

**PENGEMBANGAN ARSITEKTUR HIBRIDA EFFICIENTNET-
VIT DENGAN EXPLAINABLE AI (XAI) UNTUK DETEKSI
PENYAKIT TANAMAN PISANG**

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1-Informatika



disusun oleh

RASYIF IBNU HASRI

22.11.4710

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**PENGEMBANGAN ARSITEKTUR HIBRIDA EFFICIENTNET-
VIT DENGAN EXPLAINABLE AI (XAI) UNTUK DETEKSI
PENYAKIT TANAMAN PISANG**

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi S1-Informatika



disusun oleh

RASYIF IBNU HASRI

22.11.4710

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

SKRIPSI

**PENGEMBANGAN ARSITEKTUR HIBRIDA EFFICIENTNET-VTT
DENGAN EXPLAINABLE AI (XAI) UNTUK DETEKSI PENYAKIT
TANAMAN PISANG**

yang disusun dan diajukan oleh

RASYTE IBNU HASRI

22.11.4710

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 18 Juni 2026

Dosen Pembimbing,



Ferian Fauzi Abdulloh, S.Kom., M.Kom.

NIK. 190302276

HALAMAN PENGESAHAN
SKRIPSI
PENGEMBANGAN ARSITEKTUR HIBRIDA EFFICIENTNET-VIT
DENGAN EXPLAINABLE AI (XAI) UNTUK DETEKSI PENYAKIT
TANAMAN PISANG

yang disusun dan diajukan oleh

RASYIF IBNU HASRI

22.11.4710

Telah dipertahankan di depan Dewan Penguji
pada tanggal 18 Juni 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Yudi Sutanto, S.Kom., M.Kom.,
NIK. 190302039

Yudha Riwanto, M.Kom.,
NIK. 190302702

Ferian Fauzi Abdulloh, S.Kom., M.Kom.,
NIK. 190302276



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 18 Juni 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.,
NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Rasyif Ibnu Hasri
NIM : 22.11.4710

Menyatakan bahwa Skripsi dengan judul berikut:

**PENGEMBANGAN ARSITEKTUR HIBRIDA EFFICIENTNET-VIT
DENGAN EXPLAINABLE AI (XAI) UNTUK DETEKSI PENYAKIT
TANAMAN PISANG**

Dosen Pembimbing : Ferian Fauzi Abdulloh, S.Kom., M.Kom.

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

Yogyakarta, 18 Juni 2026

Yang Menyatakan,



Rasyif Ibnu Hasri

HALAMAN PERSEMBAHAN

Puji dan syukur senantiasa penulis panjatkan ke hadirat Allah SWT atas segala rahmat, karunia, dan kemudahan yang diberikan, sehingga penulis dapat menyelesaikan skripsi ini dengan baik. Dengan rasa syukur yang mendalam, karya ilmiah ini penulis persembahkan kepada:

1. Keluarga tercinta, khususnya Ayah, Ibu, Adik dan keluarga besar, atas segala bentuk doa yang tidak pernah putus, kasih sayang, serta dukungan moral maupun material yang menjadi fondasi utama keberhasilan penulis hingga dapat menyelesaikan studi.
2. Almamater tercinta, Program Studi S1 Informatika Universitas Amikom Yogyakarta, sebagai tempat penulis menimba ilmu, berproses, dan mengembangkan diri.
3. Seluruh pihak, termasuk jajaran dosen pembimbing maupun penguji, rekan mahasiswa, dan kerabat yang telah memberikan bimbingan, bantuan, serta dukungan nyata dalam penyelesaian skripsi ini, yang tidak dapat penulis sebutkan satu per satu.

Semoga Allah SWT senantiasa memberikan balasan kebaikan, perlindungan, dan keberkahan kepada kita semua.

KATA PENGANTAR

Segala puji syukur penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat dan karunia-Nya, sehingga skripsi yang berjudul **"PENGEMBANGAN ARSITEKTUR HIBRIDA EFFICIENTNET-VIT DENGAN EXPLAINABLE AI (XAI) UNTUK DETEKSI PENYAKIT TANAMAN PISANG"** dapat diselesaikan dengan baik. Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana pada Program Studi Informatika, Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta.

