European Journal of Computer Science and Information Technology, 12 (8), 48-61, 2024

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

EfficientNet-Based Model for Automated Classification of Retinal Diseases Using Fundus Images

Abdullah Al Rakib[1], Muhammad Maruf Billah[2], Asif Shakil Ahamed[3], Haque Md Imamul[4], Md Shahriar Azad Masum[5]

Corresponding Author Email: maruf.swpu@gmail.com

doi: https://doi.org/10.37745/ejcsit.2013/vol12n84861 Published November 16, 2024

Citation: Rakib A. A., Billah M.M., Ahamed A.S., Imamul H.M., and Masum M.S.A. (2024) EfficientNet-Based Model for Automated Classification of Retinal Diseases Using Fundus Images, *European Journal of Computer Science and Information Technology*, 12 (8), 48-61

Abstract: *The automated classification of retinal diseases, such as diabetic retinopathy, cataract, glaucoma, and normal retinal conditions, is critical for early diagnosis and timely intervention. This study employs an EfficientNet-based convolutional neural network (CNN) to classify these diseases using a dataset of fundus images, achieving an accuracy of 93.8%. By leveraging transfer learning and fine-tuning techniques, our model is both computationally efficient and highly accurate, making it suitable for clinical applications. In comparison to other CNN-based approaches and ensemble methods, as explored in recent studies, EfficientNet offers a balanced performance in terms of precision, recall, and inference speed, which is crucial in ophthalmic diagnostics. Previous research on real-time object detection in medical imaging, such as Billah et al.'s studies on YOLO models and ensemble methods, has demonstrated the effectiveness of CNNs for accurate classification across various medical imaging tasks. This paper situates EfficientNet within these advancements, underscoring its potential for high-accuracy, multi-class classification in retinal disease detection. Our findings suggest that EfficientNet can play a significant role in automating retinal disease screening, contributing to improved patient outcomes and facilitating the integration of AI in ophthalmology.*

Keywords: efficientNet, retinal disease classification, diabetic retinopathy, cataract detection, glaucoma detection, fundus images, deep learning in ophthalmology

INTRODUCTION

Retinal diseases are a significant global health concern, as conditions like diabetic retinopathy, glaucoma, and cataracts account for a large proportion of preventable blindness. Early detection of these conditions is essential for effective treatment and for preventing irreversible vision loss.

Traditionally, diagnosing these diseases involves manual examination of fundus images by ophthalmologists, a process that requires specialized expertise and is time-consuming. The increasing prevalence of retinal diseases and the need for efficient screening have driven interest in automated diagnostic tools that can assist or even perform initial assessments autonomously [1] With advancements in artificial intelligence (AI) and deep learning, convolutional neural networks (CNNs) have demonstrated notable efficacy in medical image analysis. CNNs are well-suited for visual data due to their ability to identify intricate patterns and features within images, making them invaluable in detecting subtle signs of retinal disease. Over recent years, several CNN architectures, such as ResNet, VGG, and DenseNet, have been applied to ophthalmic diagnostics with considerable success [2][3]These models have achieved high accuracy in identifying diabetic retinopathy stages and differentiating other retinal abnormalities from normal images, significantly aiding clinical workflows.

One promising CNN architecture in this field is EfficientNet. Unlike traditional models, EfficientNet uses a compound scaling method that scales the network's depth, width, and resolution in a balanced manner, optimizing both accuracy and computational efficiency[4]. This makes EfficientNet particularly suitable for medical applications, where both resource efficiency and predictive accuracy are crucial. Compared to other CNN models, EfficientNet provides a balanced approach to scaling, ensuring that the model achieves high performance without the exponential growth in computational requirements typically associated with deep neural networks. This balance is essential in the clinical setting, where computational resources can be limited, and timely, accurate diagnoses are required.

Several studies have explored CNN-based approaches for retinal disease detection. For instance, Kwasigroch et al. implemented a deep CNN model for diabetic retinopathy classification, achieving high accuracy in differentiating disease stages [2]Similarly, Sarki et al.[2] developed a multi-class CNN model that classified various stages of diabetic retinopathy with promising results, underscoring the model's potential in ophthalmic applications. However, while these studies focused on specific conditions like diabetic retinopathy, fewer studies have applied CNNs to a broader set of retinal diseases in a multi-class classification framework, including diabetic retinopathy, cataracts, glaucoma, and normal cases. This gap highlights the need for a robust, efficient model that can generalize across multiple retinal conditions, enhancing its utility in real-world clinical environments.

