

Multi-Class Classification of Rice Leaf Diseases Using MobileNetV2 Architecture with Transfer Learning

Muhammad Bisri Mustofa^{1*}, Muhammad Arifin^{2*}, Shofiatuz Zulfia^{3*}

*Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Nahdlatul Ulama Sunan Giri
bisri3522@gmail.com¹, muhhammadarifin3054@gmail.com², shofiatuzzulfia@gmail.com³

Article Info

Article history:

Received 2026-02-15

Revised 2026-03-07

Accepted 2026-04-10

Keyword:

Deep Learning,
Image Classification,
Mobilenetv2,
Rice Leaf Disease,
Transfer Learning.

ABSTRACT

Rice leaf diseases are a major cause of yield loss and remain a persistent challenge in agriculture. Conventional diagnosis through visual inspection is subjective and time-consuming, necessitating an accurate and efficient automated system for early detection. This study develops and evaluates a multi-class rice leaf disease classifier based on digital images using MobileNetV2 with transfer learning and fine-tuning of the last 20 layers. The dataset comprises 6,180 images distributed equally across four classes: brown spot, leaf blast, leaf blight, and normal with an imbalance ratio of 1.00:1, split into 70% training, 15% validation, and 15% testing using stratified random sampling. The methodology incorporates image normalization, adaptive data augmentation, and class weighting as preventive regularization strategies. The model achieves 97.20% overall test accuracy, with precision of 97.30%, recall of 97.09%, F1-score of 97.19%, and AUC of 99.72%, converging optimally at epoch 25 across 30 training epochs. Per-class accuracy ranges from 96.12% (brown spot) to 97.84% (leaf blast and normal), with misclassifications primarily occurring between visually similar disease classes. With a compact model size of 9.89 MB (FP32) reducible to 2.71 MB via TFLite conversion, and an average inference time of 144.80 ms per image, the proposed model demonstrates high efficiency and strong generalization, making it well-suited for deployment on mobile devices as a practical tool for real-time early rice disease detection in field settings.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Padi (*Oryza sativa* L.) merupakan komoditas pangan utama di Indonesia dan menjadi sumber penghidupan bagi sebagian besar masyarakat yang bergantung pada sektor pertanian [1], [2]. Keberlanjutan produktivitas padi menjadi aspek krusial dalam menjaga ketahanan pangan nasional [3], [4]. Namun, produktivitas padi masih menghadapi berbagai tantangan, salah satunya adalah serangan penyakit daun yang dapat menyebabkan penurunan hasil panen secara signifikan [5].

Penyakit daun padi seperti *brown spot*, *leaf blast*, dan *leaf blight* dapat menghambat fotosintesis, merusak jaringan daun, dan mengganggu pertumbuhan tanaman [6]. Serangan penyakit tersebut tidak hanya menurunkan kualitas tanaman, tetapi juga berdampak langsung pada kuantitas hasil panen [7]. Deteksi dini menjadi kunci untuk mencegah penyebaran penyakit yang lebih luas dan mengurangi kerugian petani [8].

Identifikasi penyakit daun padi secara manual melalui pengamatan visual memiliki sejumlah keterbatasan, antara lain sangat bergantung pada pengalaman dan subjektivitas pengamat, memerlukan waktu yang relatif lama, serta berpotensi menimbulkan kesalahan akibat kemiripan gejala antar jenis penyakit [9], [10]. Kondisi ini menjadi kurang efektif, terutama pada lahan pertanian yang luas dengan intensitas serangan tinggi.

Perkembangan deep learning, khususnya *Convolutional Neural Network* (CNN), membuka peluang besar dalam pengembangan sistem otomatis untuk klasifikasi penyakit tanaman berbasis citra digital [11], [12]. Namun, arsitektur CNN konvensional umumnya memiliki jumlah parameter yang besar dan membutuhkan sumber daya komputasi tinggi, sehingga kurang optimal untuk perangkat dengan keterbatasan memori dan daya komputasi. Sebagai solusi, MobileNetV2 diperkenalkan sebagai arsitektur CNN yang dirancang khusus untuk perangkat dengan sumber daya

terbatas melalui teknik *depthwise separable convolution* dan *inverted residuals*, sehingga menghasilkan model yang lebih ringan dan efisien tanpa mengorbankan performa secara signifikan[13]. Dibandingkan arsitektur seperti ResNet-50 (~25 juta parameter) dan EfficientNet-B7 (~66 juta parameter), MobileNetV2 hanya memiliki sekitar 3,4 juta parameter dengan ukuran model yang jauh lebih kecil, sehingga lebih sesuai untuk skenario deployment pada perangkat mobile dengan keterbatasan sumber daya komputasi di lingkungan pertanian.

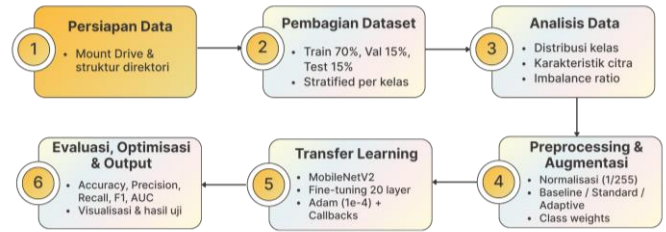
Penelitian terdahulu menunjukkan efektivitas MobileNetV2 dalam klasifikasi penyakit daun tanaman. Pada penyakit daun kentang, MobileNetV2 mencapai akurasi 95,31% dengan precision, recall, dan F1-score di atas 95% [14], sedangkan pada penyakit daun bawang mencapai akurasi pelatihan 99,71% dan validasi 97,37% [15]. Meskipun demikian, tantangan seperti variasi kondisi visual lapangan, risiko *overfitting*, dan kebutuhan akan generalisasi model pada data yang beragam masih menjadi perhatian utama dalam pengembangan sistem klasifikasi penyakit tanaman berbasis *deep learning*.

