



Peningkatan Akurasi Deteksi Penyakit Daun Padi Menggunakan Augmentasi Data Berbasis *Generative Adversarial Networks* (GAN)

Nasya Sunia Tubuon^{1*}, Nancy Jeane Tuturoong², Salaki Reynaldo Joshua³

^{1,2,3} Teknik Informatika Fakultas Teknik Universitas Sam Ratulangi

Informasi Artikel

Sejarah Artikel:
 Submit: 11 May 2026
 Revisi: 18 Juni 2026
 Diterima: 07 Juli 2026
 Diterbitkan: 10 Juli 2026

Kata Kunci

Data augmentation, rice leaf disease, GAN, EfficientNetB0, StyleGAN2-ADa

Korespondensi

E-mail:

nasyatubuon026@student.unsrat.ac.id*

A B S T R A C T

Rice is a staple food crop for more than half of the global population, with Asia contributing approximately 90% of global production. Leaf diseases, particularly Blast and Bacterial Blight, are major factors contributing to reduced rice productivity in Indonesia. Conventional detection methods based on visual observation are often subjective and time-consuming, highlighting the need for more reliable automated detection systems. This study aims to implement StyleGAN2-ADA to generate synthetic rice leaf disease images and evaluate its impact on the performance of Convolutional Neural Network-based classification. This research employed a quantitative experimental approach by comparing two scenarios: Baseline without GAN augmentation and Proposed with StyleGAN2-ADA synthetic image augmentation. Two CNN architectures, EfficientNetB0 and ResNet50, were evaluated using accuracy, precision, recall, F1-Score, and confusion matrix metrics. The quality of synthetic images was assessed using the Fréchet Inception Distance. The results demonstrated that StyleGAN2-ADA augmentation improved the overall F1-Score of EfficientNetB0 from 97.62% to 98.20%, with the largest improvement observed in the Blight class, increasing by 3.11%. For ResNet50, the overall F1-Score increased from 97.60% to 98.20%, although the Blast class showed no performance improvement after augmentation. GAN augmentation provided the most consistent benefits for the minority Blight class, while its impact on the Blast class varied across metrics. In EfficientNetB0, improvements in precision and F1-Score were accompanied by a decrease in recall. These findings indicate that model evaluation should consider class-specific performance and the trade-off between precision and recall rather than relying solely on aggregate metrics.

Abstrak

Padi merupakan tanaman pangan pokok bagi lebih dari separuh populasi dunia, dengan Asia menyumbang sekitar 90% produksi global. Penyakit daun, terutama Blast dan Bacterial Blight, menjadi penyebab penting penurunan produktivitas padi di Indonesia. Metode deteksi konvensional berbasis pengamatan visual cenderung subjektif dan lambat, sehingga dibutuhkan sistem deteksi otomatis yang lebih andal. Penelitian ini bertujuan mengimplementasikan StyleGAN2-ADA untuk menghasilkan citra sintetis penyakit daun padi serta mengevaluasi pengaruhnya terhadap performa klasifikasi berbasis Convolutional Neural Network. Penelitian menggunakan pendekatan eksperimental kuantitatif dengan membandingkan skenario Baseline tanpa augmentasi GAN dan Proposed dengan augmentasi citra sintetis StyleGAN2-ADA. Arsitektur yang diuji adalah EfficientNetB0 dan ResNet50 dengan metrik accuracy, precision, recall, F1-Score, dan confusion matrix. Kualitas citra sintetis dievaluasi menggunakan Fréchet Inception Distance. Hasil menunjukkan bahwa augmentasi StyleGAN2-ADA meningkatkan F1-Score keseluruhan EfficientNetB0 dari 97,62% menjadi 98,20%, dengan

peningkatan terbesar pada kelas Blight sebesar 3,11%. Pada ResNet50, F1-Score keseluruhan meningkat dari 97,60% menjadi 98,20%, meskipun performa kelas Blast tidak berubah. Augmentasi memberikan manfaat paling konsisten pada kelas minoritas Blight, sedangkan pengaruhnya pada kelas Blast berbeda antar metrik. Pada EfficientNetB0, peningkatan precision dan F1-Score disertai penurunan recall. Temuan ini menunjukkan bahwa evaluasi model perlu mempertimbangkan performa setiap kelas dan trade-off antara precision dan recall, bukan hanya metrik agregat.

This is an open access article under the CC-BY-SA license



1. Pendahuluan

Padi merupakan tanaman pangan pokok bagi lebih dari separuh populasi dunia, di mana Asia memberikan kontribusi sebesar 90% dari total produksi global [1]. Sebagai bagian integral dari wilayah tersebut, di Indonesia padi memegang peranan krusial sebagai pilar *triple security* yang meliputi aspek keamanan pangan, ekonomi, dan nasional. Hal ini disebabkan oleh kontribusi padi yang memenuhi hingga 96% kebutuhan kalori dan protein masyarakat, serta menjadi tumpuan 25,4 juta rumah tangga petani [2]. Meskipun produksi padi nasional terus diupayakan meningkat, tantangan di lapangan akibat Organisme Pengganggu Tanaman tetap signifikan. Berdasarkan data Kementerian Pertanian, luas serangan penyakit Blas mencapai 19.176 hektar, sementara penyakit Hawar Daun memberikan dampak yang lebih besar dengan luas serangan mencapai 27.721 hektar pada kategori terkena dan puso [3]. Salah satu penyebab utama penurunan produktivitas tanaman padi adalah penyakit daun, khususnya Blas dan Hawar Daun Bakteri, yang banyak ditemukan di lahan pertanian padi di Indonesia. Penyakit Blas (*Rice Blast*) disebabkan oleh jamur *Pyricularia oryzae* dengan ciri khas bercak berbentuk belah ketupat pada daun [4]. Adapun Hawar Daun Bakteri (*Bacterial Blight*) merupakan infeksi sistemik pada jaringan pembuluh (vaskular) yang disebabkan oleh patogen *Xanthomonas oryzae* *pv.* *oryzae* (Xoo), yang menyerang saluran xilem sehingga mengganggu distribusi nutrisi dan air pada seluruh bagian tanaman [5].

Deteksi dini terhadap penyakit daun padi adalah kunci untuk menekan penyebarannya. Namun, metode deteksi konvensional yang mengandalkan pengamatan visual oleh petani sering kali bersifat subjektif dan lambat. Penerapan teknologi visi komputer berbasis *deep learning*, khususnya *Convolutional Neural Networks* (CNN), menawarkan potensi besar untuk mengotomatisasikan proses identifikasi penyakit daun padi secara lebih akurat dan efisien. Teknik augmentasi data konvensional seperti rotasi, translasi, dan *flipping* sering digunakan untuk memperbanyak dataset secara artifisial, namun variasi yang dihasilkan cenderung terbatas dan belum mampu menangani ketidakseimbangan data secara optimal karena tidak menghasilkan fitur-fitur visual yang cukup beragam [6]. Sebagai solusi yang lebih efektif, *Generative Adversarial Networks* (GAN) menawarkan pendekatan augmentasi data yang lebih unggul. GAN bekerja melalui mekanisme *adversarial* antara *generator* dan *discriminator* sehingga mampu mempelajari distribusi data asli dan menghasilkan citra sintetis baru yang menyerupai data nyata secara lebih realistis dan beragam [7]. Pada penelitian ini, pendekatan *Style-based Generative Adversarial Network 2 with Adaptive Discriminator Augmentation* (StyleGAN2-ADA) diusulkan sebagai metode augmentasi data untuk meningkatkan performa model klasifikasi CNN dalam mendeteksi penyakit Blas dan Hawar Daun Bakteri pada tanaman padi.

Beberapa penelitian terkait telah dilakukan dalam bidang deteksi penyakit tanaman berbasis *deep learning* dan augmentasi data generatif. Stephen et al. mengeksplorasi penggunaan arsitektur ResNet34 dengan *Self-Attention* mampu mencapai akurasi puncak sebesar 98,54%, mengungguli model CNN standar lainnya. Saputra et al. melaporkan akurasi 92% untuk tiga kelas penyakit daun padi menggunakan CNN, namun menunjukkan gejala *overfitting* yang umum terjadi saat model dilatih pada dataset kecil dan kurang bervariasi [8]. Zeng et al. menerapkan model *Inception-v3* untuk mendeteksi

tingkat keparahan penyakit *Huanglongbing* (HLB) pada daun jeruk menggunakan 5.406 citra dengan akurasi awal sebesar 74,38%. Setelah dilakukan augmentasi data menggunakan DCGAN hingga jumlah mencapai 14.056, akurasi model meningkat signifikan menjadi 92,60 [9]. Haruna et al. melaporkan implementasi *StyleGAN2-ADA* yang menghasilkan citra sintetis berkualitas tinggi, ditunjukkan oleh nilai *Frechet Inception Distance* (FID) sebesar 26,67 dan *Kernel Inception Distance* (KID) sebesar 0,08. Citra sintetis tersebut digabungkan dengan dataset asli untuk melatih *Faster-RCNN* dan *Single Shot Detector* (SSD), menghasilkan peningkatan nilai *mean Average Precision* (mAP) hingga 0,93 dan 0,91 secara berturut-turut, dengan hasil uji statistik yang signifikan dibandingkan augmentasi standar [10]. Selain itu, penelitian oleh Agustin et al. menunjukkan bahwa augmentasi DCGAN memberikan peningkatan akurasi sebesar 5,44% lebih tinggi dibandingkan augmentasi klasik yang hanya memberikan kenaikan 2,56%[11].

