IES International Journal of Multidisciplinary Engineering Research

Eye Disease Detection Enhancement Using a Multi-Stage Deep Learning Approach

Nowreen Mohamedmon[1], Nibin P N[2], Sahal M A[3], Sreehari E[4], Neethu Prabhakaran[5]

[1,2,3,4] Student, CSE, IES College of Engineering, Kerala,

[5] Assistant Professor, CSE, IES College of Engineering, Kerala,

Email_id: nourinmohamedmon17@gmail.com, nibinpn5@gmail.com, sahalma448@gmail.com, sreehariies67@gmail.com, neethuprabhakaranp@iesce.info

Abstract

Eye disease detection enhancement using a multi-stage deep learning approach introduces a lightweight and multi-stage deep learning framework for the automatic detection of eye diseases using retinal OCT and fundus images. The proposed model is designed to achieve a balance between accuracy and computational efficiency, making it ideal for real-world deployment, particularly in resource-limited clinical environments. The system begins with a robust preprocessing phase that normalizes and enhances the input images while addressing issues related to rotation, translation, and illumination variations. This step ensures that the model remains invariant to positional and intensity distortions, thereby improving its robustness and consistency across diverse imaging conditions. The architecture of the model is composed of three sequential stages that work together to deliver highly accurate diagnostic predictions. In the first stage, fine-grained spatial and textural features are extracted from retinal images, allowing the network to capture subtle structural variations that often indicate early disease progression. The second stage integrates a dual-branch structure that combines convolutional feature blocks with identity mapping branches. This design helps the network simultaneously learn both low-level detailed features and high-level semantic representations, enabling a deeper and more comprehensive understanding of retinal abnormalities. The extracted features from both branches are then fused to retain discriminative information and eliminate redundancy before being passed into the final stage. In the classification stage, the merged feature representations are processed through fully connected layers and a softmax-based output unit to generate precise disease predictions. This multi-stage feature integration improves the model's discriminative power and enhances its overall accuracy compared to conventional deep learning approaches. Experimental evaluations conducted on three benchmark datasets OCT 2017, Dataset-101, and Retinal OCT C8 demonstrated that the proposed model consistently outperforms existing architectures in terms of accuracy, precision, and recall. The results confirm that the combination of spatial, frequency, and contextual features contributes to better generalization across different types of retinal disorders. Moreover, the lightweight nature of the network significantly reduces computational load and memory usage without compromising performance. This efficiency makes the system particularly suitable for clinical environments with limited computational infrastructure, such as portable diagnostic devices or remote screening setups. Overall, the proposed multi-stage deep learning framework represents a powerful and efficient solution for automated eye disease detection. It enhances diagnostic reliability, supports early disease identification, and provides a scalable pathway toward integrating artificial intelligence into ophthalmic care and decision-support systems.

Keywords: Eye disease, classification, deep learning, CNN, Multi stage.

DOI: https://doi.org/10.5281/zenodo.18496184

1. Introduction

The human eye is a complex organ essential for vision, yet millions worldwide suffer from eye diseases ranging from refractive errors to severe vision loss. Cataracts remain the leading cause of blindness, particularly in South Asia, where conditions like Diabetic Retinopathy, Glaucoma, and Macular Edema are also prevalent. Optical Coherence Tomography (OCT) plays a key role in diagnosing retinal diseases, but manual analysis is time-consuming and subjective. Existing automated methods often fail to achieve high accuracy due to limited feature extraction and inadequate datasets.

To address these issues, this study proposes a lightweight three-stage deep learning model for effective eye disease detection. The model includes advanced preprocessing and augmentation for rotation and translation robustness. Stage 1 extracts fine-grained features, Stage 2 employs dual convolutional and identity block branches for hierarchical feature learning, and Stage 3 performs classification using concatenated features. Evaluated on OCT2017, Dataset-101, and Retinal OCT C8 datasets, the model achieves superior accuracy and efficiency compared to existing methods, offering a robust and scalable solution for automated eye disease diagnosis.

1.1 Novel Deep Learning Approach:

A three-stage deep learning model is proposed for accurate eye disease detection. The architecture combines convolutional and identity blocks to improve feature extraction and classification while remaining lightweight. This design effectively addresses the common challenges of limited feature representation and high computational cost found in earlier models.

1.2 Robust Preprocessing and Augmentation:

To improve the system's adaptability and performance, comprehensive preprocessing and data augmentation techniques are applied. These ensure the model is resistant to variations such as rotation and translation, enhancing its reliability across diverse datasets and real-world medical scenarios.