Berdasarkan uraian di atas, penelitian ini bertujuan untuk mengembangkan dan mengevaluasi sistem klasifikasi multi-kelas penyakit daun padi berbasis arsitektur MobileNetV2 dengan pendekatan *transfer learning*. Secara khusus, penelitian ini menggunakan Rice Leaf FN dataset yang tersedia pada repositori publik Kaggle [16] yang terdiri atas 6.180 citra daun padi dengan empat kelas (*brown spot*, *leaf blast*, *leaf blight*, dan daun normal) yang mencakup variasi kondisi visual lapangan untuk meningkatkan representativitas model. Model dilatih dengan menerapkan strategi *adaptive data augmentation* untuk meningkatkan keragaman data pelatihan dan mengurangi risiko *overfitting*. Kinerja model kemudian dievaluasi dari dua aspek utama, yaitu akurasi klasifikasi (*accuracy*, *precision*, *recall*, dan *F1-score*) serta efisiensi komputasi melalui pengukuran waktu inferensi guna menilai kelayakan implementasi pada perangkat mobile. Berbeda dari penelitian sebelumnya yang umumnya berfokus pada peningkatan akurasi tanpa mempertimbangkan efisiensi model secara bersamaan, penelitian ini mengoptimalkan trade-off antara akurasi klasifikasi dan kelayakan deployment melalui kombinasi augmentasi adaptif, fine-tuning selektif, dan konversi model ke format TFLite, sehingga menghasilkan model yang kompetitif secara performa sekaligus efisien untuk implementasi nyata pada perangkat mobile. Oleh karena itu, penelitian ini diharapkan dapat memberikan kontribusi terhadap pengembangan sistem deteksi penyakit tanaman berbasis *deep learning* yang tidak hanya akurat tetapi juga efisien untuk implementasi pada perangkat dengan sumber daya terbatas.

II. METODE

Alur metodologi penelitian ini disajikan secara ringkas dan sistematis dalam bentuk *flowchart*, sebagaimana ditunjukkan pada Gambar 1. *Flowchart* tersebut menggambarkan keseluruhan tahapan eksperimen yang dilakukan, mulai dari

persiapan data hingga evaluasi kinerja model klasifikasi citra berbasis *transfer learning*.

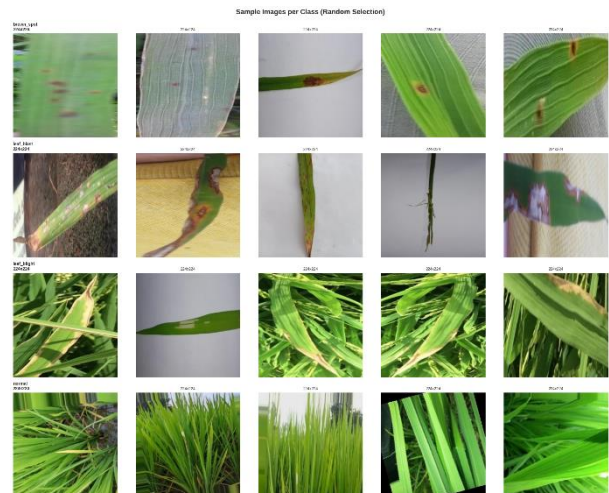


Gambar 1. Flowchart alur penelitian

A. Persiapan dan Sumber Data

Penelitian ini menggunakan dataset publik “Rice Leaf FN dataset” dari Kaggle [16], terdiri atas 6.180 citra daun padi dalam empat kelas: *brown spot*, *leaf blast*, *leaf blight*, dan daun normal. Citra dikumpulkan dalam kondisi lapangan nyata dengan variasi pencahayaan, latar belakang, dan orientasi, sehingga merepresentasikan tantangan visual riil. Dataset dimuat ke *Google Colab* melalui *Google Drive* dan disusun dalam struktur folder berlabel per kelas, memungkinkan integrasi langsung dengan *flow_from_directory* [17].

Sebelum pelatihan, dilakukan verifikasi format, resolusi (di-*resize* ke 224×224 piksel), dan distribusi kelas untuk meminimalkan bias [18]. Gambar 2 menampilkan contoh citra dari dataset untuk memberikan gambaran visual mengenai karakteristik setiap kelas penyakit yang digunakan dalam penelitian ini.



Gambar 2. Contoh citra dari Rice Leaf FN dataset

Citra dalam Gambar 2 menunjukkan variasi gejala penyakit yang khas: (a) *brown spot* dengan bercak coklat bulat, (b) *leaf blast* dengan lesi elips berwarna abu-coklat, (c) *leaf blight* dengan garis kuning memanjang sepanjang tulang daun, dan (d) daun normal tanpa gejala penyakit. Variasi latar belakang dan pencahayaan mencerminkan kondisi pengambilan citra lapangan nyata.

B. Pembagian Dataset

Dataset dibagi secara *stratified* menjadi 70% latih (4.324 citra), 15% validasi (928), dan 15% uji (928), dengan *random_state* = 42 untuk menjaga proporsi kelas. Partisi disimpan di direktori terpisah menggunakan *shutil.copy* guna mencegah *data leakage*. Data uji tidak digunakan selama pelatihan dan hanya untuk evaluasi akhir [19], [20]. Metode *single split* dipilih atas pertimbangan efisiensi komputasi dengan justifikasi bahwa distribusi kelas yang sepenuhnya seimbang meminimalkan risiko bias akibat partisi tunggal, meskipun penerapan *k-fold cross-validation* pada penelitian lanjutan tetap direkomendasikan.

C. Analisis Karakteristik Dataset

Analisis awal menunjukkan dataset memiliki variasi tinggi dalam pencahayaan, kontras, dan orientasi karakteristik umum pada citra pertanian yang dapat memengaruhi stabilitas pembelajaran CNN jika tidak ditangani [21]. Meski demikian, format dan resolusi citra konsisten, mendukung efisiensi pemrosesan tanpa transformasi berlebihan [22]. Hasil analisis ini menjadi dasar perancangan strategi *preprocessing* dan augmentasi.

D. Preprocessing dan Strategi Augmentasi Data

Seluruh citra dinormalisasi ke rentang [0,1] melalui *rescaling* $1/255$ [23]. Tiga strategi augmentasi diterapkan hanya pada data latih:

- Baseline: hanya normalisasi tanpa transformasi tambahan
- Standard: rotasi ($\pm 30^\circ$), pergeseran spasial ($\pm 20\%$), zoom (0,8–1,2 \times), flip horizontal dan vertikal, shear (15%), kecerahan (0,8–1,2 \times)
- Adaptive: rotasi ($\pm 40^\circ$), pergeseran spasial ($\pm 25\%$), zoom (0,75–1,25 \times), flip horizontal dan vertikal, shear (20%), channel shift ($\pm 10\%$), kecerahan (0,7–1,3 \times) [24].

