



Eye Disease Detection and Classification Optimization Using EfficientNet-B5 with Emphasis on Data Augmentation and Fine-Tuning

Anggi Muhammad Rifai¹, Muhammad Fatchan², Ahmad Turmudi Zy³, Donny Maulana⁴, Sufajar Butsianto⁵
^{1,2,3,4,5}Department of informatics engineering, Faculty of engineering, Pelita Bangsa University, Bekasi, Indonesia

¹anggimuhammad@pelitabangsa.ac.id, ²fatchan@pelitabangsa.ac.id, ³turmudi@pelitabangsa.ac.id,

⁴donny.maulana@pelitabangsa.ac.id, ⁵sufajar@pelitabangsa.ac.id

Abstract

Eye diseases significant global health challenges, underscoring the need for efficient and accurate diagnostic. This study employed the EfficientNet-B5 model to enhance the detection and classification of eye diseases by incorporating advanced data augmentation and fine-tuning techniques. The research utilizes the ODIR dataset, consisting of 4,217 fundus images categorized into four classes: normal, glaucoma, cataract, and diabetic retinopathy. The methodology comprises three phases: baseline model training, model training with data augmentation, and fine-tuning. The baseline model achieved an accuracy of 60.43%, which improved to 63.03% with data augmentation an increase of 2.6 percentage points. Fine-tuning further elevated the accuracy to 93.23%, representing a notable improvement of 33.8 percentage points over the baseline. Model performance was evaluated using standard classification metrics. These findings demonstrate the technical efficacy of combining augmentation and fine-tuning to enhance model generalization. This result approach offers a robust framework for developing dependable AI-driven diagnostic tools to support early detection and facilitate informed clinical decision-making.

Keywords: Automated Diagnosis; Data Augmentation; EfficientNet-B5; Eye Disease Detection; Fine-tuning

How to Cite: Anggi Muhammad Rifai, Muhammad Fatchan, Ahmad Turmudi Zy, Donny Maulana, and Sufajar Butsianto, "Eye Disease Detection and Classification Optimization Using EfficientNet-B5 with Emphasis on Data Augmentation and Fine-Tuning", *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 5, pp. 1100 - 1108, Oct. 2025.

Permalink/DOI: <https://doi.org/10.29207/resti.v9i5.6519>

Received: April 6, 2025

Accepted: September 20, 2025

Available Online: October 24, 2025

This is an open-access article under the CC BY 4.0 License

Published by Ikatan Ahli Informatika Indonesia

1. Introduction

Eye diseases as glaucoma, cataracts, and diabetic retinopathy continue to pose significant global health challenges, impacting millions of people across the world [1], [2]. Without timely detection and appropriate treatment, these disorders can lead to serious visual impairment or permanent blindness [3]. Therefore, early diagnosis and precise classification play a crucial role in enabling effective medical intervention and preventing further vision deterioration [4]. However, conventional diagnostic approaches frequently depend on manual assessments conducted by medical professionals [5]. his process is often labor-intensive and resource-demanding, while also being vulnerable to subjective variability particularly in regions where access to qualified professionals is limited [6].

In recent years, rapid advancements in artificial intelligence (AI) and deep learning have reshaped the field of medical image analysis [7]. These technologies offer promising solutions of automation for detection

and classification eye diseases, enabling more accurate and consistent diagnoses [8]. Among various artificial intelligence techniques, deep learning models. Convolutional neural networks (CNN) excel at analyzing and interpreting complex retinal imagery [9]. Nevertheless, issues such as limited dataset diversity, overfitting, and inadequate model generalization continue to pose significant challenges, emphasizing the importance of further research aimed at improving the robustness and reliability of AI-driven diagnostic frameworks [10].

Numerous studies have comprehensively examined the application of deep learning models for ocular disease detection, producing encouraging outcomes. [11]. CNN like ResNet and InceptionNet have shown exceptional performance in classifying retinal images due to their ability to capture intricate visual patterns [12]. These models have been pivotal in advancing automated eye disease diagnosis, proving their effectiveness in analyzing complex medical image datasets [13].

However, high processing costs and fine-tuning complexity often preclude deployment in resource-scarce environments [14].

Among current deep learning architectures, the EfficientNet family has emerged as a state-of-the-art framework, maximizing accuracy while minimizing computational expense [15]. Using a compound scaling approach that jointly optimizes network depth, width, and resolution, EfficientNet models achieve exceptional adaptability and are especially well-suited for medical image analysis applications [16]. Despite their advantages, challenges such as overfitting, lack of dataset diversity, and suboptimal generalization still pose significant barriers [17]. These limitations underscore the necessity for advanced optimization techniques, such as data augmentation and fine-tuning, to enhance the reliability of these models in real-world clinical settings [18].

Previous research by [19] employed a system based on InceptionV3 for diabetic retinopathy detection, achieving an accuracy of 90.3%. The model demonstrated high sensitivity and specificity, comparable to ophthalmologists, yet required extensive labeled datasets for training.

Another research by [20] applied CNN for retinal disease classification, obtaining an accuracy of 89.81%, showing that models can achieve high performance with effective computational resources. Research by [21] implemented data augmentation on ResNet152, improving classification accuracy to 89.17%, demonstrating the importance of dataset diversity in mitigating overfitting.

This research utilizes the EfficientNet-B5 model, a scalable and computationally efficient deep learning architecture, to enhance eye disease detection and classification. Its balance between accuracy and

efficiency makes it well-suited for medical image analysis, particularly for retinal disease classification [22]. This research employs advanced augmentation with fine-tuning to strengthen generalization and reduce overfitting. By systematically evaluating these methods, this research provides new insights into their combined impact, contributing to the development of more robust AI-based diagnostic tools [6], [9], [10].

