Journal of Science Engineering Technology and Management Science
Volume 02, Issue 04, April 2025 ISSN: 3049-0952
www.jsetms.com DOI:10.63590/jsetms.2025.v02.i04.94-104

A DEEP LEARNING MODEL WITH SMART PADDING FOR EARLY RETINAL DISEASE DIAGNOSIS

U. Satya Narayana, B. Navitha, Ch. Lathasri, M.V.P. Praneetha, N. Sukanya Reddy

Department of Computer Science and Engineering (AI&ML), Geethanjali Institute of Science and Technology, Nellore, Andhra Pradesh, India.

To Cite this Article

U. Satya Narayana, B. Navitha, Ch. Lathasri, M.V.P. Praneetha, N. Sukanya Reddy," A Deep Learning Model With Smart Padding For Early Retinal Disease Diagnosis", Journal of Science Engineering Technology and Management Science, Vol. 02, Issue 04, April 2025,pp: 94-104, DOI: http://doi.org/10.63590/jsetms.2025.v02.i04.pp94-104

Submitted: 09-03-2025 Accepted: 18-04-2025 Published: 26-04-2025

ABSTRACT

The project introduces an AI-powered automated diagnosis system designed to accurately and efficiently detect eye diseases from retinal images. By combining advanced deep learning methodologies with an intuitive graphical user interface, the system provides clinicians and researchers with a reliable tool for early disease detection and in-depth analysis. The application begins with robust dataset management capabilities, allowing users to upload retinal images and categorize them into specific diagnostic classes, including Diabetic Retinopathy (DR), Macular Hole (MH), Normal, and Other Diseases/Conditions (ODC). Essential preprocessing steps such as image resizing, normalization, and advanced data augmentation are incorporated to increase dataset diversity and enhance model generalization. At the core of the system are two deep learning model architectures. The first is a pre-existing Deep Neural Network (DNN) trained using the Stochastic Gradient Descent (SGD) optimizer. This model is composed of multiple convolutional layers, batch normalization, pooling, and dense layers, enabling it to effectively extract and learn complex features from retinal images. The second model is a newly proposed Convolutional Neural Network (CNN) utilizing the Adam optimizer with a specific configuration of valid padding (AVP). This architecture incorporates additional dropout layers and batch normalization to reduce overfitting and improve generalization capabilities. Both models are rigorously evaluated using key performance metrics such as accuracy, precision, recall, and F1-score. Detailed classification reports and confusion matrices offer comprehensive insights into the models' diagnostic performance. Experimental results show that while both models perform reliably, the proposed CNN with AVP outperforms the DNN in terms of diagnostic accuracy and the ability to distinguish between subtle disease features. The system's modular architecture and extensive use of data augmentation make it adaptable to various clinical datasets and scalable for real-world implementation. This research not only demonstrates the effectiveness of deep learning in the medical imaging domain but also delivers a practical, user-focused solution for automated retinal disease diagnosis contributing significantly to the development of next-generation computer-aided diagnostic systems.

Keywords: Retinal Disease Diagnosis, Macular Hole, Medical Image Analysis, Deep Learning, Computer-Aided Diagnosis.

1. INTRODUCTION

The retinal disease (RD) classification involves the categorization of retinal images to identify signs of the condition in patients with diabetes. Traditionally, this process requires patients to visit hospitals or clinics for screening, where healthcare professionals manually examine the images [1]. However, this method is time-consuming, often leading to delays in treatment initiation due to the requirement for human interpretation. To address this issue, researchers are implementing AI based classification systems for RD. These systems utilize advanced algorithms, such as Deep Learning (DL), and Machine Learning (ML), to analyse retinal images and detect signs of RD automatically. Doctors utilize various diagnostic tools such as fundus photography and optical coherence tomography (OCT) to capture detailed images of the retina. These images provide critical information about the presence and extent of retinal damage caused by diabetes [2]. Through careful examination of these images, doctors can identify characteristic signs of RD, including microaneurysms, haemorrhages, exudates, and neovascularization. Each of these signs plays a vital in determining the severity of the condition and guiding treatment decisions. The process of RD classification involves interpreting retinal images and assigning them to appropriate categories depends on the severity of retinal abnormalities observed. This classification system helps doctors determine the most suitable course of action for each patient [3], whether it involves regular monitoring, lifestyle modifications, or referral for further treatment. By accurately classifying RD, doctors can intervene early to prevent or delay vision loss in affected individuals. While traditional methods of RD classification rely heavily on manual interpretation of retinal images by trained ophthalmologists, recent advancements in AI have introduced automated classification systems [4]. These AI-based systems leverage ML algorithms to analyze retinal images and classify them depends on predefined criteria. By automating the classification process, AI offers the potential to expedite diagnosis and improve the efficiency of RD screening programs.