Augmentasi bertujuan meningkatkan generalisasi tanpa merusak fitur patologis utama. Selain itu, *class weights* dihitung menggunakan pendekatan *balanced* dan diterapkan selama pelatihan. Karena distribusi kelas pada dataset ini bersifat *perfectly balanced*, seluruh kelas memperoleh bobot seragam sebesar 1,0000 sehingga tidak menghasilkan koreksi diferensial, namun tetap dipertahankan sebagai langkah preventif apabila diimplementasikan pada distribusi data yang berbeda [25]. Validasi kualitatif menunjukkan augmentasi adaptif menghasilkan variasi visual yang kaya namun tetap mempertahankan identitas kelas [26].

E. Transfer Learning Menggunakan MobileNetV2

Penelitian ini menerapkan pendekatan *transfer learning* dengan arsitektur MobileNetV2 sebagai model dasar untuk klasifikasi penyakit daun padi. Pemilihan MobileNetV2 didasarkan pada efisiensi komputasi, jumlah parameter yang relatif kecil, serta stabilitasnya pada tugas klasifikasi citra dengan keterbatasan sumber daya, khususnya pada data pertanian berbasis citra lapangan. MobileNetV2 dibangun

menggunakan konsep *depthwise separable convolution* dan *inverted residual blocks*, yang memungkinkan ekstraksi fitur visual secara efisien tanpa mengorbankan kemampuan representasi, menjadikannya sesuai untuk dataset dengan variasi visual tinggi namun ukuran data yang terbatas.

1) Konfigurasi Model Dasar MobileNetV2

Model MobileNetV2 dipilih berdasarkan tiga pertimbangan: (1) secara empiris, efisien menangkap fitur morfologis daun pada klasifikasi citra pertanian, (2) secara teoretis, *transfer learning* memanfaatkan pengetahuan visual umum untuk meningkatkan stabilitas pembelajaran pada dataset berukuran menengah, dan (3) secara praktis, mudah direproduksi dan kompatibel dengan keterbatasan komputasi [27]. Model digunakan dalam kondisi *pretrained* dengan bobot ImageNet, dengan 134 layer awal dibekukan untuk mempertahankan representasi fitur dasar, sementara lapisan klasifikasi kustom ditambahkan dan dilatih sesuai kelas target. Lapisan kustom tersebut terdiri atas *GlobalAveragePooling2D*, *BatchNormalization* dan *Dropout* (0,5), *Dense* (256 unit), *BatchNormalization* dan *Dropout* (0,3), serta *Dense* (4 unit, *softmax*) sebagai output klasifikasi, yang memungkinkan adaptasi domain spesifik tanpa mengganggu struktur fitur stabil pada layer awal [28]. Data latih diproses dengan augmentasi adaptif, sedangkan data validasi dan uji hanya melalui normalisasi tanpa transformasi tambahan untuk mencegah *data leakage*. *Class weights* diterapkan sebagai langkah preventif dengan nilai seragam 1,0000 karena distribusi dataset bersifat *perfectly balanced* [29].

2) Strategi Fine-tuning dan Hyperparameter

Pelatihan dilakukan dalam dua fase: (1) pelatihan awal dengan seluruh 134 layer konvolusional MobileNetV2 terbekukan, hanya melatih lapisan klasifikasi kustom yang ditambahkan, dan (2) *fine-tuning* dengan membuka 20 layer terakhir yang mencakup blok *inverted residual* tingkat tinggi untuk adaptasi fitur semantik spesifik terhadap karakteristik visual penyakit daun padi.

Model dilatih menggunakan optimizer Adam dengan *learning rate* 1×10^{-4} , dipilih karena kemampuan adaptasi *learning rate* per parameter yang mempercepat konvergensi pada proses *fine-tuning* [30]. Konfigurasi pelatihan menggunakan *batch size* 32 dengan maksimal 30 epoch, dimonitor dengan dua *callback* utama: (1) *EarlyStopping* dengan *patience* 3 epoch untuk menghentikan pelatihan apabila tidak terjadi peningkatan *validation accuracy*, dan (2) *ModelCheckpoint* untuk menyimpan bobot terbaik berdasarkan *validation accuracy* tertinggi. Seluruh proses pelatihan dilakukan pada lingkungan Google Colaboratory dengan penyimpanan dataset dan model melalui Google Drive.

F. Evaluasi Model Klasifikasi

Evaluasi dilakukan pada data uji tanpa augmentasi menggunakan lima metrik: (1) *accuracy*, (2) *precision*, (3)

recall, (4) *F1-score*, dan (5) *AUC (Area Under the ROC Curve)* untuk mengukur kemampuan diskriminasi model tanpa bergantung pada *threshold* klasifikasi. Metrik dihitung per kelas dan agregat (*macro-average*, *weighted-average*). *Confusion matrix* digunakan untuk menganalisis pola kesalahan antar kelas. Pendekatan multi-metrik memberikan penilaian komprehensif tanpa ketergantungan pada satu indikator.

G. Output dan Dokumentasi Eksperimen

Seluruh output eksperimen didokumentasikan secara sistematis untuk mendukung prinsip reproduktibilitas dan transparansi penelitian. Output meliputi model terlatih, konfigurasi *preprocessing* dan augmentasi, serta artefak evaluasi seperti *confusion matrix* dan laporan klasifikasi.

Parameter eksperimen disimpan dalam berkas konfigurasi terstruktur sehingga setiap tahapan dapat direplikasi dan ditelusuri kembali secara konsisten. Pendekatan ini memastikan bahwa keseluruhan proses eksperimen terdokumentasi dengan baik dan dapat digunakan untuk analisis lanjutan.

III. HASIL DAN PEMBAHASAN

A. Karakteristik Dataset

Dataset yang digunakan dalam penelitian ini adalah Rice Leaf Disease Dataset (Rice Leaf FN) yang bersumber dari repositori publik Kaggle [16] terdiri atas 6.180 citra daun padi berformat JPEG beresolusi 224×224 piksel yang terbagi ke dalam empat kelas: brown spot, leaf blast, leaf blight, dan normal. Citra mencakup variasi pencahayaan alami dan latar belakang yang beragam, merepresentasikan kondisi pengambilan gambar di lapangan secara nyata. Meskipun demikian, dataset ini belum mencakup variasi geografis lintas wilayah atau varietas padi yang berbeda, sehingga generalisasi model pada kondisi agroekologi yang sangat berbeda perlu diverifikasi lebih lanjut.

