AI Tool for Automated Diagnosis of Eye Diseases from Retinal Images

Gudimilla Pallavi[1*], Bongoni Siddartha[2], Vemuru Bhuvaneshwary [2], Kondeti Shiva[2], Revuri Karthik Reddy[2]

[1]Associate Professor, [2]UG Student, [1,2]Department of Computer Science and Engineering (Data Science), Vaagdevi College of Engineering, Bollikunta, Warangal, Telangana.

*Corresponding Email: pallavi.gudimilla@gmail.com

ABSTRACT

This project presents an AI-based automated diagnosis system designed to detect eye diseases from retinal images with high accuracy and efficiency. It combines advanced deep learning techniques with a user-friendly graphical interface, providing clinicians and researchers with a powerful tool for early disease detection and analysis. The system begins with comprehensive dataset management, allowing users to upload retinal images, map them to specific target classes such as Diabetic Retinopathy (DR), Macular Hole (MH), Normal, and Other Diseases/Conditions (ODC), and perform essential preprocessing operations. Preprocessing includes image resizing, normalization, and data augmentation using state-of-the-art methods to enhance dataset variability and improve model robustness. Two model architectures are at the core of the system. The first is an existing deep neural network (DNN) with a Stochastic Gradient Descent (SGD) optimizer. This model features multiple convolutional layers, batch normalization, pooling, and dense layers to effectively learn intricate features from retinal images. The second is a proposed Convolutional Neural Network (CNN) that employs the Adam optimizer and valid padding (AVP). This model includes additional dropout and batch normalization layers to mitigate overfitting and enhance generalization. Both models are evaluated through performance metrics like accuracy, precision, recall, and F1 score, with confusion matrices providing diagnostic insights. Experimental results show that both models achieve reliable performance, but the proposed CNN with AVP demonstrates superior diagnostic accuracy and better differentiation of subtle pathological features. The system's modular design, coupled with extensive data augmentation, ensures scalability and adaptability for real-world clinical applications, contributing to the advancement of computer-aided diagnostic tools in medical imaging.

Keywords: Retinal imaging, Diabetic retinopathy, Predictive analytics, Supervised learning, Deep neural networks, Convolutional neural networks.

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.

doi: 10.48047/ijiee.2025.15.4.20

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.

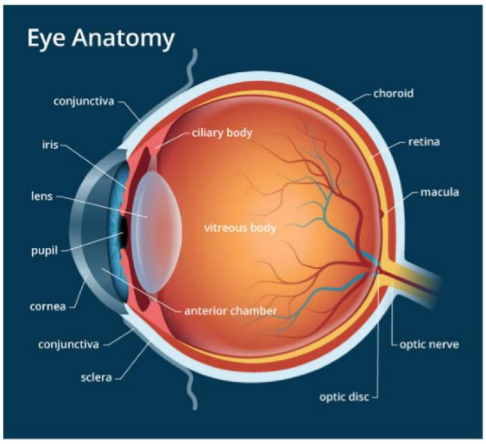


Fig. 1: Eye anatomy.

Implementing AI-based RD classification requires integrating ML models into existing healthcare infrastructure and workflows. This involves training the AI approaches utilizing advanced datasets of labeled retinal images to teach them to recognize patterns associated with different stages of RD. Once trained, these AI approaches accurately classify retinal images, providing timely and consistent assessments without the requirement for extensive manual intervention.

2. LITERATURE SURVEY

In the realm of medical diagnostics, the application of ML and ML has significantly transformed. Presently, as patients search for medical care for their eye health, the traditional diagnostic process involving hospital visits, doctor consultations, and time-consuming screenings tends to impede the swift initiation of treatment. However, researchers have harnessed the power of AI to streamline this process by developing robust ML and DL models for the automated classification of RD. These models analyze retinal images, capturing subtle details and patterns that escape the naked eye, enabling swift and accurate identification of RD severity levels.The implementation of AI-based RD classification not only expedites the diagnostic phase. This was improving patient outcomes and optimizes the allocation of healthcare resources. In essence, the ML integration and DL in RD classification represents a tangible advancement in the medical field, offering a more efficient and timely approach to addressing the complexities of this sight-threatening condition.