Berdasarkan kajian terhadap penelitian-penelitian terkait di atas, terdapat celah yang nyata. Meskipun GAN telah terbukti efektif sebagai metode augmentasi data dan *StyleGAN2-ADA* telah menunjukkan hasil yang menjanjikan, penerapannya secara spesifik untuk meningkatkan performa model klasifikasi CNN pada penyakit *Blast* dan *Blight* tanaman padi masih sangat terbatas. Sebagian besar penelitian yang ada menerapkan *StyleGAN2-ADA* pada model deteksi objek seperti *Faster-RCNN*, bukan pada tugas klasifikasi citra daun yang relevan dengan kondisi pertanian di Indonesia.

Oleh karena itu, penelitian ini bertujuan untuk: (1) mengimplementasikan model *StyleGAN2-ADA* yang mampu menghasilkan citra sintetis penyakit daun padi yang realistis; dan (2) mengevaluasi sejauh mana peningkatan akurasi model klasifikasi CNN setelah dilatih menggunakan dataset yang diperbanyak dengan citra hasil augmentasi GAN. Hasil penelitian ini diharapkan dapat menjadi solusi praktis bagi pendeteksian dini penyakit daun padi secara otomatis, serta memberikan kontribusi metodologis pada pengembangan sistem pertanian cerdas di Indonesia.

2. Metode Penelitian

2.1. Tahapan Penelitian

Penelitian ini merupakan penelitian eksperimental dengan pendekatan kuantitatif yang bertujuan mengevaluasi pengaruh augmentasi data berbasis *Generative Adversarial Network* (GAN) terhadap kinerja model klasifikasi penyakit citra daun padi berbasis *Convolutional Neural Network* (CNN). Variabel independen dalam penelitian ini adalah penggunaan augmentasi data berbasis GAN, sedangkan variabel dependennya adalah akurasi model klasifikasi penyakit daun padi [12].

Pendekatan eksperimental diimplementasikan melalui dua skenario komparasi sebagai berikut:

1. Skenario A (*Baseline*): Model CNN dilatih hanya menggunakan dataset asli tanpa augmentasi berbasis GAN.
2. Skenario B (Model Usulan): Model CNN dilatih menggunakan dataset yang telah diperbanyak melalui penambahan citra sintetis hasil augmentasi *StyleGAN2-ADA*.

Perbedaan performa antara kedua skenario dianalisis menggunakan confusion matrix, dan selanjutnya dihitung metrik evaluasi berupa *accuracy*, *precision*, *recall*, dan *F1-score* sebagai dasar penilaian efektivitas augmentasi data berbasis GAN.

Alur penelitian ini menggunakan pendekatan paralel, di mana dataset hasil pra-pemrosesan dibagi ke dalam dua jalur komparasi. Jalur pertama (Skenario A) berfungsi sebagai *Baseline*, sedangkan jalur kedua mengintegrasikan arsitektur *StyleGAN2-ADA* untuk menghasilkan data sintetis yang kemudian digabungkan dengan data asli. Kedua output model tersebut kemudian dievaluasi untuk mengukur sejauh mana peningkatan akurasi. Alur penelitian secara lengkap ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

2.2 Preprocessing untuk CNN

Sebelum pelatihan, termasuk pelatihan StyleGAN2-ADA, seluruh data dibagi dalam tiga partisi menggunakan rasio 80:10:10, yaitu 80% untuk data latih (*training set*), 10% untuk data validasi (*validation set*), dan 10% untuk data uji (*test set*). Setiap kelas (*Blast*, *Blight*, dan *Healthy*) dipisahkan dalam subdirektori masing-masing partisi secara terstruktur, sehingga label dapat ditetapkan secara otomatis berdasarkan nama *folder*.

Seluruh citra diubah ukurannya menjadi resolusi seragam 224×224 piksel, sesuai persyaratan input dari arsitektur ResNet50 dan EfficientNetB0. Normalisasi piksel dilakukan berbeda untuk masing-masing arsitektur, sesuai dengan standarnya. Pipeline data dioptimasi menggunakan tiga mekanisme TensorFlow Data API.

2.3 Preprocessing StyleGAN2-ADA

Pelatihan pada tahap ini hanya menggunakan data *training set* (80% dari total dataset) yang telah dipisahkan sebelumnya. *Validation set* dan *Test set* tidak digunakan dalam proses pelatihan GAN untuk menghindari kontaminasi data. Preprocessing untuk StyleGAN2-ADA mencakup dua tahap. Secara manual, seluruh citra dikonversi ke format RGB, di-crop menjadi rasio 1:1, di-resize ke resolusi 256×256 piksel menggunakan interpolasi LANCZOS, lalu dikemas dalam format ZIP. Secara otomatis, StyleGAN2-ADA melakukan normalisasi piksel ke rentang $[-1, 1]$, horizontal flip acak (50%), dan

Adaptive Discriminator Augmentation (ADA) secara dinamis untuk mencegah *overfitting* pada diskriminator.

2.4 Augmentasi Data dengan StyleGAN2-ADA

Model generatif yang digunakan adalah StyleGAN2-ADA berbasis PyTorch yang dikembangkan oleh Karras et al. [13]. Pemilihan model ini didasarkan pada beberapa pertimbangan: (1) mekanisme ADA dirancang khusus untuk kondisi limited data seperti dataset penelitian ini (3.307 citra); (2) model ini menghasilkan citra dengan skor *Fréchet Inception Distance* (FID) lebih rendah dibandingkan arsitektur GAN sebelumnya pada dataset kecil; (3) tersedia *pre-trained* model untuk *rice leaf disease* yang dapat dimanfaatkan untuk *transfer learning*; dan (4) implementasi PyTorch yang terdokumentasi dengan baik [14].

Penelitian ini memanfaatkan bobot *pre-trained* model yang dipublikasikan oleh Yumusa, telah dilatih pada 5.932 citra penyakit daun padi mencakup empat kategori: *Bacterial Blight*, *Tungro*, *Brown Spot*, dan *Rice Blast* [15]. *Transfer learning* dipilih karena: mempercepat konvergensi, mengurangi risiko *overfitting* pada dataset kecil, serta kesamaan domain (*rice leaf disease*) yang memastikan transfer pengetahuan yang optimal. Proses *transfer learning* meliputi: (1) memuat model *pre-trained* sebagai *checkpoint* awal menggunakan parameter *--resume* dalam format .pkl; (2) *fine-tuning* pada seluruh lapisan dengan *learning rate* 0,002, *optimizer* Adam ($\beta_1=0$, $\beta_2=0,99$, $\epsilon=10^{-8}$), *batch size* 4, dan R1 regularisasi $\gamma=10$; (3) pelatihan dilanjutkan hingga 120 king atau hingga konvergensi tercapai; serta (4) penyimpanan *checkpoint* setiap 5 king mencakup *network snapshot*, sampel citra hasil generasi, dan evaluasi skor FID.

Citra sintetis yang dihasilkan oleh StyleGAN2-ADA hanya akan ditambahkan ke dalam *training set*, sehingga *validation set* dan *test set* hanya terdiri dari citra asli yang belum pernah dilihat oleh model GAN maupun CNN selama proses pelatihan. Dengan demikian, tidak terjadi *data leakage* antara data augmentasi dan data evaluasi.

2.5 Pelatihan Model CNN

Pelatihan model CNN dilakukan pada dua arsitektur, yaitu ResNet50 dan EfficientNetB0, menggunakan strategi pelatihan dua tahap. Pada Tahap 1 (*Feature Extraction*), seluruh lapisan *backbone* dibekukan (*frozen*) dan pelatihan hanya berfokus pada lapisan klasifikasi baru dengan *learning rate* 10^{-4} menggunakan *optimizer* Adam selama 15 *epoch* awal. Pada Tahap 2 (*Fine-Tuning*), sebagian lapisan terakhir *backbone* dibuka: pada ResNet50 dilakukan *unfreeze* pada ± 30 hingga 34 lapisan akhir, sedangkan pada EfficientNetB0 pada 20 lapisan terakhir, dengan *learning rate* yang diturunkan menjadi 10^{-5} dan batas maksimal 50 *epoch*.