Setiap kelas memiliki 1.545 citra (rasio 1,00:1), mengkonfirmasi dataset bersifat perfectly balanced sehingga potensi bias akibat dominasi kelas tertentu dapat dieliminasi. Dari aspek kualitas visual, nilai kecerahan berada pada rentang 36,52–170,16 (rata-rata $125,24 \pm 21,67$) dan kontras berkisar 15,22–75,22 (rata-rata $41,39 \pm 11,72$), mengindikasikan variabilitas kondisi pencahayaan dan tekstur antarcitra yang berpotensi memengaruhi stabilitas ekstraksi fitur CNN dalam membedakan gejala penyakit dari elemen visual non-relevan. Ukuran berkas rata-rata 14,8 KB mendukung efisiensi pemrosesan pada lingkungan komputasi terbatas.

Dataset dibagi menggunakan stratified random sampling menjadi data latih (70%), validasi (15%), dan uji (15%) dengan *random_state* = 42. Metode single split dipilih atas pertimbangan efisiensi komputasi, dengan justifikasi bahwa distribusi kelas yang sepenuhnya seimbang meminimalkan risiko bias akibat partisi tunggal; meskipun demikian, penerapan k-fold cross-validation pada penelitian lanjutan tetap direkomendasikan untuk estimasi generalisasi yang

lebih robust. Data uji disimpan terpisah tanpa augmentasi guna memastikan evaluasi generalisasi model berlangsung secara objektif. Rincian distribusi dataset disajikan pada Tabel I.

TABEL I
DISTRIBUSI DATASET HASIL STRATIFIED RANDOM SAMPLING

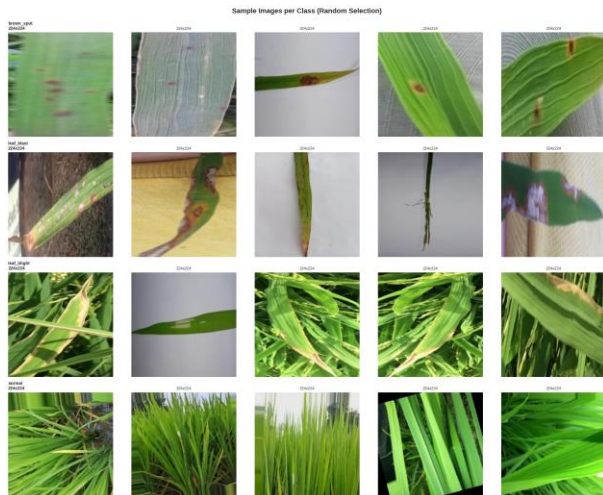
No	Class	Train	Val	Test	Total
1	Brown Spot	1.081	232	232	1.545
2	Leaf Blast	1.081	232	232	1.545
3	Leaf Blight	1.081	232	232	1.545
4	Normal	1.081	232	232	1.545
Total		4.324	928	928	6.180

B. Preprocessing dan Perbandingan Strategi Augmentasi

Meskipun dataset relatif seimbang secara kuantitatif, kompleksitas visual citra daun padi yang dikumpulkan pada kondisi lapangan nyata menuntut penerapan strategi preprocessing dan augmentasi yang tepat. Variasi pencahayaan, sudut pengambilan gambar, dan latar belakang berpotensi memengaruhi konsistensi fitur visual yang dipelajari oleh model CNN.

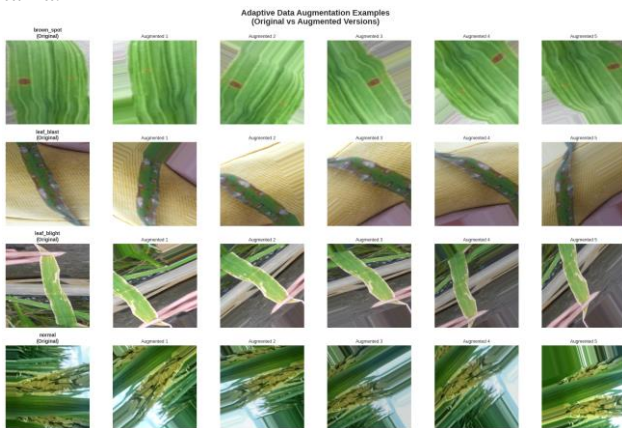
Seluruh citra dinormalisasi ke rentang nilai piksel [0,1] melalui rescaling 1/255, dengan resolusi input dipertahankan pada 224×224 piksel dan batch size 32 sesuai kebutuhan arsitektur MobileNetV2. Untuk menjaga objektivitas evaluasi, augmentasi hanya diterapkan pada data latih, sedangkan data validasi dan uji digunakan dalam kondisi asli tanpa transformasi tambahan.

Tiga strategi preprocessing diterapkan: baseline (tanpa augmentasi) sebagai pembanding, augmentasi standar dengan rotasi $\pm 30^\circ$, pergeseran spasial $\pm 20\%$, shear range 15%, zoom 0,8–1,2×, horizontal dan vertical flip, serta variasi kecerahan 0,8–1,2×; dan augmentasi adaptif dengan parameter yang lebih luas, yaitu rotasi $\pm 40^\circ$, pergeseran spasial $\pm 25\%$, shear range 20%, zoom 0,75–1,25×, variasi kecerahan 0,7–1,3×, dan channel shift $\pm 10^\circ$, guna meningkatkan ketahanan model terhadap kondisi visual ekstrem di lapangan. Karena dataset bersifat perfectly balanced dengan distribusi merata 25% per kelas, augmentasi adaptif diterapkan secara seragam pada seluruh kelas (intensitas 1,00×) dengan class weights bernilai 1,0000, sehingga tidak memberikan pengaruh korektif yang berbeda antarkelas. Perbandingan efek ketiga strategi ditunjukkan pada Gambar 3, sedangkan contoh hasil augmentasi adaptif pada seluruh kelas ditampilkan pada Gambar 4, yang memperlihatkan peningkatan keragaman visual tanpa menghilangkan karakteristik patologis utama masing-masing kelas.