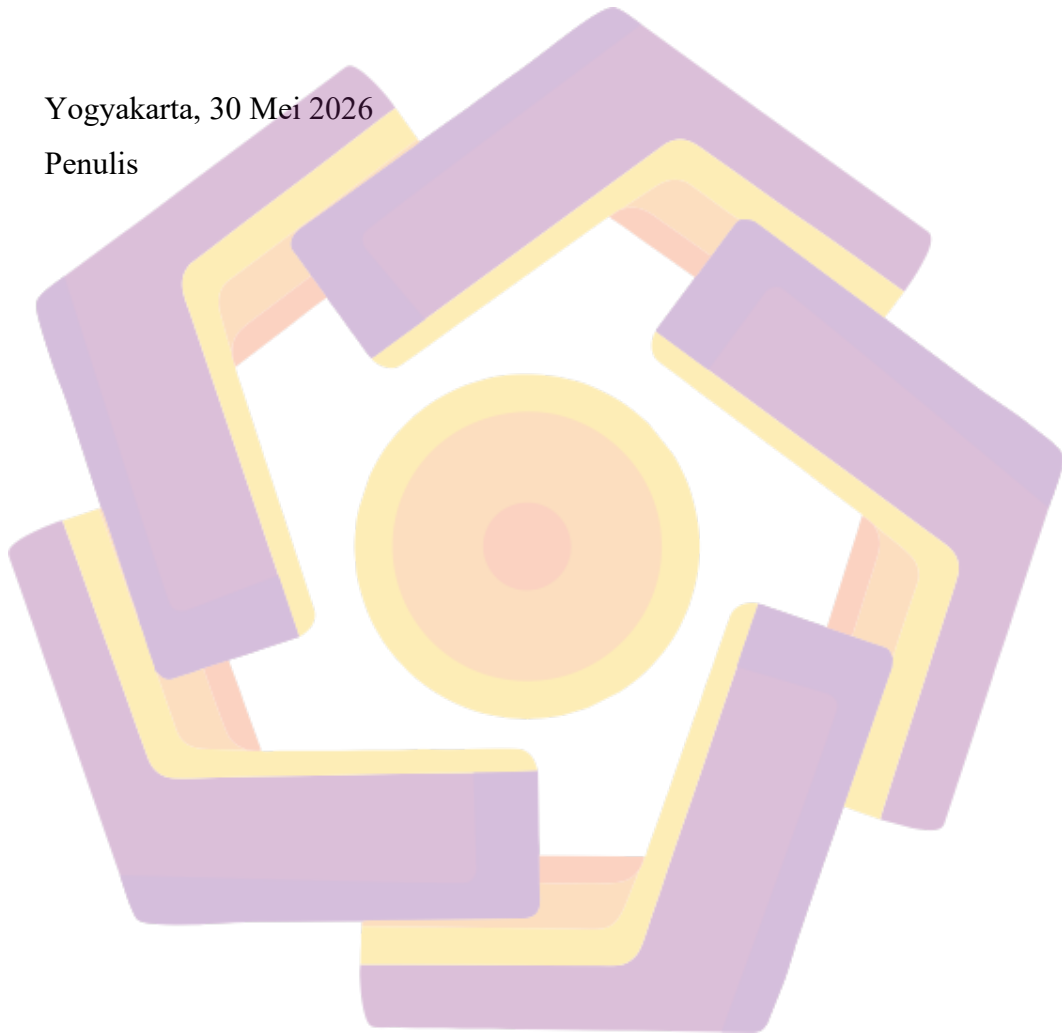
Proses penyusunan skripsi ini tidak terlepas dari bantuan, dukungan, dan bimbingan berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Allah SWT, atas segala petunjuk, kemudahan, dan kekuatan yang diberikan.
2. Bapak Haryono dan Ibu Sri Rustati selaku orang tua penulis atas doa, kasih sayang, dan dukungan yang tidak terhingga.
3. Bapak Ferian Fauzi Abdulloh, S.Kom., M.Kom. selaku dosen pembimbing penulis atas waktu, arahan, bimbingan, dan masukan berharga selama proses penelitian dan penulisan.
4. Seluruh dosen dan staf Fakultas Ilmu Komputer, Universitas AMIKOM Yogyakarta, atas ilmu dan fasilitas yang telah diberikan selama perkuliahan.
5. Rekan-rekan seperjuangan di Program Studi Informatika, atas kebersamaan, motivasi, dan dukungan yang tulus.
6. Tim Slepet FC Jogja-Sleman, atas kebersamaan, hiburan, motivasi, serta dukungannya.
7. Bapak Junaidi beserta keluarga, selaku pemilik indekos, atas kebaikan, dukungan, serta doa yang telah diberikan selama penulis menempuh masa studi.
8. Semua pihak yang tidak dapat disebutkan satu per satu, atas bantuan dan dukungan yang diberikan selama penyusunan skripsi ini.