Perbandingan *hyperparameter* kedua arsitektur disajikan pada Tabel 1.

Tabel 1. Perbandingan Hyperparameter Model

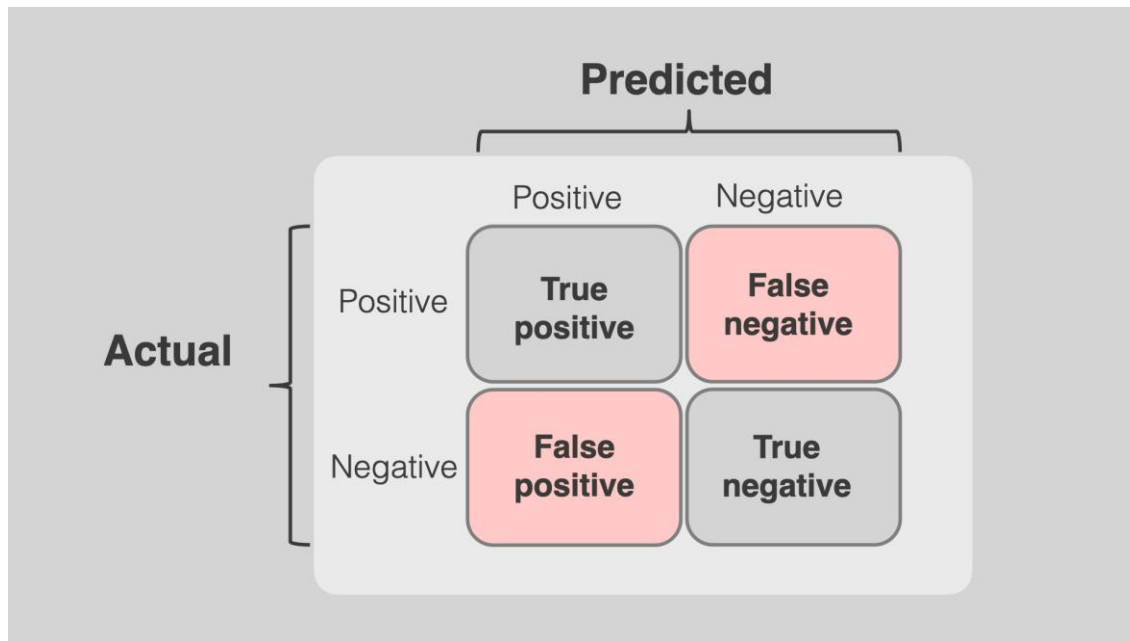
Parameter	EfficientNetB0	ResNet50
<i>Input Size</i>	224 x 224 piksel	224 x 224 piksel
<i>Batch Size</i>	32	32
<i>Optimizer</i>	Adam	Adam
<i>Loss Function</i>	Categorical Cross-Entropy	Categorical Cross-Entropy
Tahap 1 <i>Learning Rate</i>	1×10^{-5} (<i>Backbone Frozen</i>)	1×10^{-4} (<i>Backbone Frozen</i>)
Tahap 2 <i>Learning Rate</i>	1×10^{-5} (<i>Fine-Tuning</i>)	1×10^{-5} (<i>Fine-Tuning</i>)
<i>Dropout Rate</i>	0.3	0.3
<i>Fine-Tuning Setup</i>	Unfreeze last 45 layer	Unfreeze last 30-34 layer
<i>Early Stopping</i>	Patience 10	Patience 10
<i>LR Scheduler</i>	ReduceLROnPlateau (factor 0.2)	ReduceLROnPlateau (factor 0.2)

Untuk memastikan efisiensi pelatihan, diterapkan tiga mekanisme *callbacks*: (1) *Early Stopping*, pelatihan dihentikan otomatis jika *validation loss* tidak menurun selama 10 *epoch* berturut-turut; (2) *Checkpoint*, menyimpan bobot model terbaik berdasarkan akurasi validasi tertinggi; dan (3) *Reduce*

Learning Rate on Plateau, *learning rate* diturunkan otomatis dengan faktor 0,2 jika performa terhambat selama 5 *epoch*.

2.6 Evaluasi Kinerja Model

Evaluasi kinerja model dilakukan secara kuantitatif menggunakan *confusion matrix* yang menampilkan empat komponen: *True Positive (TP)*, *False Positive (FP)*, *True Negative (TN)*, dan *False Negative (FN)*, seperti yang diilustrasikan pada Gambar 2 [16].



Gambar 2. Struktur Confusion Matrix

Berdasarkan nilai *confusion matrix*, dihitung empat evaluasi utama sebagai berikut:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Selain metrik kuantitatif, dilakukan analisis kurva *loss* dan *accuracy* selama proses pelatihan dan validasi untuk memantau konvergensi dan mendeteksi gejala *overfitting*. Analisis perbandingan akhir dilakukan dengan mengukur selisih nilai akurasi dan stabilitas kurva pembelajaran antar skenario A (*Baseline*) dan Skenario B (*Proposed*) pada kedua arsitektur *EfficientNetB0* dan *ResNet50*.

3. Hasil Dan Pembahasan

3.1 Dataset Asli

Dataset yang digunakan dalam penelitian ini terdiri dari 3.307 citra daun padi yang diperoleh dari dataset publik Kaggle dan Mendeley Dataset. Datset mencakup dua kategori penyakit dan satu kelas daun sehat.

Tabel 1. Distribusi Dataset per Kelas

Kelas	Jumlah Citra	Persentase
-------	--------------	------------

<i>Blast</i>	1016	30,72
<i>Blight</i>	624	18,87
<i>Healthy</i>	1667	50,41

Berdasarkan Tabel 2, dataset terdiri dari kelas *Blast* dengan citra sebanyak 1016, *Blight* sebanyak 624, dan *Healthy* sebanyak 1667.

Tabel 3. Distribusi Dataset Berdasarkan Kondisi Latar Belakang

Kelas	<i>Controlled</i>	<i>Uncontrolled</i>	Total
<i>Blast</i>	421	595	1016
<i>Blight</i>	382	242	624
<i>Healthy</i>	1454	213	1667

Berdasarkan Tabel 3, sebagian citra dalam dataset merupakan citra *controlled* sebanyak 2257 citra, yang didominasi oleh kelas *Healthy* sebanyak 1454. Kemudian dataset *uncontrolled* sebanyak 1050 citra, yang didominasi oleh kelas *Blast*. Variasi kondisi latar belakang ini mencerminkan keberagaman kondisi pengambilan gambar di lapangan yang berpotensi mempengaruhi kemampuan generalisasi model dalam mengenali penyakit daun padi pada kondisi nyata.

3.2 Dataset Split

Dataset untuk seluruh eksperimen dibagi menjadi tiga subset menggunakan *stratified split* untuk menjaga proporsi kelas, dengan subset *Training Set* untuk melatih model, *Validation Set* untuk *Tuning Hyperparameter* dan *Early Stopping*, serta *Test Set* untuk evaluasi final performa. Seluruh data dibagi dalam tiga partisi menggunakan rasio 80:10:10, yaitu 80% untuk data latih (*training set*), 10% untuk data validasi (*validation set*), dan 10% untuk data uji (*test set*).

Tabel 4. Distribusi Dataset Skenario *Baseline*

Kelas	<i>Trainin</i>	<i>Validation</i>	<i>Test</i>	Total
<i>Blast</i>	812	101	103	1016
<i>Blight</i>	499	62	63	624
<i>Healthy</i>	1333	166	168	1667
<i>Total</i>	2644	329	334	3307

Tabel 5. Distribusi Dataset Skenario *Proposed*

Kelas	<i>Trainin</i>	<i>Validation</i>	<i>Test</i>	Total
<i>Blast</i>	1333	101	103	1016
<i>Blight</i>	1333	62	63	624
<i>Healthy</i>	1333	166	168	1667
<i>Total</i>	3.999	329	334	3307

3.3 Konfigurasi Eksperimen

Seluruh eksperimen dijalankan pada Google Colaboratory dengan akselerasi GPU Tesla T4 menggunakan Python dan framework TensorFlow/Keras. Kedua model menerapkan strategi transfer learning dua tahap yang identik: tahap pertama membekukan seluruh backbone untuk melatih lapisan klasifikasi, dan tahap kedua membuka sebagian lapisan akhir backbone dengan learning rate lebih kecil untuk fine-tuning. Detail hyperparameter disajikan pada Tabel 4.