Gambar 3. Perbandingan Efek Tiga Strategi Augmentasi pada Sampel Kelas brown spot

Efek dari ketiga strategi augmentasi tersebut ditunjukkan pada Gambar 3, yang memperlihatkan perbandingan citra asli, hasil tanpa augmentasi, augmentasi standar, dan augmentasi adaptif pada satu sampel kelas *brown spot*. Terlihat bahwa augmentasi adaptif menghasilkan variasi orientasi, pencahayaan, dan pergeseran kanal warna yang lebih beragam dibandingkan strategi lain, namun tetap mempertahankan pola lesi penyakit sebagai informasi visual utama.



Gambar 4. Contoh Hasil Augmentasi Adaptif pada Keempat Kelas Dataset

Sementara itu, Gambar 4 menampilkan contoh augmentasi adaptif pada seluruh kelas dataset, yaitu *brown spot*, *leaf blast*, *leaf blight*, dan *normal*. Visualisasi ini menunjukkan bahwa pendekatan augmentasi adaptif mampu diterapkan secara konsisten pada berbagai karakteristik penyakit maupun daun sehat, tanpa menimbulkan distorsi yang berpotensi mengaburkan identitas kelas.

C. Hasil Klasifikasi Menggunakan MobileNetV2

Kinerja model MobileNetV2 dengan augmentasi adaptif diukur menggunakan data uji sebanyak 928 citra. Hasil pengujian menunjukkan bahwa model mencapai akurasi keseluruhan sebesar 97,20%, yang mengindikasikan

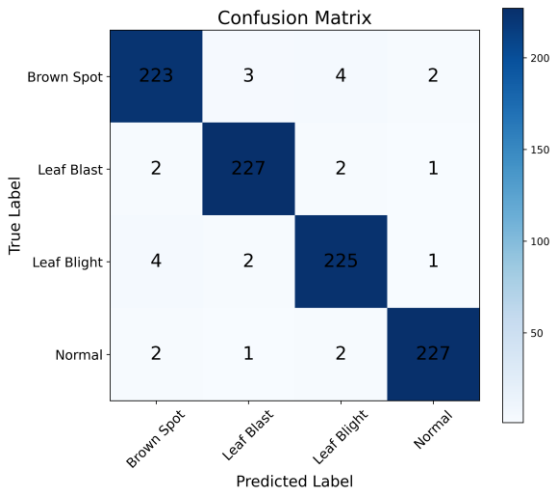
kemampuan klasifikasi yang sangat baik pada kondisi data lapangan yang kompleks.

Proses fine-tuning dilakukan dengan membekukan seluruh layer MobileNetV2 kecuali 20 layer terakhir yang dibiarkan dapat diperbarui bobotnya. Strategi ini dipilih untuk mempertahankan representasi fitur umum yang telah dipelajari dari dataset ImageNet pada layer awal, sekaligus mengadaptasi fitur tingkat tinggi terhadap karakteristik visual spesifik penyakit daun padi. Model dilatih menggunakan optimizier Adam dengan learning rate konstan 1×10^{-4} , dengan ModelCheckpoint menyimpan bobot terbaik pada epoch ke-25 (val_accuracy 97,41%) dan EarlyStopping dengan patience 3 epoch sebagai pengaman konvergensi. Pendekatan ini membantu proses optimasi mencapai konvergensi yang lebih stabil serta mengurangi risiko osilasi parameter pada fase akhir pelatihan.



Gambar 5. Kurva Pelatihan Model MobileNetV2 dengan Augmentasi Adaptif dan Fine-Tuning Selektif

Kurva pelatihan yang ditunjukkan pada Gambar 5 memperlihatkan peningkatan performa yang stabil pada metrik akurasi, presisi, recall, dan AUC baik pada data latih maupun validasi. Model berhasil mencapai val_accuracy tertinggi sebesar 97,41% pada epoch ke 25, dengan val_AUC mencapai 0,9979 dan val_loss yang terus menurun hingga 0,4554. Perbedaan yang kecil antara kurva latih dan validasi serta tren penurunan nilai loss yang konsisten mengindikasikan bahwa model memiliki kemampuan generalisasi yang baik dan tidak mengalami overfitting secara signifikan. Fluktuasi minor yang terlihat pada beberapa epoch akhir merupakan respons wajar terhadap dinamika optimasi, namun model berhasil mempertahankan konvergensi yang stabil sepanjang 30 epoch pelatihan.



Gambar 6. Confusion Matrix Hasil Klasifikasi pada Data Uji Menggunakan MobileNetV2

Analisis lebih rinci terhadap hasil prediksi pada data uji disajikan melalui *confusion matrix* pada Gambar 6. Hasil tersebut menunjukkan bahwa kelas *leaf blast* dan kelas normal sama-sama mencapai akurasi per kelas tertinggi sebesar 97,84% dengan masing-masing hanya 5 sampel yang salah diklasifikasikan dari total 232 sampel. Kelas *leaf blast* terbukti paling mudah dikenali oleh model berkat karakteristik visualnya yang khas dan berbeda dari kelas lainnya, sehingga menghasilkan nilai recall sebesar 97,84% dan presisi sekitar 97%. Kelas normal juga menunjukkan performa yang sangat baik dengan presisi sekitar 98% dan recall 97,84%, yang berarti model mampu membedakan daun sehat dari daun yang terinfeksi secara konsisten sebuah pencapaian penting dalam konteks deteksi dini agar tidak terjadi kesalahan penanganan di lapangan.

Sementara itu, kelas *brown spot* mencatat akurasi per kelas terendah sebesar 96,12% dengan 9 sampel yang salah diklasifikasikan, di mana sebagian di antaranya terkonfusi dengan kelas *leaf blight* akibat adanya kemiripan pola bercak secara visual antara kedua jenis penyakit tersebut. Kelas *leaf blight* sendiri memperoleh akurasi per kelas sebesar 96,98% dengan 7 sampel yang keliru diprediksi, sebagian di antaranya juga terkonfusi dengan *brown spot*, mengonfirmasi adanya tumpang tindih karakteristik gejala di antara kedua kelas tersebut.

