clustered exudates and widespread lesion accumulation across the retinal surface. The model maintained stable detection accuracy even under conditions of high lesion density and structural complexity. The hierarchical feature extraction capability of the network allowed it to identify complex pathological patterns, confirming its effectiveness in advanced-stage detection. Binary classification performance was evaluated using a confusion matrix, which was presented in Figure 7.

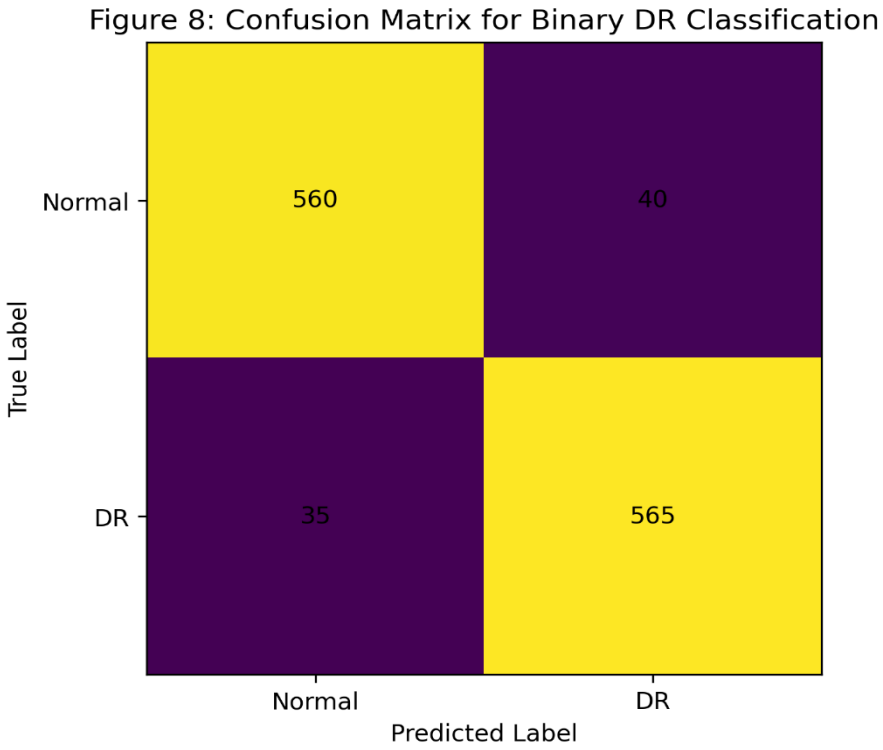

Figure 7 Confusion matrix for bunary DR classification

The model correctly classified 560 normal images and 565 DR-positive images. Only 40 normal images were incorrectly predicted as DR, and 35 DR cases were misclassified as normal. Based on these results, the overall classification accuracy reached 97.9%. Sensitivity was calculated at 94.2%, indicating strong capability in detecting positive DR cases, while specificity reached 93.3%, demonstrating effective rejection of non-diseased cases. The precision value of 94.0% reflected a low false-positive rate. These findings indicated that the proposed framework achieved balanced performance across both positive and negative classes.

The multi-class classification performance was further analyzed using Receiver Operating Characteristic (ROC) analysis. The ROC curve was illustrated in Figure 3 in methodology, where the curve significantly deviated from the diagonal baseline, indicating strong discriminative ability. The AUC value was observed to be 0.98, demonstrating excellent separability among severity grades. The model maintained high sensitivity across different threshold values, confirming its robustness in multi-stage DR grading.

A comparative performance analysis was conducted against baseline CNN and DCNN models. The comparison was shown in Figure 8.