Tabel 6. Hyperparameter Eksperimen CNN

Hyperparameter	Nilai
Ukuran citra input	224×224
<i>Batch size</i>	32
<i>Learning rate</i> (Tahap 1)	1×10^{-4}
<i>Learning rate</i> (Tahap 2)	1×10^{-5}
<i>Minimum Learning rate</i>	1×10^{-7}
Maksimum <i>epoch</i> Tahap 1	15

Maksimum <i>epoch</i> Tahap 2	50
<i>Optimizer</i>	Adam
<i>Loss Function</i>	Categorical Crossentropy
<i>Early stopping patience</i>	10 <i>epoch</i>
<i>ReduceLROnPlateau patience</i>	5 <i>epoch</i>

3.4 Hasil EfficientNetB0 Baseline

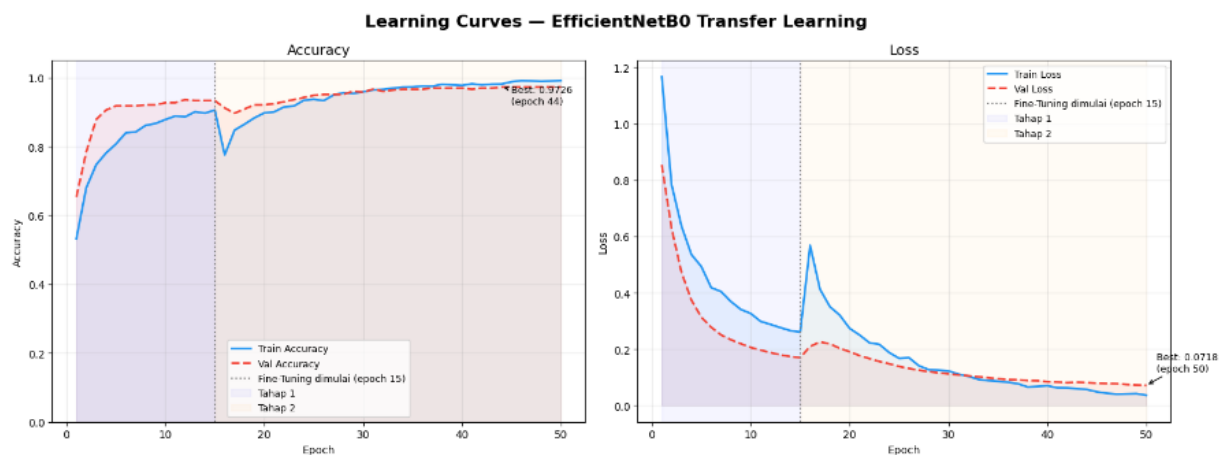
Pengujian model EfficientNetB0 *Baseline* dilakukan menggunakan dataset asli tanpa augmentasi GAN. Hasil evaluasi kuantitatif model disajikan pada Tabel 7.

Tabel 7. Metrik Performa EfficientNet *Baseline*

Kelas	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)	Support
<i>Blast</i>	97,09	96,15	97,09	96,62	103
<i>Blight</i>	95,24	93,75	95,24	94,49	63
<i>Healthy</i>	98,81	1.00	98,81	99,40	168
<i>Overall</i>	97,60	97,60	97,60	97,62	334

Berdasarkan tabel hasil pengujian di atas, model EfficientNetB0 *Baseline* menunjukkan performa yang baik dalam mendeteksi penyakit daun padi. Secara keseluruhan, model ini berhasil mencapai nilai *F1-Score* sebesar 97,62% dari total 334 data uji. Namun, berdasarkan distribusi kelas yang tidak seimbang (*imbalanced data*), kelas *Blight* menjadi kelas dengan performa terendah, yakni sebesar 94,49%.

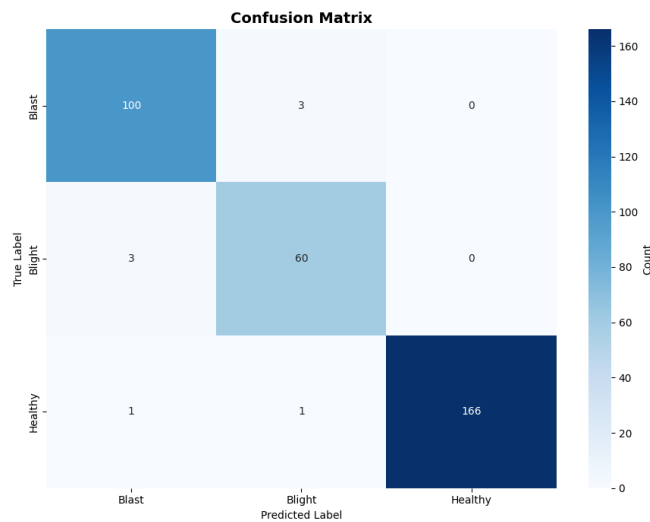
Untuk menganalisis proses konvergensi model selama pelatihan, kurva *learning* disajikan pada Gambar 3.



Gambar 3. Learning Curves EfficientNetB0 *Baseline*

Model menunjukkan peningkatan cepat pada 5 epoch pertama, diikuti fluktuasi singkat pada *epoch* ke-16 akibat proses *unfreezing backbone*. Setelah fase adaptasi, model konvergen secara stabil dengan akurasi validasi terbaik 97,26% pada epoch ke-4.

Distribusi hasil prediksi model secara lebih rinci dapat dilihat melalui *confusion matrix* pada Gambar 4.



Gambar 4. Confusion Matrix EfficientNetB0 Baseline

Kelas *Healthy* merupakan kelas dengan performa paling optimal, di mana 166 berhasil diklasifikasikan dengan benar dan hanya terdapat 2 sampel yang salah prediksi, 1 ke kelas *Blast* dan 1 ke kelas *Blight*. Pada kelas *Blast* model berhasil memprediksi 100 sampel uji dengan 3 kesalahan prediksi pada kelas *Blight*. Pada kelas *Blight*, meskipun kelas ini minoritas dengan jumlah data paling sedikit (63 sampel), model tetap mampu mengenali 60 sampel dengan benar, namun terdapat 3 kesalahan klasifikasi sebagai *Blight*.

3.5 Hasil ResNet50 Baseline

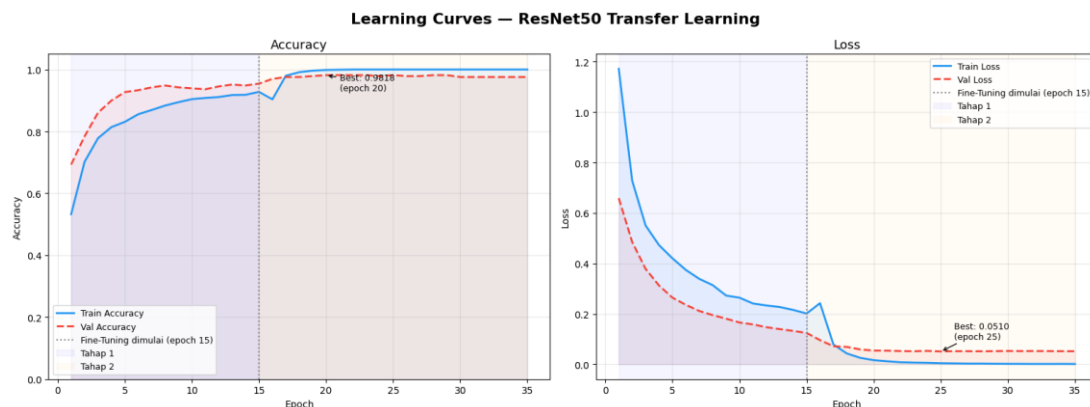
Pengujian model ResNet50 Baseline dilakukan menggunakan dataset asli tanpa augmentasi GAN. Hasil evaluasi kuantitatif model disajikan pada Tabel 8.

Tabel 8. Metrik Performa ResNet50 Baseline

Kelas	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)	Support
Blast	97,09	98,04	97,09	97,56	103
Blight	95,24	96,77	95,24	96,00	63
Healthy	98,81	97,65	98,81	98,22	168
Overall	97,60	97,60	97,60	97,60	334

Pada kelas *Healthy* yang merupakan kelas mayoritas dengan *support* sebanyak 168 citra, kelas ini mencatat *F1-Score* tertinggi sebesar 98,22%. Kelas *Blast* mencapai nilai *F1-score* sebesar 97,56%. Sementara itu, kelas *Blight* mencatat nilai *F1-Score* sebesar 96,00%, nilai yang lebih rendah, mengingat kelas ini merupakan kelas minoritas yang sampelnya hanya 63, terlihat jelas bahwa model memiliki keterbatasan dalam melakukan generalisasi dibandingkan kelas lainnya.