TABEL II
PERFORMA KLASIFIKASI PER KELAS MENGGUNAKAN MOBILENETV2

Kelas	Precision (%)	Recall (%)	F1-Score (%)	Support
Brown Spot	98,00	96,12	97,05	232
Leaf Blast	97,00	97,84	97,42	232
Leaf Blight	96,00	96,98	96,49	232
Normal	98,00	97,84	97,92	232
Overall Accuracy	—	97,20%	—	928

Untuk menilai stabilitas hasil evaluasi, perhitungan confidence interval 95% dilakukan terhadap akurasi keseluruhan model pada data uji menggunakan metode

Wilson score interval. Dengan akurasi 97,20% pada 928 sampel uji, diperoleh confidence interval sebesar [95,97%, 98,14%], yang mengindikasikan bahwa kinerja model bersifat konsisten dan tidak bergantung pada fluktuasi partisi data tertentu. Nilai ini memperkuat validitas hasil evaluasi yang dilaporkan.

Untuk memberikan gambaran performa yang lebih spesifik, Tabel II menyajikan metrik klasifikasi per kelas. Kelas *leaf blast* dan normal sama-sama mencapai recall tertinggi sebesar 97,84%, diikuti oleh kelas *leaf blight* sebesar 96,98%, sementara kelas *brown spot* memperoleh recall terendah sebesar 96,12%. Dari sisi presisi, kelas *brown spot* dan normal mencapai nilai tertinggi sekitar 98%, yang menandakan minimnya *false positive* pada kedua kelas tersebut. Nilai *precision* keseluruhan pada test set sebesar 97,30% dan *recall* sebesar 97,09% mencerminkan keseimbangan yang baik antara kemampuan deteksi dan ketepatan prediksi model. Nilai AUC pada test set sebesar 0,9972 mengkonfirmasi kemampuan diskriminasi model yang sangat tinggi secara keseluruhan.

Dengan akurasi test set sebesar 97,20%, ukuran model 9,89 MB (FP32) dan 2,71 MB (TFLite), serta waktu inferensi rata-rata 144,80 ms per citra yang diukur pada CPU Intel Core i3 generasi ke-10 dengan RAM 8 GB tanpa akselerasi GPU, model yang diusulkan menunjukkan kelayakan teknis yang tinggi untuk implementasi berbasis perangkat mobile. Pengukuran dilakukan menggunakan model hasil konversi TensorFlow Lite sehingga merepresentasikan skenario komputasi ringan yang mendekati kondisi implementasi nyata.

TABEL III
PERBANDINGAN EFISIENSI DAN KINERJA MODEL

Peneliti	Arsitektur	Kelas	Akurasi (%)	Ukuran Model	Parameter
Ghosal & Sarkar (2020)	VGG16-based CNN	4	92,46	~528 MB	~138 M
Sahputra et al. (2025)	MobileNet	5	95,00	~16 MB	~4,2 M
Sobuj et al. (2024)	EfficientNet-B7 + HOG	4	97,00	~256 MB	~66 M

Tabel III menyajikan perbandingan kinerja dan efisiensi model yang diusulkan terhadap penelitian terdahulu berdasarkan akurasi, ukuran model, dan jumlah parameter. Ghosal dan Sarkar menggunakan arsitektur VGG16-based CNN dan memperoleh akurasi 92,46%, namun dengan kompleksitas sangat tinggi, yaitu sekitar 138 juta parameter dan ukuran model ~528 MB sehingga kurang sesuai untuk implementasi pada perangkat mobile [30]. Sahputra et al. menunjukkan bahwa arsitektur MobileNet mampu mencapai akurasi 95,00% dengan kompleksitas lebih rendah (~4,2 juta parameter dan ukuran model ~16 MB), yang menegaskan keunggulan arsitektur lightweight CNN dalam klasifikasi penyakit tanaman [31]. Sementara itu, Sobuj et al. mengombinasikan EfficientNet-B7 dengan fitur HOG dan

mencapai akurasi 97,00%, namun dengan kompleksitas yang jauh lebih besar (~66 juta parameter dan ukuran model ~256 MB) sehingga kurang efisien untuk deployment pada perangkat terbatas [32].

Model yang diusulkan, yaitu MobileNetV2 dengan augmentasi adaptif, mencapai akurasi tertinggi sebesar 97,20% dengan AUC 0,9972 hanya menggunakan ~3,4 juta parameter dan ukuran model 2,71 MB dalam format TFLite. Dibandingkan EfficientNet-B7 dan VGG16, model ini masing-masing sekitar 94 kali dan 195 kali lebih ringan dengan selisih akurasi yang minimal. Dengan waktu inferensi rata-rata 144,80 ms per citra, kombinasi MobileNetV2, augmentasi adaptif, dan fine-tuning selektif pada 20 layer terakhir terbukti mampu mengoptimalkan keseimbangan antara akurasi dan efisiensi, sehingga lebih layak untuk implementasi deteksi dini penyakit daun padi berbasis perangkat mobile. Temuan ini menegaskan bahwa arsitektur lightweight tidak harus mengorbankan akurasi, dan bahwa strategi augmentasi adaptif berperan signifikan dalam memaksimalkan potensi MobileNetV2 pada dataset kondisi lapangan.

D. Keterbatasan Penelitian dan Saran untuk Penelitian Mendatang

Meskipun model MobileNetV2 yang diusulkan mampu mencapai kinerja klasifikasi yang baik pada empat kelas penyakit daun padi, penelitian ini masih memiliki beberapa keterbatasan. Pertama, seluruh data berasal dari satu sumber publik (*Rice Leaf FN dataset* dari Kaggle), sehingga variasi geografis, varietas padi, kondisi pencahayaan alami, dan latar belakang lingkungan belum sepenuhnya terwakili. Kondisi ini berpotensi memengaruhi kemampuan generalisasi model ketika diterapkan pada skenario lapangan yang lebih beragam. Kedua, penelitian ini hanya memanfaatkan satu arsitektur deep learning, yaitu MobileNetV2, tanpa melakukan perbandingan langsung dengan arsitektur lain seperti EfficientNet, ResNet, atau *Vision Transformer*, sehingga keunggulan relatif model terhadap pendekatan lain belum dapat dievaluasi secara komprehensif. Ketiga, model belum diuji pada perangkat Android aktual sehingga performa inferensi dalam kondisi deployment nyata, termasuk variasi perangkat keras, sistem operasi, dan kondisi pencahayaan saat pengambilan citra langsung, belum dapat dikonfirmasi sepenuhnya.