This study investigates how incorporating data augmentation and fine-tuning techniques can improve the classification performance of the EfficientNet-B5 model in detecting eye diseases. These optimization methods are developed to mitigate prevalent deep learning challenges such as overfitting and insufficient dataset diversity by improving the model's generalization to unseen data. Through harnessing the architectural strengths of EfficientNet-B5, this research seeks to demonstrate that targeted optimization approaches can lead to significant advancements in performance for medical image analysis tasks.

2. Methods

Figure 1 illustrates the research methodology, which consists of four key stages: (a) Data Collection, (b) Data Preprocessing, (c) Model, and (d) Model Evaluation. The dataset, collected for eye disease detection, is analyzed for class distribution and divided into training, validation, and testing. Preprocessing includes noise reduction, data augmentation, and normalization to enhance image quality. The architecture uses EfficientNet-B5 with pre-trained weights, followed by GlobalAveragePooling2D, a 512-unit ReLU dense head, and a Softmax layer for multi-class prediction. Training uses the Adam optimizer with fine-tuning and a 0.001 learning rate, while evaluation employs a Confusion Matrix, Classification Report, and Misclassification Analysis to measure model performance.

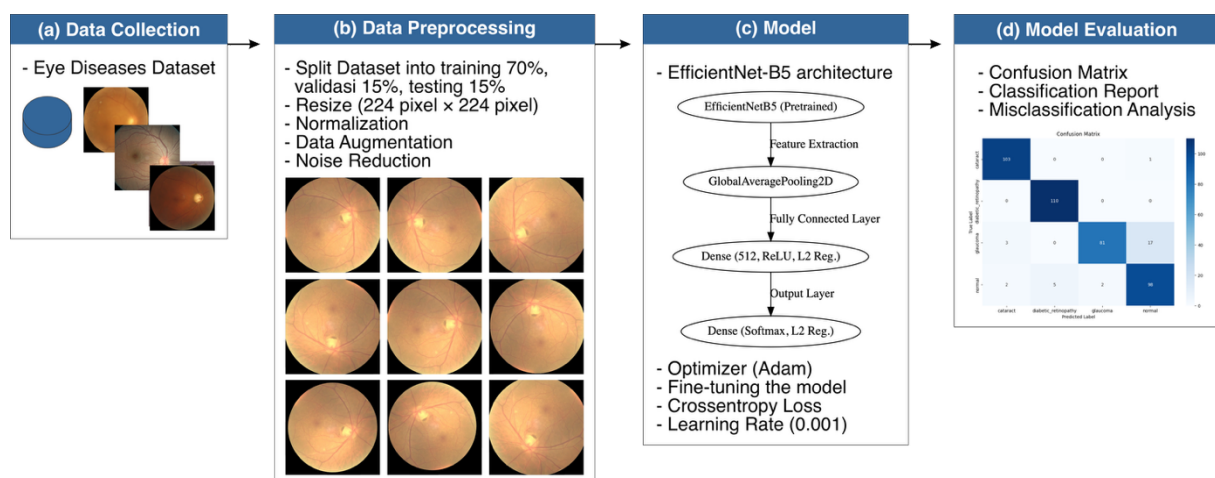


Figure 1. Research Flow

2.1 Data Collection

The Ocular Disease Intelligent Recognition (ODIR) dataset serves as a comprehensive ophthalmic resource

comprising medical records from approximately 5,000 patients. It includes color fundus image of both right and left eyes, accompanied by diagnostic annotations

provided by medical specialists. The dataset was compiled from various hospitals and medical institutions across China under the supervision of Shanggong Medical Technology Co., Ltd., ensuring that it reflects real-world clinical environments [23]. Fundus images were obtained using several commercial retinal imaging devices such as those produced by Canon, Zeiss, and Kowa resulting in variations in image resolution. To guarantee accurate labeling, a thorough annotation process was carried out by trained experts who classified the images into four diagnostic categories: Normal, Diabetic Retinopathy, Cataract, and Glaucoma, all under strict quality assurance protocols show in Figure 2.

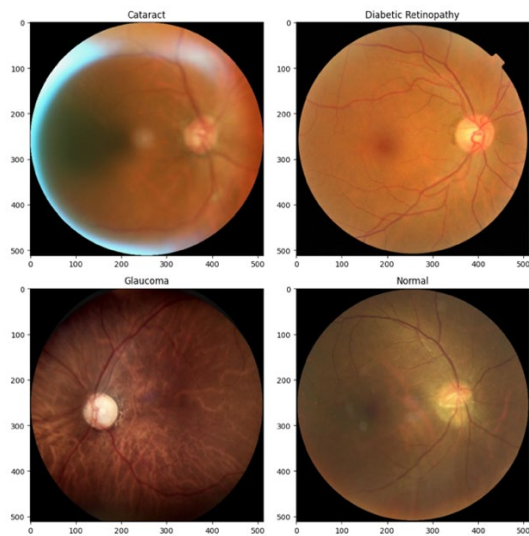


Figure 2. Class Dataset Collection

2.2 Data Preprocessing

Dataset is split into three class: training use 70%, validation use 15%, and test use 15%, each containing fundus images from four diagnostic categories: Normal, Glaucoma, Cataract, and Diabetic Retinopathy show in figure 3. The distribution of images across these categories is depicted in the bar charts, showcasing a relatively balanced dataset across all subsets.