2.1 Survey on RD Classification Methods

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

doi: 10.48047/ijiee.2025.15.4.20

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. Bali, Akanksha, et al. [16] presented a multi-class, multi-label classification framework for ophthalmological fundus images depends on an optimized deep feature space evolutionary model. The proposed methodology combines deep learning with evolutionary algorithms to optimize the feature extraction process, enhancing the model's capability to differentiate among various retinal diseases. The evolutionary model iteratively refines the deep feature space, selecting the most relevant features that contribute to accurate classification. This approach not only improves the diagnostic performance but also reduces computational complexity, making it suitable for deployment in clinical settings where resources may be limited.

Chavan, et al. [17] introduced diabetic disease detection methodology depends on optic disc and blood vessel analysis utilizing an enhanced Long-Short-Term Memory (LSTM) network. This approach

doi: 10.48047/ijiee.2025.15.4.20

focuses on extracting features from the optic disc and retinal blood vessels, which are critical indicators of diabetic retinopathy progression. The enhanced LSTM network is tailored to handle sequential data, capturing temporal dependencies that are critical for accurate disease detection. By integrating these anatomical features into the diagnostic process, the proposed method significantly improves the sensitivity and specificity of diabetic disease detection in retinal fundus images. ATLAN, et al. [18] explored the impact of noise removal filters on the classification accuracy of different types of medical images, including retinal images. The authors evaluated various filtering methods to improvise image quality by reducing noise, which is a common issue in medical imaging that can obscure important diagnostic features. By systematically analyzing the effect of different noise removal filters, machine learning models in medical image classification. The findings suggest that optimal noise reduction is critical for achieving high diagnostic accuracy in retinal disease detection. Sivaz, et al. [19] combined EfficientNet with a Machine Learning Decoder (ML-Decoder) classification head for multi-label retinal disease classification. The EfficientNet architecture is utilized for its high parameter efficiency and strong feature extraction capabilities, while the ML-Decoder is designed to handle the complexities of multi-label classification tasks. This combination allows for the effective identification of multiple retinal diseases from a single image, leveraging the strengths of both deep learning and traditional machine learning methods. The integrated model demonstrates improved accuracy and robustness in multi-label classification compared to standard CNN-based approaches.

Du, Jiawei, et al. [20] introduced RET-CLIP, a foundation model for retinal image analysis pre-trained with clinical diagnostic reports. RET-CLIP leverages a large-scale dataset of retinal images paired with corresponding clinical diagnoses to adopt rich feature descriptions that are clinically relevant. The pre-training process involves contrastive learning, which helps the model analyse the subtle variations in retinal images that correspond to different disease states. RET-CLIP is designed to be fine-tuned on specific diagnostic tasks, providing a versatile foundation for developing specialized models for various retinal disease detection applications. Subramaniam, et al. [21] proposed an enhanced retinal fundus image classification methodology utilizing an Active Gradient Deep CNN combined with Red Spider Optimization (RSO). This approach focuses on optimizing the deep learning model's gradient descent process to improve convergence and accuracy in disease detection tasks. The RSO is employed to fine-tune the hyperparameters of the CNN, enhancing its performance on multi-label classification tasks. This combination results in a robust and efficient model capable of accurately classifying diverse retinal diseases in fundus images. Choi, Joon Yul, et al. [22] developed generative adversarial networks (GANs) to synthesize high-quality retinal images, which were utilized to train a DL classifier for epiretinal membranes classification. The model improves its capability to generalize to new and unseen cases. The approach demonstrates significant improvements in detection accuracy, particularly in challenging cases where the epiretinal membrane is not easily distinguishable from the surrounding tissue.

Chavan, et al. [23] presented a comparative study on retinal fundus images class identification utilizing two deep learning architectures, VGG16 and Inception V3. They were used in diagnosing various retinal diseases such as diabetic retinopathy, glaucoma, and AMD. The study highlights the assets and debilities of each architecture in accuracy, computational efficiency, and robustness to image quality variations. The findings suggest that while Inception V3 offers superior performance in accuracy, VGG16 provides a more computational environment capable solution for real-time clinical applications.

Müller, Dominik, et al. [24] The research focuses on multi-disease detection in retinal imaging utilizing an ensemble of heterogeneous deep learning models. The proposed methodology combines different types of neural networks, such as CNNs and recurrent neural networks (RNNs), to leverage their complementary strengths in feature extraction and temporal data analysis. This ensemble approach

doi: 10.48047/ijiee.2025.15.4.20

enhances the model's capability to detect multiple diseases in a single retinal image, providing a more comprehensive diagnostic tool. The study demonstrates that the ensemble model outperforms individual models in reliability and robustness, particularly in complex cases with multiple co-existing retinal conditions. Xu, Ziwen, et al. [25] proposed a deep network guided by dark and bright channel priors for retinal image quality assessment. The network is designed to estimate the performance of retinal images by identifying artifacts and poor lighting conditions that could hinder accurate disease detection. The dark and bright channel priors help the model distinguish among well-lit and shadowed areas, providing a more accurate assessment of image quality. This evaluation tool was utilized as a pre-processing step to filter out low-quality images, ensuring that only high-quality data is utilized for disease detection tasks.

