

Evaluation of Data Augmentation Techniques for Rice Leaf Disease and Pest Classification Using EfficientNetV2B3

Afis Julianto¹, Miftahul Jannah²,

¹Teknik Informatika, Politeknik Negeri Bengkalis, Indonesia
afisjulianto@polbeng.ac.id¹, miftahuljannah@polbeng.ac.id²

Abstract

Rice is a primary staple crop in Indonesia that is highly susceptible to diseases and pests such as blast, brown spot, and planthopper infestations. Early detection supported by advanced technology is essential to strengthen national food security. This study aims to enhance the classification accuracy of rice leaf diseases and pests using the EfficientNetV2B3 architecture through the application of various data augmentation methods. Three approaches were employed: non-augmentation, traditional augmentation techniques such as rotation, shift, zoom, shear, and flip, as well as StyleGAN2-ADA-based augmentation. The model was trained on Google Colab Pro+ using an NVIDIA Tesla T4 GPU for 50 epochs. The results indicate that the non-augmentation method achieved a validation accuracy of 78.80%, while traditional augmentation improved it to 88.80%. The StyleGAN2-ADA approach yielded the best performance, with a validation accuracy of 97.58%, test accuracy of 98%, and an F1-score of 98%, albeit requiring longer computational time. GAN-based augmentation proved effective in enhancing the model's generalization capability and stability. The StyleGAN2-ADA approach was found to be optimal for improving the performance of EfficientNetV2B3 in classifying rice leaf diseases and pests.

Keywords : Classification, Data Augmentation, StyleGAN2-ADA, EfficientNetV2B3, Rice Leaf and Pest Diseases

1. INTRODUCTION

Rice (*Oryza sativa*) is a major staple food commodity in Indonesia and serves as a key pillar of national food security. However, rice productivity often declines due to diseases and pest infestations such as blast, brown spot, and planthoppers (Octaviana & Ekawati, 2022; Wagiyanti et al., 2024). Early detection of these diseases and pests is crucial to enable timely and effective preventive and control measures. Conventional methods that rely on manual observation by agricultural experts are often subjective, time-consuming, and inefficient on a large scale. Therefore, the utilization of deep learning technology, particularly Convolutional Neural Networks (CNNs), has emerged as a potential solution to facilitate rapid and accurate classification of rice leaf diseases and pest infestations.

CNN models have been proven effective in agricultural image classification. Among them, the EfficientNetV2B3 architecture has demonstrated strong performance in balancing computational efficiency and high accuracy in image (Liu et al., 2022; Tan & Le, 2021). This model is capable of accurately identifying various crop diseases. The success of deep learning models, however, largely depends on the availability of high-quality, diverse, and balanced image datasets. In real-world agricultural practices, limited data and inter-class imbalance often pose major challenges that hinder a model's generalization capability (Pushpa Athisaya Sakila Rani & Suresh Singh, 2025).

To address this issue, data augmentation has become an essential approach for expanding and enriching datasets without the need for costly and time-consuming new data collection. Traditional augmentation techniques—such as rotation, zoom, shear range, flipping, and shifting—are widely used to increase data diversity and reduce the risk of overfitting in plant classification models (Ayuni et al., 2023). Nevertheless, traditional augmentation methods have limitations in generating image variations and complexities that truly reflect natural conditions in the field, such as variations in lighting, texture, and leaf morphology caused by specific diseases or pests.

To overcome these limitations, modern approaches based on Generative Adversarial Networks (GANs) have been increasingly applied in agricultural research. StyleGAN2-ADA, in particular, has shown the ability to generate high-quality synthetic images with realistic

texture and lighting variations, even under conditions of limited real data (Woodland et al., 2022). A study by (Chen et al., 2024) demonstrated that augmentation using StyleGAN2-ADA improved the classification accuracy of cotton weed images by up to 2.30% compared to traditional augmentation methods.

This study to evaluate the effectiveness of various data augmentation techniques, both traditional and GAN-based (StyleGAN2-ADA), in improving the classification accuracy of rice leaf diseases and pest infestations using the EfficientNetV2B3 architecture. The EfficientNetV2B3 model was selected for its ability to achieve a balance between computational efficiency and high accuracy on resource-constrained devices. The findings of this study are expected to provide insights into the most optimal augmentation strategy to support the development of more accurate and adaptive deep learning-based plant disease detection systems in real agricultural environments.