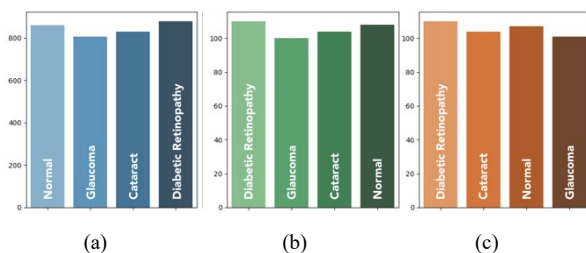


Figure 3. Split Dataset Collection. (a) Train Data, (b) Validation Data, (c) Test Data

In the training set, the number of images per category varies slightly, with Diabetic Retinopathy having the highest count and Glaucoma the lowest. The validation and test sets follow a similar distribution pattern, ensuring consistency in model evaluation. The overall dataset consists of 4,217 images, distributed as follows: Glaucoma (1,007 images), Diabetic Retinopathy (1,098

images), Cataract (1,038 images), and Normal (1,074 images). The uniform distribution across categories reduces bias during model training and enhances generalizability for disease classification.

Preprocessing data plays a vital role in maintaining the consistency, quality, and reliability of the dataset prior to model training. In this study, all fundus images were resized to a uniform resolution of 224×224 pixels to ensure sample consistency while retaining essential ophthalmic details. [24]. Given that the original images were acquired using various camera models with different resolutions, this resizing step was necessary to ensure compatibility with deep learning architectures. We normalized pixel intensities to a fixed range, improving numerical stability and speeding up training convergence [25]. To augment data diversity and mitigate the risk of overfitting, data augmentation techniques were employed, encompassing rotation (up to 30 degrees), horizontal and vertical flipping, and brightness adjustments within the range of 0.5 to 1.5 [6]. These augmentation strategies produced disciplined variations in the training data, enhancing generalization to previously unseen examples, as illustrated in Figure 4.

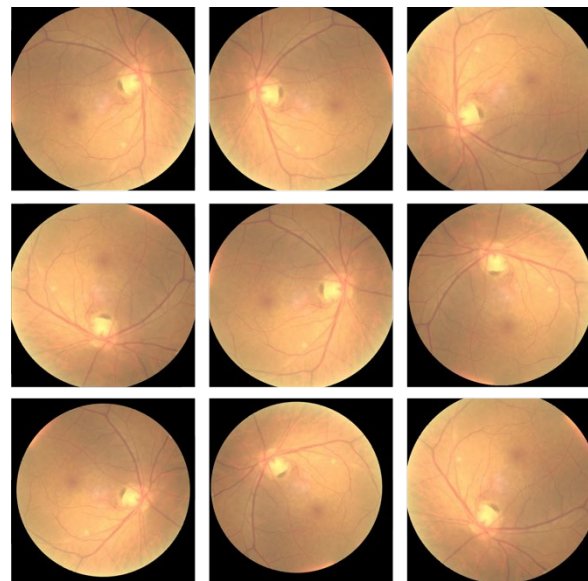


Figure 4. Data Augmentation

In addition to data augmentation, noise reduction techniques were incorporated to suppress unwanted artifacts and enhance the clarity of diagnostically relevant visual features, ensuring accurate disease classification. The preprocessing pipeline was implemented using ImageDataGenerator, where augmentation transformations were exclusively applied to the training dataset, whereas the validation and test datasets remained unaltered to preserve their integrity for unbiased evaluation.

The dataset was subsequently structured for multi-class classification, facilitating the differentiation between normal and pathological conditions such as diabetic retinopathy, glaucoma, and cataract. Furthermore, an

efficient batch-loading mechanism was utilized to optimize memory usage and ensure systematic data shuffling during training. Collectively, these preprocessing steps played a pivotal role in enhancing dataset quality, improving model generalization, and ensuring robust classification performance in ocular disease recognition.

2.3 Model

The proposed deep learning framework is based on the EfficientNet-B5 architecture, a pretrained CNN recognized for its excellent balance between computational efficiency and classification accuracy. This equilibrium is attained using a compound scaling technique that simultaneously optimizes the network depth, width, and input resolution, as illustrated in Figure 5.

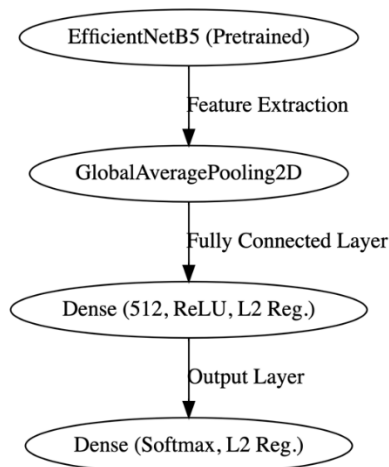


Figure 5. Model Architecture

EfficientNet-B5 was selected over other variants in the EfficientNet family due to its superior performance on medical imaging tasks and its ability to capture more complex visual features while maintaining a manageable model size and training time.

In this study, EfficientNet-B5 was fine-tuned using the ODIR dataset with specific hyperparameter configurations: a batch size of 32, a learning rate of 0.0001, a dropout rate of 0.5 to reduce overfitting, and early stopping applied during 60 training epochs based on validation loss split in without augmentation, augmentation, and fine-tuning. These parameters were optimized empirically to balance model performance and training stability.

The features are extracted, a GlobalAveragePooling2D layer is applied to transform the spatial feature maps into a compact feature vector, which helps reduce computational complexity while retaining essential discriminative information [26]. Mathematically, the global average pooling (GAP) operation can be expressed as Formula 1.

$$Z_k = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_{ijk} \quad (1)$$

Z_k is the output for the k -th feature map, x_{ijk} represents the activation at spatial location (i, j) in the k -th feature map, and $H \times W$ denotes the spatial dimensions of the feature map. This operation effectively computes the mean activation of each feature map, ensuring spatial invariance.