Sengar, Neha, et al. [26] the study introduced EyeDeep-Net, a DNN framework for multi-class diagnosis of retinal diseases. This architecture was specifically designed to handle the complexities of retinal disease classification, including the requirement for accurate feature extraction and disease differentiation. EyeDeep-Net integrates multiple layers of convolutional, pooling, and fully connected nodes to build a robust feature hierarchy that can recognise among diverse retinal conditions. The framework also includes procedures for data augmentation and regularization, which help enhance model's generalization capabilities and reduce overfitting. Verma, , et al. [27] presented a method retinal diseases class identification utilizing a deep learning model that combines both image-based and clinical data. The proposed model integrates image data from retinal fundus photography with clinical variables, such as patient age and medical history, to enhance reliability of disease detection. This multimodal approach leverages the depths of information types, providing a more holistic view of the patient's condition. The study demonstrates that incorporating clinical data into the diagnostic process significantly enhances the model's performance, particularly in cases where image data alone is insufficient for accurate classification. Jothi, , et al. [28] focused on a multi-disease detection framework utilizing a hybrid CNN with feature optimization for fundus images. The proposed framework integrates CNN-based feature extraction and selection methods to optimize the data input feature set, reducing redundancy and improving model efficiency. This hybrid approach enhances the model's capability to detect multiple diseases in retinal images, providing a more comprehensive diagnostic tool. The study highlights the importance of feature optimization in improving both the accuracy and computational efficiency of deep learning models in retinal disease detection.

2.2 Research Gaps

A survey of existing methods for RD grading reveals several critical research gaps that underscore the requirement for advancement in this field. One significant gap is the lack of automation in traditional grading methods, which are heavily reliant on manual examination by ophthalmologists. This manual approach introduces inefficiencies and inconsistencies, as grading accuracy depends on the individual expertise of the grader. The variability in results among different graders highlights a critical requirement for AI methods that can provide consistent and objective assessments. Current research is also limited in addressing the scalability of RD screening; manual approaches were not flexible to large-scale or remote screenings, leaving gaps in accessibility and timely care for underserved populations. Additionally, there was a requirement for improved integration of advanced technologies into RD grading systems. Existing methods often do not fully utilize the potential of emerging technologies such as IoMT devices and AI algorithms. Research was not necessarily explored how these technologies was integrated to improvise the efficiency and accuracy of RD grading. For example, while IoMT devices can collect real-time data, analysing this data effectively and provide actionable insights. The integration of AI for automated grading and predictive modeling is still underexplored, with current

doi: 10.48047/ijiee.2025.15.4.20

research lacking comprehensive solutions that combine these technologies to address the limitations of manual methods and improve patient outcomes.

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 2. 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.

Step 1: RFMID Dataset: The first process involves acquiring and preparing the RFMID dataset, which was a collection of RD images specifically annotated for grading. This dataset is essential for training and evaluating CNNs in RD detection. The dataset should be comprehensive, containing a diverse range of images that represent various stages and types of RDs. Proper data preprocessing is critical at this stage to ensure that the images are of high quality and appropriately labelled. This includes steps such as ensuring consistent image resolution, color balance, and format.

Step 2: Image Processing: In this step, it contains diverse analysis like normalization, contrast adjustment, and noise reduction to ensure that the images are clear and the features relevant to RD are prominently visible. Image processing methods help in reducing artifacts and standardizing the input data, which improves the performance of the CNN models. Additionally, methods such as segmentation was applied to isolate regions of interest, such as blood vessels or lesions, which were critical for accurate RD assessment.

Step 3: Image Augmentation: This step helps in increasing the diversity of the dataset and prevents overfitting by introducing variations such as rotation, scaling, flipping, and brightness adjustments. Augmentation methods improvise the CNN model, accepting it to generalize better to new data. By simulating different conditions and perspectives, augmentation ensures that the model can handle diverse data variations and improve its predictive accuracy.