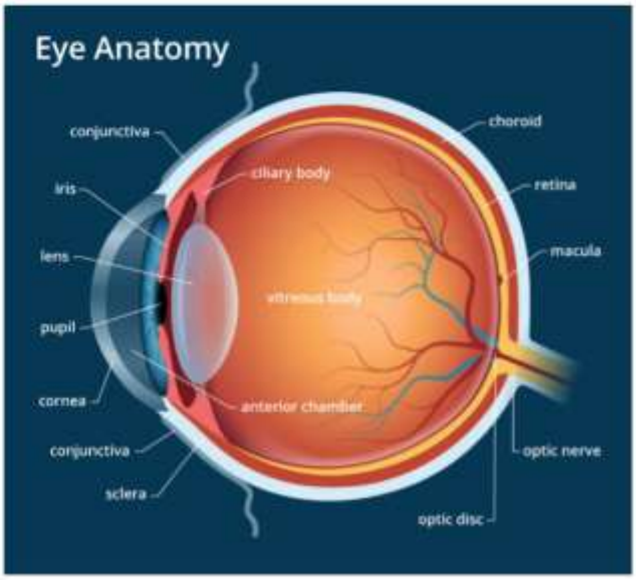


Fig. 1: Eye anatomy.

2. LITERATURE SURVEY

Pachade, Samiksha, et al. [11] proposed the Retinal Fundus Multi-disease Image Dataset (RFMiD), which is specifically designed to advance research in multi-disease detection within retinal imaging. The dataset includes 3,200 retinal fundus images, capturing various pathological conditions such as diabetic retinopathy (DR), age-related macular degeneration (AMD), and glaucoma. The dataset is structured to facilitate multi-label classification tasks, enabling the development of algorithms. RFMiD also supports research on image segmentation and enhancement, providing annotated images that help improve the accuracy and robustness of automated diagnostic models. The dataset's diversity in disease representation and image quality makes it a comprehensive tool for developing and validating new methodologies in retinal disease detection.

Siswadi, et al. [12] introduced a multi-modality and multi-label detection approach for ocular abnormalities utilizing a Transformer-based semantic dictionary learning framework. The multiple modalities such as fundus photography and OCT, enabling a comprehensive analysis of various retinal conditions. The semantic dictionary learning component allows the model to analyse context and relationships among different features, improving its capability to detect multiple diseases simultaneously. By leveraging the Transformer architecture, the approach benefits from its strong capabilities in capturing long-range dependencies and contextual information within the data, which enhances the overall detection accuracy for complex ocular pathologies.

Inan, et al. [13] presented an adaptive multiscale retinal diagnosis methodology utilizing a hybrid trio-model approach for comprehensive fundus multi-disease detection. This work leverages transfer learning and Siamese networks to improvise detection capabilities across different scales of retinal images. The adaptive multiscale approach allows the model to effectively capture features at various resolutions, which is critical for identifying diverse retinal conditions that manifest at different scales. The trio-model integrates three distinct models that specialize in different aspects of feature extraction and classification, combining their strengths to achieve superior diagnostic performance. This hybrid model is particularly effective in handling the variability in retinal image quality and disease manifestation.

Elsayed, et al. [14] developed computer-aided multi-label retinopathy diagnosis framework that incorporates inter-disease graph regularization. This methodology models the relationships among different retinal diseases utilizing a graph-based approach, which helps in understanding the co-occurrence patterns of diseases. The graph regularization technique enhances the model's capability to adopt the correlations among multiple diseases, thereby improving its performance in multi-label classification tasks. By leveraging this inter-disease dependency, the framework is capable of providing more accurate and comprehensive diagnostic predictions, particularly in cases where multiple retinal diseases are present in the same patient.

Vemparala, Yoshita, et al. [15] introduced OcuVision, CNN-powered framework for analyzing retinal images to diagnose diseases. The proposed methodology utilizes advanced CNN architectures to automatically extract features from retinal images, which are then utilized to classify different retinal conditions such as diabetic retinopathy, glaucoma, and AMD. OcuVision is designed to handle large-scale datasets and is optimized for high-speed and accurate image processing, making it suitable for real-time clinical applications. The framework also includes a mechanism for continuous learning, accepting it to improve its diagnostic accuracy over time as more data becomes available.

3. PROPOSED METHODOLOGY

The integration of ML in RD grading addresses critical gaps in traditional manual methods, offering enhanced efficiency and accuracy. Traditional manual grading of RD images was time-consuming and subject to variability depends on grader's expertise, leading to inconsistencies and potential errors. DL algorithms, particularly CNN, can automate the analysis of retinal images. This automation not only speeds up the grading process but also reduces the variability associated with human interpretation, ensuring more consistent and reliable assessments across different healthcare settings.