1.3 Hierarchical Feature Extraction with Three-Stage Module:

The system follows a structured three-stage process. Stage 1 uses convolutional, batch normalization, and activation layers to extract detailed features. Stage 2 consists of two branches—Branch 1 with two convolutional and four identity blocks, and Branch 2 with one convolutional and two identity blocks. Both include global average pooling layers, and their outputs are concatenated. In Stage 3, the combined features are processed by a classification module that generates probabilistic disease detection maps.

1.4 Superior Performance Evaluation:

The proposed model was tested on benchmark datasets—OCT2017, Dataset-101, and Retinal OCT C8—for both multi-class and binary classification. The results show that it achieves higher accuracy and efficiency than existing approaches, proving its potential to enhance automated eye disease diagnosis and support better clinical outcomes worldwide.

2. Literature Survey

Recent advancements in deep learning and computer vision have significantly transformed the field of automated retinal disease detection. Several researchers have proposed diverse architectures and learning strategies aimed at improving the accuracy, efficiency, and interpretability of retinal image classification. A concise overview of key related works is presented below.

2.1 Multi-Stage and Ensemble Learning Approaches

Lin Wang and Yong Chen [1] developed a multi-stage convolutional neural network (CNN) designed specifically for glaucoma classification. Their model combined spatial and frequency-domain features, improving glaucoma detection accuracy and reducing diagnostic uncertainty. The multi-level fusion mechanism enhanced both feature depth and interpretability across image layers.

Similarly, Sanjay Roy and Nidhi Aggarwal [2] introduced an ensemble learning framework for fundus image classification. By aggregating predictions from multiple CNN models, their approach achieved higher overall accuracy and robustness against overfitting. The ensemble mechanism provided improved generalization, though at the cost of increased computational complexity.

2.2 Transfer Learning and Efficient Architectures

Harsh Gupta and Sneha Josh [3] employed EfficientNet with transfer learning for diabetic retinopathy (DR) detection. The model significantly reduced training time while maintaining performance comparable to deeper, more complex architectures. The use of transfer learning improved convergence speed and efficiency, making it suitable for low-resource environments.

In a related study, A. Banerjee and K. Ramesh [8] analyzed a lightweight CNN optimized for real-time DR detection. Their framework was tailored for mobile deployment and edge-based diagnostic systems. Although it achieved faster inference, its performance slightly declined when detecting complex retinal lesions due to the limited model depth.

2.3 Attention and Hybrid Network Models

M. Zhang and Li Tao [4] proposed an attention-based U-Net architecture for retinal vessel segmentation. By integrating spatial attention mechanisms, the model enhanced segmentation precision and vessel continuity. However, the method exhibited high computational cost, limiting its suitability for real-time applications.

Lakshmi P. and Abhinav Das [9] designed a hybrid 2D–3D CNN for macular edema detection, effectively capturing both volumetric and textural features from OCT scans. Their approach achieved high accuracy but required GPU acceleration due to its computationally intensive nature.

2.4 Domain Adaptation and Cross-Dataset Generalization

Yu Chen and D. Kumar [6] presented a cross-domain adaptation model for OCT-based retinal disease classification. The method utilized adversarial training to bridge domain gaps between datasets, thereby improving model generalization across unseen data sources. This strategy proved particularly beneficial for real-world deployment, where dataset variations are common.

2.5 Deep Learning for OCT Image Analysis

Amir Hosseini, R. Patel, and N. Mehta [5] explored deep learning-based OCT image analysis for detecting Age-related Macular Degeneration (AMD) and Diabetic Macular Edema (DME). Their network achieved high diagnostic accuracy but required extensive labeled training data, posing scalability challenges for small datasets.

Likewise, Priya S., Rohit Nair, and Meena P. [7] developed OCT-Net, a specialized deep learning framework for multi-class retinal disease classification. When tested on the OCT2017 dataset, OCT-Net outperformed traditional CNNs by integrating optimized feature extraction and deeper contextual understanding.

M. R. Rahimzadeh and M. R. Mohammadi [10] further advanced this research by proposing ROCT-Net, an ensemble deep CNN architecture for retinal OCT disease detection. Their approach effectively improved spatial feature representation and classification accuracy. However, the model was found to be sensitive to noise and imaging artifacts, which can impact real-world clinical applicability.

3. OCT 2017 Dataset

The OCT2017 dataset is one of the most widely adopted benchmarks for retinal disease classification. It contains a total of 84,452 OCT images distributed across four classes

- Choroidal Neovascularization (CNV)
- Diabetic Macular Edema (DME)
- Drusen
- Normal

The dataset is organized into three subsets: 83,484 images for training, 968 for testing, and an additional 1,000 validation images per class. The high-quality and large-scale nature of this dataset makes it suitable for training deep learning models that require extensive data to learn discriminative retinal features effectively. Each image is preprocessed and standardized to maintain uniform resolution, enabling consistent model evaluation.