Step 4: Train-Test Splitting: Once the dataset has been processed and augmented, it is split into training and test sets. A usual process was to allocate a majority of information to training information data and a smaller portion to the test set, ensuring that the model has enough data to adopt from while reserving a representative subset for unbiased evaluation.

doi: 10.48047/ijiee.2025.15.4.20

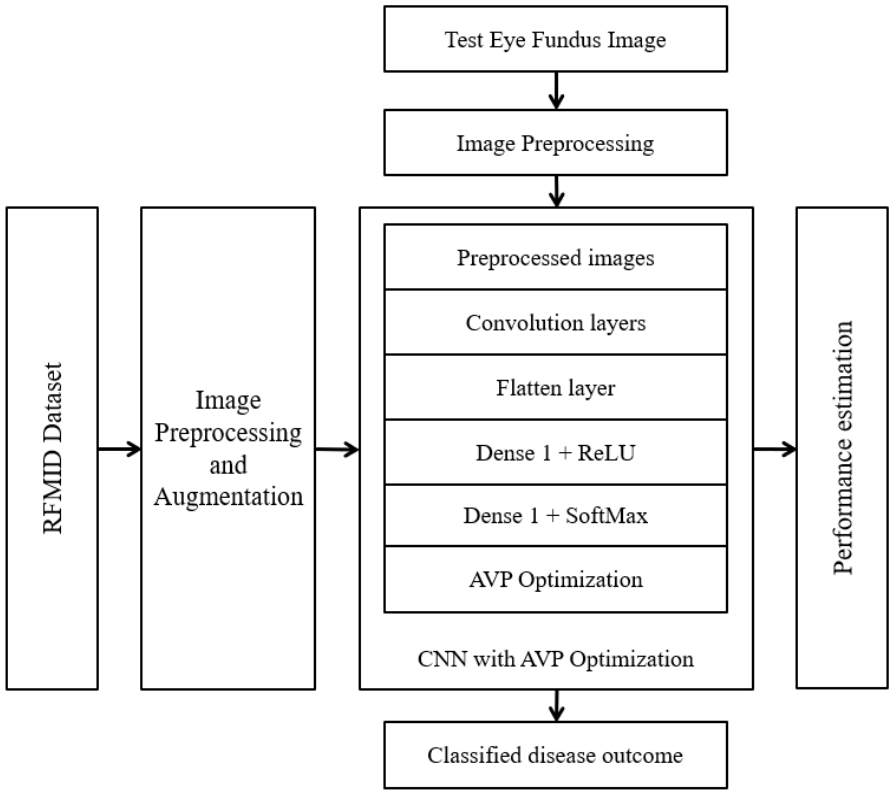


Fig.2: Proposed RD grading system architecture using CNN with AVP model.

Step 5: Existing CNN with SGD Optimizer: At this stage, a baseline CNN is implemented utilizing the Stochastic Gradient Descent (SGD) optimizer. The existing CNN was trained on processed and augmented dataset to establish a performance benchmark. Although effective, SGD have limitations in converging to optimal solutions, particularly in complex models with large datasets.

Step 6: Proposed CNN with Adam-Valid Padding Optimization: The proposed CNN model is introduced, incorporating Adam optimizer and valid padding (AVP) methods. The AVP contains adaptive with fast learning rates and is often more effective than SGD in converging to optimal solutions. Valid padding, as opposed to same padding, involves cropping the edges of input data to maintain spatial dimensions, which enhance model's focus on the central regions of interest. Finally, training the proposed CNN with these enhancements to evaluate their impact on model performance.

Step 7: Prediction from Test Image: With the proposed CNN model trained, predictions are made on the test images to assess the model's capability to catalogue and grade RD accurately. The test data were applied to trained model, and the outputs are compared to the ground truth labels to evaluate the model's performance. So, proposed model generalizes to new data and how accurately it can identify, and grade RD compared to the baseline model.

Step 8: Performance Comparison: The final step involves contrasting performance of existing CNN with SGD optimizer and the proposed CNN with Adam-Valid Padding Optimization. This comparison highlights the improvements brought by the proposed optimizations and supplies a quantitative assessment of their effectiveness in enhancing RD grading. The results were explored to determine which model offers superior performance.

doi: 10.48047/ijiee.2025.15.4.20

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 3. 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.

4. RESULTS AND ANALYSIS

This research creates a graphical user interface (GUI) for building, training, and evaluating deep learning models for the automated diagnosis of retinal (eye) diseases.