Penulis menyadari sepenuhnya bahwa skripsi ini masih jauh dari sempurna. Oleh sebab itu, kritik dan saran yang bersifat membangun akan diterima dengan terbuka demi penyempurnaan di masa mendatang. Besar harapan penulis agar skripsi ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan.

Yogyakarta, 30 Mei 2026

Penulis



DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	viii
DAFTAR TABEL.....	xi
DAFTAR GAMBAR.....	xii
DAFTAR LAMBANG DAN SINGKATAN	xiii
DAFTAR ISTILAH	xiv
INTISARI	xv
ABSTRACT.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah.....	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	4
BAB II TINJAUAN PUSTAKA	6
2.1 Studi Literatur	6
2.2 Dasar Teori.....	11

2.2.1	Penyakit pada Daun Pisang.....	11
2.2.2	Pengolahan Citra Digital.....	12
2.2.3	<i>Convolutional Neural Network (CNN)</i>	12
2.2.4	Arsitektur EfficientNetV2.....	13
2.2.5	Vision Transformer (ViT).....	14
2.2.6	Arsitektur Hibrida (EfficientNet-ViT).....	16
2.2.7	Augmentasi Data.....	16
2.2.8	<i>Explainable AI (Grad-CAM)</i>	17
2.2.9	Evaluasi Kinerja.....	18
BAB III METODE PENELITIAN		20
3.1	Objek Penelitian.....	20
3.2	Alur Penelitian.....	20
3.3	Alat dan Bahan.....	25
3.3.1	Data Penelitian.....	25
3.3.2	Alat dan Instrumen.....	25
BAB IV HASIL DAN PEMBAHASAN		27
4.1	Pengumpulan Dataset.....	27
4.1.1	Dataset Sekunder	27
4.1.2	Dataset Primer.....	29
4.2	Prapemrosesan Data.....	30
4.2.1	Pembersihan Data (<i>Data Cleaning</i>).....	30
4.2.2	Penyesuaian Dimensi Citra (<i>Resizing</i>).....	33
4.3	Pembagian Data (<i>Data Splitting</i>).....	34
4.4	Augmentasi dan Data Generator.....	36
4.4.1	Augmentasi Data (<i>Offline Augmentation</i>)	36