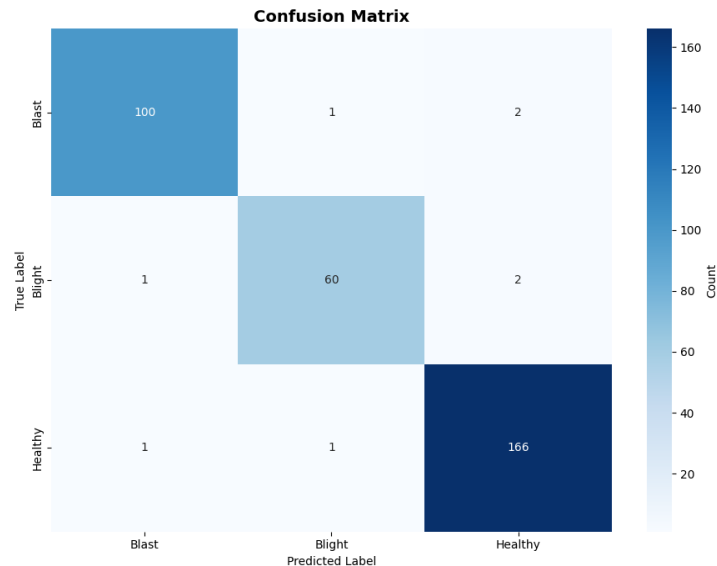
Untuk menganalisis proses konvergensi model selama pelatihan, kurva *learning* disajikan pada Gambar 5.



Gambar 5. Learning Curves ResNet50 Baseline

Memasuki tahap kedua (*fine-tuning*), terjadi penurunan akurasi dan lonjakan *loss* sebagai respons terhadap pembukaan lapisan *backbone*. Setelahnya, model menunjukkan kemampuan pemulihan yang sangat cepat, karena hanya dalam beberapa epoch setelah transisi, akurasi kembali meningkat. Model mencapai Akurasi Validasi terbaik sebesar 0,9818 pada epoch ke-20, dan mengalami penurunan *loss* yang stabil hingga mencapai *Loss* Validasi terbaik sebesar 0,0510 pada epoch ke-25.

Distribusi hasil prediksi model secara lebih rinci dapat dilihat melalui *confusion matrix* pada Gambar 6.



Gambar 6. Confusion Matrix ResNet50 Baseline

Pada kelas *Healthy*, dengan total sampel 168 model berhasil mengklasifikasikan 166 sampel secara akurat hanya dengan 2 kesalahan klasifikasi yaitu 1 sampel sebagai *Blast* dan 1 sebagai *Blight*. Performa kelas *Blast* berhasil memprediksi 100 sampel dengan benar dari total 103 data uji, 1 misklasifikasi sebagai *Blast* dan 1 sebagai *Blight*. Adapun pada kelas *Blight* yang merupakan kelas minoritas, model tetap mampu mengenali 60 sampel dengan benar, 1 sampel diklasifikasikan sebagai *Blast* dan 2 sebagai *Healthy*.

3.6 Konfigurasi Model StyleGAN2-ADA

Augmentasi citra sintetis dilakukan menggunakan model *pretrained* StyleGAN2-ADA dengan bobot *rice_leaf_disease.pkl* yang diperoleh dari Mendeley Data [15]. Model ini sebelumnya telah dilatih menggunakan 5.932 citra penyakit daun padi yang mencakup empat kelas penyakit, sehingga representasi fitur visualnya relevan dengan dataset penelitian ini. Proses pelatihan dilanjutkan menggunakan dataset penelitian ini yang terdiri dari dua kelas (*Blast* dan *Blight*), menggunakan arsitektur StyleGAN2-ADA yang diimplementasikan melalui repositori StyleGAN3 (Karras et al., n.d.) dengan konfigurasi - - `cfg=stylegan2`. Pendekatan ini dipilih karena repositori StyleGAN2-ADA Pytorch mengalami inkompatibilitas dengan versi CUDA yang tersedia pada platform komputasi yang digunakan. Pendekatan resume dari model pretrained dipilih untuk memanfaatkan representasi fitur visual penyakit daun padi yang telah dipelajari sebelumnya, sehingga proses konvergensi dapat dicapai lebih cepat meskipun jumlah data pelatihan terbatas.. Parameter yang digunakan dalam pelatihan ini disajikan pada Tabel 9.

Tabel 9. Parameter Model StyleGAN2-ADA

Hyperparameter	Nilai
Pre-trained	Rice_leaf_disease.pkl
Resolusi output	256 x 256 piksel
Truncation psi	0,7
Device	GPU T4 x2

3.7 Kualitas Citra Sintetis StyleGAN2-ADA

Kualitas citra sintetis dievaluasi menggunakan metrik *Fréchet Inception Distance* (FID), yang mengukur jarak distribusi statistik antara citra asli dan citra sintetis. Hasil evaluasi FID per kelas disajikan pada Tabel 10.

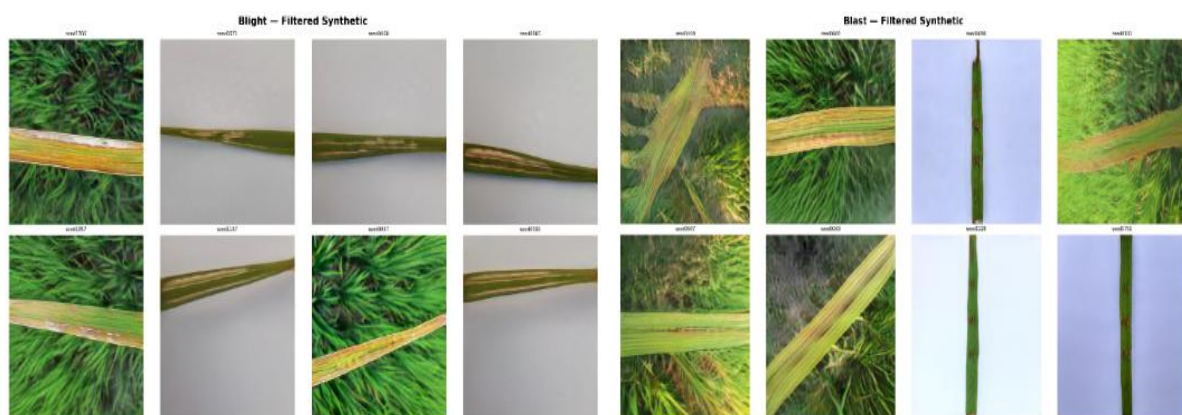
Tabel 10. Hasil Evaluasi FID per Kelas

Kelas	FID Score
<i>Blight</i>	79,08
<i>Blast</i>	103,20

Evaluasi kualitas citra sintetis menggunakan metrik FID menunjukkan perbedaan yang cukup besar antar kelas. Kelas *Blight* mencapai FID sebesar 79,08 sedangkan kelas *Blast* sebesar 103,20. Perbedaan ini mengindikasikan bahwa model *StyleGAN2-ADA* lebih berhasil dalam mempelajari distribusi visual kelas *Blight* dibandingkan kelas *Blast*. Kompleksitas morfologi lesi pada kelas *Blast* yang lebih bervariasi lebih sulit dimodelkan secara akurat oleh generator.

3.8 Visualisasi Citra Sintetis

Gambar 7 menyajikan contoh citra sintetis yang dihasilkan oleh model *StyleGAN2-ADA* untuk masing-masing kelas. Pada kelas *Blight*, citra sintetis menunjukkan karakteristik morfologi yang relevan, seperti pola bercak dan perubahan warna pada daun yang konsisten dengan ciri visual masing-masing kelas penyakit. Sementara itu, pada kelas *Blast*, beberapa citra sintetis menunjukkan artefak visual berupa pola yang tidak terstruktur dan kurang merepresentasikan karakteristik morfologi kelas *Blast* secara akurat. Hal ini sejalan dengan nilai FID kelas *Blast* yang lebih tinggi (103,20) dibandingkan kelas *Blight* (79,08), yang mengindikasikan bahwa model mengalami kesulitan lebih besar dalam mempelajari distribusi visual kelas *Blast*.



(a) Sampel Citra Sintetis kelas *Blight*

(b) Sampel Citra Sintetis kelas *Blast*

Gambar 7. Visualisasi Citra Sintetis Per Kelas

3.9 Komposisi Dataset Augmentasi

Dari proses augmentasi, dihasilkan 3000 citra sintetis untuk kelas *Blast*. Dan 2000 citra sintetis untuk kelas *Blight*. Citra tersebut kemudian melalui dua tahap penyaringan, yaitu (1) filter *Laplacian* yang digunakan untuk menyisihkan citra yang mengandung keburaman secara otomatis, sehingga jumlah citra berkurang menjadi 2790 untuk kelas *Blast*, 1317 untuk kelas *Blight*, dan (2) seleksi visual manual untuk menyaring citra yang mengandung artefak hasil proses generasi GAN. Hasil seleksi akhir menghasilkan 521 citra untuk kelas *Blast* dan 834 citra untuk kelas *Blight*. Seluruh citra sintetis yang telah melewati seleksi visual digabungkan dengan dataset asli untuk membentuk dataset augmentasi pada Skenario B. Seluruh citra sintetis yang telah melewati seleksi visual digabungkan dengan dataset asli untuk membentuk dataset augmentasi pada Skenario B. Komposisi akhir dataset disajikan pada Tabel 11.