Moreover, DL methods can improve the scalability of RD screening, making it feasible to implement widespread and remote screenings. By leveraging DL models, healthcare systems can process and understands a large volume of RD images quickly and accurately, which is particularly beneficial in underserved or resource-limited areas where access to trained ophthalmologists is limited. Additionally, DL models were continuously updated and refined with new data, enabling them to adapt to evolving patterns and trends in RD. This dynamic capability ensures that the grading system remains effective over time and was integrated into comprehensive screening programs to provide early and precise detection of RD, ultimately enhancing patient outcomes and reducing vision problems.

The proposed approach aims to improvise RD grading utilizing DL methods by leveraging the RFMID as presented in Figure 4.1. This approach demonstrates the potential of advanced CNN architectures and optimization methods in enhancing the reliability of RD grading, offering a scalable solution for widespread clinical application.

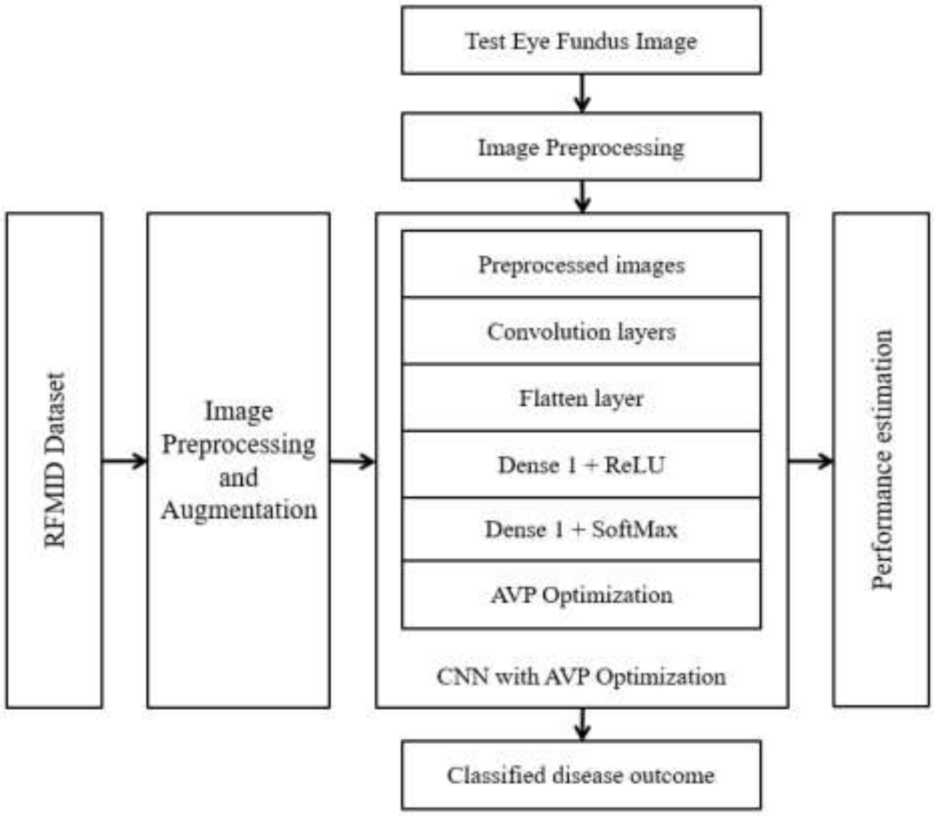


Fig. 2: Proposed RD grading system architecture using CNN with AVP model.

Proposed CNN Classifier with AVP Optimization

In deep CNN for RD classification from OCT images, several key layers and components are commonly employed as presented in the Figure 4.2. The convolution layer plays a pivotal role in extracting features from the input images. To build feature maps that emphasize various patterns and textures within the picture data, it entails sliding a tiny filter or kernel over the input image, executing element-wise multiplications and summations, and then showing the result.

It is common practice to employ a max-pooling layer after the convolution layer in order to down sampling the feature maps, while still preserving the most substantial information. Through a concentration on the most important characteristics, this layer contributes to the reduction of computational complexity and the prevention of overfitting. The ReLU algorithm essentially introduces a threshold for activation by forcing all negative values to be equal to zero. After the convolutional and pooling layers, a flatten layer is applied to the feature maps in order to transform them into a one-dimensional vector. This makes it possible to extract higher-level characteristics and patterns from the flattened feature vector. When doing multi-class classification tasks, such as RD classification, it is common practice to apply a SoftMax classifier after the dense classification layer.

To enable overall architecture to generate a probability distribution that encompasses many classes, SoftMax computes the probabilities of each class and then normalizes those probabilities. It is usual practice to utilize the AVP optimization method during training in order to repeatedly adjust the weights and biases of overall architecture. The AVP makes dynamic adjustments to the learning rate and ensures that all the parameters with unique adaptive learning rates.