2. REVIEW OF LITERATURE

This section discusses the literature review relevant to the research topic to provide a theoretical foundation and an understanding of previous studies. The review encompasses prior research related to the main aspects of this study, thereby establishing the position and contribution of the present research within the existing body of knowledge.

2.1. Data Augmentation

Data augmentation is an essential technique for improving the performance of image classification models, particularly when the original dataset is limited or imbalanced. Traditional augmentation techniques such as rotation, flipping, zooming, and lighting adjustments are commonly employed to increase the diversity of training data without the need for additional data collection. Moreover, generative augmentation methods, such as those utilizing Generative Adversarial Networks (GANs), have recently been applied to generate realistic synthetic images that can significantly enhance both the quality and quantity of training data. The study by (Agustin et al., 2025) demonstrated that combining classical augmentation with DCGAN improved CNN accuracy for rice disease classification to over 98% compared to the original dataset.

2.2. Traditional Augmentation

Traditional augmentation is a method of expanding training data through simple techniques such as rotation, flipping, zooming, and lighting adjustments. This approach is commonly used in rice disease and pest classification to enrich data variability without the need for new image acquisition (Agustina et al., 2024). Traditional augmentation enables deep learning models to learn more robust feature representations and reduces the risk of overfitting, thereby significantly improving classification accuracy (Kusuma et al., 2025).

A study (Sapitri et al., 2023) investigated the impact of data augmentation techniques on the ability of deep learning models to recognize Komering script characters. Three augmentation techniques random rotation, height shift, and width shift were employed to increase the diversity of a limited dataset. The results showed that the CNN model was still able to recognize the augmented images with an average accuracy of 80.05%, precision of 98%, and F1-score of 88%. Among these methods, random rotation provided the most stable performance compared to vertical and horizontal shifts, demonstrating the effectiveness of augmentation in enhancing the generalization capability of character recognition models.

2.3. StyleGAN2-ADA

StyleGAN2-ADA is a generative model developed from StyleGAN2 with the addition of adaptive discriminator augmentation (ADA) to improve training stability when working with limited datasets (Karras et al., 2020). This model is capable of producing high-quality

synthetic images that closely resemble real data, making it an effective data augmentation technique across various domains, including medical imaging and agriculture.

A study (Joo et al., 2024) demonstrated that the use of StyleGAN2-ADA for data augmentation significantly enhances classification performance under data-scarce conditions by reducing overfitting and improving the accuracy of CNN models such as ResNet50 and ResNeXt101. Furthermore, this model can generate synthetic images that are both visually and statistically similar to real data, as verified using the Frechet Inception Distance (FID) metric and expert evaluation. Other studies have also shown that this technique excels in producing high-quality medical images without requiring complex hyperparameter tuning, making it suitable for data augmentation in plant disease classification tasks, which typically involve limited datasets (Woodland et al., 2022).

Similarly (Sahib et al., 2025) proposed the use of StyleGAN2-ADA to enhance the classification accuracy of medicinal plant species in small datasets through synthetic image augmentation. The dataset consisted of 438 images from 10 species of Iraqi medicinal plants and was evaluated using six transfer learning models, including BiT_Sr50x1, ResNet152, and MobileNetV2. The results indicated an improvement in F1-score ranging from 3.56% to 19.42% compared to conventional augmentation methods, with the highest performance achieved by the BiT_Sr50x1 model at 96.47%. This study confirms the effectiveness of GAN-based augmentation in expanding data diversity and improving the generalization capability of plant classification models.

2.4. EfficientNetV2B3

EfficientNetV2 adalah arsitektur CNN yang terkenal dengan efisiensi parameter dan kecepatan pelatihan yang tinggi (Tan & Le, 2021). EfficientNetV2B3 merupakan versi pengembangan dari EfficientNet yang dirancang untuk memberikan keseimbangan antara akurasi tinggi dan efisiensi komputasi. Model ini menggunakan compound scaling yang mengoptimalkan kedalaman, lebar, dan resolusi input secara bersamaan yang memungkinkan penggunaan sumber daya yang lebih efisien. Berbeda dengan generasi sebelumnya, EfficientNetV2B3 mengadopsi teknik Fused-MBConv blocks yang mempercepat proses pelatihan dan inferensi (Liu et al., 2022). Arsitektur ini menunjukkan performa superior dalam klasifikasi berbagai dataset citra seperti medis dan pertanian.