Tabel 11. Komposisi Dataset Augmentasi

Kelas	Citra Asli (Training)	Citra Sintetis GAN	Citra filter <i>Laplacian</i>	Citra Akhir
<i>Blast</i>	499	2000	1317	834
<i>Blight</i>	812	3000	2790	521
Total	1.311	5000	4107	1355

3.10 Konfigurasi Eksperimen Skenario *Proposed*

Skenario *Proposed* menggunakan konfigurasi pelatihan yang identik dengan skenario *Baseline*. Satu-satunya perbedaan terletak pada data latih yang digunakan, yaitu dataset yang telah diperbanyak dengan citra sintetis hasil StyleGAN2-ADA yang dijelaskan pada bagian 3.3. Konfigurasi ini tetap dipertahankan agar perbandingan antar skenario bersifat adil dan perbedaan performa yang teramati dapat didistribusikan pada pengaruh augmentasi GAN.

3.11 Hasil EfficientNetB0 *Proposed*

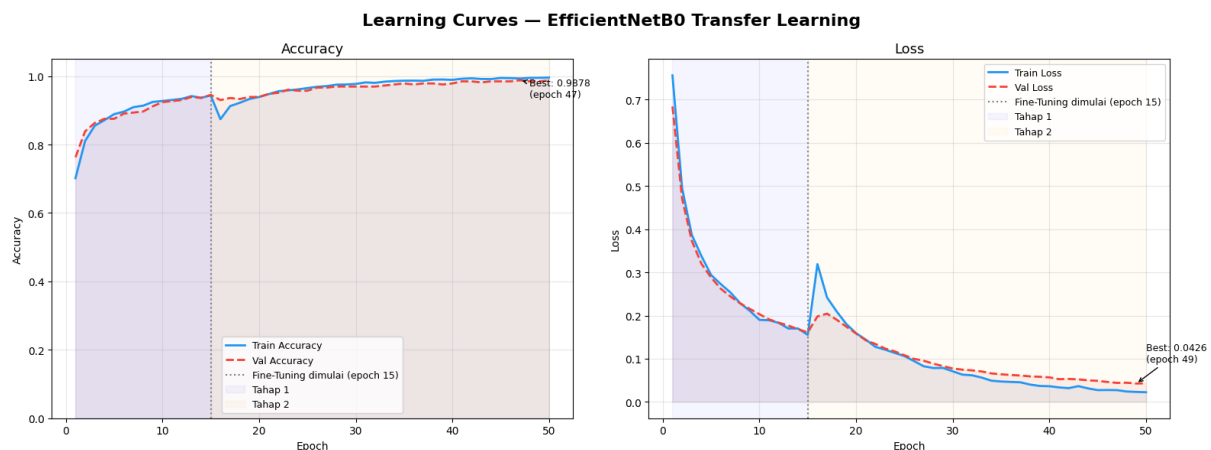
Pengujian model EfficientNetB0 *Proposed* dilakukan menggunakan dataset yang telah diseimbangkan melalui teknik augmentasi StyleGAN2-ADA. Hasil evaluasi kuantitatif model disajikan pada Tabel 11.

Kelas Blast mencatat nilai metrik F1-Score di angka 97,06%, hal ini mengindikasikan bahwa model mampu mempertahankan akurasi klasifikasi yang tinggi pada kelas ini. Kelas *Blight* menunjukkan peningkatan performa yang terbesar dibandingkan kedua kelas lainnya. Jika pada model Baseline kelas ini mendapat nilai F1-Score sebesar 94,49%, pada model Proposed ini nilai metriknya meningkat menjadi 97,60%.

Tabel 12. Hasil EfficientNetB0 *Proposed*

Kelas	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)	Support
<i>Blast</i>	96,12	98,02	96,12	97,06	103
<i>Blight</i>	96,83	98,39	96,83	97,60	63
<i>Healthy</i>	100	98,25	100	99,12	168
Overall	98,20	98,20	98,20	98,20	334

Untuk menganalisis proses konvergensi model selama pelatihan, kurva *learning* disajikan pada Gambar 7.

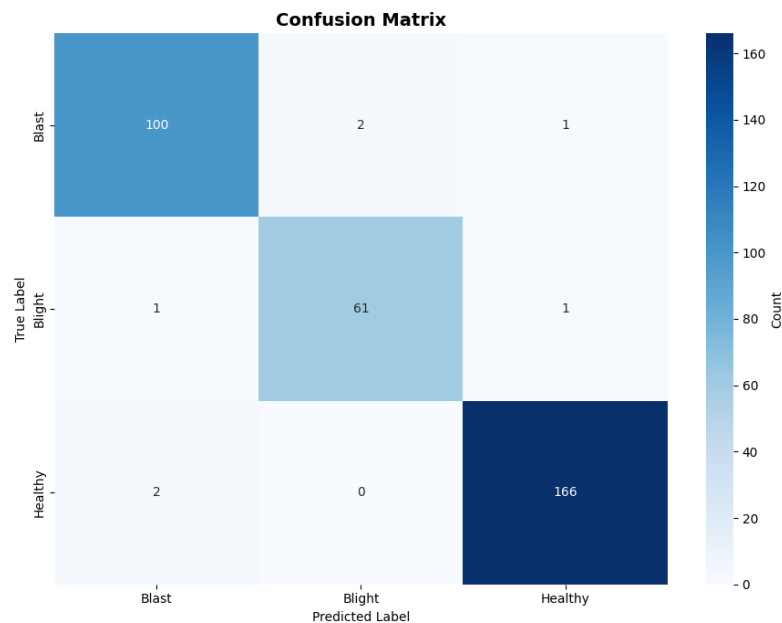


Gambar 8. Learning Curves EfficientNetB0 *Proposed*

Pada tahap *Transfer Learning*, model menunjukkan peningkatan pembelajaran yang meningkat. Akurasi validasi meningkat tajam pada kisaran 0,70 hingga 0,94 pada akhir *epoch* ke-15. Pada *epoch* ke-16 terdapat penurunan akurasi sesaat dan lonjakan kurva loss. Hal ini merupakan respons terhadap dimulainya tahap *fine-tuning*, di mana sebagian lapisan backbone dibuka (*unfreeze*) untuk dilatih kembali. Setelah beradaptasi, model kembali menunjukkan tren positif yang sangat stabil. Akurasi validasi terus naik hingga mencapai titik akurasi terbaik sebesar 0,9878 pada *epoch* ke-47. Begitupun

dengan *loss* validasi terus menurun hingga mencapai titik terendah atau *best loss* sebesar 0,0426 pada *epoch* ke-49.

Distribusi hasil prediksi model secara lebih rinci dapat dilihat melalui *confusion matrix* pada Gambar 9.



Gambar 9. *Confusion Matrix EfficientNetB0 Proposed*

Kelas *Blast* berhasil diklasifikasikan dengan benar sebanyak 100 dari 103 sampel dengan 2 misklasifikasi sebagai *Blight* dan 1 sebagai *Healthy*. . Kelas *Blight* sendiri menunjukkan hasil yang sangat stabil dengan memprediksi 61 sampel benar dari total 63 data uji, dan hanya terdapat 2 sampel salah prediksi, 1 sebagai *Blast* dan 1 sebagai *Healthy*. Pada kelas *Healthy*, model berhasil mengidentifikasi 166 sampel daun secara tepat dari total 168 sampel. Terdapat kesalahan berupa 2 sampel yang terdeteksi sebagai *Blast*. Performa ini membuktikan bahwa model mampu mempertahankan akurasi tinggi pada kelas mayoritas tanpa terganggu oleh penambahan data sintetis pada kelas lainnya.