4.1 Dataset Description

The Retinal Fundus Multi-disease Image Dataset (RFMiD) is a comprehensive collection designed to support the development of automated methods for the classification of both common and rare ocular diseases. The dataset includes 1,300 high-quality RD images, which was meticulously captured utilizing three different types of fundus cameras, ensuring diversity in image acquisition conditions and enhancing the generalizability of models trained on this dataset. RFMiD stands out due to its extensive coverage of 46 distinct ophthalmic conditions, including prevalent diseases such as RD, glaucoma, and age-related macular degeneration, as well as rare pathologies that are often underrepresented in other datasets. The images are expertly annotated through a rigorous adjudicated agreement procedure, ensuring the accuracy and reliability of the annotations. This dataset is particularly valuable because it is the first of its kind to offer such a broad spectrum of retinal diseases in a single, publicly accessible repository, providing a unique opportunity for researchers to develop and test robust and generalizable models for retinal screening in routine clinical settings. By facilitating the development of ML algorithms that can handle a wide array of conditions, RFMiD aims to improve the early detection and management of ocular diseases, ultimately contributing to better patient outcomes in ophthalmology.

The dataset includes four distinct classes representing various retinal conditions and a healthy state: Diabetic Retinopathy (DR), Macular Hole (MH), Normal, and Optic Disc Cupping (ODC). Each class signifies a unique characteristic or pathology that affects the retina or the optic nerve head, impacting vision to varying degrees.

4.2 Results Analysis

Confusion matrix: In the confusion matrix for RD classification, each row represents the actual classes of retinal images, while each column represents the predicted classes by the AI-based classification

doi: 10.48047/ijiee.2025.15.4.20

system. For instance, the first row corresponds to images without RD, while the second row corresponds to images with RD. The first column represents images predicted as non-RD, while the second column represents images predicted as having RD. The values within the matrix indicate the number of images falling into each category: true negatives (TN) are correctly predicted non-RD images, false positives (FP) are non-RD images incorrectly classified as having RD, false negatives (FN) are RD images incorrectly classified as non-RD, and true positives (TP) are correctly predicted RD images.

Accuracy: The accuracy of the classifier's predictions is a key parameter that quantifies the overall correctness of the predictions made by the classifier. Specifically, it is determined by dividing the total number of occurrences. The accuracy of a classifier gives an overall impression of how well it performs across all classes of RD, regardless of the distribution of the classes on the RD.

$$\text{Accuracy} = ((TP + TN) / \text{Total}).$$

Precision: It is another essential statistic that focuses on the classifier's capability to properly identify instances of a certain class among all examples that are categorized as belonging to that class.

$$\text{Precision} = (TP / (TP + FP))$$

Recall: It is sometimes referred to as sensitivity, and it is a measurement of the classifier's capability to accurately identify instances of a certain class among all cases that belong to that class.

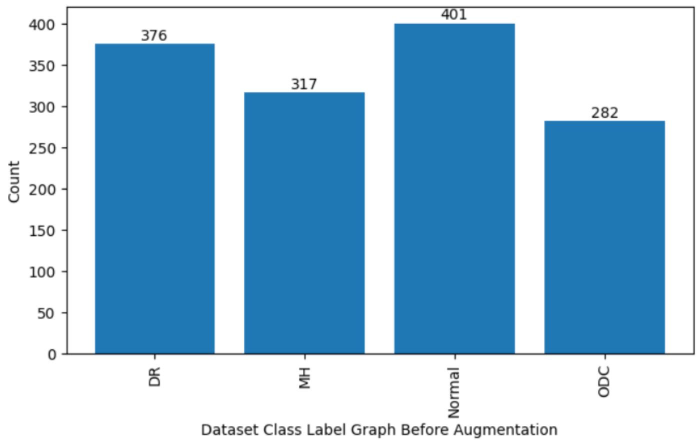


Fig. 10: Dataset labels versus number of sample count before augmentation.