Berdasarkan keterbatasan tersebut, penelitian selanjutnya disarankan untuk memperluas dan memperkaya dataset dengan citra daun padi yang diperoleh dari berbagai lokasi, musim tanam, dan kondisi lingkungan guna meningkatkan robustness model. Selain itu, evaluasi komparatif dengan beberapa arsitektur deep learning lainnya serta penerapan teknik optimasi lanjutan, seperti fine-tuning adaptif dan augmentasi berbasis kondisi lapangan, perlu dilakukan untuk memperoleh model dengan kinerja dan generalisasi yang lebih optimal. Pengembangan sistem deteksi berbasis perangkat mobile atau edge computing juga menjadi arah penelitian yang potensial agar model dapat

diimplementasikan secara real-time dan praktis bagi petani di lapangan.

IV. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem klasifikasi multi-kelas penyakit daun padi menggunakan arsitektur MobileNetV2 dengan augmentasi adaptif dan fine-tuning selektif. Model menghasilkan akurasi 97,20%, presisi 97,30%, recall 97,09%, F1-Score 97,19%, dan AUC 0,9972 pada data uji sebanyak 928 citra. Analisis confusion matrix menunjukkan bahwa kesalahan klasifikasi terdistribusi relatif merata tanpa bias kelas, dengan kerancuan utama terjadi antara brown spot dan leaf blight akibat kemiripan karakteristik visual intrinsik pada pola bercak daun. Model mencapai konvergensi optimal pada epoch ke-25 dengan nilai *val_accuracy* 97,41% dan *val_AUC* 0,9979 tanpa indikasi overfitting yang signifikan. Dari sisi efisiensi, model memiliki ukuran 9,89 MB (FP32) dan berkurang menjadi 2,71 MB setelah konversi ke format TFLite (kompresi 72,6%), dengan waktu inferensi rata-rata 144,80 ms per citra pada CPU Intel Core i3 generasi ke-10 tanpa akselerasi GPU. Hasil ini menunjukkan bahwa kombinasi MobileNetV2, augmentasi adaptif, dan fine-tuning selektif mampu mengoptimalkan trade-off antara akurasi klasifikasi dan efisiensi komputasi, sehingga model layak diimplementasikan sebagai sistem deteksi dini penyakit daun padi pada perangkat mobile di lingkungan pertanian.

DAFTAR PUSTAKA

- [1] C. Leiwakabessy, F. Inayati, E. Jambormias, J. Patty, and R. E. Ririhena, "Ketahanan Enam Varietas Padi Terhadap Penyakit Blas (*Pyricularia oryzae* Cav.) pada Lahan Sawah Irigasi dan Sawah Tadah Hujan," *Jurnal Budidaya Pertanian*, vol. 16, no. 2, pp. 147–156, 2020, doi: 10.30598/jbdp.2020.16.2.147.
- [2] B. H. Isnawan and A. N. Gumilar, "Selection of Intermittent Irrigation to Increase Growth and Yield of Some Local Rice Varieties (*Oryza sativa* L.) in the Rainy Season," *Jurnal Ilmiah Multidisiplin Indonesia*, vol. 2, 2023, doi: 10.58471/esaprom.
- [3] M. Irvai and D. Mahdalena, "Penerapan Teknik Masking dan Augmentasi pada Arsitektur CNN dan MobileNetV2 untuk Klasifikasi Fase Pertumbuhan Tomat," *Jurnal Teknik Informatika Unika ST. Thomas (JTIUST)*, vol. 10, no. 1, pp. 60–67, 2025.
- [4] R. A. Safitri and E. Suhartono, "The Effect of Harvest Area, Rice Productivity, and Population Density on Food Security in NTT Province 2018–2023," *Jambura Equilibrium Journal*, vol. 7, no. 2, pp. 71–78, 2025, doi: 10.37479.
- [5] Ratnasari and D. Feriyanto, "Deteksi Penyakit Daun Padi Menggunakan Deep Learning untuk Mendukung Produktivitas dan Pertanian Berkelanjutan," *Jurnal Algoritma*, vol. 22, no. 2, pp. 716–727, 2025, doi: 10.33364/algoritma/v.22-2.2900.
- [6] B. W. Sari, D. Prabowo, Y. Prityanto, and A. Aminuddin, "Transfer Learning-Based Convolutional Neural Network for Accurate Detection of Rice Leaf Disease in Precision Agriculture," *Journal of Information Systems Engineering and Business Intelligence*, vol. 11, no. 3, pp. 420–432, 2025, doi: 10.20473/jisebi.11.3.420-432.
- [7] M. E. Rana, V. A. Hameed, I. K. Y. Eng, H. K. Tripathy, and S. Mallik, "Harnessing Artificial Intelligence for Sustainable Rice Leaf Disease Classification," *Front. Plant Sci.*, vol. 16, pp. 1–27, 2025, doi: 10.3389/fpls.2025.1594329.
- [8] P. Novantara, R. L. Firmansyah, and M. Arismawati, "Deteksi Hama Penyakit Daun Padi Dengan Menggunakan Teknik Optimasi

