Lightweight Convolutional Neural Network for Multi-Class Classification of Retinal Fundus Images

Baburam Adhikari
Faculty of Computer Science & Tech.
Algoma University
Sault Ste. Marie, ON, Canada
baadhikari@algomau.ca

Sagar Phuyal
Faculty of Computer Science & Tech.
Algoma University
Sault Ste. Marie, ON, Canada
saphuyal@algomau.ca

Ujjwal Uprety
Faculty of Computer Science & Tech.
Algoma University
Sault Ste. Marie, ON, Canada
uuprety@algomau.ca

Khalil Al-Hussaeni
Associate Professor of Computing Sciences
Rochester Institute of Technology
Dubai, United Arab Emirates
kxacad@rit.edu

Rashid Hussain Khokhar
Faculty of Computer Science and Technology
Algoma University
Sault Ste. Marie, ON, Canada
rashid.khokhar@algomau.ca

***Abstract*—Eye diseases such as Diabetic Retinopathy, Glaucoma, and Cataract are leading causes of blindness worldwide, underscoring the importance of early detection. Manual examination of retinal fundus images is labor-intensive and prone to oversight, motivating the need for automated solutions. We present a lightweight Convolutional Neural Network (CNN) designed to classify fundus images into four categories: Normal, Cataract, Glaucoma, and Diabetic Retinopathy. Unlike prior approaches that depend on deep architectures or transfer learning, our model emphasizes efficiency with only four convolutional layers while maintaining strong performance. It achieves a validation accuracy of 91.16%, surpassing both a Squeeze-and-Excitation (SE) block based CNN with batch normalization and the pre-trained EfficientNetB0. The key contribution of this work is demonstrating that a compact, application-ready CNN can achieve competitive accuracy with reduced complexity, making it practical for portable screening tools and deployment in resource-limited healthcare environments.**

***Index Terms*—Convolutional Neural Network, Eye Disease Classification, Lightweight Deep Learning, Retinal Fundus Images**

I. Introduction

Our eyes are very important because they help us perceive the world, but many diseases can harm vision and even cause blindness. Cataracts are one of the main reasons people lose their sight worldwide, especially in low-resource countries where access to treatment is limited. Another serious eye problem is diabetic retinopathy, which occurs when diabetes damages the tiny blood vessels at the back of the eye. It is the leading cause of blindness among adults in the United States, but it can be prevented or treated early through proper blood sugar control and regular eye checkups. Glaucoma is also a common cause of permanent blindness, affecting an estimated 80 million people worldwide by 2025. It often has no early symptoms, so many individuals do not realize they have it until significant vision loss has occurred [1].

Cataracts are one of the leading causes of visual impairment and blindness globally. Around 65 million number of people are affected by this and the number was expected to rise to around 70 million by the year 2020. Many people with cataracts are still not diagnosed and makes it a serious public health issue. [2].

Diabetic retinopathy causes continuous damage to the blood vessels of the retina, the light-sensitive tissue at the back of the eye necessary for good vision. Its risks can be reduced by controlling blood sugar, blood pressure, and cholesterol. Although diabetic retinopathy can be treated in its early stages, around 50% of patients do not receive regular eye examinations for timely detection. An estimated 4.1 million Americans have diabetic retinopathy, and 899,000 have vision-threatening retinopathy [1].

Glaucoma is a group of diseases that can damage the optic nerve and lead to vision loss and blindness. It typically occurs when the fluid pressure inside the eye gradually increases; however, recent findings show that glaucoma can also occur with normal eye pressure. It has became one of the cause of irreversible cause of blindness all over the world. [3].

As of 2025, glaucoma remains one of the leading causes of permanent blindness worldwide. About 80 million people are living with the disease globally, and this number is expected to grow as populations age. The most common form, open-angle glaucoma, affects around 60 million people. Glaucoma accounts for approximately 9%–12% of all cases of blindness, meaning that nearly 5.9 million people have already lost their sight because of it. This makes glaucoma the second leading cause of blindness after cataracts [4].