After pooling, the classification head applies a 512-unit Dense layer with ReLU activation [27]. Which is mathematically expressed as Formula 2.

$$f(x) = \max(0, x) \quad (2)$$

This activation function introduces non-linearity into the model, ensuring the learning of complex patterns [28]. Additionally, L2 regularization is applied to the fully connected layers to prevent overfitting. L2 regularization, also known as weight decay, penalizes large weight values and is given as formula 3.

$$L_{reg} = \lambda \sum_i w_i^2 \quad (3)$$

The weights are denoted by w , and λ specifies the regularization intensity. The last layer a Dense layer with Softmax produces the class probability outputs [29]. The Softmax function is defined as Formula 4.

$$P(y_i) = \frac{e^{z_i}}{\sum_j e^{z_j}} \quad (4)$$

$P(y_i)$ represents the probability of class i and z_i is the output score for class i . The Softmax function guarantees that the probabilities of all classes sum to one, making it well-suited for multi-class classification tasks.

By integrating an EfficientNet-B5 backbone for feature extraction with global average pooling for dimension reduction, a dense head for feature mapping, and a Softmax classifier, the model attains a strong accuracy–efficiency trade-off. Moreover, L2 regularization curbs overfitting and improves generalization to unseen data.

2.4 Data Evaluation

The model performance is evaluated using key assessment metrics, including the confusion matrix, classification report, and misclassification analysis. The confusion matrix serves as a core method for evaluating classification effectiveness, offering a structured overview of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) [6]. This matrix can be formally represented as Formula 5.

$$\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix} \quad (5)$$

Where TP indicates correctly classified positive instances, FP refers to incorrectly predicted positive samples, TN denotes correctly identified negative cases, and FN represents missed positive detections. From these values, several key performance metrics can be derived, including F1-score, precision, accuracy, and recall. Accuracy is computed as Formula 6.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Although accuracy is widely used, it often fails to fully represent model performance, particularly in scenarios involving class imbalance. Therefore, additional evaluation metrics like recall, precision, and F1-score are crucial to provide a more thorough and comprehensive performance assessment.

Precision or Positive Predictive Value is the fraction of correctly predicted positives among all positive predictions, as defined in Formula 7.

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

Recall also termed Sensitivity indicates the model ability to retrieve actual positives correctly show in Formula 8.

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

Providing a single trade-off metric, the F1-score is defined as the harmonic mean of precision and recall show in Formula 9.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

A comprehensive classification report lists these metrics per class, revealing how well the model separates categories. Furthermore, misclassification analysis is performed to investigate the specific cases where the model fails. This involves identifying

samples with high misclassification rates, analyzing patterns of errors, and determining whether they arise due to factors such as class overlap, insufficient training data, or misleading features. Such analysis is crucial for refining the model, adjusting decision thresholds, or implementing additional preprocessing techniques to enhance classification performance.

3. Results and Discussions

This section analyzes model performance through Model Training and Evaluation, Result Visualization, and Comparison Result. Training strategies, including baseline, augmentation, and fine-tuning, are evaluated based on accuracy and loss. Comparative analysis highlights improvements and limitations in accuracy, generalization, and efficiency, providing insights into the model's robustness for eye disease.

3.1 Model Training and Evaluation

In the initial training phase, EfficientNet-B5 was trained on the eye disease dataset without data augmentation for 20 epochs show in Figure 6. The model started with a training accuracy of 61.82% and validation 57.11% in the first epoch. Over the course of training, the training accuracy improved to 76.73%, but validation accuracy remained unstable, peaking at 63.27% in epoch 13 before declining to 58.29% in the final epoch. The increasing validation loss, from 1.09 in epoch 4 to 1.16 in epoch 20, indicated potential overfitting due to insufficient dataset variability.

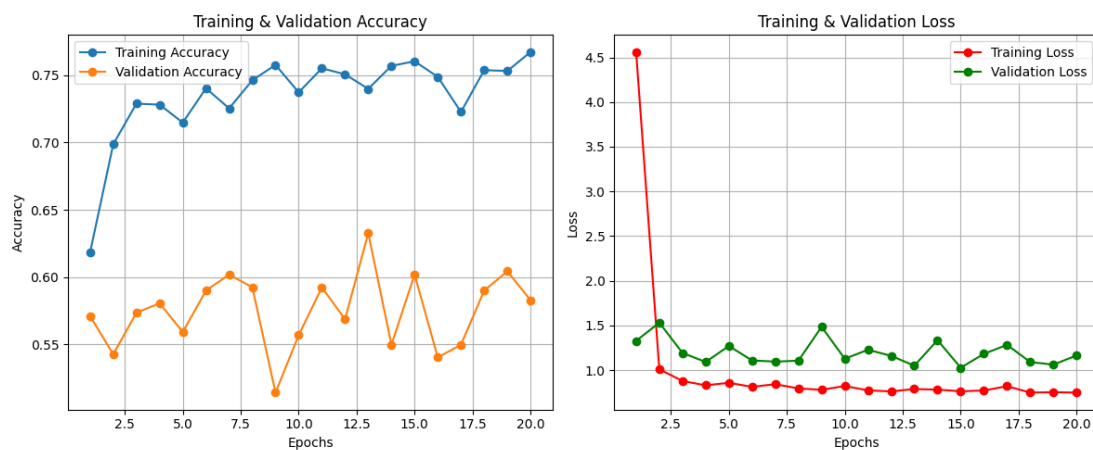


Figure 6. Result Without Augmentation

To improve generalization, data augmentation techniques such as random rotation, flipping, and zooming were applied, as illustrated in Figure 7. This resulted in a slight improvement in the initial accuracy (63.30% training, 56.40% validation) and led to more stable learning dynamics. By epoch 20, training accuracy reached 76.80%, and validation accuracy improved to 63.03%, with validation loss stabilizing at 1.03%. These results suggest that augmentation helped mitigate overfitting and improved pattern recognition, though the gains in validation accuracy were still limited, highlighting the need for further optimization.