4.4.2	<i>Data Generator</i> dan Normalisasi	38
4.5	Perancangan Multi-Skenario Model	40
4.5.1	Skenario A: Model EfficientNetV2 Tunggal	40
4.5.2	Skenario B: Model Vision Transformer Tunggal	42
4.5.3	Skenario C: Model Hibrida (EfficientNetV2 - ViT).....	44
4.6	Kompilasi dan Pelatihan Model	47
4.6.1	Dinamika Pelatihan Skenario A (EfficientNetV2)	49
4.6.2	Dinamika Pelatihan Skenario B (Vision Transformer).....	50
4.6.3	Dinamika Pelatihan Skenario C (Hibrida EfficientNetV2-ViT)....	52
4.7	Pengujian dan Evaluasi	53
4.7.1	Hasil Evaluasi Data Uji (<i>Test Set</i>).....	53
4.7.2	Pengujian Dataset Baru (External Validation).....	59
4.8	Analisis <i>Explainable AI</i> (XAI) Menggunakan Grad-CAM	64
4.8.1	Analisis Grad-CAM pada Data Uji Lokal (Internal)	67
4.8.2	Analisis Grad-CAM pada Dataset Baru (Eksternal)	71
4.9	Rekapitulasi Analisis Komparatif	74
4.10	Implementasi dan Deployment Model	77
BAB V	PENUTUP	78
5.1	Kesimpulan	78
5.2	Saran	79
REFERENSI		81

DAFTAR TABEL

Tabel 2.1. Keaslian Penelitian	8
Tabel 4.1. Distribusi Dataset Awal Hasil Akuisisi	28
Tabel 4.2. Distribusi Jumlah Citra pada Dataset Uji Lapangan	29
Tabel 4.3. Laporan Pembersihan Data	33
Tabel 4.4. Distribusi Pembagian Data (<i>Data Splitting</i>)	35
Tabel 4.5. Parameter Konfigurasi ImageDataGenerator	36
Tabel 4.6. Rekapitulasi Skor <i>Loss</i> dan Akurasi Keseluruhan	54
Tabel 4.7. Perbandingan Metrik Evaluasi Per Kelas Kondisi pada Data Uji	57
Tabel 4.8. Rekapitulasi Skor <i>Loss</i> dan Akurasi pada Dataset Eksternal	59
Tabel 4.9. Perbandingan Metrik Evaluasi Per Kelas pada Dataset Eksternal	62
Tabel 4.10. Perbandingan Karakteristik dan Performa Model	75



DAFTAR GAMBAR

Gambar 2.1. Kondisi Daun Pisang	12
Gambar 2.2. Ilustrasi Arsitektur <i>Convolutional Neural Network</i>	13
Gambar 2.3. Struktur Blok Fused-MBConv pada EfficientNetV2	14
Gambar 2.4. Mekanisme <i>Self-Attention</i> pada ViT	15
Gambar 2.5. Diagram Skematik Arsitektur Hibrida EfficientNet-ViT	16
Gambar 2.6. Contoh Visualisasi Penerapan Teknik Augmentasi Data	17
Gambar 2.7. Contoh Visualisasi <i>Heatmap</i> Grad-CAM pada Penyakit Pisang	18
Gambar 3.1. Alur Penelitian	21
Gambar 4.1. Kode Program Pemindaian Direktori	31
Gambar 4.2. Kode Program Pembangkitan serta Pendeteksian <i>Perceptual Hash</i>	31
Gambar 4.3. Kode Program Pemindahan Data Duplikat	32
Gambar 4.4. Ilustrasi Mekanisme Proses <i>Perceptual Hashing</i> (pHash)	32
Gambar 4.5. Kode Program Fungsi Penyesuaian Dimensi Citra	34
Gambar 4.6. Visualisasi Hasil Augmentasi Data	38
Gambar 4.7. Kode Program Normalisasi dan <i>Data Generator</i>	39
Gambar 4.8. Kode Program Arsitektur Skenario A (EfficientNetV2)	41
Gambar 4.9. Kode Program Arsitektur Skenario B (Vision Transformer)	44
Gambar 4.10. Kode Program Arsitektur Skenario C (Hibrida)	47
Gambar 4.11. Kode Program Kompilasi dan Mekanisme <i>Callbacks</i> Pelatihan	49
Gambar 4.12. Grafik Akurasi dan <i>Loss</i> Pelatihan Skenario A (EfficientNetV2)	50
Gambar 4.13. Grafik Akurasi dan <i>Loss</i> Pelatihan Skenario B (ViT)	51
Gambar 4.14. Grafik Akurasi dan <i>Loss</i> Pelatihan Skenario C (Hibrida)	52
Gambar 4.15. <i>Confusion Matrix</i> Skenario A (EfficientNetV2)	55
Gambar 4.16. <i>Confusion Matrix</i> Skenario B (ViT)	56
Gambar 4.17. <i>Confusion Matrix</i> Skenario C (Hibrida)	56
Gambar 4.18. <i>Confusion Matrix</i> Dataset Baru Skenario A (EfficientNetV2)	60
Gambar 4.19. <i>Confusion Matrix</i> Dataset Baru Skenario B (ViT)	60
Gambar 4.20. <i>Confusion Matrix</i> Dataset Baru Skenario C (Hibrida)	61
Gambar 4.21. Implementasi Fungsi Pembuatan <i>Heatmap</i> Grad-CAM	66
Gambar 4.22. Visualisasi Grad-CAM Skenario A pada Data Uji Lokal	68
Gambar 4.23. Visualisasi Grad-CAM Skenario C pada Data Uji Lokal	69
Gambar 4.24. Visualisasi Grad-CAM Skenario A pada Dataset Eksternal	72
Gambar 4.25. Visualisasi Grad-CAM Skenario C pada Dataset Eksternal	73