The fundus, the back portion of the eye, includes the retina,

macula, and blood vessels, which are critical for vision. Many conditions, such as diabetic retinopathy and glaucoma, can be detected by examining the fundus. An estimated 93 million people worldwide have diabetic retinopathy, a disease that can cause vision loss if not treated. Because the fundus allows physicians to directly observe the eye's tiny blood vessels without invasive procedures, regular checkups are essential for early detection and prevention of vision-threatening diseases [5].

Since eye diseases such as cataracts are the global cause of blindness of around 45%, diabetic retinopathy is the leading cause of blindness in the US, and glaucoma affects millions of people worldwide, early detection is crucial. Traditional methods heavily rely on manual screening, which is time-consuming and prone to human error. Deep learning CNN models have therefore become a practical solution for image-based diagnosis [6].

While several CNN-based methods have been developed for eye disease classification (EDC), most prioritize accuracy gains through deeper networks or transfer learning at the expense of computational efficiency. These models often require substantial resources, making them impractical for portable or low-cost screening systems. In this work, we address this gap by designing a lightweight CNN with only four convolutional layers that maintains competitive accuracy while significantly reducing computational overhead.

The contributions of this paper are as follows:

- We propose a compact CNN architecture optimized for multi-class classification of retinal fundus images (Normal, Cataract, Glaucoma, Diabetic Retinopathy) that balances accuracy and efficiency.
- We demonstrate that our lightweight model achieves higher accuracy than both EfficientNetB0 and Squeeze-and-Excitation (SE) block based networks, despite having far fewer parameters.
- We show that low-complexity, dense architectures can effectively capture subtle retinal disease patterns, making them suitable for mobile and resource-constrained healthcare applications.

II. Literature Review

In traditional eye disease diagnosis, doctors rely on clinical examinations by observing symptoms and conducting various medical tests [7]. Deep learning models, particularly convolutional neural networks (CNNs), have been developed to detect and classify diabetic retinopathy (DR), cataract, or glaucoma disease using retinal images [8]. These models often achieve higher accuracy through automated feature extraction; however, their ability to generalize decreases when image quality is poor or when applied to diverse patient datasets with varying demographics. A comprehensive review of deep learning techniques for diabetic retinopathy detection and classification from retinal fundus images [9], they analyzed different range of CNN-based and transfer-learning models like ResNet, VGG, and EfficientNet architectures. They highlighted that most state-of-the-art systems achieve accuracies between 85 % and 96 % on datasets such as EyePACS and Messidor. Since most of the models are computationally intensive and lack generalization across imaging devices, this motivates the need for lightweight architectures that are suitable for portable or real-time screening environments.

Significant progress has been made in applying deep learning to eye disease diagnosis. Researchers have shown that CNN-based models can approach radiologist-level accuracy in identifying conditions such as DR, glaucoma, and cataracts. Despite these advances, practical implementation in clinical settings remains challenging due to high computational demands and limited generalization across different patient populations [10].

A. CNN-Based Diagnosis of Major Eye Diseases

Deep learning models have shown significant success in different eye diseases like Diabetic Macular Edema (DME), Choroidal Neovascular Membranes (CNM), and Age related Macular Degeneration (AMD). A DL model achieved an accuracy of 94% [10]. An attention mechanism using CNN [4] was applied to detect glaucoma, reducing false negatives by 15% compared to traditional methods, but it struggled to detect subtle changes in the early stages. A transfer learning-based cataract detection and grading system achieved an accuracy of 92%, yet it failed to classify early-stage cases, making it more suitable for post-operative analysis than pre-diagnostic screening [11]. A multi-class image classification method using two transfer learning models was also proposed to identify abnormal cases [12]. Another study introduced a custom vision transformer for ocular disease classification, using a specialized ViT architecture with 14 transformer layers. However, it achieved only an F1-score of 83–84% due to dataset limitations [13].