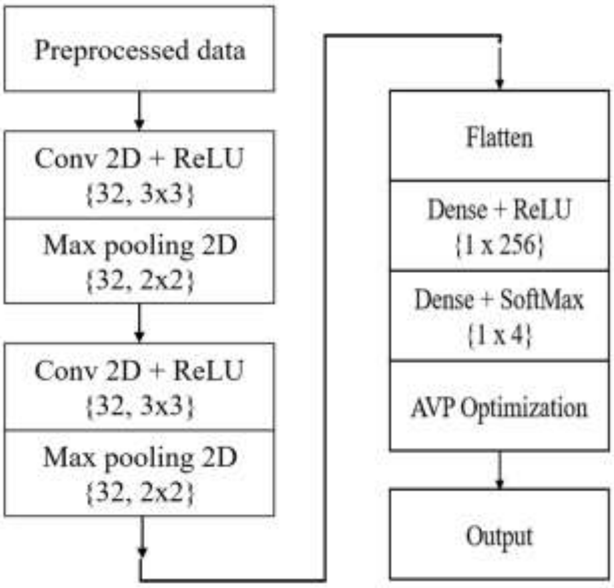


Fig. 3: DCNN Classifier

It is common practice to use binary cross-entropy loss reduction as function of loss while doing classification jobs. So, it assesses the difference among the anticipated probability distribution and the actual distribution of the labels, it is ideal for RD classification. So, training with multiple epochs entails going over the complete dataset numerous times throughout the training process, with each iteration being referred to as an epoch. Through the use of this approach, the network is able to gradually acquire knowledge from the data, so enhancing its performance and refining its weights over the course of succeeding epochs.

Convolution Layer

In the realm of RD diagnosis, CNN employ convolution layers as essential components in their architecture. These convolution layers serve a critical role in extracting features from input data, such as medical images as presented in the Figure 4.3. Each convolution layer contains a series of learnable filters or kernels that systematically convolve across the input data, identifying intricate edges and feature patterns indicative of different RD.

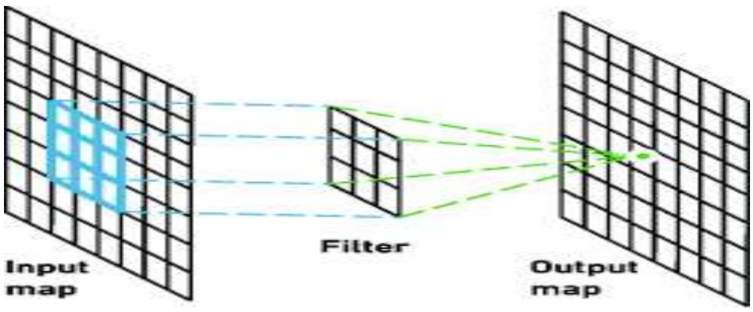


Fig. 4: Convolutional layer.

Through successive convolution layers, the CNN progressively learns classified descriptions of input information, capturing nuanced details critical for accurate diagnosis. By leveraging convolution layers, CNNs efficiently capture spatial dependencies and patterns within the input, enabling the automatic learning of relevant features without the requirement for traditional feature engineering. This hierarchical approach enhances reliability of RD diagnosis by enabling the CNN to discern progressively complicated feature correlated with different eye conditions. The utilization of convolutional layers in CNNs enables the model to efficiently capture spatial dependencies and patterns within the input data, critical for discerning subtle differences associated with various RD.

4. RESULTS AND ANALYSIS

Fig. 5 shows the GUI after the preprocessing stage, indicating that images have been resized, normalized, and that a visual representation of class distribution is provided before data augmentation takes place.

Dataset loaded
Retinal Diseases Class Labels found in dataset : ['DR', 'MH', 'Normal', 'ODC']

Dataset Preprocessing Completed

Total images found in dataset Before Augmentation : 1376

Fig. 5: Illustration of GUI application after performing dataset preprocessing operation.

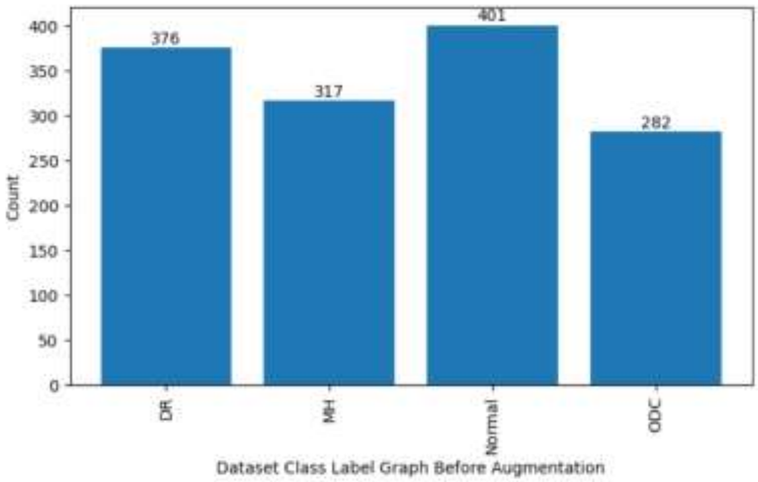


Fig. 6: Dataset labels versus number of sample count before augmentation.