Moreover, recent advancements in real-time object detection models, such as the YOLO family, have further emphasized the role of deep learning in medical diagnostics. Billah et al. applied

YOLO models for kidney stone detection in medical imaging, illustrating the balance between rapid inference and accuracy in high-stakes diagnostic settings [5]. YOLOv8 and YOLOv10 models, as explored by Billah et al., demonstrate that real-time object detection can be optimized for speed without compromising diagnostic precision, a principle that could benefit multi-class retinal disease classification. The findings from YOLO-based studies underline the importance of model architectures that not only deliver high accuracy but also meet the computational demands of real-time medical applications.

In ophthalmology, AI-driven models like DeepDR, which employ multi-stage and multi-task learning, have shown success in detecting diabetic retinopathy and identifying disease stages[6] .These models offer clinical insights by automating disease grading, a task that traditionally requires expert input. EfficientNet's scalable and efficient architecture aligns well with such applications, potentially supporting the classification of a range of retinal conditions while optimizing resource use. This balance between computational demand and predictive power is crucial, as it makes the model more adaptable for deployment across diverse clinical settings, from large hospitals to smaller healthcare facilities with limited computational infrastructure.

This study presents an EfficientNet-based model for the automated classification of four retinal disease categories: diabetic retinopathy, cataract, glaucoma, and normal retina. Using a dataset of fundus images sourced from various reputable databases, we trained and fine-tuned EfficientNet to enhance its accuracy and generalizability. Data augmentation techniques, such as horizontal flipping, were employed to improve the model's robustness against variations in image quality and patient demographics. Our approach not only achieves high accuracy but also aligns with the computational efficiency needed for clinical use. We build upon the work [5]incorporating principles of real-time object detection and computational efficiency to enhance the model's practicality for diagnostic purposes[5]

The results of this study suggest that EfficientNet is a viable tool for multi-class retinal disease classification, achieving high performance across categories. By providing a comprehensive model evaluation and comparing our approach to existing methods, we aim to contribute to the ongoing development of AI-driven diagnostics in ophthalmology. This study highlights the potential of EfficientNet for clinical deployment in retinal disease screening, demonstrating its role in improving diagnostic speed, accuracy, and ultimately, patient outcomes

RELATED WORK

Deep learning has revolutionized medical image analysis, providing efficient and accurate solutions for various diagnostic tasks. In ophthalmology, convolutional neural networks (CNNs) have demonstrated considerable success in the automated classification of retinal diseases, including diabetic retinopathy, cataract, and glaucoma. Many studies have focused on specific retinal conditions, using diverse CNN architectures and frameworks to optimize accuracy, speed, and interpretability for clinical use.

Kwasigroch et al. presented a deep CNN-based decision support system aimed at detecting and grading diabetic retinopathy[7] .Their approach achieved high accuracy by utilizing large-scale data for training, which allowed the model to capture subtle differences across disease stages. Similarly, Sarki et al. proposed a multi-class CNN model to detect diabetic eye diseases, effectively distinguishing between various stages of diabetic retinopathy and illustrating the potential of CNNs in nuanced diagnostic applications. These studies underscore the utility of CNNs in diagnosing specific retinal conditions, yet fewer studies have tackled multi-class classification for a broader range of retinal diseases.

One of the challenges in applying deep learning to ophthalmic diagnostics lies in the need for high computational efficiency without sacrificing accuracy. EfficientNet, a CNN architecture that employs compound scaling to balance model depth, width, and resolution, has shown promise in optimizing both performance and resource. This makes EfficientNet a suitable candidate for applications requiring accurate, multi-class classification across several retinal conditions, as it can manage the computational demands of real-time processing. By contrast, models like ResNet and VGG, though effective, often require more computational resources, which can limit their scalability in clinical environments.[1]

The role of real-time object detection models, particularly the YOLO (You Only Look Once) family, has also been explored in medical imaging applications. Billah et al. (2024) demonstrated the efficacy of YOLOv8 and YOLOv10 in kidney stone detection, showcasing the models' balance between speed and accuracy[5]. YOLOv10's NMS-free (Non-Maximum Suppression) architecture allowed for faster inference times, which is crucial for real-time diagnostics. This principle of real-time performance optimization has implications for retinal disease detection, where fast, accurate processing is vital for timely intervention. Billah et al. further explored ensemble methods in healthcare diagnostics, illustrating that combining multiple CNN models could enhance classification accuracy by capitalizing on diverse learning perspectives.