3.12 Hasil ResNet50 Proposed

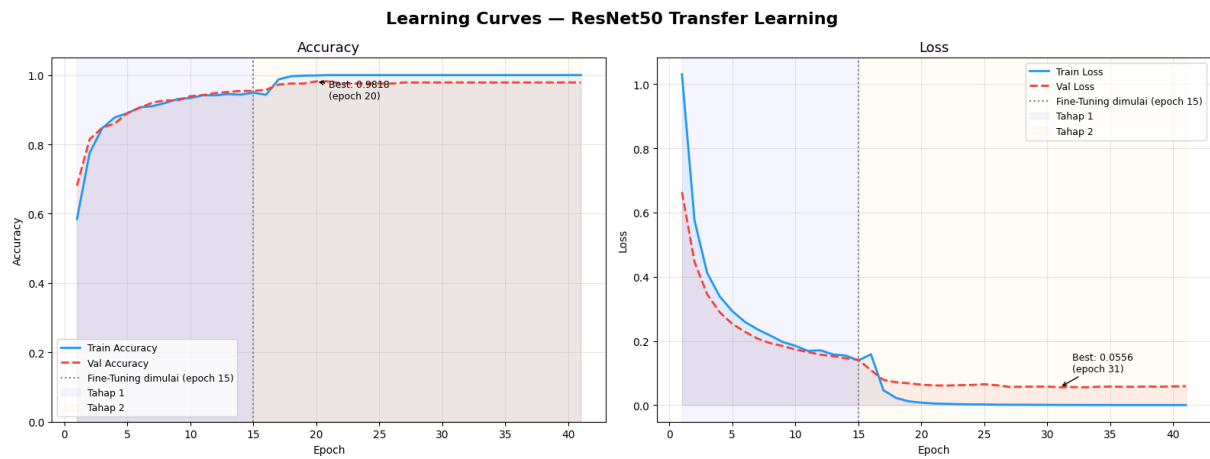
Pengujian model *ResNet50 Proposed* dilakukan menggunakan dataset yang telah diseimbangkan melalui teknik augmentasi *StyleGAN2-ADA*. Hasil evaluasi kuantitatif model disajikan pada Tabel 12.

Tabel 13. Hasil *ResNet50 Proposed*

Kelas	Accuracy(%)	Precision(%)	Recall(%)	F1-Score(%)	Support
<i>Blast</i>	97,09	98,04	97,56	97,56	103
<i>Blight</i>	95,24	1.00	95,24	97,56	63
<i>Healthy</i>	99,20	97,67	1.00	98,82	168
<i>Overall</i>	98,20	98,23	98,20	98,20	334

Berdasarkan Tabel 13, Kelas *Healthy* masih menjadi kelas dengan performa paling unggul dengan nilai F1-Score mencapai 98,82%, serta nilai akurasi sebesar 99,20%. Kelas *Blast* mencatat performa stabil dengan F1-Score sebesar 97,56%. Peningkatan performa pada kelas *Blight* terjadi pada nilai F1-Score yang pada skenario baseline sebesar 96,00% menjadi 97,56% setelah penambahan augmentasi data.

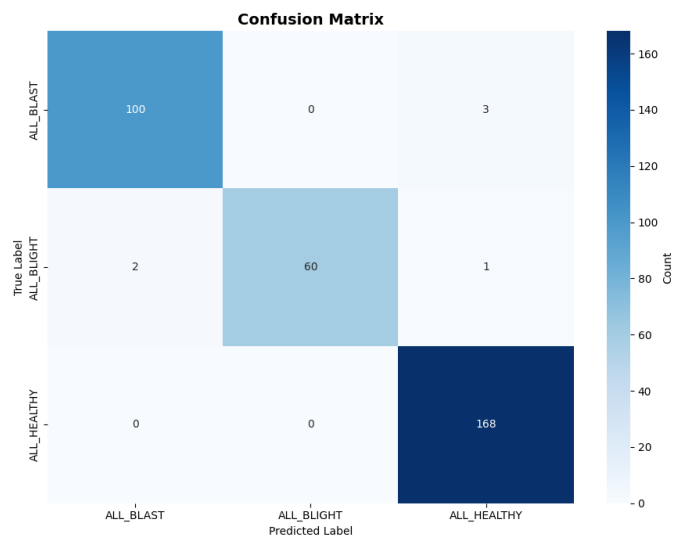
Untuk menganalisis proses konvergensi model selama pelatihan, kurva *learning* disajikan pada Gambar 10.



Gambar 10. Learning Curves ResNet50 Proposed

Terlihat pada tahap 1, model menunjukkan peningkatan akurasi dan penurunan *loss* yang stabil. Saat memasuki tahap 2 pada epoch ke-16, terdapat fluktuasi berupa lonjakan *loss* dan penurunan akurasi sesaat, di mana hal ini merupakan fase adaptasi ulang bobot backbone terhadap dataset setelah lapisan tersebut dibuka. Setelahnya, model pulih dengan cepat dan mencapai titik konvergensi stabil mulai epoch ke-20 hingga akhir pelatihan.

Distribusi hasil prediksi model secara lebih rinci dapat dilihat melalui *confusion matrix* pada Gambar 11.



Gambar 11. Confusion Matrix ResNet50 Proposed

Kelas Blast berhasil mengklasifikasi 100 sampel dengan benar, terdapat 3 sampel yang salah terprediksi sebagai *Blight*. Kelas *Blight* mampu mengenali 60 sampel secara akurat dengan 2 prediksi ke kelas Blast dan 1 ke kelas *Healthy*. Adapun pada kelas *Healthy* mencatat hasil yang stabil dengan memprediksi 168 sampel secara tepat.

3.13 Dampak Augmentasi GAN terhadap Performa EfficientNetB0

Pada skenario Baseline, model mencapai akurasi keseluruhan sebesar 97,60%, namun terdapat ketimpangan performa yang mencolok pada kelas minoritas. Kelas *Blight* hanya mampu mencapai nilai akurasi sebesar 95,24%, dengan nilai F1-Score rendah 94,49%, jauh di bawah kelas *Healthy* (99,40%) dan kelas Blast (96,62%). Ketimpangan ini sejalan dengan kondisi distribusi data awal, di mana kelas *Blight* hanya memiliki 499 sampel training dibandingkan dengan kelas *Healthy* yang memiliki 1333 sampel. Ketidakseimbangan kelas (class imbalanced) yang signifikan diketahui dapat menyebabkan model cenderung bias terhadap kelas mayoritas dan mengurangi sensitivitas terhadap kelas minoritas (Johnson & Khoshgoftaar, 2019) [6]. Setelah augmentasi GAN menyeimbangkan distribusi ketiga kelas

menjadi masing-masing 1333 sampel untuk *training*, model EfficientNetB0 *proposed* mencapai akurasi keseluruhan 98,20% (+0,6%), dengan peningkatan yang paling besar pada kelas *Blight* dari 96,83% menjadi 96,83% (+1,59%). Hal ini mengindikasikan bahwa citra sintetis yang dihasilkan StyleGAN2-ADA berhasil memberikan variasi fitur visual tambahan yang membantu model dalam mempelajari pola morfologi kelas *Blight* secara lebih komprehensif. Sementara itu, kelas *Blast* justru mengalami sedikit penurunan akurasi dari 97,09% menjadi 96,12%, namun dengan peningkatan nilai F1-Score dari 96,62% menjadi 97,06%.

Temuan ini konsisten dengan penelitian yang menunjukkan bahwa augmentasi berbasis GAN efektif dalam meningkatkan performa klasifikasi pada kelas minoritas, terutama ketika data latih asli memiliki distribusi yang tidak seimbang [10].

Perbandingan performa EfficientNetB0 antara kedua skenario disajikan secara ringkas pada Tabel 13.

Tabel 14. Perbandingan performa EfficientNetB0 Skenario A dan B

Kelas	F1-Score Baseline (%)	F1-Score Proposed (%)	Selisih (%)
<i>Blast</i>	96,62	97,06	+0,44
<i>Blight</i>	94,49	97,60	+3,11
<i>Healthy</i>	99,40	99,12	-0,28
<i>Overall</i>	97,62	98,20	+0,58

3.14 Dampak Augmentasi GAN terhadap Performa ResNet50

Pada skenario Baseline, ResNet50 mencapai akurasi keseluruhan yang tinggi, yaitu sebesar 97,60% dengan F1-Score keseluruhan 97,60%. Peningkatan terjadi setelah penambahan augmentasi data, yaitu nilai akurasi menjadi 98,20% dan nilai F1-Score meningkat ke 98,20%. Akurasi kelas *Blight* dari kedua skenario tidak mengalami perubahan, namun peningkatan terjadi pada nilai precision yang awalnya 97,77% menjadi 100% dan nilai F1-Score dari 96,00% menjadi 97,56%. Pada kelas *Blast* sendiri tidak terjadi perubahan pada semua metrik, yang menunjukkan bahwa performa kelas ini masih stabil.