B. Optimization Techniques and Lightweight Models

Researchers have attempted to address computational limitations by shrinking large models.One study [14] applied a MobileNetV2 which consisted approximately 3.3 million parameters using depth-wise separable convolutions in order to classify five stages of diabetic retinopathy and was able to achieve an overall accuracy of 93.9%. This shows that the lightweight model architecture for DR along with the optimization and efficient-model can be achieved. One study depthwise separable CNN-based multi-disease detection model [15] was proposed, but it performed poorly when training data was scarce for certain diseases. These findings suggest that reducing model size to improve speed often leads to decreased accuracy, as smaller models may fail to detect rare eye diseases or generalize well across diverse patient groups.

C. Feature Extraction and Data Challenges

One of the critical challenges in deep learning models is effective feature extraction. To detect the early signs of DR in eye scans, older methods such as Haar wavelets and LBPs (basic image analysis tools) often miss important

details. Another major challenge is the limited availability of training datasets, which can lead to generalization issues. Similarly, class imbalance in datasets is a leading cause of overfitting [16].

D. Critical Research Gaps

There remains a significant gap between academic research and clinical practice, as many findings stay at the research stage and fail to translate into real-world clinical scenarios due to their resource-intensive nature. Several issues persist in eye disease classification. For example, some pretrained models underperform compared to custom lightweight CNN models, while heavier architectures such as ResNet are not well-suited for mobile health applications. Models are also prone to overfitting due to class imbalance in datasets and the lack of mechanisms for detecting subtle features in the early stages of disease development. Furthermore, there is uncertainty about cost-effectiveness, as very little research has evaluated whether deep learning models can be economically viable compared to traditional diagnostic methods. If deep learning models are not cost-effective, adoption in healthcare systems will remain limited unless lightweight alternatives are developed.

III. Methodology

A. Dataset and Preprocessing

The dataset consists of 4217 color retinal fundus images taken from Kaggle, distributed across four categories: Normal, Cataract, Glaucoma, and Diabetic Retinopathy (as shown in Fig. 1) and each class has 1074, 1038, 1007 and 1098 images respectively. The collected images were initially of varying dimension which includes (512, 512), (2464, 1632), (2592, 1728), (1848, 1224), (256, 256). To ensure consistency in training model all images were resized to 256×256 pixels and normalized to [0,1]. We split the dataset in three part training (80%), testing (10%) and validation (10%). Then, data augmentation was done only in training set to improve robustness of the model, the images were rescaled so that the pixel value ranged between 0 and 1 followed by random flipping (vertical and horizontal), slight rotation, and small zoom by 0.1. The dataset remained balanced after augmentation and simplifying per-class evaluation. No additional pre-processing (eg. denoising or contrast enhancement) was applied, as the raw images were sufficient features for CNN training.

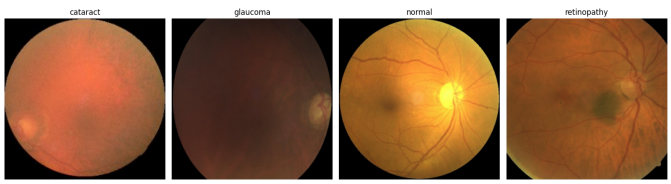

Fig. 1: Fundus Images

B. CNN Architecture

We tested three different CNN variants in addition to the baseline model. The first included SE blocks, batch normalization, spatial dropout, and focal loss to enhance feature attention and address class imbalance. This network goes through four convolutional layer with increasing filter (32, 64, 64, 64) and each layer has a combination of batch normalization and ReLu activation function for stable and efficient training. MaxPooling layer was added in each block to select the maximum element from the region of the feature map covered by the filter. To improve feature learning Squeeze and Excitation block was included, which helps the model focus on the most important channels [17]. After feature extraction the output is flattened and passed through a dense layer of 64 unit with the ReLu activation function then a dropout layer of 0.5 was applied to reduce the overfitting. In final layer we used softmax layer for the classification of the images into target class. For training purpose, we applied custom focal loss function, which gives more weight to difficult and misclassified samples and to handle class imbalance more efficiently [18].