Figure 6 shows distribution of the images across the four classes: DR, MH, Normal, and ODC. The graph displays a bar or column chart with the class names on the x-axis and the number of images on the y-axis. Before augmentation, the graph reflects a notable imbalance among the classes, with 'Normal' being the most represented class (401 images) and 'ODC' being the least (282 images). This disparity indicates potential challenges in training a ML model, as classes with fewer samples (such as ODC) lead to underfitting for those categories, while the model overfit on classes with more samples (such as Normal).

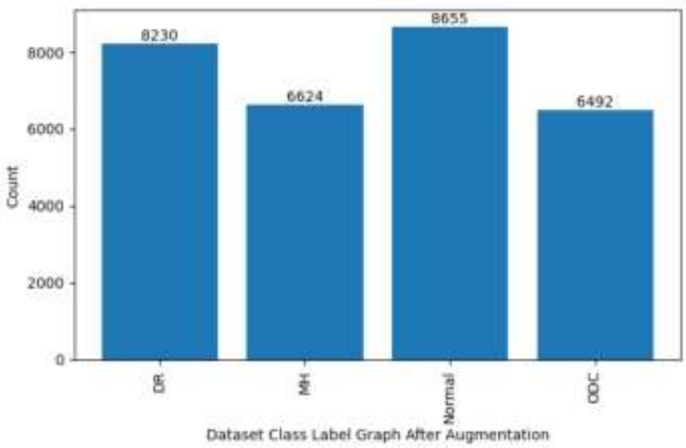


Fig. 7: Bar graph obtained after augmentation (dataset class labels versus count).

In contrast, Fig. 7 demonstrates the outcome of image information augmentation on the dataset. The graph illustrates a more balanced distribution across the four classes after augmentation. The ranging from 6,492 for ODC to 8,655 for Normal. This balanced distribution helps to reduce model bias and advances the model's capability, resulting in better generalization to unseen data. The augmentation strategy effectively addresses the initial imbalance, which is critical for accurate model predictions and fair performance across all classes.

Table 1: Class specific image count analysis.

Class	Before Augmentation	After Augmentation
DR	376	8230
MH	317	6624
Normal	401	8655
ODC	282	6492
Total	1376	30001

Table 1 provides a comprehensive analysis of allocation of images across four different retinal disease classes: DR, MH, Normal, and ODC, both before and after augmentation. The table shows the initial count of images accessible for each class and the increased count after applying data augmentation methods.

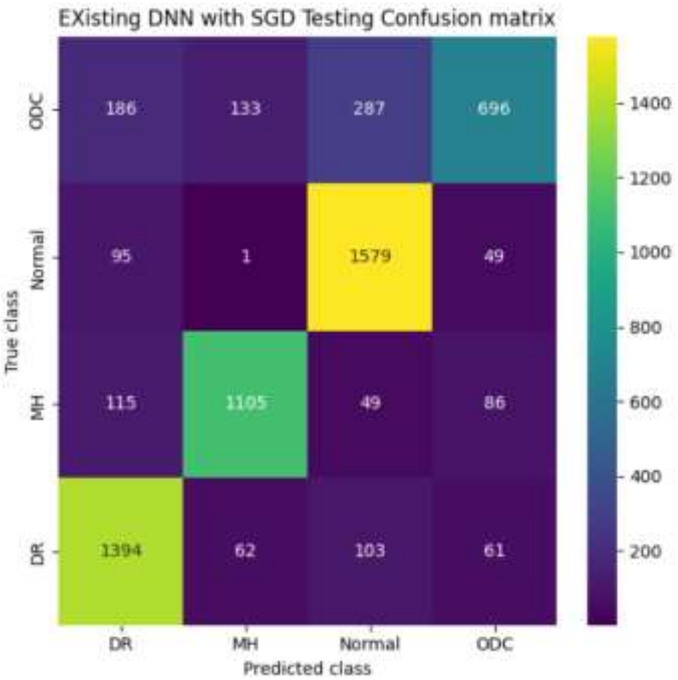


Fig. 8: Confusion matrix obtained using existing DNN with SGD approach.

Fig. 8 complements this by displaying the confusion matrix for the existing model, visually summarizing how well each class was predicted and highlighting misclassifications.