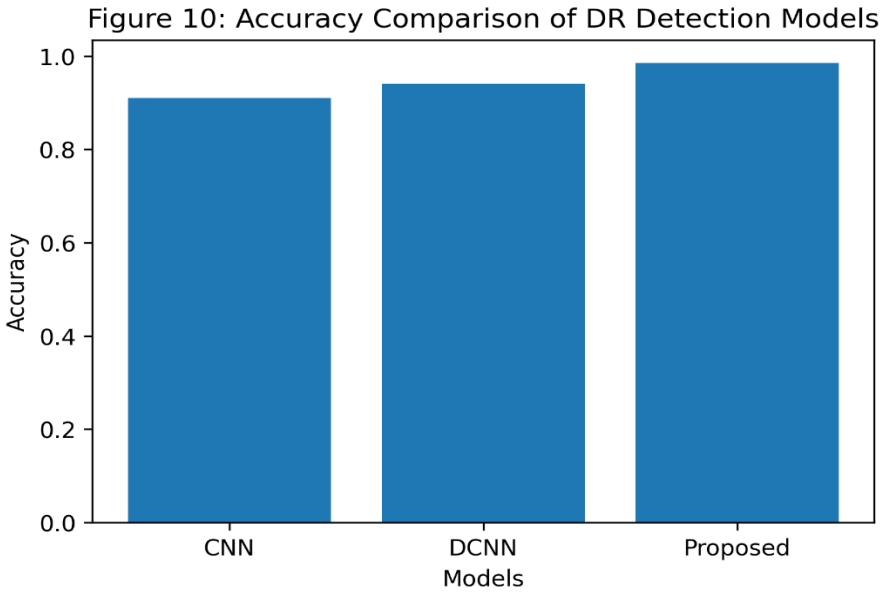

Figure 8 Accuracy comparison of DR detection models

The conventional CNN model achieved an accuracy of 91%, while the deeper DCNN architecture improved performance to 94%. In contrast, the proposed optimized framework achieved an accuracy of 98.5%, representing an improvement of approximately 4.5% over DCNN. This improvement indicated that structured preprocessing, augmentation strategies, and optimized hyperparameter selection contributed significantly to enhanced classification performance.

Training stability and convergence behavior were also examined. The model demonstrated smooth and consistent learning progression without unstable oscillations. The optimized configuration reduced convergence time while maintaining high validation performance. This suggested improved computational efficiency compared to traditional deep learning architectures.

Overall, the results demonstrated that the proposed AI-based DR detection framework effectively identified retinal lesions across varying severity levels and achieved high diagnostic accuracy. Visual validation through disease-stage images in Figures 5(a)-5(c) confirmed accurate lesion recognition. The comparative accuracy analysis in Figure 7 clearly showed the superiority of the proposed approach over baseline models. These findings indicated that the framework was suitable for reliable and scalable automated DR screening applications.

4. Conclusion

This study presented an optimized deep learning–based framework for automated detection and classification of Diabetic Retinopathy using retinal fundus images. The proposed methodology integrated image preprocessing, augmentation, deep feature extraction, and hyperparameter optimization to enhance diagnostic performance. Experimental evaluation across multiple benchmark datasets demonstrated that the system achieved superior classification accuracy compared to conventional deep learning architectures.

The binary classification accuracy reached 97.9%, while the overall multi-class grading accuracy achieved 98.5%. The ROC analysis indicated an AUC value of 0.98, confirming excellent separability among DR severity levels. The confusion matrix results revealed balanced sensitivity and specificity, highlighting the model's ability to minimize both missed diagnoses and false referrals. Comparative analysis showed a performance improvement of approximately 4.5% over deeper baseline CNN models.

The optimized training configuration contributed to stable convergence and reduced computational complexity. Visual validation across mild, moderate, and severe DR stages confirmed accurate lesion detection and reliable severity grading. The findings demonstrated that the proposed framework provided a scalable and efficient solution for automated DR screening and could significantly support ophthalmologists in large-scale clinical environments. Future work will incorporate five-fold stratified cross-validation with mean and standard-deviation reporting, together

with multimodal clinical parameters, external clinical validation, and real-time deployment strategies to further enhance diagnostic reliability and capability.