In the second model, we applied transfer learning with fine-tuning using EfficientNetB0 as the backbone. At first, the EfficientNetB0 network pre-trained on ImageNet was used as a frozen feature extractor (trainable = False), meaning its weights were not updated during the initial training phase. Then the input image were passed through data augmentation block to improve generalization, then into the base model. A Global Average Pooling layer was used which reduces the feature map to a compact vector, followed by a dense layer of 128 unit and dropout of 0.4 to prevent overfitting. Finally a softmax output layer was used with 4 neurons that gives the class probabilities. The model was first trained for 10 epochs using Adam Optimizer. Then fine-tuning was applied, top 30 layer of EfficientNetB0 are unfrozen for further training, while remaining was kept frozen. In last stage the model was recompiled with a learning rate of 1e-5 to avoid large updates that might damage pretrained weights and then it was trained for 10 epochs. In short, this was a two-step transfer learning approach: first, EfficientNetB0 was used as a frozen feature extractor, and then its top layers were unfrozen and trained with a smaller learning rate to better adapt to the dataset and improve accuracy.

In the third model, we created a custom 4-layer CNN to classify fundus images into four disease categories. In the first two layer 32 filters with 3×3 kernels and ReLU activation was applied to extract basic features such as edges and textures, followed by 64 and 128 filter in 3rd and 4th layer. After each convolution layer max-pooling of (2X2) window was used to reduced the spatial size which helped the model to focus on important details and reduce complexity.

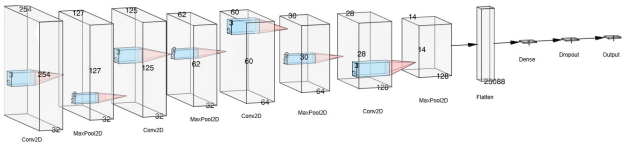

Fig. 2: CNN architecture.

The extracted features were flattened and passed through fully connected layer with 128 neurons using ReLU activation. A dropout layer with a 30% rate was included to mitigate overfitting. The final softmax layer, with four output neurons was applied which generates the class probabilities for the target categories. The model contained approximately 12.9 million trainable parameters, the majority of which are concentrated in the dense layers—particularly in the connection to the Dense(128) layer, which alone contributes over 10 million weights. The model was implemented in TensorFlow/Keras and trained from scratch using an 80-10-10 train-test-validation split for 50 epochs. We deployed the Adam optimizer (learning rate = 0.001), categorical cross-entropy loss, and a batch size of 32. The was training completed in all 50 epochs as performance gradually increased. On-the-fly data augmentation—via random flips and random rotations using Keras was applied only to the training set to improve generalization. Training was conducted on a Google Colab A100 GPU, with each epoch taking less then 10 seconds.

IV. Result and Discussion

In this section, we will see the experimental results obtained from the all CNN models and compare their performance. The evaluation includes overall accuracy, macro-averaged precision, recall and F1-scores as well as class wise results to highlight strengths and weakness around different disease categories. We will compare the output from the all model and identify the most effective model for the classification task.

A. Overall Model Comparison and Selection

The comparative analysis of the three CNN model showed clear differences in their overall performance. The first SE CNN model with focal loss achieved an overall accuracy of 0.66 and a macro-F1 of 0.66. This model performed satisfactorily in identifying cataract and retinopathy, but gave lower recall for glaucoma and reduced precision when classifying normal class. The transfer learning model based on EfficientNetB0 with fine-tuning provided a considerable improvement, reaching an accuracy of 88.9% and a macro-F1 of 0.884, with balanced gains across all classes. However, the best performance was given by the simple custom CNN, with 91.16% accuracy and macro-F1 = 0.901. The model performed high and stable for all the classes, labeling retinopathy (F1 = 1.0) correctly and obtaining good results for the cataract, glaucoma, and normal classes. Meeting both efficiency and effectiveness conditions, the simple CNN was selected as the final model because it achieved the best balance among accuracy, stability, and resource usage. Table I below shows the overall comparison.