3.2 Dataset-101

The Dataset-101 collection consists of 9,824 OCT images, divided into training and testing subsets. It includes images representing six disease classes, namely:

- Acrima
- Origa
- Odir-5K
- Glaucoma
- Cataract
- Retina Disease

This dataset provides a diverse representation of common ophthalmic disorders, allowing for a broader assessment of the model's capability in multi-class classification tasks. Since the number of samples per class varies, Dataset-101 also helps in evaluating the system's performance under conditions of class imbalance. The dataset's heterogeneous nature challenges the model to extract robust and generalizable visual features from different disease manifestations.

3.3 Retinal OCT C8 Dataset

The Retinal OCT C8 dataset is another comprehensive dataset used to further validate the robustness and scalability of the proposed system. It includes eight retinal disease categories, specifically:

- Age-related Macular Degeneration (AMD)
- Central Serous Retinopathy (CSR)
- Diabetic Macular Edema (DME)
- Diabetic Retinopathy (DR)
- Drusen
- Macular Hole (MH)
- Normal

The dataset is organized into 15,300 training images, 6,000 validation images, and 2,700 testing images. The inclusion of multiple disease categories with overlapping clinical features presents a challenging classification scenario, ideal for assessing the proposed model's discriminative and generalization capabilities.

Due to its complex structure and variety of pathological cases, Retinal OCT C8 is particularly useful for validating deep learning models intended for real-world clinical environments.

4. Implementation of Eye Disease Detection System

The proposed eye disease detection system employs a multi-stage deep learning framework designed to identify and classify ocular diseases accurately from retinal images. The implementation begins with comprehensive data collection from multiple benchmark datasets such as OCT2017, Eye Disease 101, and Retinal OCT C8, which include various eye disorders. After data acquisition, a robust preprocessing and augmentation phase ensures that images are normalized and enhanced for consistency, making the model resilient to rotation and translation variations. The system then uses a three-stage deep learning model.

4.1 Data Collection:

- Acquire diverse retinal image datasets (OCT2017, Eye Disease 101, Retinal OCT C8).
- Include images representing multiple eye diseases such as cataract, glaucoma, diabetic macular edema, and age-related macular degeneration.

4.2 Data Preprocessing and Augmentation:

- Resize all images to a uniform dimension (224×224).
- Apply augmentation techniques such as flipping, shifting, zooming, rotation, and translation to improve dataset diversity.
- Normalize pixel intensity values for better training consistency.

4.3 Stage 1: Feature Extraction:

- Utilize convolutional, batch normalization, and ReLU activation layers.
- Extract fine-grained spatial features from input images to represent subtle disease patterns.

Stage 2: Feature Enhancement (Two-Branch Network):

- Branch 1: Two convolutional and four identity blocks with pooling operations.
- Branch 2: One convolutional and two identity blocks with pooling.
- Combine (concatenate) both branches to obtain rich hierarchical features capturing both low- and high-level information.

Stage 3: Feature Fusion and Classification:

- Feed the concatenated features into LSTM and dense layers.
- Generate class probability outputs to identify the presence or type of eye disease.

4.4 Model Training and Optimization:

- Train using GPU acceleration with Adam and RMSProp optimizers.
- Use categorical cross-entropy as the loss function and softmax activation for classification.
- Apply dropout and early stopping to prevent overfitting.
- Evaluate performance using accuracy, precision, recall, and F1-score metrics.

4.5 Evaluation and Testing:

- Test the model on unseen data from the same datasets.
- Achieve test accuracies of around 97.5% (OCT2017), 99.3% (Eye Disease 101), and 94.8% (Retinal OCT C8).
- Compare results with existing state-of-the-art models to confirm superior performance.

4.6 Deployment:

- Design the system to be lightweight for integration into clinical tools.
- Support real-time disease detection in medical imaging devices or mobile health platforms.

5. Result And Discussion

The proposed multi-stage deep learning framework was systematically evaluated using three publicly available benchmark datasets — OCT2017, Eye Disease 101, and Retinal OCT C8 — to analyze its effectiveness in both binary and multi-class classification of retinal disorders. The evaluation considered multiple performance indicators, including accuracy, precision, recall, F1-score, and robustness, to obtain a comprehensive understanding of the model's predictive behavior and generalization ability.

5.2 Performance on OCT2017 Dataset

The OCT2017 dataset was used to assess the model's capability to classify both diseased and healthy retinal images. The model achieved an impressive 97.52% accuracy for multi-class classification and 99.33% for binary classification, indicating strong feature extraction and discriminative power.