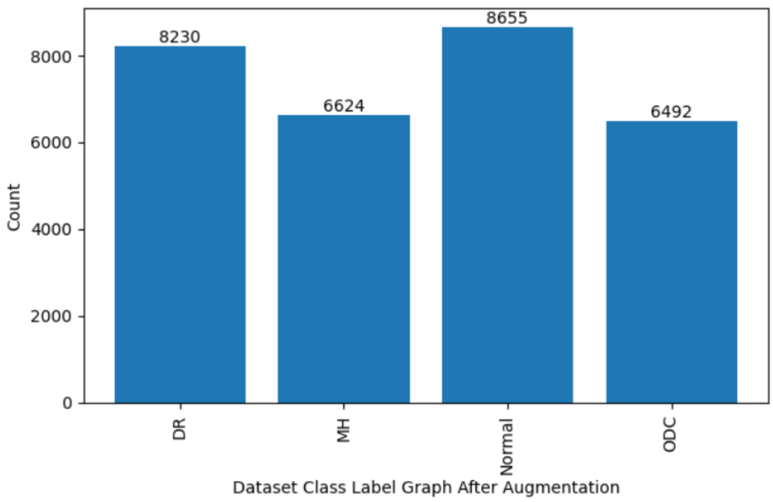
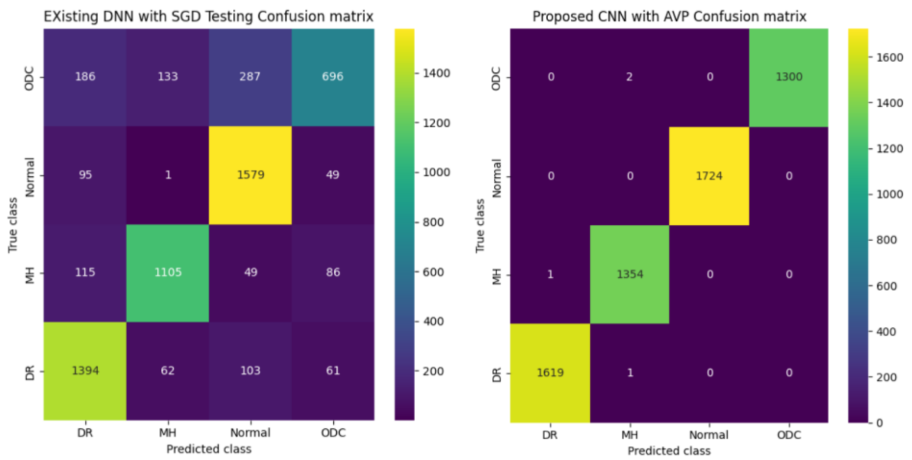


Fig. 11: Bar graph obtained after augmentation (dataset class labels versus count).

doi: 10.48047/ijiee.2025.15.4.20

Fig. 10 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). In contrast, Fig. 11 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. Fig. 12(a) complements this by displaying the confusion matrix for the existing model, visually summarizing how well each class was predicted and highlighting misclassifications. Fig. 12(b) 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. Fig. 13 visually represents the practical application of the model in predicting specific classes, providing a direct view of how well the model performs in real-world scenarios. Fig. 13 (a) showcases an example where the model predicted a test image as Normal. It demonstrates the model's capability to correctly identify Normal from test data. In Fig. 13 (b), the model's prediction of a test image as Macular Hole (MH) is displayed. This indicates the model's performance in accurately distinguishing MH from other conditions. Fig. 13 (c) represents the model's prediction of a test image as RD. It highlights the model's capability to correctly classify images with significant retinal abnormalities as RD. Fig. 13 (d) represents the model's prediction on a test image as ODC.



(a) (b)

Fig. 12: Confusion matrix obtained using (a) DNN with SGD approach. (b) proposed CNN with AVP approach.