EfficientNetV2B3 berhasil meningkatkan keakuratan klasifikasi citra medis dan identifikasi penyakit dengan transfer learning dari dataset ImageNet yang besar sehingga mampu mengekstrak fitur penting secara lebih optimal (Ibrahim et al., 2024). Penelitian lainnya mengusulkan model hybrid EfficientNetV2B3–Vision Transformer (ViT) untuk deteksi penyakit tanaman kentang yang memadukan ekstraksi fitur CNN dan ViT secara paralel. Model ini mencapai akurasi 85,06% meningkat 11,43% dibanding studi sebelumnya, serta terbukti efektif dan andal untuk deteksi penyakit tanaman pada kondisi lapangan nyata (Sinamenye et al., 2025). Penerapan EfficientNetV2B3 secara khusus pada klasifikasi penyakit daun dan hama pada tanaman menunjukkan hasil yang menjanjikan dengan peningkatan akurasi dan efisiensi pelatihan.

3. METHOD

This section describes the research workflow for evaluating data augmentation techniques in the classification of rice leaf diseases and pests, starting from dataset acquisition, data augmentation, data splitting, classification using the EfficientNetV2B3 architecture, and result evaluation, as illustrated in Figure 1.

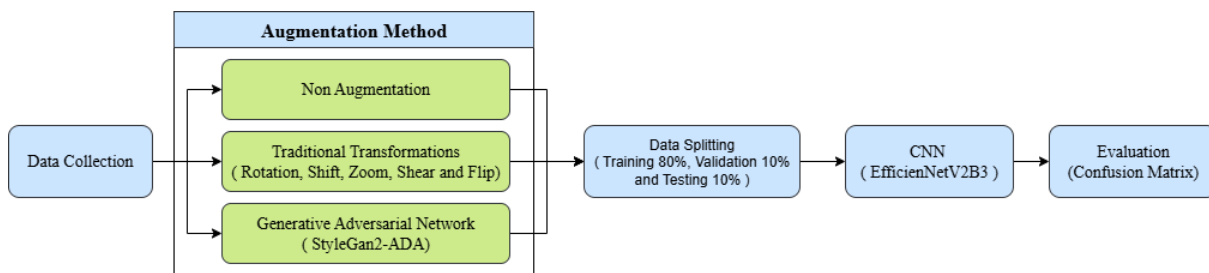


Fig. 1. Research Flow

3.1. Dataset Collection

In this study, the dataset was obtained from Kaggle and Mendeley Data, consisting of images of rice leaf diseases and pests with a total of 13 classes. The rice leaf disease dataset contains 1,105 images comprising four disease classes and one healthy class, as illustrated in Figure 2. Meanwhile, the rice pest dataset consists of 2,680 images representing eight classes, as shown in Figure 3. All image data used in this study were stored in JPG format.