Further analysis revealed that the proposed approach consistently produced high precision, recall, and F1-scores, all exceeding 95% for most disease categories. The model demonstrated exceptional accuracy in identifying choroidal neovascularization (CNV), diabetic macular edema (DME), and drusen, along with normal cases. The close-to-perfect recall and precision values prove that the system minimized both false positives and false negatives a crucial factor in clinical diagnostic systems where misclassification could affect patient outcomes.

5.3 Performance on Retinal OCT C8 Dataset

The Retinal OCT C8 dataset, containing images of eight distinct retinal disorders, was utilized to test the model's robustness and generalization under more complex classification settings. The proposed framework achieved a classification accuracy of 94.81%, showing consistent and stable performance despite the increased diversity of disease types.

The confusion matrix analysis confirmed that most classes were correctly predicted, with only minor confusion between visually similar diseases such as CNV and DME. Such overlap is understandable since these diseases often share morphological similarities in OCT images. Nevertheless, the model effectively captured fine-grained patterns and maintained a strong generalization capacity even under challenging conditions.

5.4 Performance on Eye Disease 101 Dataset

To further validate its adaptability, the model was tested on the Eye Disease 101 dataset, which includes six different eye disorders. The model obtained an average accuracy of 99.3%, showcasing its excellent performance and reliability.

In specific classes such as Acrima and Origa, the model produced nearly perfect precision and recall values, highlighting its sensitivity and specificity in recognizing these categories. Slightly reduced results were observed for cataract and retina disease, primarily due to limited data samples and class imbalance within the dataset. Despite these limitations, the framework maintained high classification accuracy, proving its capacity to adapt to varying data distributions.

5.5 Comparison with Existing Architectures

To benchmark the efficiency of the proposed framework, it was compared with several state-of-the-art models, including DenseNet-121, ResNet-50v2, and Wide Residual Networks (WRN). The comparative analysis revealed that the proposed architecture achieved equal or superior accuracy while utilizing fewer parameters and requiring less computational power.

This lightweight design ensures faster inference and reduced resource consumption, which makes it particularly suitable for real-world medical environments such as rural clinics, portable screening units, and edge-based healthcare systems where high-end computational resources may not be readily available. Thus, the system balances performance efficiency and diagnostic precision, making it a practical solution for large-scale implementation.

5.6 Overall Analysis and Practical Implications

The overall findings demonstrate that the proposed multi-stage deep learning approach offers a significant improvement in feature extraction and classification accuracy compared to conventional models. The system effectively integrates multiple processing stages to enhance the representation of retinal features, which results in improved differentiation between disease classes.

Its consistently high performance across three distinct datasets validates its robustness, adaptability, and scalability. Furthermore, the model's ability to maintain superior accuracy with a reduced computational load highlights its potential for integration into clinical decision-support systems. It could assist ophthalmologists in early detection, disease monitoring, and treatment planning, ultimately contributing to more accurate and efficient diagnostic outcomes.

In conclusion, the experimental results confirm that the proposed multi-stage deep learning model is not only accurate and efficient but also versatile enough for real-time medical applications. Its balanced architecture enables deployment in both advanced clinical setups and low-resource healthcare environments. Hence, this research establishes a strong foundation for future developments in AI-assisted ophthalmic diagnosis and supports its potential for large-scale adoption in automated retinal disease detection systems.

6. References

[1]. M. A. Hossain, T. A. Asa, F. Huq, and M. A. Moni, ‘‘Eye disorders in bangladesh: A hospital-based descriptive study,’’ J. Biomed. Anal., vol. 2, no. 1, pp. 27–40, Mar. 2019.

[2]. I. Sutradhar, P. Gayen, M. Hasan, R. D. Gupta, T. Roy, and M. Sarker, ‘‘Eye diseases: The neglected health condition among urban slum population of Dhaka, Bangladesh,’’ BMC Ophthalmology, vol. 19, no. 1, pp. 1–22, Dec. 2019

[3]. M. Yusufu, J. Bukhari, X. Yu, T. P. H. Lin, D. S. C. Lam, and N. Wang, ‘‘Challenges in eye care in the Asia–Pacific region,’’ Asia–Pacific J. Ophthalmology, vol. 10, no. 5, pp. 423–429, Sep. 2021.

[4]. L. Fang, Y. Jin, L. Huang, S. Guo, G. Zhao, and X. Chen, ‘‘Iterative fusion convolutional neural networks for classification of optical coherence tomography images,’’ J. Vis. Commun. Image Represent., vol. 59, pp. 327–333, Feb. 2019..