Fine-tuning was then applied to allow deeper layers of EfficientNet-B5 to adapt to the specific characteristics of the dataset. This phase significantly increased computational demands, with training time rising to 1,491 seconds per epoch. The model demonstrated substantial performance improvements, achieving 67.57% accuracy in the first epoch. By epoch 5, training accuracy surged to 89.88%, while validation accuracy reached 91.15% in epoch 20. Validation loss also decreased significantly to 0.3425, indicating enhanced and robustness in distinguishing eye disease patterns.

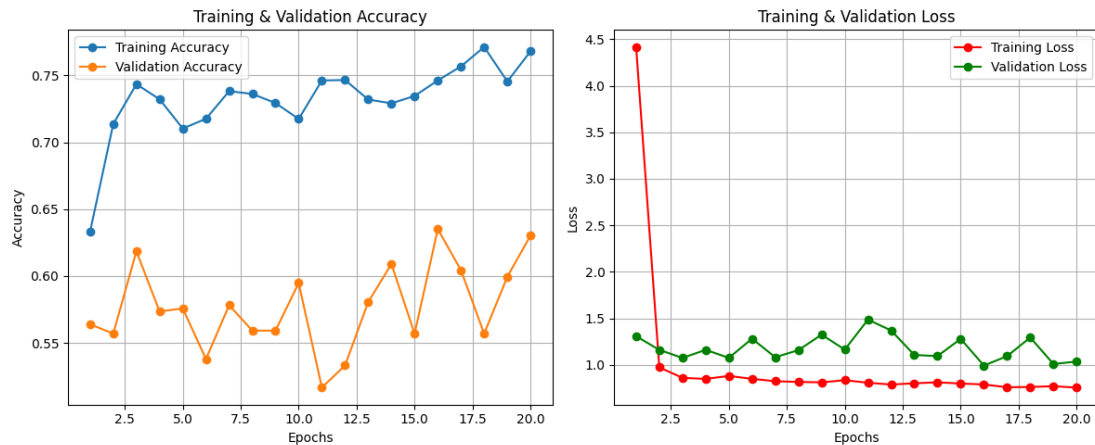


Figure 7. Result Augmentation

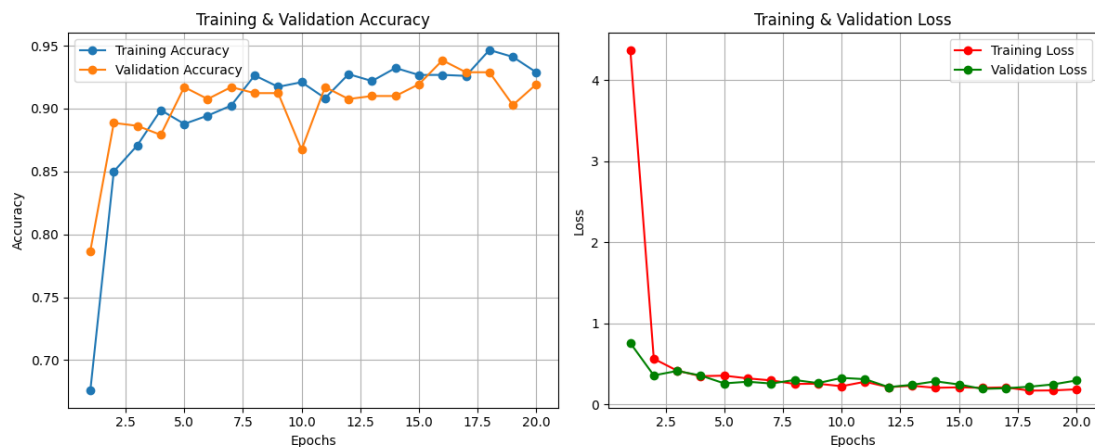


Figure 8. Result Fine-tuning

These results highlight the crucial role of fine-tuning in optimizing the performance of the EfficientNet-B5 model for eye disease classification, as depicted in Figure 8. While data augmentation contributed to better feature extraction and reduced overfitting, fine-tuning played a crucial role in leveraging pre-trained representations, ultimately leading to a more accurate and stable model.

Table 1. Result of Model

Metrics	Result
Accuracy	93.23%
Precision	93.0%
Recall	93.0%

Table 1 showcases the model's comprehensive performance following the fine-tuning process, highlighting notable enhancements in both accuracy and reliability. The model attained a high accuracy rate of 93.23%, signifying strong overall predictive power across various categories. Furthermore, the precision, recall, and F1-score each reached 93.0%, indicating consistent and balanced performance in correctly identifying positive cases while reducing misclassifications. These outcomes underscore the impact of fine-tuning in refining feature extraction and boosting the model's ability to generalize effectively in eye disease classification tasks.

3.2 Result Visualization

The Result Visualization section analyzes the confusion matrices from three training scenarios: without augmentation, with augmentation, and after fine-tuning shown in Figure 9. Without augmentation, the model struggled to differentiate between glaucoma and normal cases, misclassifying 30 glaucoma samples as diabetic retinopathy and correctly identifying only 21 normal cases. Cataract classification showed relatively good accuracy (72 correctly classified), while diabetic retinopathy had minor misclassifications, with 8 cases predicted as normal. These results indicate the model's limited generalization due to insufficient dataset variability.