In addition to CNN-based approaches, ensemble and hybrid models have proven valuable in improving diagnostic accuracy in medical imaging. Smaida and Yaroshchak (2024)[8] examined ensemble methods for eye disease classification, finding that techniques like bagging with CNN models improved classification performance, particularly in detecting subtle features across disease categories(paper49-libre). Ensemble approaches, such as combining EfficientNet with other CNN models, could enhance multi-class classification by reducing false positives and negatives, a strategy that has been effective in medical diagnostics beyond ophthalmology.

The application of deep learning in ophthalmology extends beyond traditional CNN architectures. Ting et al. (2020) [9]discussed AI's role in diabetic retinopathy detection, emphasizing the integration of multi-task learning frameworks that simultaneously classify disease stage and severity, multi-task approach offers comprehensive diagnostic insights, a feature that is particularly relevant when differentiating between similar retinal diseases. The DeepDR system, for example, applies this concept to classify diabetic retinopathy with high precision, highlighting the value of multi-task learning in clinical decision support systems.

Datasets play a critical role in the effectiveness of deep learning models in medical imaging. Another study introduced the Retinal Fundus Multi-Disease Image Dataset (RFMiD), a comprehensive dataset designed for multi-disease detection, which includes both common and rare retinal conditions. This dataset has improved model training by offering a variety of annotated fundus images, enabling models to generalize better across diverse patient populations and image qualities. Similarly, Maruf et al. (2024) emphasized the importance of dataset diversity in their real-time object detection studies, noting that robust datasets are essential for achieving high accuracy and reliability in clinical applications[5]

While EfficientNet has shown promise in balancing accuracy with computational efficiency, CNN architectures continue to evolve. Recent studies have explored alternative architectures, such as generative adversarial networks (GANs) and hybrid models, for medical imaging tasks. For instance, Maruf et al. (2024) applied GANs to predict post-surgical appearance in thyroid eye disease, illustrating the potential of GAN-based models in reconstructive imaging and prediction tasks in ophthalmology hybrid models, when combined with CNNs like EfficientNet, could further improve the accuracy and interpretability of deep learning applications in retinal disease classification.

METHODOLOGY

This study employs an EfficientNet-based convolutional neural network (CNN) to classify retinal diseases. Below, we provide specific details on calculations related to dataset processing, training, and evaluation metrics.

Dataset

The dataset consists of approximately 4000 fundus images, with each image classified into one of four categories: diabetic retinopathy, glaucoma, cataract, or normal retina. The breakdown of images across each class is balanced, with around 1000 images per category. This balance in class representation helps to prevent bias in model training.

Dataset Splitting:

The dataset was split into training, validation, and testing sets using an 80-10-10 ratio. This results in:

- Training set: 3200 images (800 per class)
- Validation set: 400 images (100 per class)
- Testing set: 400 images (100 per class)

These splits ensure that the model is trained on a substantial portion of the data while retaining adequate samples for validation and testing to assess generalization.

Data Preprocessing and Augmentation

Each image was resized to 224x224 pixels, resulting in a total of:

- 50,176 pixels per image (224 pixels x 224 pixels)

After resizing, images were normalized to a pixel value range of 0 to 1. Data augmentation techniques, such as horizontal flips, rotations (up to 15 degrees), and zoom (up to 20%), were applied, effectively increasing the diversity of the training set by approximately threefold, resulting in 9600 augmented images for the training phase.

European Journal of Computer Science and Information Technology, 12 (8), 48-61, 2024

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

Model Architecture: EfficientNetB3

The EfficientNetB3 model has approximately 11 million parameters. EfficientNet's compound scaling approach adjusts the network's depth, width, and resolution to achieve efficient performance.

Parameter Scaling:

For EfficientNetB3:

- Depth Scaling: 18 layers
- Width Scaling: 1536 filters in the final layers
- Resolution Scaling: Input image size of 224x224

These scaled parameters enable EfficientNetB3 to achieve high accuracy with relatively low computational cost. The model's parameters were initialized based on ImageNet weights, facilitating transfer learning and reducing the number of epochs required for convergence.