Fig. 2. Rice Leaf Disease

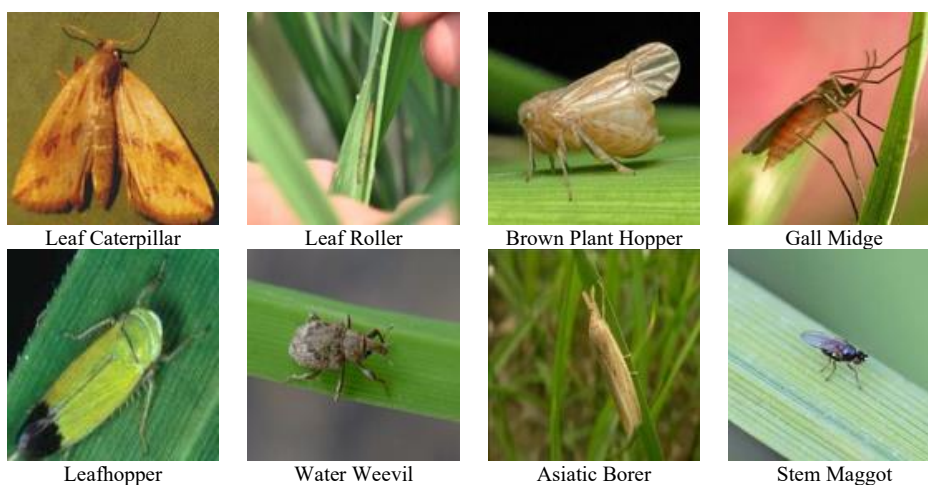


Fig. 3. Rice Pests

3.2. Data Augmentation Method

The image dataset of rice leaf diseases and pests was divided into three subsets: 80% for training, 10% for validation, and 10% for testing. Each image was then resized to 128×128 pixels to ensure uniformity during the model training process. Traditional data augmentation techniques were applied to the training set to increase both the diversity and the quantity of training samples. The applied augmentation techniques included a rotation range of 20, width shift range of 0.1, height shift range of 0.1, zoom range of 0.1, shear range of 0.1, and horizontal flip set to True. Through this process, the number of training images increased significantly compared to the original dataset.

In addition, a StyleGAN2-ADA-based augmentation method was employed to generate new synthetic images while preserving the original visual characteristics of rice

leaves. This augmentation process produced 2,000 images per class, which were further divided into 1,600 for training, 200 for validation, and 200 for testing, as presented in Table 1.

Table 1. Details of Rice Leaf Disease Dataset and Data Augmentation Results

Rice Leaf Disease and Pest	Original Image			Traditional Augmentation			StyleGan2-ada Augmentation		
	Train (80%)	Val (10%)	Test (10%)	Train (80%)	Val (10%)	Test (10%)	Train (80%)	Val (10%)	Test (10%)
Healthy	480	60	60	2400	60	60	1600	200	200
Blast	64	8	8	320	8	8	1600	200	200
Blight	64	8	8	320	8	8	1600	200	200
Tungro	64	8	8	320	8	8	1600	200	200
Stripes	212	26	27	1060	26	27	1600	200	200
Asiatic Borer	400	50	50	2000	50	50	1600	200	200
Brown Plant Hopper	276	35	34	1380	35	34	1600	200	200
Gall Midge	174	22	21	870	22	21	1600	200	200
Leaf Caterpillar	122	15	16	610	15	16	1600	200	200
Leaf Roller	573	72	71	2865	72	71	1600	200	200
Leafhopper	195	24	25	975	24	25	1600	200	200
Stem Maggot	72	9	9	360	9	9	1600	200	200
Water Weevil	332	42	41	1660	42	41	1600	200	200
Total	3785			12869			26000		

3.3. EfficientNetV2B3 Architecture

EfficientNetV2 is a convolutional neural network (CNN) architecture well known for its parameter efficiency and high training speed (Tan & Le, 2021). EfficientNetV2B3 is an enhanced version of the EfficientNet family, designed to achieve a balance between high accuracy and computational efficiency. The model employs *compound scaling*, which simultaneously optimizes network depth, width, and input resolution, enabling more efficient use of computational resources. Unlike its predecessor, EfficientNetV2B3 adopts *Fused-MBConv blocks*, which accelerate both training and inference processes (Ibrahim et al., 2024; Liu et al., 2022). The overall architecture of EfficientNetV2B3 is illustrated in Figure 4.

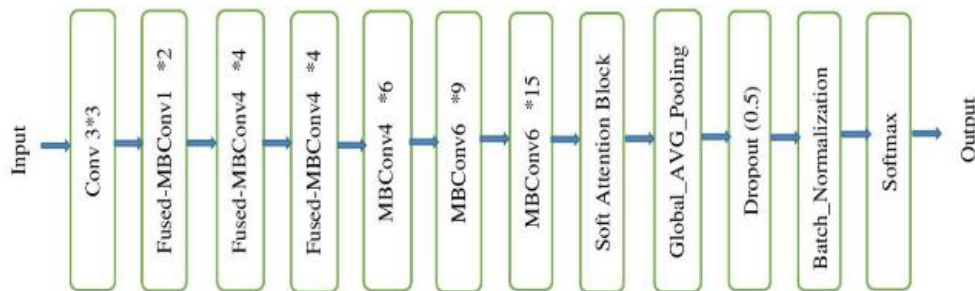


Fig. 4. Architecture EfficientNetV2B3 (Ibrahim et al., 2024)