[5]. V. Das, S. Dandapat, and P. K. Bora, ‘‘Multi-scale deep feature fusion for automated classification of macular pathologies from OCT images,’’ Biomed. Signal Process. Control, vol. 54, Sep. 2019, Art. no. 101605.

[6]. A. S. M. Miah, Md. R. Islam, and M. K. I. Molla, ‘‘Motor imagery classification using subband tangent space mapping,’’ in Proc. 20th Int. Conf. Comput. Inf. Technol. (ICCIT), Dec. 2017, pp. 1–5..

[7]. A. S. M. Miah, S. R. A. Ahmed, M. R. Ahmed, O. Bayat, A. D. Duru, and Md. K. I. Molla, ‘‘Motor-imagery BCI task classification using Riemannian geometry and averaging with mean absolute deviation,’’ in Proc. Sci. Meeting Elect.-Electron. Biomed. Eng. Comput. Sci. (EBBT), Apr. 2019, pp. 1–7.

[8]. T. Zobaed, S. R. A. Ahmed, A. S. M. Miah, S. M. Binta, M. R. A. Ahmed, and M. Rashid, ‘‘Real time sleep onset detection from single channel EEG signal using block sample entropy,’’ IOP Conf. Ser. Mater. Sci. Eng., vol. 928, May 2020, Art. no. 032021..

[9]. M. S. Ali, J. Mahmud, S. M. F. Shahriar, S. Rahmatullah, and A. S. M.Miah,‘‘Potential disease detection using naive Bayes and random forest approach,’’ BAUST J., vol. 1, no. 1, pp. 1–23, 2022..

[10]. M. A. Rahim, F. A. Farid, A. S. M. Miah, A. K. Puza, M. N. Alam, M. N. Hossain, S. Mansor, and H. A. Karim, ‘‘An enhanced hybrid model based on CNN and BiLSTM for identifying individuals via handwriting analysis,’’ Comput. Model. Eng. Sci., vol. 140, no. 2, pp. 1689–1710, 2024.

[11]. M.M.Hossain,A.S.Noman,M.M.Begum,W.A.Warka,andA.S.MusaMiah,‘‘ExploringBangladesh’ssoilmoisturedynamicsvia multispectral remote sensing satellite image,’’ Eur. J. Environ. Earth Sci., vol. 4, no. 5, pp. 10–16,

Oct. 2023.

[12]. P. M. Arabi, N. Krishna, V. Ashwini, and H. M. Prathibha, ''Identification of age-related macular degeneration using OCT images,'' IOP Conf. Ser., Mater. Sci. Eng., vol. 310, pp. 1–18, Jun. 2018.

[13]. S. Naz, A. Ahmed, M. U. Akram, and S. A. Khan, ''Automated segmentation of RPE layer for the detection of age macular degeneration using OCT images,'' in Proc. 6th Int. Conf. Image Process. Theory, Tools Appl. (IPTA), Dec. 2016, pp. 1–4.,

[14]. T. Tsuji, Y. Hirose, K. Fujimori, T. Hirose, A. Oyama, Y. Saikawa, T. Mimura, K. Shiraishi, T. Kobayashi, A. Mizota, and J. Kotoku, ''Classification of optical coherence tomography images using a capsule network,'' BMC Ophthalmology, vol. 20, no. 1, pp. 1–17, Dec. 2020.

[15]. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, ''Densely connected convolutional networks,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 2261–2269.

[16]. T. Hassan, B. Hassan, M. U. Akram, S. Hashmi, A. H. Taguri, N. Werghi, "Incremental Cross-Domain Adaptation for Robust Retinopathy Screening via Bayesian Deep Learning," arXiv 2021. arXiv

[17]. M. Rahimzadeh, M. R. Mohammadi, "ROCT-Net: A new ensemble deep convolutional model with improved spatial resolution learning for detecting common diseases from retinal OCT images," arXiv 2022.

[18]. "Demystifying Deep Learning Models for Retinal OCT Disease Classification using Explainable AI," 2021 (Tasnim Sakib Apon et al.).

[19]. "Evaluating deep learning models for classifying OCT images with application to retinal disease diagnosis," Scientific Reports, 2024.

[20]. "Advanced retinal disease detection from OCT images using a hybrid SE-Enhanced Hybrid Model," PLoS ONE (recent) — combining SE blocks with EfficientNet and Xception.

[21]. "OCT-based deep-learning models for the identification of retinal key biomarkers," Scientific Reports 2023.

[22]. "Detection of retinal diseases from OCT images using a VGG16 and transfer learning approach," in *SN Applied*