TABLE I: Overall performance comparison of different CNN models.

Model	Accuracy	Macro Precision	Macro Recall	Macro F1
SE-CNN + Focal Loss	0.660	0.700	0.660	0.660
EfficientNetB0 (Fine-tuned)	0.889	0.893	0.883	0.884
Simple CNN (Proposed)	**0.9116**	**0.9025**	**0.901**	**0.901**

B. Training and Validation Performance

As shown in Fig. 3, the training and validation accuracy curves increase gradually and level off near their highest points. The loss curves also decrease consistently, with no sign of overfitting. This confirms that the model trained effectively and the chosen design and training approach worked well.

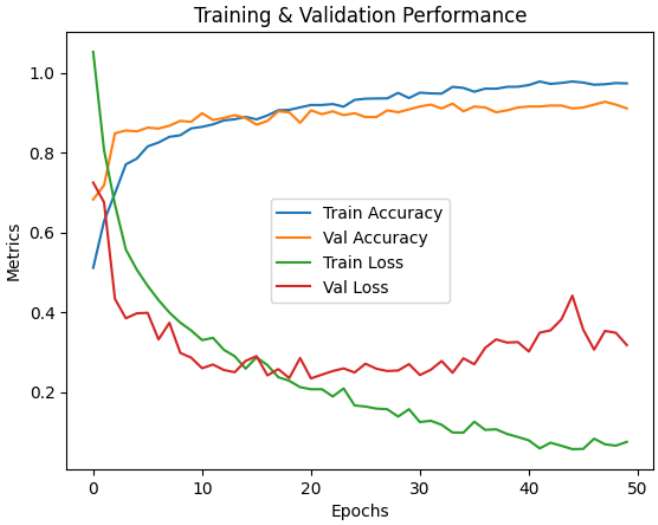

Fig. 3: Training and validation accuracy and loss curves of the final CNN model

C. Evaluation Metrics and Overall Performance

After training, the model was evaluated on a held-out validation set (10% of the data), scoring an overall accuracy of 91.16%, meaning more then 91% of fundus images were accurately classified. This performance matches state-of-the-art models like EfficientNetB0, showing that our custom CNN despite a simpler architecture was very effective. To evaluate class-wise performance, we calculated precision, recall, and F1-score for each category (refer to Table I). The model achieved the highest recall for retinopathy (1.00) followed by cataract (0.9206), which means that their results were rarely missed. Table II summarizes these metrics:

TABLE II: Class-wise performance metrics with support counts

Class	Precision	Recall	F1-score	Support
Cataract	0.9355	0.9206	0.9280	126
Glaucoma	0.8250	0.7765	0.8000	85
Normal	0.8496	0.9057	0.8767	106
Retinopathy	1.0000	1.0000	1.0000	124
Accuracy	–	–	**0.9116**	441
Macro Average	0.9025	0.9007	0.9012	441
Weighted Average	0.9117	0.9116	0.9112	441

The normal class also showed high recall (0.9057) but slightly lower precision (0.8496), meaning some diseased eyes were incorrectly labeled as normal, suggesting a mild tendency to over-predict the healthy class when pathology is barely noticeable. In contrast, glaucoma showed relatively lower recall (0.7765), indicating that this class was more challenging for the model, though performance was still robust with an F1-score of 0.80.