- Deep Learning Convolutional Neural Network,” *bit-Tech*, vol. 7, no. 3, pp. 975–983, 2025, doi: 10.32877/bt.v7i3.2284.
- [9] D. Lizard, S. Dimara, B. Rahmat, and H. Maulana, “Identifikasi Penyakit Daun Padi dengan Metode Transfer Learning MobileNet–Support Vector Machine,” 2025.
- [10] R. Suciani, D. A. Anugra, and E. Faisal, “Deteksi Penyakit Daun Padi Menggunakan Deep Learning dengan Arsitektur CNN,” *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 8, no. 5, pp. 1451–1463, 2025, doi: 10.23960/jitc.v8i5.1451.
- [11] B. N. Kumar and S. Sakthivel, “Rice Leaf Disease Classification Using a Fusion Vision Approach,” *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-87800-3.
- [12] N. Istiqomah and M. Murinto, “Klasifikasi Penyakit Tanaman Padi Berbasis Citra Daun Menggunakan Convolutional Neural Network (CNN),” *JSTIE (Jurnal Sarjana Teknik Informatika)*, vol. 12, no. 1, p. 18, 2024, doi: 10.12928/jstie.v12i1.27314.
- [13] D. A. Pangestu, O. Q. Aziz, and C. Crysdian, “Klasifikasi Penyakit pada Tanaman Berdasarkan Citra Daun Menggunakan Metode Convolutional Neural Network,” *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 10, no. 2, pp. 235–248, 2025.
- [14] R. Gunawan, F. Salim, A. I. Wahyudhy, A. Y. Wibowo, G. Yordan, and R. F. Filamori, “Klasifikasi Penyakit Daun Kentang dengan Transfer Learning Menggunakan CNN Optimalisasi Arsitektur MobileNetV2,” *COSCITECH*, vol. 6, no. 2, pp. 254–258, 2025, doi: 10.37859/coscitech.v6i2.8599.
- [15] D. P. Pamungkas and M. F. Amrulloh, “Analisis Hasil Klasifikasi Penyakit Daun Bawang Merah Menggunakan CNN Arsitektur Exception,” *JUPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 10, no. 1, pp. 359–366, 2025, doi: 10.29100/jupi.v10i1.5875.
- [16] thntrngtnh, “Data Rice Leaf FN,” 2020. [Online]. Available: <https://www.kaggle.com/datasets/thntrngtnh/data-rice-leaf-fn>
- [17] G. Gunawan, S. A. Wibowo, and W. Andriani, “Evaluasi Model Deep Learning pada Pola Dataset Biomedis,” *Jurnal Sainstekom: Sains, Teknologi, Komputer dan Manajemen*, vol. 14, no. 2, pp. 195–207, 2024, doi: 10.33020/sainstekom.v14i2.738.
- [18] F. S. Maulana, V. S. Kusuma, W. M. Baihaqi, and P. Subarkah, “Optimasi Klasifikasi Gaya Belajar Mahasiswa Inklusif Berdasarkan Model VAK dengan Stratified Split dan Multilayer Perceptron,” *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 12, no. 5, pp. 1163–1172, 2025.
- [19] M. Purba, V. Ayumi, S. Rahayu, U. Salamah, I. Handriani, and I. Farida, “Penerapan Metode Random Contrast dan Brightness Range Pada Dataset Citra Daun Fitomedisin,” *JSAL: Journal Scientific and Applied Informatics*, vol. 8, no. 2, pp. 582–587, 2025, doi: 10.36085/jsai.v8i2.582.
- [20] M. Ilham, S. Ernawati, and M. Indra, “Improving Image Classification Accuracy with Oversampling and Data Augmentation Using Deep Learning,” *Jurnal Riset Informatika*, vol. 6, no. 4, pp. 201–209, 2024, doi: 10.34288/jri.v6i4.348.
- [21] M. Rijal, A. M. Yani, and A. Rahman, “Deteksi Citra Daun untuk Klasifikasi Penyakit Padi Menggunakan Pendekatan Deep Learning dengan Model CNN,” *Jurnal Teknologi Terpadu*, vol. 10, no. 1, pp. 56–62, 2024.
- [22] R. Kusumastuti, T. D. Putra, and Z. Z. Yudam, “Klasifikasi Citra Penyakit Daun Jagung Menggunakan Algoritma CNN EfficientNet,” *Multitek Indonesia Jurnal Ilmiah*, vol. 17, no. 2, pp. 143–153, 2023.
- [23] K. Prasetyo and R. Mahendra, “Performance Analysis of Baseline Convolutional Neural Networks for Identifying Potato Disease Types,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 5, no. 2, pp. 609–615, Apr. 2025, doi: 10.57152/malcom.v5i2.1722.
- [24] A. N. Zulfa, J. Jasril, M. Irsyad, F. Yanto, and S. Sanjaya, “Optimasi Convolutional Neural Network NASNetLarge Menggunakan Augmentasi Data untuk Klasifikasi Citra Penyakit Daun Padi,” *Jurnal Media Informatika Budidarma*, vol. 7, no. 2, pp. 696–706, Apr. 2023, doi: 10.30865/mib.v7i2.6056.
- [25] A. Bahtiar, M. I. P. Hutomo, A. Widiyanto, and S. Khomsah, “Class Weighting Approach for Handling Imbalanced Data on Forest Fire Classification Using Efficient-Net-B1,” *Jurnal Informatika Sunan Kalijaga (JISKA)*, vol. 10, no. 1, pp. 63–73, Jan. 2025, doi: 10.25181/jiska.v10i1.2987.
- [26] Sariah, N. Suarna, I. Ali, and D. Solihudin, “Penerapan Convolutional Neural Network (CNN) untuk Prediksi Penyakit Tanaman Padi melalui Citra Daun,” *Jurnal Komtika (Komputasi dan Informatika)*, vol. 9, no. 1, 2025, doi: 10.31603/komtika.v9i1.12852.
- [27] H. P. Arianda and T. Y. Hadiwandura, “Implementation of Deep Learning for Disease Classification in Oil Palm Leaves Using the MobileNetV2 Architecture,” *Jurnal Computer Science and Information Technology (CoSciTech)*, vol. 6, no. 3, pp. 453–462, 2025, doi: 10.37859/coscitech.v6i3.10306.
- [28] O. V. Putra, M. Z. Mustaqim, and D. Muriatmoko, “Transfer Learning untuk Klasifikasi Penyakit dan Hama Padi Menggunakan MobileNetV2,” *Techno.COM*, vol. 22, no. 3, pp. 562–575, 2023, doi: 10.33508/techno.v22i3.562-575.
- [29] T. B. Sasongko, Haryoko, and A. Amrullah, “Analisis Efek Augmentasi Dataset dan Fine-Tune pada Algoritma Pre-Trained Convolutional Neural Network (CNN),” *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIK)*, vol. 10, no. 4, pp. 763–768, 2023, doi: 10.25126/jtiik.2023106583.
- [30] S. Ghosal and K. Sarkar, “Rice Leaf Diseases Classification Using CNN With Transfer Learning,” in *2020 IEEE Calcutta Conference (CALCON)*, IEEE, 2020. doi: 10.1109/CALCON49167.2020.9106423.
- [31] I. Sahputra, A. F. Ulva, B. A. Putri, and C. Y. Eviyan, “Comparative Study of VGG16 and MobileNet Architectures for Rice Leaf Disease Classification Using CNN,” *Journal of Informatics and Telecommunication Engineering*, 2025.
- [32] Md. S. I. Sobuj, Md. I. Hossen, and Md. F. Mahmud, “Leveraging Pre-trained CNNs for Efficient Feature Extraction in Rice Leaf Disease Classification,” *arXiv preprint*, 2024.