3.4. Evaluation Metrics

The evaluation metrics in this study include accuracy, precision, recall, and F1-score to assess the performance of the multiclass classification model. The confusion matrix is an evaluation tool used to measure classification performance by displaying the number of correct and incorrect predictions categorized by actual and predicted classes. In multiclass cases, the confusion matrix takes the form of an $n \times n$ matrix, where n represents the number of classes. Each matrix element x_{ij} represents the number of samples from the actual class i that were predicted as class j . The accuracy (A), precision (P), recall (R), and F1-score (F1) for each class i are calculated using Equations (1)–(4) (Manliguez, 2016):

$$A_i = \frac{TTP_{all}}{(Total\ Number\ of\ Entries)} \quad (1)$$

$$P_i = \frac{(TP_{all})}{(TP_{all} + TF_{Pi})} \quad (2)$$

$$R_i = \frac{(TP_{all})}{(TP_{all} + TF_{Ni})} \quad (3)$$

$$F_i = \frac{2(P_i)(R_i)}{P_i + R_i} \quad (4)$$

4. RESULT & DISCUSSION

Processes in this study were conducted using Google Colab Pro+, equipped with an NVIDIA Tesla T4 GPU accelerator, 25 GB of RAM, and a 166 GB runtime disk. The model training process was efficiently executed in a cloud-based environment with sufficient computational resources to accelerate processing. Utilizing the T4 GPU provided a stable platform capable of handling high computational workloads. The complete results of the experiments conducted are presented below.

Table 2. Accuracy, loss and time computing training and validation of EfficienNetV2B3 after 50 epochs

Augmentation Methods	Train Acc (%)	Val Acc (%)	Train Loss	Val Loss	Time (Minutes)
Non Augmentation	76,56	78,80	0,65	0,78	55
Traditional Augmentation	94,02	88,80	0,17	0,41	211
StyleGan2-Ada	96,32	97,58	0,10	0,08	380

Table 2 presents the evaluation results of the EfficientNetV2B3 model during training and validation using three different augmentation approaches: Non-Augmentation, Traditional Augmentation, and StyleGAN2-ADA-based Augmentation. The analyzed parameters include training accuracy, validation accuracy, training loss, validation loss, and computation time over 50 epochs. In the Non-Augmentation method, the model achieved a training accuracy of 76.56% and a validation accuracy of 78.80%, with a training loss of 0.65 and a validation loss of 0.78. The relatively low accuracy and high loss values indicate that the model struggled to learn meaningful patterns from the limited dataset, resulting in weak generalization capability.

In contrast, the Traditional Augmentation method significantly improved model performance, achieving a training accuracy of 94.02% and a validation accuracy of 88.80%, with the validation loss reduced to 0.41. Augmentation techniques such as rotation, shifting, zooming, shearing, and flipping effectively increased data diversity, enabling the model to recognize visual patterns more robustly. However, this improvement came with a longer computation time of 211 minutes.

The StyleGAN2-ADA method yielded the most optimal results, with a training accuracy of 96.32% and a validation accuracy of 97.58%, accompanied by a training loss of 0.10 and a validation loss of 0.08. This performance improvement can be attributed to the capability of Generative Adversarial Networks (GANs) to produce realistic and diverse synthetic data, significantly expanding the training data distribution. Although it required the longest computation time 380 minutes the results demonstrate that StyleGAN2-ADA increased validation accuracy by nearly 9% compared to the non-augmented approach. These findings confirm that GAN-based augmentation is effective in enhancing the model's generalization ability without causing overfitting.