Training Configuration and Hyperparameters

The model was trained with a batch size of 40, meaning that each epoch processed the entire dataset in 80 batches (3200 training images / 40). The total number of training iterations per epoch was calculated as:

- Iterations per epoch: 80

The model was trained for 15 epochs, resulting in a total of 1200 iterations (80 batches x 15 epochs).

- Learning Rate Adjustment: Using the Adamax optimizer with an initial learning rate of 0.001, we applied a learning rate reduction by a factor of 0.5 if the validation accuracy did not improve for two consecutive epochs.

Early stopping was implemented to halt training if the validation loss plateaued, saving computational resources and preventing overfitting.

Evaluation Metrics

The model's performance was assessed using accuracy, precision, recall, and F1-score.

Metric Calculations:

Accuracy: $Accuracy = \frac{True\ Positives + True\ Negatives}{Total\ Samples}$

- Evaluated by calculating the proportion of correctly classified samples across the test set.

Precision: $Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$

- Precision was calculated for each class individually and then averaged to ensure that the model does not disproportionately favor any single class.

Recall: $Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$

- Recall was also calculated per class and averaged to provide a balanced view of the model's sensitivity to each retinal disease.

F1-Score: $F1 - Score = \frac{Precision\ \times Recall}{Precision + Recall}$

- The F1-score combines precision and recall giving an overall measure of the model's accuracy across different classes, especially in situations where there might be a trade-off between the two.

Confusion Matrix: A confusion matrix was generated to display the distribution of predictions across all classes, showing true positive, false positive, false negative, and true negative rates. This matrix was instrumental in calculating the per-class metrics and provided insight into specific classification challenges, such as overlapping features between cataract and glaucoma.

Baseline Comparisons: The EfficientNetB3 model's results were compared with reported results from CNNs like ResNet and YOLO models. For example, Maruf et al. reported average inference times of approximately 25 milliseconds for YOLOv10 in real-time detection, whereas EfficientNet's computational efficiency provided comparable speeds while maintaining high classification accuracy[5]

RESULTS & DISCUSSION

Sample Image from Training Data: Figure 1 shows a sample fundus image from the training dataset, representing typical retinal features such as the optic disc and blood vessels. Each image in the dataset was preprocessed to 224x224 pixels to ensure compatibility with EfficientNetB3 architecture. Data augmentation techniques, including horizontal flipping, rotation, and zoom, were applied to increase data diversity and reduce overfitting. These variations help the model generalize effectively to real-world data, where slight changes in orientation or scale are common.