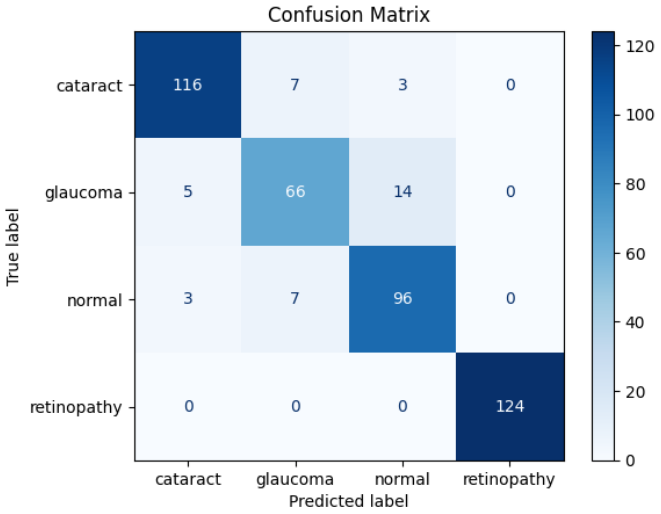

Fig. 4: Confusion matrix of the four-class classification results

The confusion matrix (as shown in Fig. 4) validates the performance analysis, showing most predictions correctly aligned along the diagonal, with only 39 misclassifications out of 441 samples. Most errors occurred when early-stage or mild cases were predicted as normal, especially in glaucoma and cataract. For instance, 14 glaucoma and 3 cataract images were misclassified as normal. This suggests the model has high specificity but slightly lower sensitivity for detecting subtle disease features.

Surprisingly, the model never confused one disease with another, as retinopathy was always predicted correctly, and no cases were misclassified as glaucoma or vice versa. This indicates that the model learned strong class-specific features. This means that the model effectively learned the unique characteristics of each class. The result is consistent with previous findings and indicates that the model can be optimized for early screening environments where high sensitivity is particularly crucial.

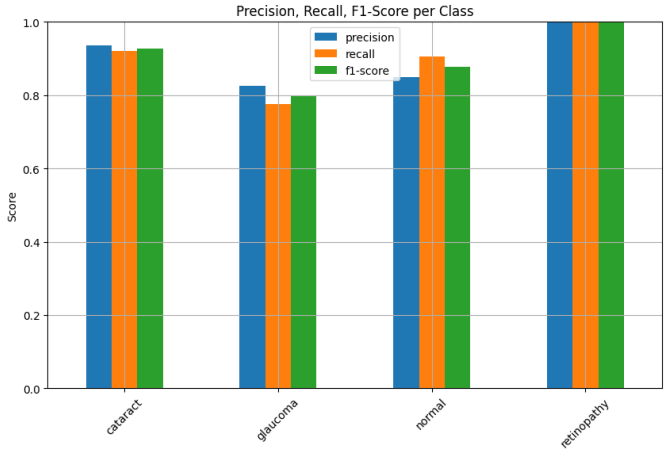

Fig. 5: Class-wise Precision, Recall, and F1-Score Comparison

The bar chart (shown in Fig. 5) compares the precision, recall, and F1-score for each eye disease class. From the chart and the corresponding confusion matrix, we can observe the following results:

1) Cataract: 116 out of 126 cataract images were correctly classified, with a recall of 0.9206. 7 images were misclassified as glaucoma and 3 as normal. These misclassifications suggest that in some cases, the model might have confused overlapping features—such as brightness at the optic disc or mild opacities—with features seen in glaucomatous or normal eyes. Positively, no cataract images were mistaken for retinopathy, showing good discrimination against retinal pathology.

2) Glaucoma: The model worked well in classifying 66 out of 85 glaucoma images, recalling 0.7765 and having an F1-score of 0.8000—the poorest among all classes. 14 glaucoma images were misclassified as normal, and 5 as cataract, implying the model is possibly not so great at detecting very faint signs of early glaucoma. However, no glaucoma cases were confused with retinopathy, showing exemplary class-specific learning.

3) Normal: Out of 106 normal images, 96 were correctly classified with a recall of 0.9057. 3 were misclassified as cataract, and 7 as glaucoma. These errors may be due to small artifacts or small features that resemble pathology. Despite that, the model performed well with an F1-score of 0.8767.

4) Retinopathy: The model achieved perfect performance for this class: it correctly predicted all 124 retinopathy images, with precision, recall, and F1-score of 1.0000. This confirms the high capacity of the model to detect retinal lesions and classify retinopathy very well from other eye disorders.