Tabel 15. Hasil EfficientNetB0 *Proposed*

Kelas	F1-Score Baseline (%)	F1-Score Proposed (%)	Selisih (%)
<i>Blast</i>	97,56	97,56	-
<i>Blight</i>	96,00	97,56	+1,56
<i>Healthy</i>	98,22	98,82	+0,6
<i>Overall</i>	97,60	98,20	+0,6

3.15 Analisis Komparatif dan Implikasi Augmentasi GAN

Hasil penelitian ini menunjukkan bahwa efektivitas augmentasi berbasis GAN memberikan dampak positif pada kedua arsitektur yang diuji. EfficientNetB0 menunjukkan peningkatan F1-Score overall dari 97,62% menjadi 98,20%, sementara ResNet50 mengalami peningkatan dari 97,60% menjadi 98,20%, dengan kontribusi peningkatan yang terjadi juga pada kelas *Healthy*. Meskipun demikian, kelas *Blast* pada ResNet50% tidak menunjukkan perubahan performa setelah augmentasi diterapkan. Pola ini mengindikasikan bahwa *diminishing return* augmentasi sintetis tidak hanya bergantung pada arsitektur model, melainkan juga pada karakteristik masing-masing kelas, di mana hal ini selaras dengan argumen Shorten & Khoshgoftaar bahwa manfaat augmentasi menurun ketika distribusi data suatu kelas sudah cukup terwakili [6].

Perbedaan nilai FID antar kelas, *Blight* (79,08) dan *Blast* (103,20), mengisyaratkan bahwa kualitas citra sintetis bervariasi antar kelas dan berpotensi menjadi salah satu faktor yang memengaruhi konsistensi hasil augmentasi pada kelas *Blast* yang stagnan. Meskipun demikian, hubungan kausal antara nilai FID dengan performa klasifikasi tidak dapat disimpulkan hanya dari dua titik observasi ini.

Proses pelatihan GAN yang lebih lama atau seleksi citra sintetis yang lebih ketat berpotensi meningkatkan kualitas augmentasi [10].

Pada kelas *Blight* dengan sampel paling sedikit, augmentasi menunjukkan dampak positif terutama pada model EfficientNetB0. Hasil ini mengindikasikan potensi StyleGAN2-ADA sebagai solusi augmentasi dalam kondisi keterbatasan data untuk kelas tertentu, dengan catatan bahwa efektivitasnya bergantung pada model dan kualitas data sintetis yang dihasilkan.

4. KESIMPULAN

Arsitektur StyleGAN2-ADA berhasil diimplementasikan untuk menghasilkan citra sintetis penyakit daun padi yang realistis dan beragam. Model dilatih secara terpisah untuk dua kelas minoritas, yaitu *Blast* dan *Blight*, dengan memanfaatkan mekanisme Adaptive Discriminator Augmentation (ADA) yang memungkinkan pelatihan stabil meskipun pada kondisi terbatas. Kualitas citra sintetis dievaluasi menggunakan metrik Fréchet Inception Distance (FID), di mana kelas *Blight* mencapai nilai FID sebesar 79,08 dan kelas *Blast* sebesar 103,20. Nilai FID yang lebih rendah pada kelas *Blight* mengindikasikan bahwa distribusi visual citra sintetisnya lebih mendekati distribusi citra asli, sementara nilai FID yang tinggi pada kelas *Blast* menunjukkan masih adanya kesenjangan distribusi yang perlu ditingkatkan pada penelitian selanjutnya. Pada kelas *Blight* dengan sampel paling sedikit, augmentasi menunjukkan dampak positif terutama pada model EfficientNetB0. Hasil ini mengindikasikan potensi StyleGAN2-ADA sebagai solusi augmentasi dalam kondisi keterbatasan data untuk kelas tertentu, dengan catatan bahwa efektivitasnya bergantung pada model dan kualitas data sintetis yang dihasilkan.

Penerapan augmentasi StyleGAN2-ADA menunjukkan dampak yang berbeda antar model dan antar kelas. Pada EfficientNetB0, penambahan citra sintetis pada dua kelas yang dilatih dengan GAN menunjukkan dampak yang positif, dengan peningkatan F1-Score overall dari 97,62% menjadi 98,20%. Peningkatan nilai F1-Score terbesar pada kelas *Blight* dari yang 94,49% menjadi 97,60%. Pada model ResNet50, F1-Score overall juga meningkat dari 97,60% menjadi 98,20%, dengan kontribusi peningkatan yang terjadi juga pada kelas *Healthy*. Meskipun demikian, kelas *Blast* pada ResNet50 tidak menunjukkan perubahan performa setelah augmentasi diterapkan.

Berdasarkan temuan ini, augmentasi GAN menggunakan StyleGAN2-ADA terbukti efektif meningkatkan performa pada kedua model, meskipun peningkatan konsistensi yang berbeda antar kelas, dengan manfaat paling konsisten terlihat pada kelas minoritas *Blight*. Sementara itu, peningkatan pada kelas *Blast* tidak konsisten antar metrik, yang ditunjukkan oleh kenaikan F1-Score dan precision yang diiringi dengan penurunan recall pada EfficientNetB0. Hal mengindikasikan bahwa augmentasi sintetis dapat memengaruhi *trade-off* antara *precision* dan *recall* secara berbeda pada setiap kelas. Temuan ini menunjukkan bahwa pemilihan model terbaik tidak dapat disederhanakan hanya berdasarkan satu metrik agregat, melainkan perlu mempertimbangkan karakteristik performa per kelas sesuai kebutuhan aplikasi.

Daftar Pustaka

- [1] N. K. FUKAGAWA and L. H. ZISKA, "Rice: Importance for Global Nutrition," J. Nutr. Sci. Vitaminol. (Tokyo), vol. 65, no. Supplement, pp. S2-S3, Oct. 2019, doi: 10.3177/jnsv.65.S2.
- [2] F. Rozi et al., "Indonesian market demand patterns for food commodity sources of carbohydrates in facing the global food crisis," Heliyon, vol. 9, no. 6, Jun. 2023, doi: 10.1016/j.heliyon.2023.e16809.
- [3] Sekretariat Jenderal, "Statistik Iklim, Organisme Pengganggu Tanaman (OPT) dan Dampak Perubahan Iklim (DPI) Tahun 2023 - 2025," Kementerian Pertanian.
- [4] H. Semangun, Penyakit-penyakit tanaman pangan di Indonesia. Gadjah Mada University Press, 2008.
- [5] S. S. Gnanamanickam, V. Brindha Priyadarisini, N. N. Narayanan, and S. Kavitha, "An overview of bacterial Blight disease of rice and strategies for its management," 1999.
- [6] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," J. Big Data, vol. 6, no. 1, p. 60, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

- [7] I. Goodfellow et al., "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139-144, Oct. 2020, doi: 10.1145/3422622.
- [8] R. A. Saputra, S. Wasyianti, A. Supriyatna, and D. F. Saefudin, "Penerapan Algoritma Convolutional Neural Network Dan Arsitektur MobileNet Pada Aplikasi Deteksi Penyakit Daun Padi," *JURNAL SWABUMI*, vol. 9, no. 2, 2021, [Online]. Available: <https://archive.ics.uci.edu/ml/datasets/Rice>
- [9] Q. Zeng, X. Ma, B. Cheng, E. Zhou, and W. Pang, "GANS-based data augmentation for citrus disease severity detection using deep learning," *IEEE Access*, vol. 8, pp. 172882-172891, 2020, doi: 10.1109/ACCESS.2020.3025196.
- [10] Y. Haruna, S. Qin, and M. J. Mbyamm Kiki, "An Improved Approach to Detection of Rice Leaf Disease with GAN-Based Data Augmentation Pipeline," *Applied Sciences (Switzerland)*, vol. 13, no. 3, Feb. 2023, doi: 10.3390/app13031346.
- [11] T. Agustin, I. A. Saputro, and M. L. Rahmadi, "Optimizing Rice Plant Disease Classification Using Data Augmentation with GANs on Convolutional Neural Networks," *INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi*, vol. 9, no. 1, pp. 97-114, Feb. 2025, doi: 10.29407/intensif.v9i1.23834.
- [12] Sugiyono, *Metode penelitian kuantitatif, kualitatif, dan R&D*, 3rd ed. Bandung: Alfabeta, 2021 ©2021, 2021.
- [13] T. Karras NVIDIA and S. Laine NVIDIA, "A Style-Based Generator Architecture for Generative Adversarial Networks Timo Aila NVIDIA," 2019. [Online]. Available: <https://github.com/NVlabs/stylegan>
- [14] T. Karras, M. Aittala, J. Hellsten, S. Laine, J. Lehtinen, and T. Aila, "Training Generative Adversarial Networks with Limited Data," Oct. 2020, [Online]. Available: <http://arxiv.org/abs/2006.06676>
- [15] Y. Haruna, "Pre-trained StyleGAN2-ADA pytorch Rice Leaf Disease," Mendeley Data.
- [16] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861-874, Jun. 2006, doi: 10.1016/j.patrec.2005.10.010.
- [17] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," 2017.
- [18] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," Sep. 2020.