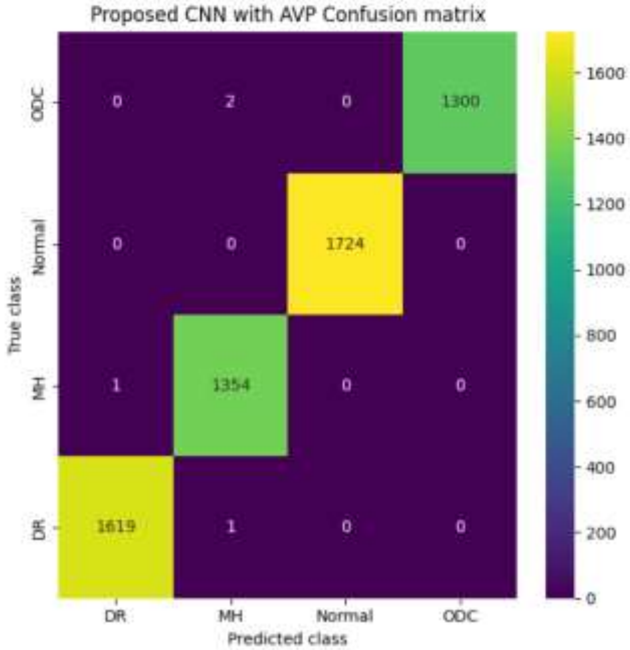
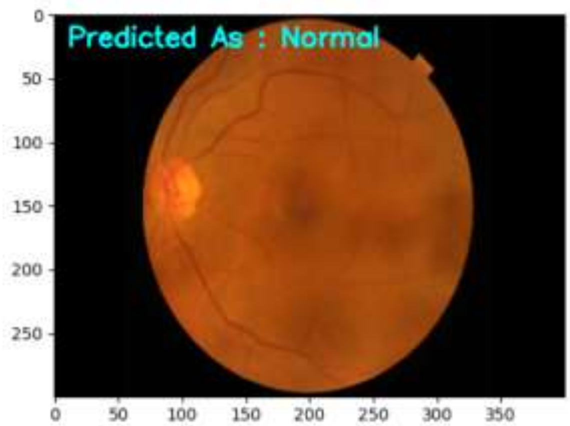
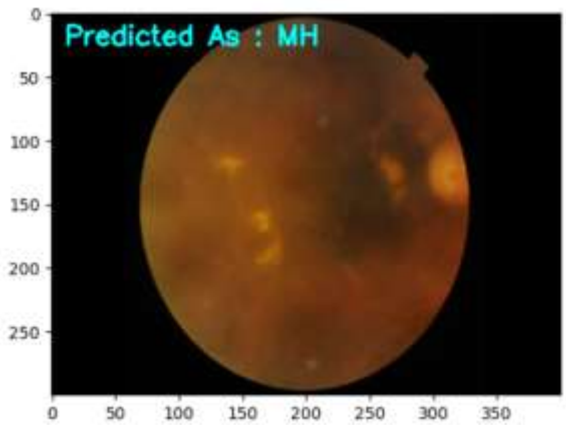


Fig. 9: Confusion matrix obtained using proposed CNN with AVP approach.

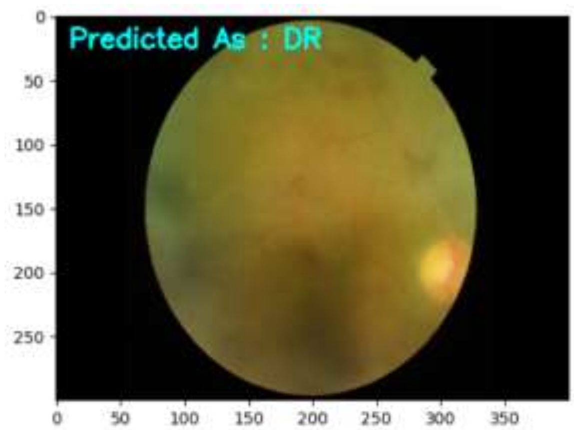
Fig. 9 provides the confusion matrix for the proposed CNN model, offering insight into the model's prediction distribution and revealing its strengths and areas for improvement. Together, these figures offer a complete view of the system's functionality, from user interaction and data handling to model training and evaluation.



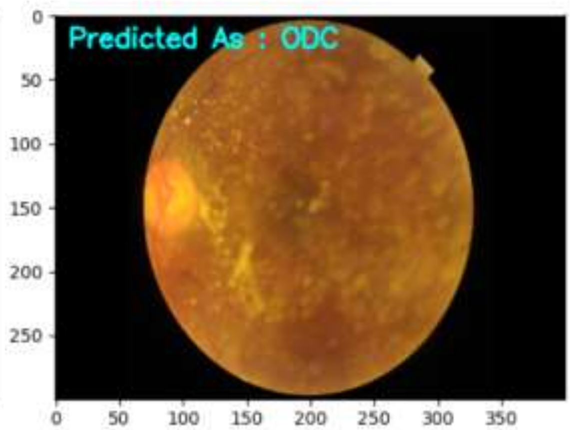
(a)



(b)



(c)



(d)

Fig. 10: Sample predictions on test images using proposed CNN with AVP model.