D. Visual Analysis of Predictions

We assessed both correct and incorrect predictions to understand the model's behavior. Accurate classifications aligned with typical visual features: cataracts presented as hazy fundus images, glaucoma showed large optic disc cupping, and diabetic retinopathy revealed lesions such as microaneurysms, indicating that the model had learned clinically essential patterns. Errors mainly involved subtle or early-stage cases,

where distinguishing features were negligible. For instance, mild cataracts or early glaucoma were misclassified as Normal, and poor image quality sometimes led to false positives.

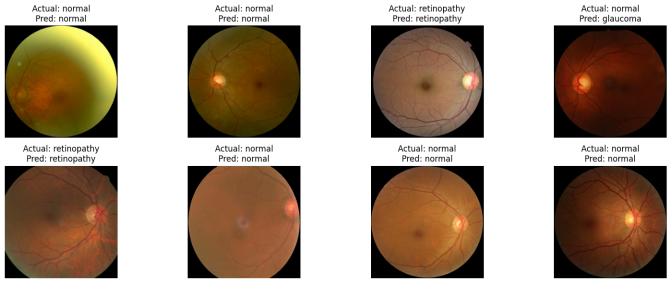

Fig. 6: Predicted vs Actual Images

Fig. 6 shows sample images along with their actual and predicted labels. Most images are correctly classified, particularly when the disease features are clear and easily visible, such as in retinopathy cases. However, one Normal image was incorrectly predicted as Glaucoma, suggesting that the model may occasionally misinterpret subtle patterns.

Overall, the results indicate that the model performs well in typical cases, particularly when disease signs are distinct. The overall accuracy (91.16%) and strong F1-scores across classes support the reliability of the approach. To improve early disease detection, where symptoms may be less visible, expanding the dataset with more diverse images or leveraging pretrained models could make the system more sensitive to subtle changes.

In general, the streamlined CNN proved to be the most effective, complex enough to learn critical retinal characteristics, yet simple enough to avoid overfitting, showing strong and balanced performance across all disease categories.

V. Conclusion

This research demonstrates that a lightweight, custom-made Convolutional Neural Network can effectively classify eye diseases such as cataract, glaucoma, diabetic retinopathy, and normal retinal fundus images with high accuracy. By using a balanced dataset, minimal preprocessing, and real-time data augmentation, the proposed four-layer CNN architecture achieved an accuracy of 91.16%. This system outperformed high-parameter, pre-trained models such as EfficientNetB0. The architecture employed convolutional layers with progressive filter sizes, max pooling, ReLU activation, and dropout for regularization. The results showed excellent performance, including perfect precision and recall for diabetic retinopathy. Overall, the model demonstrated strong generalization, computational efficiency, and practical suitability for early screening and diagnostic support, making it a promising solution for portable and resource-constrained applications.

References

[1] Centers for Disease Control and Prevention. (2023, October) Common eye disorders. Accessed: 2025-07-20. [Online]. Available: https://www.cdc.gov/vision-health/about-eye-disorders/index.html

[2] J. H. L. Goh, Z. W. Lim, X. Fang, A. Anees, S. Nusinovici, T. H. Rim, C.-Y. Cheng, and Y.-C. Tham, "Artificial intelligence for cataract detection and management," *Ophthalmology Science*, vol. 3, no. 6, p. 100414, 2023.

[3] R. N. Weinreb, T. Aung, and F. A. Medeiros, "Glaucoma," *The Lancet*, vol. 389, no. 10075, p. 136–148, 2017. [Online]. Available: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(17)31469-1/fulltext

[4] A. R. Wahab Sait, "Artificial intelligence-driven eye disease classification model," *Applied Sciences*, vol. 13, no. 20, p. 11437, 2023. [Online]. Available: https://doi.org/10.3390/app132011437

[5] O. Bernabé, E. Acevedo, A. Acevedo, R. Carreño, and S. Gómez, "Classification of eye diseases in fundus images," *IEEE Access*, vol. 9, pp. 101 267–101 276, 2021. [Online]. Available: https://doi.org/10.1109/ACCESS.2021.3094649

[6] M. Juneja, S. Singh, N. Agarwal, S. Bali, S. Gupta, N. Thakur, and P. Jindal, "Automated detection of glaucoma using deep learning convolution network (G-net)," *Multimedia Tools and Applications*, vol. 79, pp. 15 531–15 553, 2020.