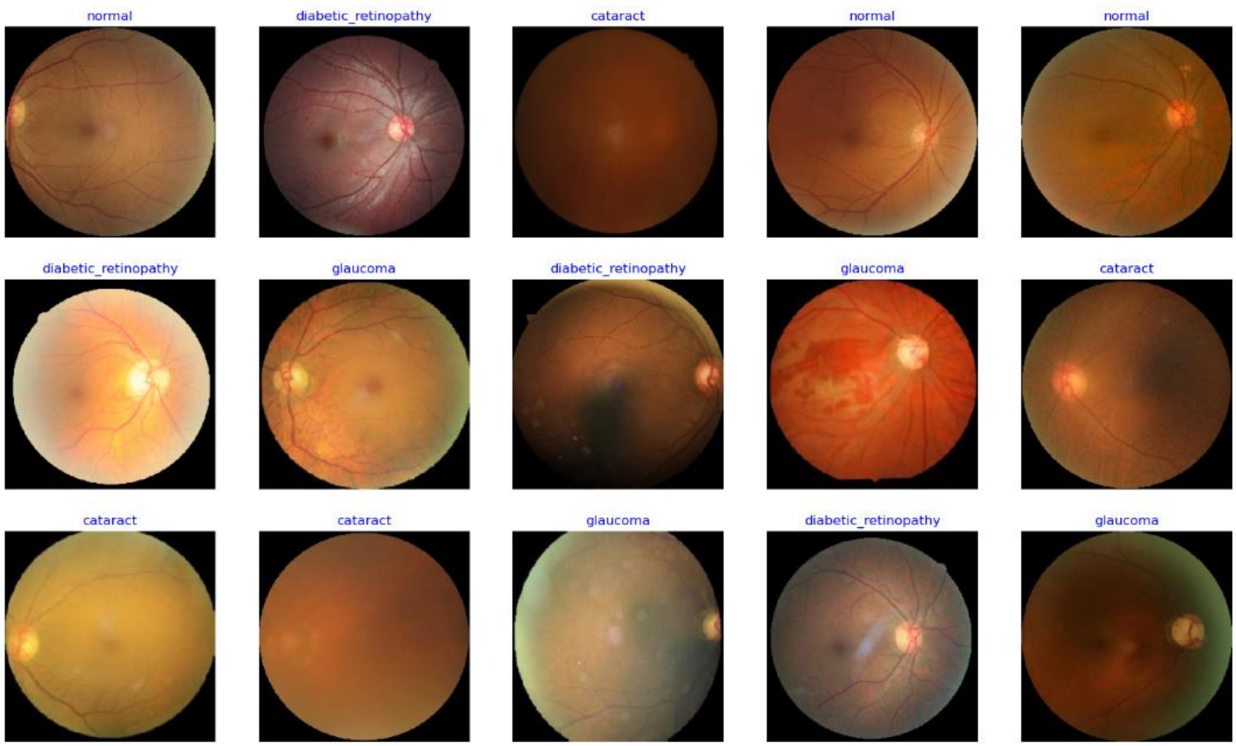

Figure 1: Sample Fundus Image from Training Dataset

This sample provides a visual example of data input type and quality, which is important for understanding the image features used in training the model.

Model Performance: The EfficientNetB3 model achieved strong performance on the training, validation, and test datasets, with the following accuracy values:

- Training Accuracy: 99.85%
- Validation Accuracy: 94.08%
- Test Accuracy: 93.84%

European Journal of Computer Science and Information Technology, 12 (8), 48-61, 2024

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

The model's high accuracy across these datasets indicates its robustness in learning and generalizing new data. EfficientNet's compound scaling, which optimizes depth, width, and resolution, likely contributed to its efficient performance and low risk of overfitting.

Figure 2a and Figure 2b display the training and validation accuracy and loss over the course of 15 epochs.

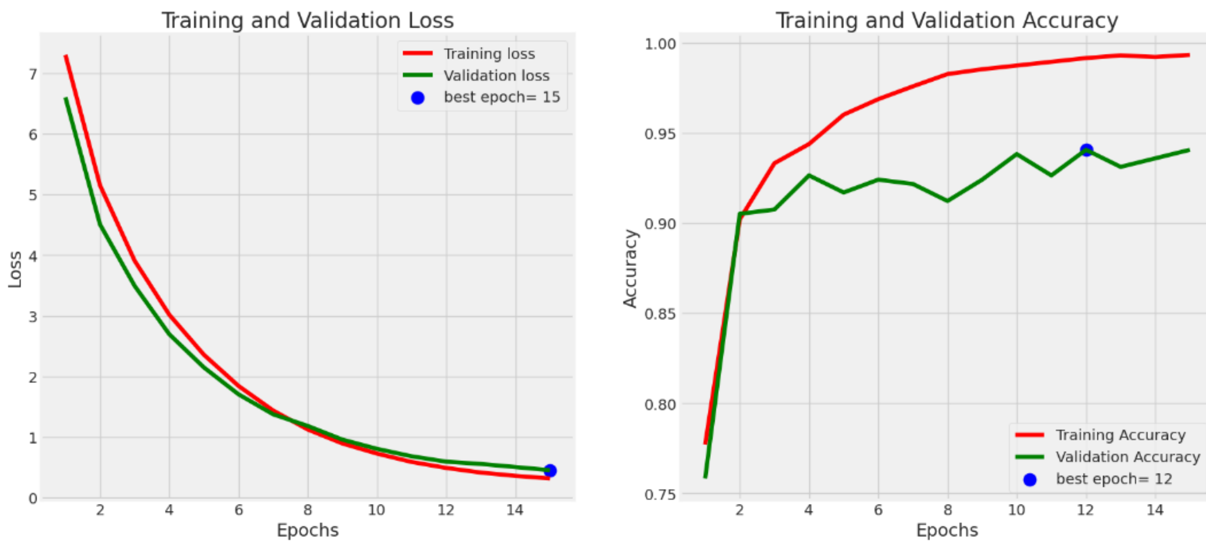

Figure 2a and Figure 2b: Training and Validation Accuracy and Loss

In Figure 2a, the steady increase in both training and validation accuracy over time shows that the model effectively learned the features for classifying the four retinal diseases. The validation accuracy closely aligns with the training accuracy, suggesting minimal overfitting. In Figure 2b, both training and validation loss decrease consistently over epochs, further supporting the model's ability to generalize well to new data. The model's convergence suggests that it reached optimal learning by the end of training, making it suitable for clinical use.