Applying data augmentation improved classification performance, particularly for normal and glaucoma cases. Cataract classification remained stable (72 correct), while diabetic retinopathy had slightly more misclassification as normal (18 cases). However, glaucoma classification improved, reducing misclassification with diabetic retinopathy (from 30 to 21). Normal cases were identified with better accuracy (38 correct vs. 21 previously), reducing confusion with other categories. These improvements suggest that augmentation helped the model learn more robust features and mitigate overfitting.

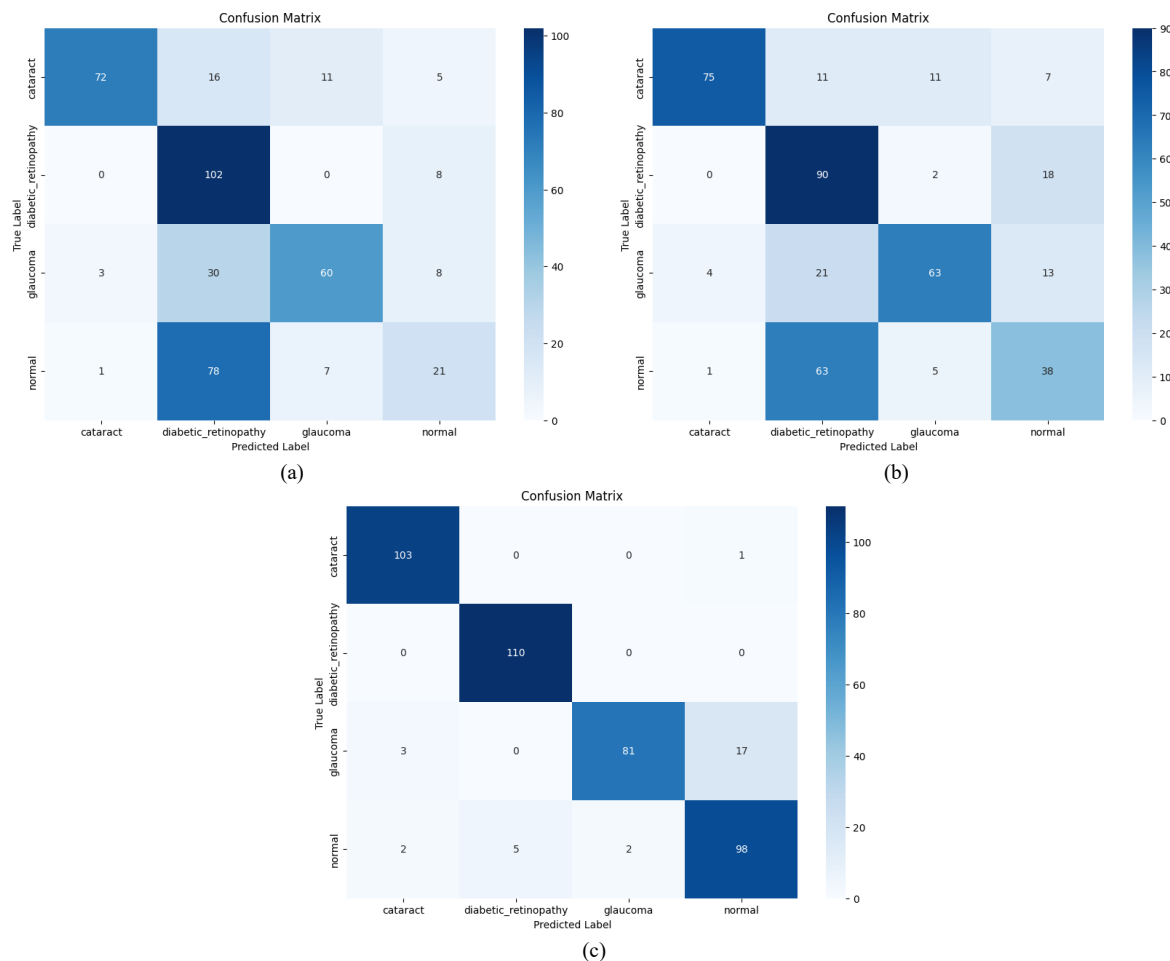


Figure 9. Confusion Matrix. (a) Without Augmentation, (b) Augmentation, (c) After Fine-tuning

Fine-tuning significantly enhanced classification accuracy across all categories. Cataract and diabetic retinopathy were nearly perfectly classified (103/104 and 110/110, respectively), while glaucoma improved to 81 correctly classified cases. Although some confusion remained between normal and glaucoma (17 misclassified cases), normal classification reached 98 correct predictions, showing strong generalization. These results indicate that fine-tuning successfully utilized the pre-trained features, resulting in a more accurate and robust model for eye disease classification.

Table 2. Classification Report

	Precision	Recall	F1-Score	Support
Cataract	95.0%	99.0%	97.0%	104
Diabetic Retinopathy	96.0%	100.0%	98.0%	110
Glaucoma	98.0%	80.0%	98.0%	101
Normal	84.0%	92.0%	88.0%	107

Diabetic retinopathy and cataract classes the highest effectiveness, with F1-scores of 98.0% and 97.0%, respectively, supported by both high precision (96.0% and 95.0%) and recall (100% and 99.0%), indicating excellent detection capability with minimal false negatives show in Table 2. The glaucoma class achieved a precision of 98.0%, showing few false positives; however, its recall was lower at 80.0%, suggesting that some glaucoma cases were missed. The

normal class achieved balanced performance with a precision of 84.0% and recall of 92.0%, resulting in an F1-score of 88.0%. Both the macro and weighted average F1-scores stood at 93.0%, reflecting a well-distributed model performance across all classes, particularly in detecting high-risk conditions such as diabetic retinopathy and cataract.