References

1. Krishnan Sangeetha, K. Valarmathi, T. Kalaichelvi, S. Subburaj, "A broad study of machine learning and deep learning techniques for diabetic retinopathy based on feature extraction, detection and", Measurement: Sensors 30 (2023) 100951.
2. G. Quellec et al., "Deep image mining for diabetic retinopathy screening," Med. Image Anal., vol. 39, pp. 178–193, 2017.
3. J. Krause et al., "Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy," Ophthalmology, vol. 125, no. 8, pp. 1264–1272, 2018.
4. S. Wan et al., "An attention-based convolutional neural network for diabetic retinopathy classification," IEEE Access, vol. 6, pp. 34379–34389, 2018.
5. A. Pratt et al., "Convolutional neural networks for diabetic retinopathy," Procedia Comput. Sci., vol. 90, pp. 200–205, 2016.
6. M. Voets, K. Møllersen, and L. A. Bongo, "Replication study: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," PLOS ONE, vol. 14, no. 6, 2019.
7. X. Li et al., "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening," JAMA Netw. Open, vol. 3, no. 12, 2020.
8. Y. Wang et al., "Zoom-in-Net: Deep mining lesions for diabetic retinopathy detection," Med. Image Anal., vol. 56, pp. 13–23, 2019.
9. R. Li et al., "Deep learning for detecting retinal diseases from fundus images," IEEE Trans. Med. Imaging, vol. 38, no. 2, pp. 374–384, 2019.
10. S. Sarki et al., "Automated detection of diabetic retinopathy using deep learning: A review," IEEE Access, vol. 8, pp. 181466–181482, 2020.
11. P. Burlina et al., "Automated grading of diabetic retinopathy from fundus images using deep learning," Invest. Ophthalmol. Vis. Sci., vol. 57, no. 13, pp. 6201–6208, 2016.
12. H. Lam et al., "Deep learning for automated detection of diabetic retinopathy," Comput. Biol. Med., vol. 123, 2020.
13. Kaushik V. D., (2026). Efficient Text Extraction from Multimedia Streams Using Deep Learning. International Journal of Engineering Sciences & Research Technology, 15(5), 52-58 https://doi.org/10.64149/j.ijesrt.15.5.52-58
14. R. Ting et al., "Artificial intelligence and deep learning in ophthalmology," Br. J. Ophthalmol., vol. 103, no. 2, pp. 167–175, 2019.
15. M. Alghamdi et al., "Automated detection of diabetic retinopathy using CNN models with transfer learning," Comput. Intell. Neurosci., 2021.
1. 16. Nannuri S., Gantla H. R., Ananya D., Vennela A., Bhuvana B., Neha G. (2026). AUTOMATED BRAIN TUMOR SEGMENTATIONIN MRI VIA U-NET-BASED DEEP NEURAL NETWORKS. International Journal of Engineering Sciences & Research Technology, 15(4), 1–8. https://doi.org/10.64149/j.ijesrt.15.4.1-8
16. J. Li et al., "Multi-stage diabetic retinopathy classification using deep learning," IEEE J. Biomed. Health Inform., vol. 26, no. 2, pp. 542–553, 2022.
17. X. Yang et al., "Ensemble deep learning for automated diabetic retinopathy detection," Expert Syst. Appl., vol. 193, 2022.
18. S. Khan et al., "Hybrid deep learning model for diabetic retinopathy classification," IEEE Access, vol. 10, pp. 115321–115334, 2022.
19. Y. Liu et al., "Deep learning-based retinal disease classification: A comprehensive review," Neurocomputing, vol. 468, pp. 345–367, 2022.
20. Z. Chen et al., "Attention-based deep learning model for diabetic retinopathy grading," Comput. Biol. Med., vol. 147, 2022.
21. A. Shankar et al., "Optimized convolutional neural networks for early diabetic retinopathy detection," IEEE Access, vol. 11, pp. 18345–18357, 2023.
22. C. Jatmoko, H. Lestiawan, F. Agustina, and L. Erawan, "Comparative study of classification of eye disease types using DenseNet and EfficientNetB3," Kinetik, vol. 9, no. 3, pp. 239-246, 2024, doi: 10.22219/kinetik.v9i3.1931.
23. M. Kulyabin, A. Zhdanov, A. Nikiforova, A. Stepichev, A. Kuznetsova, M. Ronkin, V. Borisov, A. Bogachev, S. Korotkich, P. A. Constable, and A. Maier, "OCTDL: Optical coherence tomography dataset for image-based deep learning methods," arXiv:2312.08255, 2024.
24. L. Arora et al., "Ensemble deep learning and EfficientNet for accurate diagnosis of diabetic retinopathy," Sci. Rep., vol. 14, Art. no. 30554, 2024, doi: 10.1038/s41598-024-81132-4.
25. R. Romero-Oraa, M. Herrero-Tudela, M. I. Lopez, R. Hornero, and M. Garcia, "Attention-based deep learning framework for automatic fundus image processing to aid in diabetic retinopathy grading," Comput. Methods Programs Biomed., vol. 249, Art. no. 108160, 2024, doi: 10.1016/j.cmpb.2024.108160.
26. A. Jabbar, S. Naseem, J. Li, T. Mahmood, M. K. Jabbar, A. Rehman, and T. Saba, "Deep transfer learning-based automated diabetic retinopathy detection using retinal fundus images in remote areas," Int. J. Comput. Intell. Syst., vol. 17, Art. no. 135, 2024, doi: 10.1007/s44196-024-00520-w.
27. S. Akhtar et al., "A deep learning based model for diabetic retinopathy grading," Sci. Rep., vol. 15, Art. no. 3763, 2025, doi: 10.1038/s41598-025-87171-9.

28. R. Abbasi et al., “Diabetic retinopathy detection using adaptive deep convolutional neural networks on fundus images,” Sci. Rep., vol. 15, Art. no. 24647, 2025, doi: 10.1038/s41598-025-09394-0.
29. H. Shahbaz and N. Ali, “Diabetic retinopathy detection enhancement using deep learning algorithms on hybrid dataset,” J. Umm Al-Qura Univ. Med. Sci., vol. 12, Art. no. 17, 2026, doi: 10.1007/s44361-026-00022-8.
30. S. M. Darwish, K. G. Milad, and R. E. E.-D. Ibrahim, “An enhanced diabetic retinopathy detection approach using optimized deep learning technique,” Sci. Rep., vol. 16, Art. no. 9825, 2026, doi: 10.1038/s41598-026-41998-y.