Figure 10 illustrates the practical application of the model in classifying retinal images, offering a clear visual insight into its real-world performance. Subfigure 10(a) displays an instance where the model accurately predicts a test image as Normal, showcasing its ability to correctly identify healthy retinal conditions. In subfigure 10(b), the model successfully classifies a test image as Macular Hole (MH), demonstrating its effectiveness in distinguishing MH from other retinal abnormalities. Subfigure 10(c) highlights the model's correct prediction of a test image as Retinal Detachment (RD), indicating its proficiency in detecting significant pathological features. Lastly, subfigure 10(d) presents the model's prediction of a test image as belonging to the Other Diseases/Conditions (ODC) category, reflecting its capacity to handle a broader range of retinal issues beyond the primary classes.

Metric	Existing SGD	Existing Adam	Proposed CNN-AVP
Validation Accuracy (%)	79.60	84.46	94.58
Testing Accuracy (%)	79.85	83.39	94.43
Testing Precision (%)	80.15	83.11	94.10
Testing Recall (%)	78.41	82.77	93.98
Testing F-Score (%)	78.54	82.47	93.95

Table 2: Performance Comparison of Various Models.

Table 2 presents a comparative analysis of model performance based on key evaluation metrics, including validation accuracy, testing accuracy, testing precision, recall, and F-score. The existing model trained with the SGD optimizer demonstrates comparatively lower performance, achieving a validation accuracy of 79.60%, testing accuracy of 79.85%, precision of 80.15%, recall of 78.41%, and an F-score of 78.54%. While these results indicate a moderately effective model, they also reveal opportunities for optimization. In contrast, the existing model utilizing the Adam optimizer shows notable improvements across all metrics, with a validation accuracy of 84.46%, testing accuracy of 83.39%, precision of 83.11%, recall of 82.77%, and an F-score of 82.47%, underscoring Adam's effectiveness in enhancing CNN performance. The proposed CNN model with AVP (Adam with Valid Padding) optimization delivers the highest performance, recording a validation accuracy of 94.58%, testing accuracy of 94.43%, precision of 94.10%, recall of 93.98%, and an F-score of 93.95%. These results demonstrate that the AVP-optimized model significantly enhances accuracy, precision, and overall diagnostic reliability, making it the most effective among the evaluated approaches.

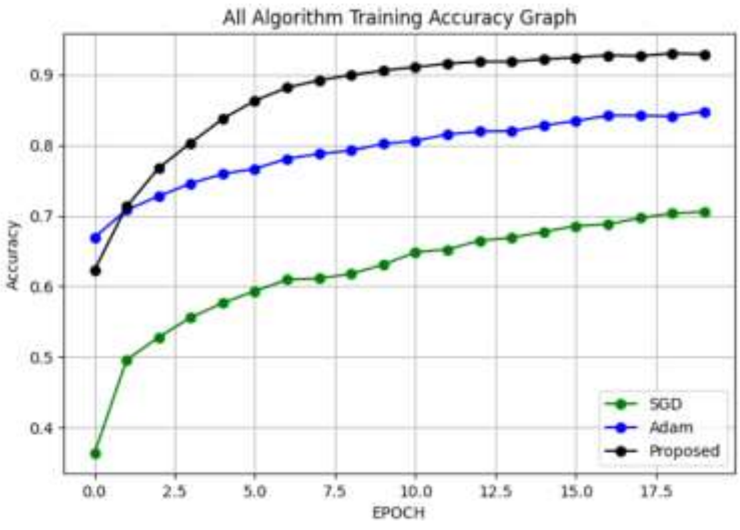


Fig. 11: Performance evaluation of existing and proposed classifiers.

Figure 11 shows how accuracy changes over 20 epochs of training for the various models. It features line plots depicting accuracy trends over time. The proposed CNN-AVP model display a more rapid and sustained increase in accuracy compared to the SGD and Adam models, reflecting its superior performance and stability.

5. CONCLUSION

The development of the AI-driven automated diagnostic system for retinal image analysis highlights a strong integration of deep learning methodologies with an intuitive graphical user interface, aimed at supporting early detection of eye diseases. In this project, two models were designed and evaluated: an existing Deep Neural Network (DNN) utilizing the Stochastic Gradient Descent (SGD) optimizer, and a proposed Convolutional Neural Network (CNN) using the Adam optimizer with Valid Padding (AVP). Both models were trained on a well-preprocessed and augmented dataset, following a systematic train-test split to ensure high-quality training data. The existing DNN model delivered a respectable level of accuracy, with its precision, recall, and F1-score reflecting consistent performance in classifying retinal images into categories such as Diabetic Retinopathy (DR), Macular Hole (MH), Normal, and Other Diseases/Conditions (ODC). However, the corresponding confusion matrix indicated minor misclassifications between visually similar classes, pointing to potential areas for enhancement. In contrast, the proposed CNN model, which incorporated a more refined architecture and an optimized loss function, achieved superior performance across key evaluation metrics. Notably, it demonstrated higher overall accuracy and improved recall for challenging categories, as shown by a more evenly distributed and accurate confusion matrix. The accompanying classification reports further supported the model's enhanced per-class accuracy, confirming that the proposed CNN with AVP offers a significant performance advantage over the baseline DNN model.