doi: 10.48047/ijiee.2025.15.4.20

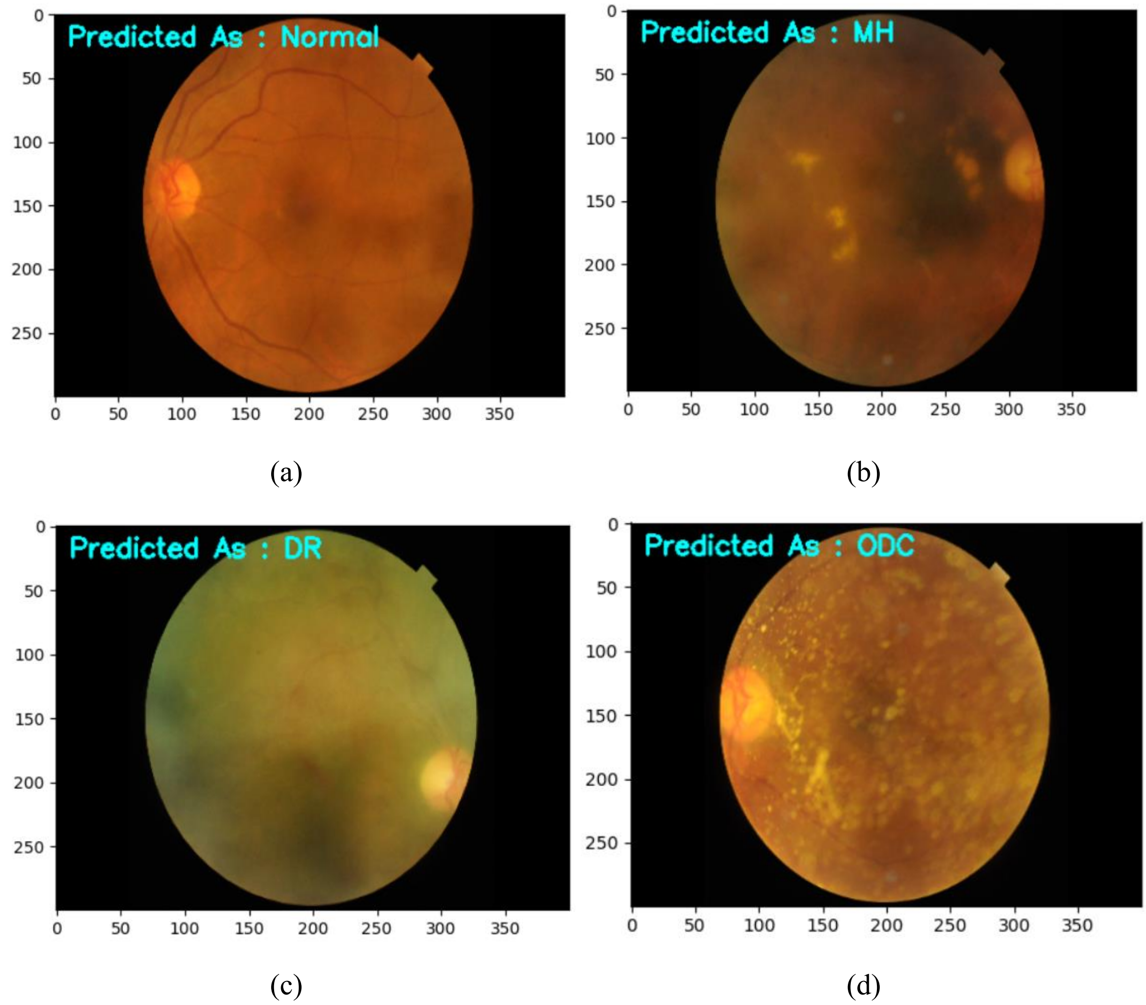


Fig. 13: Sample predictions on test images using proposed CNN with AVP model.

Table.1: Performance Comparison of Various Models.

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.1 compares the performance of different models depends on key metrics: Validation Accuracy, Testing Accuracy, Testing Precision, Testing Recall, and Testing F-Score.

Existing SGD: This model shows lower performance metrics across all aspects, with validation accuracy at 79.60%, testing accuracy at 79.85%, precision at 80.15%, recall at 78.41%, and F-Score at 78.54%. This indicates that while the model performs reasonably well, there is room for improvement.

doi: 10.48047/ijiee.2025.15.4.20

Existing Adam: The Adam optimizer significantly improves performance compared to SGD, with validation accuracy at 84.46%, testing accuracy at 83.39%, precision at 83.11%, recall at 82.77%, and F-Score at 82.47%. These improvements highlight Adam's effectiveness in optimizing the CNN model.

Proposed CNN-AVP: The proposed model with AVP optimization shows the best performance, with validation accuracy at 94.58%, testing accuracy at 94.43%, precision at 94.10%, recall at 93.98%, and F-Score at 93.95%. The significant improvements in these metrics suggest that AVP optimization greatly enhances model accuracy and overall performance.