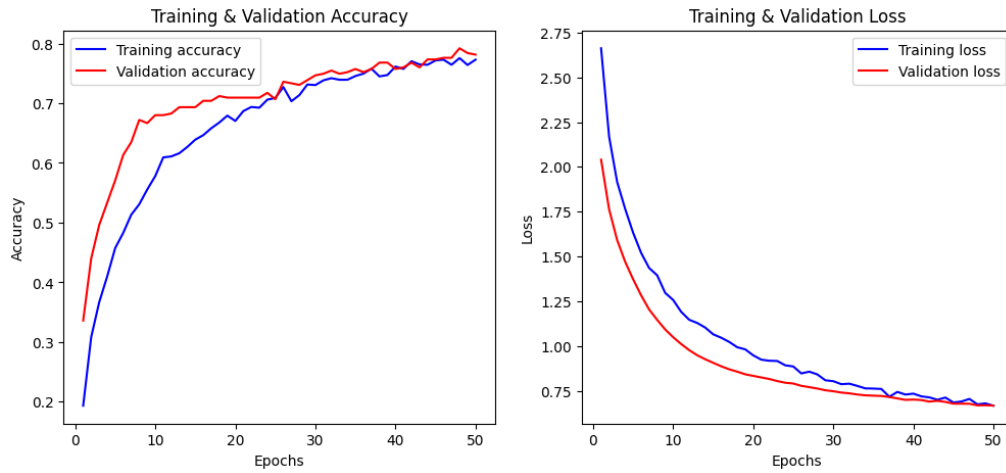


Fig. 5. Accuracy and Loss Training EfficientNetV2B3 with Non Augmentation

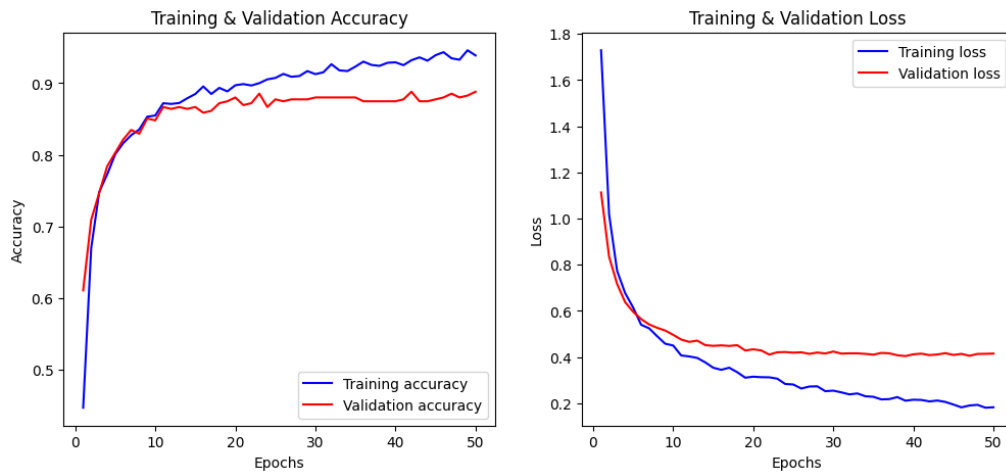


Fig. 6. Accuracy and Loss Training EfficientNetV2B3 with Traditional Augmentation

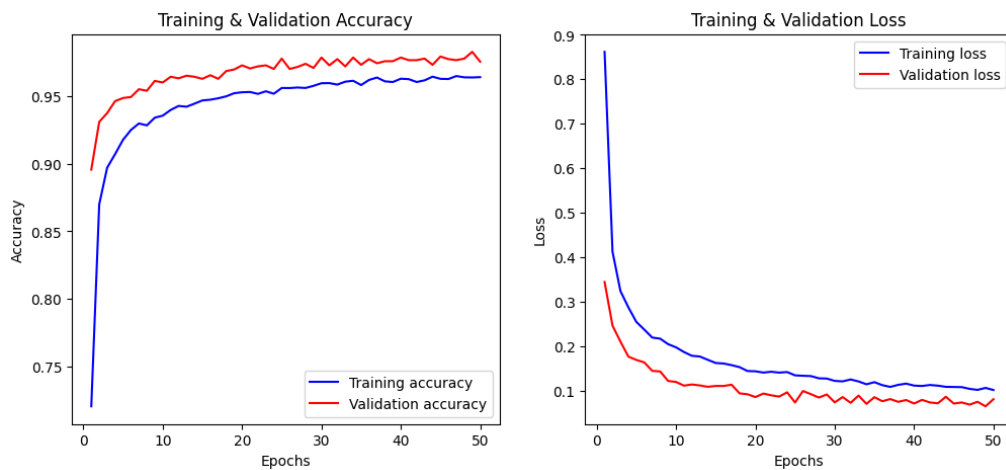


Fig. 7. Accuracy and Loss Training EfficientNetV2B3 with StyleGAN2-ADA Augmentation