REFERENCES

[1] Jolly, Lochan, Niket Amoda, and K. Mishra. "Artificial Intelligence-Enabled IOMT for Medical Application." In *Handbook of Research on Artificial Intelligence and Soft Computing Methods in Personalized Healthcare Services*, pp. 3-32. Apple Academic Press, 2024.

[2] Nadhan, Archana S., and I. Jeena Jacob. "Enhancing healthcare security in the digital era: Safeguarding medical images with lightweight cryptographic methods in IoT healthcare applications." Biomedical Signal Processing and Control 88 (2024): 105511.

[3] Pradyumna, G. R., Roopa B. Hegde, K. B. Bommegowda, Tony Jan, and Ganesh R. Naik. "Empowering healthcare with IoMT: Evolution, machine learning integration, security, and interoperability challenges." *IEEE Access* (2024).

[4] Rani, Sita, Sachin Kumar, Aman Kataria, and Hong Min. "Smart Health: An intelligent framework to secure IoMT service applications utilizing machine learning." *ICT Express* 10, no. 2 (2024): 425-430.

[5] Ioannou, Iacovos, Prabagarane Nagaradjane, Pelin Angin, Palaniappan Balasubramanian, Karthick Jeyagopal Kavitha, Palani Murugan, and Vasos Vassiliou. "GEMLIDS-MIOT: A Green Effective Machine Learning Intrusion Detection System depends on Federated Learning for Medical IoT network security hardening." *Computer Communications* 218 (2024): 209-239.

[6] Tenepalli, Deepika, and Navamani TM. "A systematic review on IoT and machine learning algorithms in e-healthcare." *International Journal of Computing and Digital Systems* 16, no. 1 (2024): 279-294.

[7] Jasim, Abdulrahman Ahmed, Layth Rafea Hazim, Hayder Mohammedqasim, Roa'A. Mohammedqasem, Oguz Ata, and Omar Hussein Salman. "e-Diagnostic system for diabetes disease prediction on an IoMT environment-based hyper AdaBoost machine learning model." *The Journal of Supercomputing* (2024): 1-26.

[8] Ameen, Aya Hamid, Mazin Abed Mohammed, and Ahmed Noori Rashid. "Enhancing Security in IoMT: A Blockchain-Based Cybersecurity Framework for Machine Learning-Driven ECG Signal Classification." *Fusion: Practice and Applications* 14, no. 1 (2024): 221-251.

[9] Jiang, Zhiyuan, Mohammed AA Al-qaness, AL-Alimi Dalal, Ahmed A. Ewess, Mohamed Abd Elaziz, Abdelghani Dahou, and Ahmed M. Helmi. "Fall Detection Systems for Internet of Medical Things Depends on Wearable Sensors: A Review." *IEEE Internet of Things Journal* (2024).

[10] Subramanian, Murugan, S. Kanaga Suba Raja, A. Sampath Kumar, Manikandan Ramachandran, and Ambeshwar Kumar. "Wearable Sensor-Based Monitoring and Classification Utilizing Deep Learning For Personalized Healthcare." In *Handbook of Research on Artificial Intelligence and Soft Computing Methods in Personalized Healthcare Services*, pp. 209-227. Apple Academic Press, 2024.

[11] Pachade, Samiksha, Prasanna Porwal, Dhanshree Thulkar, Manesh Kokare, Girish Deshmukh, Vivek Sahasrabuddhe, Luca Giancardo, Gwenolé Quellec, and Fabrice Mériaudeau. "Retinal fundus multi-disease image dataset (RFMiD): a dataset for multi-disease detection research." Data 6, no. 2 (2021): 14.

[12] Siswadi, A. A. P., Bricq, S., & Meriaudeau, F. (2024). Multi-modality multi-label ocular abnormalities detection with transformer-based semantic dictionary learning. *Medical & Biological Engineering & Computing*, 1-12.

[13] Inan, Yavuz Selim. "Adaptive Multiscale Retinal Diagnosis: A Hybrid Trio-Model Approach for Comprehensive Fundus Multi-Disease Detection Leveraging Transfer Learning and Siamese Networks." *arXiv preprint arXiv:2405.18449* (2024).

[14] Elsayed, Tasnim Samir, and Muhammad Ali Rushdi. "Computer-aided multi-label retinopathy diagnosis via inter-disease graph regularization." *Biomedical Signal Processing and Control* 96 (2024): 106516.

[15] Vemparala, Yoshita, Niveditha Manne, Sruthi Katapally, and Ramakrishna Kolikipogu. "OcuVision: CNN powered Analysis of Retinal images for Disease Diagnosis." In *2024 Second International Conference on Emerging Trends in Information Technology and Engineering (ICETITE)*, pp. 1-8. IEEE, 2024.