5. CONCLUSION

The development of the AI-based automated diagnosis tool for retinal images demonstrates a robust integration of deep learning techniques and user-friendly GUI design to facilitate the early detection of eye diseases. Throughout the project, two different models were implemented and evaluated: an existing deep neural network (DNN) with a Stochastic Gradient Descent (SGD) optimizer and a proposed Convolutional Neural Network (CNN) employing the Adam optimizer with AVP (Adam with Valid Padding). Both models underwent a series of rigorous preprocessing, data augmentation, and train-test splitting steps to ensure that the dataset was optimally prepared for training. The existing DNN model achieved a commendable overall accuracy, with performance metrics including precision, recall, and F1 score that indicated a reliable ability to classify images into categories such as Diabetic Retinopathy (DR), Macular Hole (MH), Normal, and Other Diseases/Conditions (ODC). The confusion matrix for this model revealed that while most classes were accurately predicted, there were minor misclassifications between similar categories, suggesting areas for potential refinement. On the other hand, the proposed CNN model, which featured advanced architectural configurations and incorporated a loss optimization mechanism, demonstrated improved performance in several key metrics. In particular, the proposed model showed higher accuracy and better recall rates for classes that were challenging to distinguish, as evidenced by its more balanced confusion matrix. The detailed classification reports provided clear insights into per-class performance, validating the effectiveness of the proposed improvements over the baseline 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.

doi: 10.48047/ijiee.2025.15.4.20

[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.

[16] Bali, Akanksha, and Vibhakar Mansotra. "Multiclass multilabel ophthalmological fundus image classification depends on optimised deep feature space evolutionary model." *Multimedia Tools and Applications* 83, no. 16 (2024): 49813-49843.

[17] Chavan, Rupali, and Dnyandeo Pete. "Optic Disc and Blood Vessel based Diabetic Disease Detection Utilizing Enhanced Long-Short-Term Memory." In *2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)*, vol. 2, pp. 1-6. IEEE, 2024.

[18] ATLAN, Furkan, and Ihsan PENCE. "The Effect of Noise Removal Filters on Classifying Different Types of Medical Images." Digital Signal Processing (2024): 104613.

[19] Sivaz, Orhan, and Murat Aykut. "Combining EfficientNet with ML-Decoder classification head for multi-label retinal disease classification." *Neural Computing and Applications* (2024): 1-11.

doi: 10.48047/ijiee.2025.15.4.20

[20] Du, Jiawei, Jia Guo, Weihang Zhang, Shengzhu Yang, Hanruo Liu, Huiqi Li, and Ningli Wang. "RET-CLIP: A Retinal Image Foundation Model Pre-trained with Clinical Diagnostic Reports." *arXiv preprint arXiv:2405.14137* (2024).

[21] Subramaniam, Krishnakumar, and Archana Naganathan. "Enhancing retinal fundus image classification through Active Gradient Deep Convolutional Neural Network and Red Spider Optimization." *Neural Computing and Applications* (2024): 1-13.

[22] Choi, Joon Yul, Ik Hee Ryu, Jin Kuk Kim, In Sik Lee, and Tae Keun Yoo. "Development of a generative deep learning model to improve epiretinal membrane detection in fundus photography." *BMC Medical Informatics and Decision Making* 24, no. 1 (2024): 25.

[23] Chavan, Rupali, and Dnyandeo Pete. "Classification of Retinal Fundus Images utilizing VGG16 and Inception V3." In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pp. 1-5. IEEE, 2023.

[24] Müller, Dominik, Iñaki Soto-Rey, and Frank Kramer. "Multi-disease detection in retinal imaging depends on ensembling heterogeneous deep learning models." In *German Medical Data Sciences 2021: Digital Medicine: Recognize–Analyse–Heal*, pp. 23-31. IOS Press, 2021.

[25] Xu, Ziwen, Beiji Zou, and Qing Liu. "A dark and bright channel prior guided deep network for retinal image quality assessment." *Biocybernetics and Biomedical Engineering* 42, no. 3 (2022): 772-783.

[26] Sengar, Neha, Rakesh Chandra Joshi, Malay Kishore Dutta, and Radim Burget. "EyeDeep-Net: a multi-class diagnosis of retinal diseases utilizing deep neural network." *Neural Computing and Applications* 35, no. 14 (2023): 10551-10571.

[27] Rodríguez, Manuel Alejandro, Hasan AlMarzouqi, and Panos Liatsis. "Multi-label retinal disease classification utilizing transformers." *IEEE Journal of Biomedical and Health Informatics* 27, no. 6 (2022): 2739-2750.

[28] Jiang, Yankai, Ke Xu, Xinyue Wang, Yuan Li, Hongguang Cui, Yubo Tao, and Hai Lin. "SatFormer: Saliency-Guided Abnormality-Aware Transformer for Retinal Disease Classification in Fundus Image." In *IJCAI*, pp. 987-994. 2022.

doi: 10.48047/ijiee.2025.15.4.20