

AUTOMATED RETINAL DISEASE DIAGNOSIS USING CONVOLUTIONAL NEURAL NETWORKS

S. Sundee¹, D. Balaji², V. Naga Sai², P. Raja Ravi Varma²

¹Assistant Professor, ²UG Student, ^{1,2}Department of Computer Science and Engineering (CSIT)

^{1,2}Sree Dattha Institute of Engineering and Science, Sheriguda, Ibrahimpatnam, 501510, Telangana.

Received: 09-07-2025

Accepted: 23-08-2025

Published: 30-08-2025

ABSTRACT

This project introduces an AI-driven diagnostic system designed to detect retinal diseases from fundus images with high accuracy and operational efficiency. Leveraging advanced deep learning techniques, the system integrates a user-friendly graphical interface to assist clinicians and researchers in early disease detection and classification. It begins with a robust dataset management module that allows users to upload retinal images, associate them with specific classes—including Diabetic Retinopathy (DR), Macular Hole (MH), Normal, and Other Diseases/Conditions (ODC)—and apply essential preprocessing techniques such as image resizing, normalization, and data augmentation. These steps enhance model robustness and dataset diversity. At the system's core are two distinct model architectures: the first is an existing Deep Neural Network (DNN) utilizing the Stochastic Gradient Descent (SGD) optimizer and composed of multiple convolutional, batch normalization, pooling, and dense layers; the second is a proposed Convolutional Neural Network (CNN) employing the Adam optimizer and incorporating valid padding (AVP), dropout, and additional batch normalization layers to reduce overfitting and improve generalization. Both models are trained and evaluated using standard metrics such as accuracy, precision, recall, and F1-score, with confusion matrices and classification reports highlighting diagnostic performance. Experimental results demonstrate that while both models perform effectively, the proposed CNN with AVP achieves superior accuracy and better differentiation across disease classes, particularly for subtle pathological features. Designed with modularity and scalability in mind, the system supports application across diverse clinical datasets and real-world healthcare environments. Overall, this research underscores the utility of deep learning in medical imaging and offers a practical, scalable solution for automated retinal disease diagnosis, contributing to the advancement of computer-aided diagnosis in ophthalmology.

Keywords: Retinal disease detection, deep learning, convolutional neural network (CNN), diabetic retinopathy, macular hole, fundus images, image preprocessing, automated diagnosis, Adam optimizer, SGD, valid padding, ophthalmology AI, computer-aided diagnosis, medical image classification, early disease detection.

1. INTRODUCTION

Retinal disease (RD) classification involves analyzing retinal images to detect signs of conditions such as diabetic retinopathy, a common complication of diabetes. Traditionally, this diagnostic process requires patients to visit healthcare facilities, where ophthalmologists manually examine fundus or OCT (Optical Coherence Tomography) images. While effective, this manual approach is time-consuming and often delays treatment due to the reliance on human interpretation. To address these limitations, artificial intelligence

(AI)-based classification systems have been developed. These systems leverage advanced machine learning (ML) and deep learning (DL) algorithms to automatically detect retinal abnormalities, including microaneurysms, hemorrhages, exudates, and neovascularization—key indicators used to determine disease severity and guide treatment. AI-driven classification not only accelerates diagnosis but also improves consistency and scalability in RD screening programs. Implementation involves training models on labeled datasets of retinal images,

enabling them to identify patterns associated with varying disease stages and perform accurate automated classification with minimal manual intervention. In parallel, the Internet of Medical Things (IoMT) has emerged as a transformative force in healthcare, integrating connected medical devices and cloud-based platforms to support real-time monitoring and diagnostics. IoMT-based systems for RD grading combine digital fundus cameras, OCT scanners, and AI algorithms to capture and analyze retinal images.

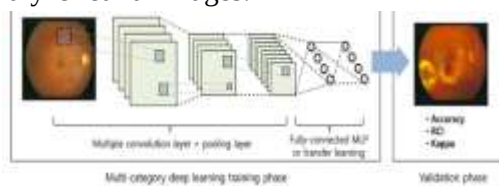


Fig. 1: Illustrating a CNN algorithm.

These systems facilitate continuous data collection and automated grading through centralized cloud infrastructure, allowing for early detection and proactive disease management—even from patients' homes. Deep learning models trained on large retinal datasets process incoming images to grade the severity of conditions such as diabetic retinopathy, glaucoma, and macular degeneration. Compared to manual assessment, IoMT-based grading provides faster, more reliable, and scalable diagnostics while reducing inter-observer variability. Integration with Electronic Health Records (EHRs) further enhances clinical decision-making by offering comprehensive insights into a patient's retinal and overall health. Collectively, AI and IoMT technologies represent a significant advancement toward efficient, accessible, and accurate eye care solutions.

2. LITERATURE SURVEY

Pachade, Samiksha, et al. [1] 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. [2] 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. [3] 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. [4] 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. [5] 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. [6] 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. [7] introduced diabetic disease detection methodology depends on optic disc and blood vessel analysis utilizing an enhanced Long-Short-Term Memory (LSTM) network. This approach 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. [8] 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 improve 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. [9] 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. [10] 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.

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.

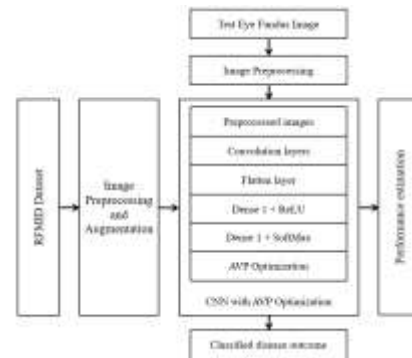


Fig. 2: Proposed RD grading system architecture using CNN with AVP model.