Table 3. Evaluation of model training with data testing

Augmentation Methods	Test Acc (%)	Precision (%)	Recall (%)	F1 Score (%)
Non Augmentation	82,00	83,00	79,00	80,00
Traditional Augmentation	89,00	90,00	88,00	89,00
StyleGan2-Ada	98,00	98,00	98,00	98,00

Table 3 presents the evaluation results of the EfficientNetV2B3 model performance during the testing phase using three different augmentation approaches. The evaluation parameters include test accuracy, precision, recall, and F1-score, which collectively represent the model's ability to accurately recognize and classify new data. In the non-augmentation approach, the model achieved a test accuracy of 82%, with a precision of 83%, recall of 79%, and an F1-score of 80%. These results indicate that the model still exhibited a relatively high error rate due to limited data variation, leading to suboptimal generalization on unseen data.

The implementation of Traditional Augmentation significantly improved the model's performance, achieving a test accuracy of 89%, precision of 90%, recall of 88%, and an F1-score of 89%. This improvement demonstrates that simple augmentation techniques such as rotation, shifting, zooming, shearing, and flipping are effective in enhancing data diversity and strengthening the model's ability to recognize visual patterns under varying conditions.

The StyleGAN2-ADA approach yielded the best results among the three methods, with test accuracy, precision, recall, and F1-score all reaching 98%. These consistently high and balanced metric values indicate that the model is not only highly accurate but also exhibits excellent equilibrium between its ability to detect positive and negative classes, minimizing the risk of bias.

5. CONCLUSION

Based on the experimental results, it can be concluded that data augmentation plays a crucial role in improving the performance of the EfficientNetV2B3 model. The model without augmentation achieved only 82% accuracy and an F1-score of 80%, indicating limited generalization capability. The application of traditional augmentation techniques such as rotation, flipping, and zooming improved performance to 89% accuracy with an F1-score of 89%, demonstrating the effectiveness of data variation in enriching the model's learning process. The best results were obtained using the StyleGAN2-ADA approach, achieving 98% for accuracy, precision, recall, and F1-score. These findings indicate that GAN-based augmentation can generate realistic and diverse synthetic data, thereby enhancing the model's ability to recognize patterns and reducing classification errors.

6. ACKNOWLEDGEMENTS

This research was supported by the Pusat Penelitian dan Pengabdian kepada Masyarakat (P3M) Politeknik Negeri Bengkalis through the PNBK Research Grant for the Year 2025 under the Beginner Lecturer Research Scheme.

7. REFERENCES

- Agustin, T., Saputro, I. A., & Rahmadi, M. L. (2025). Optimizing Rice Plant Disease Classification Using Data Augmentation with GANs on Convolutional Neural Networks. *INTENSIF: Jurnal Ilmiah Penelitian Dan Penerapan Teknologi Sistem Informasi*, 9(1), 97–114. <https://doi.org/10.29407/intensif.v9i1.23834>
- Agustina, A., Yanto, F., Budianita, E., Iskandar, I., & Syafria, F. (2024). Klasifikasi Penyakit Tanaman Padi Menggunakan Metode Cnn Arsitektur Densenet-121 Dan Augmentasi Data. *JOISIE (Journal Of Information Systems And Informatics Engineering)*, 8(1), 124–134. <https://doi.org/10.35145/joisie.v8i1.4256>
- Ayuni, D. P., Jasril, Irsyad, M., Yanto, F., & Sanjaya, S. (2023). Data Augmentation on the Implementation of Convolutional Neural Network Architecture Efficientnet-B3 for Classification of Rice Leaf Diseases. *ZONasi: Jurnal Sistem Informasi*, 5(2), 239–249.
- Chen, D., Qi, X., Zheng, Y., Lu, Y., Huang, Y., & Li, Z. (2024). Synthetic data augmentation by diffusion probabilistic models to enhance weed recognition. *Computers and Electronics in Agriculture*, 216.
- Ibrahim, S., Amin, K. M., Ibrahim Alkanhel, R., Abdallah, H. A., & Ibrahim, M. (2024). Soft