[7] Y. Peng, S. Dharssi, Q. Chen, T. D. Keenan, E. Agrón, W. T. Wong, E. Y. Chew, and Z. Lu, "DeepSeeNet: A deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs," *Ophthalmology*, vol. 126, pp. 565–575, 2019.

[8] T. Babaqi, M. Jaradat, A. E. Yildirim, S. H. Al-Nimer, and D. Won, "Eye disease classification using deep learning techniques," in *Proceedings of the IISE Annual Conference & Expo 2023*, K. Babski-Reeves, B. Eksioglu, and D. Hampton, Eds. Binghamton, NY, USA: Institute of Industrial and Systems Engineers (IISE), 2023. [Online]. Available: https://arxiv.org/abs/2307.10501

[9] N. Tsiknakis, D. Theodoropoulos, G. Manikis, E. Ktistakis, O. Boutsora, A. Berto, F. Scarpa, A. Scarpa, D. I. Fotiadis, and K. Marias, "Deep learning for diabetic retinopathy detection and classification based on fundus images: A review," *Computers in Biology and Medicine*, vol. 135, p. 104599, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010482521003930

[10] M. Elkholy and M. A. Marzouk, "Deep learning-based classification of eye diseases using convolutional neural network for OCT images," *Frontiers in Computer Science*, vol. 5, p. 1252295, 2024, open Access. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fcomp.2023.1252295/full

[11] S. Yadav and J. K. P. S. Yadav, "Automatic cataract severity detection and grading using deep learning," *Journal of Sensors*, vol. 2023, 2023, article ID 2973836, 17 pages. [Online]. Available: https://doi.org/10.1155/2023/2973836

[12] M. S. Junayed, M. B. Islam, A. Sadeghzadeh, and S. Rahman, "CataractNet: An automated cataract detection system using deep learning for fundus images," *IEEE Access*, vol. 9, pp. 128 799–128 808, 2021.

[13] S. D. Gummadi and A. Ghosh, "Classification of ocular diseases: A vision transformer-based approach," in *Innovations in Computational Intelligence and Computer Vision*, ser. Lecture Notes in Networks and Systems. Springer, 2023, pp. 37–47.

[14] H. N. Huynh, M. T. Do, G. T. Huynh, A. T. Tran, and T. N. Tran, "Classification of stages diabetic retinopathy using MobileNetV2 model," in *Proceedings of the International Symposium on Applied Science 2021 (ISAS 2021)*, T. T. Truong, T. N. Tran, T. N. Nguyen, and Q. K. Le, Eds., vol. 4. Kalpa Publications in Engineering, 2022, pp. 147–157. [Online]. Available: https://easychair.org/publications/paper/k4C9/open

[15] P. G. Subin and P. Muthukannan, "Optimized convolution neural network based multiple eye disease detection," *Computers in Biology and Medicine*, vol. 146, p. 105648, 2022. [Online]. Available: https://doi.org/10.1016/j.compbiomed.2022.105648

[16] D. Allen, "Cataract," *BMJ clinical evidence*, 2011, [Online; accessed 2025-07-20]. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3275311/

[17] J. Hu, L. Shen, S. Albanie, G. Sun, and E. Wu, "Squeeze-and-excitation networks," *arXiv preprint arXiv:1709.01507*, 2018, originally submitted to arXiv on September 5, 2017; revised May 16, 2019. [Online]. Available: https://doi.org/10.48550/arXiv.1709.01507

[18] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," *arXiv preprint arXiv:1708.02002*, 2018, originally submitted August 7, 2017; revised February 7, 2018. [Online]. Available: https://doi.org/10.48550/arXiv.1708.02002