The implementation process for retinal disease (RD) detection using convolutional neural networks (CNNs) begins with acquiring and preparing the RFMID dataset, a diverse and well-annotated collection of RD images used for training and evaluation. Initial image preprocessing ensures high-quality input through normalization, contrast adjustment, and noise reduction, while segmentation isolates critical regions such as blood vessels and lesions. Image augmentation techniques like rotation, scaling, and brightness adjustment further increase dataset variability, helping the model generalize and avoid overfitting. The dataset is then split into training and testing sets to support unbiased evaluation. An existing baseline CNN is first

trained using the Stochastic Gradient Descent (SGD) optimizer to establish performance benchmarks. Following this, a proposed CNN model is developed using the Adam optimizer with valid padding (AVP), which provides faster convergence and focuses on central image regions by trimming borders, enhancing feature learning. Once trained, the proposed model is tested on unseen data to evaluate its predictive accuracy and grading capability. The final step compares the performance of the baseline CNN with the proposed CNN, demonstrating the advantages of Adam optimization and valid padding in improving classification accuracy and robustness in RD detection.

4. RESULTS AND DISCUSSION

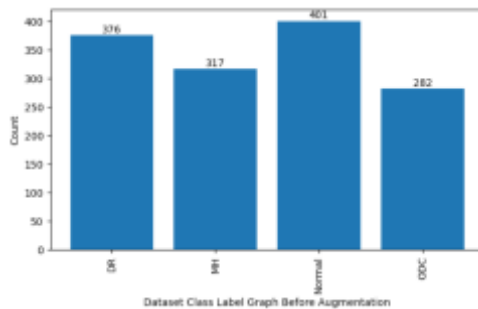


Fig. 3: Dataset labels versus number of sample count before augmentation.

Figure 3 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. 4 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.

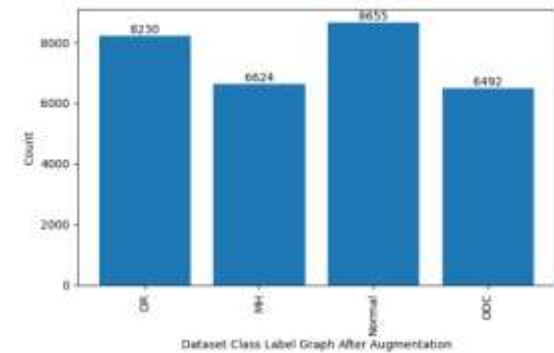
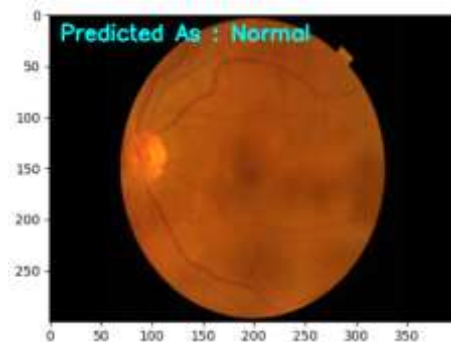
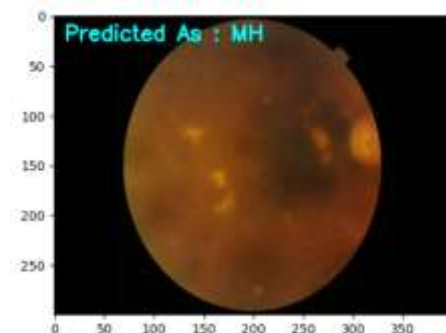


Fig. 4: Bar graph obtained after augmentation (dataset class labels versus count).

Finally, Fig. 5 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)



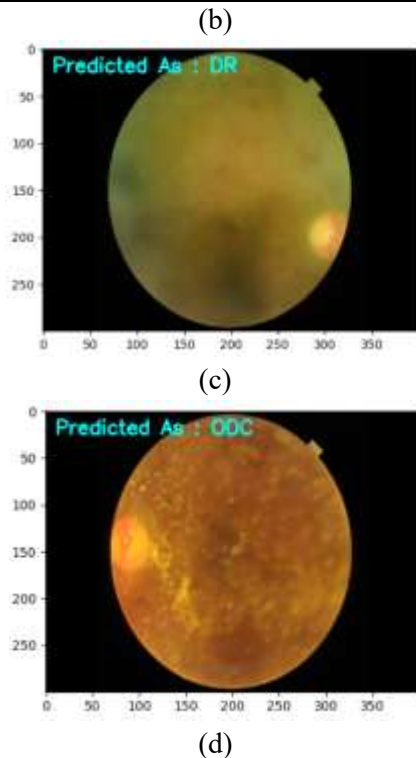


Fig. 5: Sample predictions on test case images using proposed CNN with AVP model.

5. CONCLUSION

REFERENCES

- [1] Pachade, Samiksha, Prasanna Porwal, Dhanshree Thulkar, Manesh Kokare, Girish Deshmukh, Vivek Sahasrabuddhe, Luca Giancardo, Gwenolé Quéllec, and Fabrice Mériaudeau. "Retinal fundus multi-disease image dataset (RFMiD): a dataset for multi-disease detection research." *Data* 6, no. 2 (2021): 14.
- [2] 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.
- [3] 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).
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] ATLAN, Furkan, and Ihsan PENCE. "The Effect of Noise Removal Filters on Classifying Different Types of Medical Images." *Digital Signal Processing* (2024): 104613.
- [9] 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.
- [10] 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).