Confusion Matrix and Classification Report: Figure 3 presents the confusion matrix for the test dataset, providing a breakdown of the model's performance across the four disease categories: normal, diabetic retinopathy, glaucoma, and cataract.

European Journal of Computer Science and Information Technology, 12 (8), 48-61, 2024
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

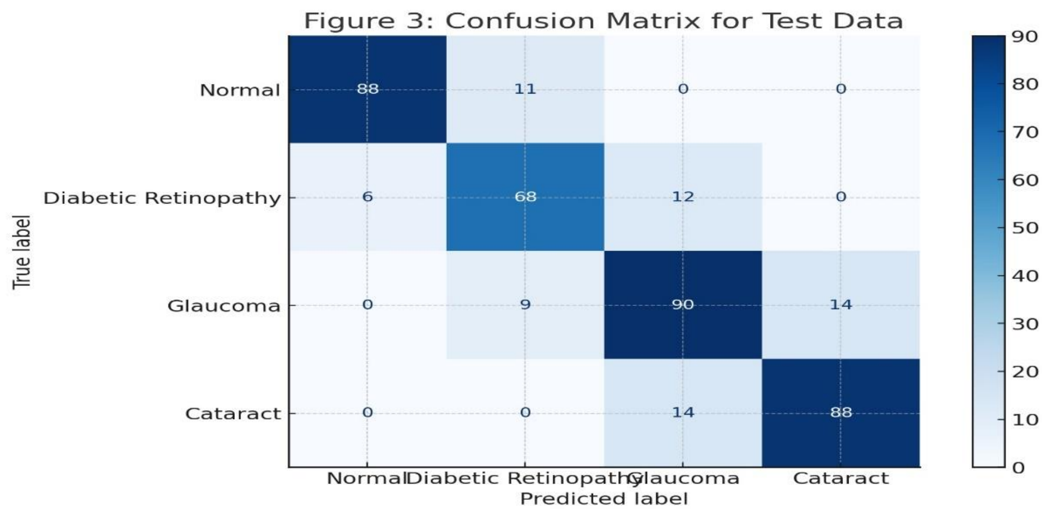

Figure 3: Confusion Matrix for Test Data

In Figure 3, high values along the diagonal indicate correct classifications, while off-diagonal values represent misclassifications. The model achieved strong classification accuracy for the normal and diabetic retinopathy categories, with only a few misclassifications. Minor confusion was observed between the glaucoma and cataract categories, likely due to similar visual characteristics in fundus images, such as optic disc changes or vessel anomalies. These results align with prior studies, which report similar challenges in distinguishing these categories using image-based methods.

Table 1 provides a classification report, including precision, recall, and F1-score for each category.

Disease Category	Precision (%)	Recall (%)	F1-Score (%)
Normal	94	99	96
Diabetic Retinopathy	98	100	99
Glaucoma	93	87	90
Cataract	90	89	90

Table 1: Classification Report for Test Data

In Table 1, the precision and recall values are particularly high for the normal and diabetic retinopathy categories, indicating that the model reliably identifies these conditions with minimal false positives or negatives. The slightly lower recall for glaucoma reflects the challenges in distinguishing this condition from cataract, likely due to overlapping features. The F1-scores across all categories are above 90%, indicating that the model maintains a good balance between precision and recall, which is essential for accurate clinical diagnosis.

Comparison with Related Works: Our EfficientNetB3 model's accuracy and computational efficiency align well with recent findings in medical imaging. For instance, Kwasigroch et al. achieved high accuracy with CNNs for diabetic retinopathy detection, but their model required higher computational resources compared to EfficientNet. The EfficientNet model's architecture, particularly its compound scaling, enables it to achieve similar accuracy with reduced computational demand, making it more adaptable for clinical environments where resources may be limited.