3.3 Comparison Result

To gain deeper insights into the performance outcomes, the proposed EfficientNet-B5 model enhanced through normalization, data augmentation, and noise reduction preprocessing is compared against several existing models. These existing approaches utilize different feature extraction techniques from those employed in this study, as seen in Table 3.

Table 3 provides a comparative analysis between the proposed method and several previous studies in the domain of eye disease classification. The work by [30], which utilized the InceptionV3 model with data augmentation but without fine-tuning, achieved an accuracy of 79.02% across four classes. Meanwhile, [21] focused on a binary classification task and employed the ResNet152 model with data augmentation and fine-tuning, reaching an accuracy of 89.17%.

Table 3. Comparison with previous work

Author	Classes	Pre-Processing	Model	Fine-tune	Accuracy
[30]	4	Data Augmentation	InceptionV3	×	79.02%
[21]	2	Data Augmentation	ResNet152	✓	89.17%
[20]	4	Converting images to arrays, Data Augmentation	CNN	✓	89.81%
Our Result	4	Normalization, Data Augmentation, Noise Reduction	EfficientNet-B5	✓	93.23%

Another study [20] used a custom CNN model with image-to-array conversion and data augmentation, along with fine-tuning, achieving an accuracy of 89.81% on four classes. In contrast, the proposed approach in this study combines normalization, data augmentation, and noise reduction as preprocessing techniques and integrates fine-tuning into the EfficientNet-B5 model. This method yielded the highest accuracy of 93.23%, demonstrating its superior performance and the effectiveness of the chosen preprocessing and fine-tuning strategies.

4. Conclusions

This study successfully demonstrates the effectiveness of the EfficientNet-B5 model in classifying eye diseases through a structured training approach involving baseline modeling, data augmentation, and fine-tuning. Starting from a baseline accuracy of 60.43%, the integration of data augmentation increased the performance to 63.03%, while fine-tuning significantly boosted accuracy to 93.23%. Compared to previous approaches using models such as InceptionV3, ResNet152, and conventional CNNs, the proposed method—enhanced with normalization, data augmentation, and noise reduction—consistently outperformed in terms of accuracy. These findings affirm the importance of fine-tuning and advanced preprocessing in improving model generalization for medical image classification tasks. This study is limited to a single dataset and controlled experimental environment. Future work should explore validation across multiple datasets and real-world clinical settings to assess generalizability and practical deployment potential of the proposed system.

Acknowledgements

This research was funded and supported by a research grant provided by the Department of Research and Community Service, Universitas Pelita Bangsa. The author declares no conflict of interest. This study did not receive any specific grant from government agencies, commercial entities, or non-profit organizations.

References

- [1] Z. L. Teo *et al.*, “Global Prevalence of Diabetic Retinopathy and Projection of Burden through 2045,” *Ophthalmology*, vol. 128, no. 11, pp. 1580–1591, Nov. 2021, doi: 10.1016/j.ophtha.2021.04.027.
- [2] G. Clare, J. H. Kempen, and C. Pavésio, “Infectious eye disease in the 21st century—an overview,” *Eye*, vol. 38, no. 11, pp. 2014–2027, Aug. 2024, doi: 10.1038/s41433-024-02966-w.
- [3] J. D. Steinmetz *et al.*, “Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study,” *Lancet Glob Health*, vol. 9, no. 2, pp. e144–e160, Feb. 2021, doi: 10.1016/S2214-109X(20)30489-7.
- [4] T. Upadhyay, R. Prasad, and S. Mathurkar, “A Narrative Review of the Advances in Screening Methods for Diabetic Retinopathy: Enhancing Early Detection and Vision Preservation,” *Cureus*, Feb. 2024, doi: 10.7759/cureus.53586.
- [5] R. Aggarwal *et al.*, “Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis,” *NPJ Digit Med*, vol. 4, no. 1, p. 65, Apr. 2021, doi: 10.1038/s41746-021-00438-z.
- [6] A. M. Rifai, S. Raharjo, E. Utami, and D. Ariatmanto, “Analysis for diagnosis of pneumonia symptoms using chest X-ray based on MobileNetV2 models with image enhancement using white balance and contrast limited adaptive histogram equalization (CLAHE),” *Biomed Signal Process Control*, vol. 90, p. 105857, Apr. 2024, doi: 10.1016/j.bspc.2023.105857.
- [7] M. Tsuneki, “Deep learning models in medical image analysis,” *J Oral Biosci*, vol. 64, no. 3, pp. 312–320, Sep. 2022, doi: 10.1016/j.job.2022.03.003.
- [8] A. R. Wahab Sait, “Artificial Intelligence-Driven Eye Disease Classification Model,” *Applied Sciences*, vol. 13, no. 20, p. 11437, Oct. 2023, doi: 10.3390/app132011437.
- [9] A. M. Rifa’i, A. T. Zy, W. Hadikristanto, Amali, and S. Butsianto, “Leveraging Data Augmentation and Dropout Layer in MobileNetV3 for Accurate Skin Cancer Detection ISIC Dataset,” in *2024 8th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, IEEE, Aug. 2024, pp. 591–596. doi: 10.1109/ICITISEE63424.2024.10730105.
- [10] A. M. Rifa’i, E. Utami, and D. Ariatmanto, “Analysis for Diagnosis of Pneumonia Symptoms Using Chest X-Ray Based on Resnet-50 Models With Different Epoch,” in *2022 6th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, IEEE, Dec. 2022, pp. 471–476. doi: 10.1109/ICITISEE57756.2022.10057805.
- [11] B. Goutam, M. F. Hashmi, Z. W. Geem, and N. D. Bokde, “A Comprehensive Review of Deep Learning Strategies in Retinal Disease Diagnosis Using Fundus Images,” *IEEE Access*, vol. 10, pp. 57796–57823, 2022, doi: 10.1109/ACCESS.2022.3178372.
- [12] R. Chandra Joshi, A. Kumar Sharma, and M. Kishore Dutta, “VisionDeep-AI: Deep learning-based retinal blood vessels segmentation and multi-class classification framework for eye diagnosis,” *Biomed Signal Process Control*, vol. 94, p. 106273, Aug. 2024, doi: 10.1016/j.bspc.2024.106273.
- [13] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, “Automatic Detection of Diabetic Eye Disease Through Deep Learning Using Fundus Images: A Survey,” *IEEE Access*, vol. 8, pp. 151133–151149, 2020, doi: 10.1109/ACCESS.2020.3015258.
- [14] G. Bai *et al.*, “Beyond Efficiency: A Systematic Survey of Resource-Efficient Large Language Models,” Dec. 2023.
- [15] Hossein. Abbasi, Ahmed. Alshaeib, Yasin. Orouskhani, Behrouz. Rahimi, and Mostafa. Shomalzadeh, “Advanced Deep Learning Models for Accurate Retinal Disease State Detection,” *International Journal of Information*