- Attention Based Efficientnetv2b3 Model for Skin Cancer's Disease Classification Using Dermoscopy Images. *IEEE Access*, 12(September), 161283–161295. <https://doi.org/10.1109/ACCESS.2024.3486153>
- Joo, J., Kim, S. Y., Kim, D., Lee, J. E., Lee, S. M., Suh, S. Y., Kim, S. J., & Kim, S. C. (2024). Enhancing automated strabismus classification with limited data: Data augmentation using StyleGAN2-ADA. *PLoS ONE*, 19(5), 1–14. <https://doi.org/10.1371/journal.pone.0303355>
- Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., & Aila, T. (2020). Training generative adversarial networks with limited data. *Advances in Neural Information Processing Systems, 2020-Decem*(NeurIPS).
- Kusuma, B., Hermanto, T. I., & Lestari, C. D. (2025). Klasifikasi Jenis Penyakit Pada Tanaman Padi Menggunakan Algoritma Convolutional Neural Network. *JIKO (Jurnal Informatika Dan Komputer)*, 9(1), 40. <https://doi.org/10.26798/jiko.v9i1.1395>
- Liu, D., Wang, W., Wu, X., & Yang, J. (2022). EfficientNetv2 Model for Breast Cancer Histopathological Image Classification. *International Conference on Electronic Communication and Artificial Intelligence (IWECAI)*, 384–387.
- Manliguez, C. (2016). *Generalized Confusion Matrix for Multiple Classes*. November, 5–7. <https://doi.org/10.13140/RG.2.2.31150.51523>
- Octaviana, I., & Ekawati, S. (2022). Inventarisasi Hama dan Musuh Alami pada Tanaman Padi di Kecamatan Pulau Laut Timur. *Jurnal Pertanian Terpadu*, 10(1), 24–36. <https://doi.org/10.36084/jpt.v10i1.379>
- Pushpa Athisaya Sakila Rani, A., & Suresh Singh, N. (2025). Classification and identification of pest, diseases and nutrient deficiency in paddy using layer based EMD phase features with decision tree. *Information Processing in Agriculture*, 12(2), 232–244. <https://doi.org/10.1016/j.inpa.2024.09.003>
- Sahib, K. A., Oleiwi, B. K., & Nasser, A. R. (2025). Improving Classification Performance of Medicinal Plants with StyleGAN2-ADA. *Ingenierie Des Systemes d'Information*, 30(3), 823–834. <https://doi.org/10.18280/isi.300325>
- Sapitri, W., Kunang, Y. N., Yadi, I. Z., & Mahmud. (2023). The Impact of Data Augmentation Techniques on the Recognition of Script Images in Deep Learning Models. *Jurnal Online Informatika*, 8(2), 169–176. <https://doi.org/10.15575/join.v8i2.1073>
- Sinamenye, J. H., Chatterjee, A., & Shrestha, R. (2025). Potato plant disease detection: leveraging hybrid deep learning models. *BMC Plant Biology*, 25(1). <https://doi.org/10.1186/s12870-025-06679-4>
- Tan, M., & Le, Q. V. (2021). EfficientNetV2: Smaller Models and Faster Training. *Proceedings of Machine Learning Research*, 139, 10096–10106.
- Wagiyanti, Hamidson, H., & Suwandi. (2024). Intensitas dan Insidensi Serangan Hama Penyakit pada Tanaman Padi di Desa Enggal Rejo, Kecamatan Air Salek. *Journal of Global Sustainable Agriculture*, 4(2), 144.
- Woodland, M. K., Wood, J., Anderson, B. M., Kundu, S., Lin, E., Koay, E., Odisio, B., Chung, C., Kang, H. C., Venkatesan, A. M., Yedururi, S., De, B., Lin, Y. M., Patel, A. B., & Brock, K. K. (2022). Evaluating the Performance of StyleGAN2-ADA on Medical Images. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 13570 LNCS, 142–153. https://doi.org/10.1007/978-3-031-16980-9_14