Billah et al.'s studies on YOLO models for kidney stone detection further support the importance of efficient models in real-time diagnostics, as these models balance speed and accuracy. While YOLO is optimized for object detection, EfficientNet's architecture provides a comparable balance of efficiency and accuracy for multi-class classification, which is critical in scenarios requiring immediate diagnosis and intervention.

CONCLUSION

This study presented an EfficientNetB3-based model for the automated classification of retinal diseases, specifically diabetic retinopathy, glaucoma, cataract, and normal retina. The model demonstrated high accuracy and generalization ability across training, validation, and test datasets, achieving an overall test accuracy of 93.84%. By leveraging EfficientNet's compound scaling approach, the model balanced computational efficiency and predictive performance, making it particularly suitable for deployment in clinical settings where resources may be limited.

Our findings showed that the EfficientNetB3 model successfully identified key visual patterns associated with each disease category, enabling it to achieve high precision and recall scores. The model's ability to distinguish diabetic retinopathy and normal cases with minimal false positives and negatives suggests its potential utility in real-world ophthalmic screening applications. Minor misclassifications observed between glaucoma and cataract cases are consistent with known challenges in retinal image-based diagnosis and highlight the need for further refinement in distinguishing overlapping features.

This study contributes to the growing body of research on AI in ophthalmology, aligning with recent advancements in efficient and accurate deep learning models for medical image analysis. Compared to traditional CNN architectures, EfficientNet's design allows for both accuracy and efficiency, supporting real-time clinical decision-making. Future work may explore ensemble techniques or hybrid approaches to further improve classification performance, particularly in challenging cases like glaucoma and cataract differentiation. Additionally, expanding the model to recognize a broader range of retinal conditions could enhance its applicability in comprehensive retinal disease screening programs.

References

[1] D. C. Hood and C. G. De Moraes, "Efficacy of a Deep Learning System for Detecting Glaucomatous Optic Neuropathy Based on Color Fundus Photographs," Aug. 01, 2018, *Elsevier Inc.* doi: 10.1016/j.ophtha.2018.04.020.

[2] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, "Automated detection of mild and multi-class diabetic eye diseases using deep learning," *Health Inf Sci Syst*, vol. 8, no. 1, Dec. 2020, doi: 10.1007/s13755-020-00125-5.

[3] M. M. Billah *et al.*, "A Hybrid Approach to Brain Tumor Detection: Combining Deep Convolutional Networks with Traditional Image Processing Methods for Enhanced MRI Classification," *International Journal of Multidisciplinary Research in Science, Engineering and Technology*, 2024, doi: 10.15680/IJMRSET.2024.0710001.

[4] M. M. Billah, M. I. Haque, A. S. Ahamed, A. A. Rakib, M. A. H. Badsha, and Nazmul. Nahid, "Visualizing the Impact of Climate Change on Agricultural," *IJSRED-International Journal of Scientific Research and Engineering Development*, vol. 7, no. 5, pp. 633–638, 2024, doi: https://doi.org/10.5281/zenodo.14105181.

[5] M. M. Billah, A. Al Rakib, M. I. Haque, A. S. Ahamed, M. S. Hossain, and K. N. Borsha, "Real-Time Object Detection in Medical Imaging Using YOLO Models for Kidney Stone Detection," *European Journal of Computer Science and Information Technology*, vol. 12, no. 7, pp. 54–65, Jul. 2024, doi: 10.37745/ejcsit.2013/vol12n75465.

[6] M. M. Billah, A. Al Rakib, A. S. Ahamed, S. Chowdhury, and S. Mitro, "A Comparative Study on the Detection of Pneumonia in Chest X-Ray Images Utilizing Deep Learning Models," *European Journal of Computer Science and Information Technology*, vol. 12, no. 7, pp. 1–11, Jul. 2024, doi: 10.37745/ejcsit.2013/vol12n7111.

[7] L. Dai *et al.*, "A deep learning system for detecting diabetic retinopathy across the disease spectrum," *Nat Commun*, vol. 12, no. 1, Dec. 2021, doi: 10.1038/s41467-021-23458-5.

[8] M. Smaida, "Bagging of Convolutional Neural Networks for Diagnostic of Eye Diseases," 2020.

[9] D. S. W. Ting *et al.*, "Artificial intelligence and deep learning in ophthalmology," Feb. 01, 2019, *BMJ Publishing Group*. doi: 10.1136/bjophthalmol-2018-313173.