- Technology and Computer Science*, vol. 16, no. 3, pp. 61–71, Jun. 2024, doi: 10.5815/ijites.2024.03.06.
- [16] Md. A. Talukder, Md. A. Layek, M. Kazi, Md. A. Uddin, and S. Aryal, “Empowering COVID-19 detection: Optimizing performance through fine-tuned EfficientNet deep learning architecture,” *Comput Biol Med*, vol. 168, p. 107789, Jan. 2024, doi: 10.1016/j.combiomed.2023.107789.
- [17] K. Bian and R. Priyadarshi, “Machine Learning Optimization Techniques: A Survey, Classification, Challenges, and Future Research Issues,” *Archives of Computational Methods in Engineering*, Mar. 2024, doi: 10.1007/s11831-024-10110-w.
- [18] T. Islam, Md. S. Hafiz, J. R. Jim, Md. M. Kabir, and M. F. Mridha, “A systematic review of deep learning data augmentation in medical imaging: Recent advances and future research directions,” *Healthcare Analytics*, vol. 5, p. 100340, Jun. 2024, doi: 10.1016/j.health.2024.100340.
- [19] V. Gulshan *et al.*, “Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs,” *JAMA*, vol. 316, no. 22, p. 2402, Dec. 2016, doi: 10.1001/jama.2016.17216.
- [20] S. Ejaz, R. Baig, Z. Ashraf, M. M. Alnfai, M. M. Alnahari, and R. M. Alotaibi, “A deep learning framework for the early detection of multi-retinal diseases,” *PLoS One*, vol. 19, no. 7, p. e0307317, Jul. 2024, doi: 10.1371/journal.pone.0307317.
- [21] T. D. Nguyen, D.-T. Le, J. Bum, S. Kim, S. J. Song, and H. Choo, “Retinal Disease Diagnosis Using Deep Learning on Ultra-Wide-Field Fundus Images,” *Diagnostics*, vol. 14, no. 1, p. 105, Jan. 2024, doi: 10.3390/diagnostics14010105.
- [22] M. Li, Y. Jiang, Y. Zhang, and H. Zhu, “Medical image analysis using deep learning algorithms,” *Front Public Health*, vol. 11, Nov. 2023, doi: 10.3389/fpubh.2023.1273253.
- [23] S. D. Khan, S. Basalamah, and A. Lbath, “A novel deep learning framework for retinal disease detection leveraging contextual and local features cues from retinal images,” *Med Biol Eng Comput*, Feb. 2025, doi: 10.1007/s11517-025-03314-0.
- [24] V. Kumar and K. Paul, “Fundus Imaging-Based Healthcare: Present and Future,” *ACM Trans Comput Healthc*, vol. 4, no. 3, pp. 1–34, Jul. 2023, doi: 10.1145/3586580.
- [25] S. Wu *et al.*, “L1 - Norm Batch Normalization for Efficient Training of Deep Neural Networks,” *IEEE Trans Neural Netw Learn Syst*, vol. 30, no. 7, pp. 2043–2051, Jul. 2019, doi: 10.1109/TNNLS.2018.2876179.
- [26] S. Rani, G. Kasana, and S. Batra, “An efficient content based image retrieval framework using separable CNNs,” *Cluster Comput*, vol. 28, no. 1, p. 56, Feb. 2025, doi: 10.1007/s10586-024-04731-w.
- [27] G. Lin and W. Shen, “Research on convolutional neural network based on improved Relu piecewise activation function,” *Procedia Comput Sci*, vol. 131, pp. 977–984, 2018, doi: 10.1016/j.procs.2018.04.239.
- [28] M. Goyal, R. Goyal, P. Venkatappa Reddy, and B. Lall, “Activation Functions,” 2020, pp. 1–30. doi: 10.1007/978-3-030-31760-7_1.
- [29] V. Deepa, C. S. Kumar, and T. Cherian, “Ensemble of multi-stage deep convolutional neural networks for automated grading of diabetic retinopathy using image patches,” *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, pp. 6255–6265, Sep. 2022, doi: 10.1016/j.jksuci.2021.05.009.
- [30] P. U. Pandey *et al.*, “Ensemble of deep convolutional neural networks is more accurate and reliable than board-certified ophthalmologists at detecting multiple diseases in retinal fundus photographs,” *British Journal of Ophthalmology*, vol. 108, no. 3, pp. 417–423, Mar. 2024, doi: 10.1136/bjo-2022-322183.