DAFTAR LAMBANG DAN SINGKATAN



AI	: <i>Artificial Intelligence</i>
CNN	: <i>Convolutional Neural Network</i>
DCT	: <i>Discrete Cosine Transform</i>
FN	: <i>False Negative</i>
FP	: <i>False Positive</i>
GELU	: <i>Gaussian Error Linear Unit</i>
Grad-CAM	: <i>Gradient-weighted Class Activation Mapping</i>
MBConv	: <i>Mobile Inverted Bottleneck Convolution</i>
MLP	: <i>Multi-Layer Perceptron</i>
OOP	: <i>Object-Oriented Programming</i>
pHash	: <i>Perceptual Hashing</i>
ReLU	: <i>Rectified Linear Unit</i>
TN	: <i>True Negative</i>
TP	: <i>True Positive</i>
ViT	: <i>Vision Transformer</i>
XAI	: <i>Explainable Artificial Intelligence</i>

DAFTAR ISTILAH

Backbone	: Arsitektur inti ekstraksi fitur.
Batch Size	: Jumlah sampel citra dalam satu iterasi komputasi.
Black-box	: Sifat model dengan proses keputusan tidak transparan.
Bottleneck	: Penyempitan dimensi yang menghambat aliran informasi.
Computer Vision	: Bidang AI pengekstraksi informasi dari citra.
Confusion Matrix	: Tabel evaluasi prediksi model terhadap label aktual.
CNN	: Arsitektur jaringan saraf pengekstraksi fitur lokal.
Data Leakage	: Kebocoran data latih ke data uji akibat duplikasi.
Deep Learning	: Pembelajaran mesin berbasis jaringan saraf berlapis.
Domain Shift	: Perbedaan visual antara data latih dan data lapangan.
Early Stopping	: Penghentian otomatis saat performa tidak meningkat.
EfficientNetV2	: Arsitektur konvolusi dengan tingkat efisiensi tinggi.
Epoch	: Satu iterasi penuh pelatihan pada seluruh data.
Explainable AI (XAI)	: Teknik visualisasi landasan pengambilan keputusan AI.
Feature Map	: Matriks keluaran hasil ekstraksi operasi konvolusi.
Global Average Pooling	: Teknik reduksi matriks menggunakan nilai rata-rata.
Grad-CAM	: Algoritma penghasil peta panas atensi keputusan model.
Inductive Bias	: Asumsi bawaan algoritma untuk memprediksi data.
Klorosis	: Gejala patologi penyakit berupa penguningan pada daun.
Lesi Nekrotik	: Area jaringan sel yang mati akibat infeksi penyakit.
Overfitting	: Salah prediksi akibat terlalu hafal dengan data latih.
Overshoot	: Fenomena sebaran peta atensi yang melebar berlebihan.
Patch	: Potongan berukuran kecil dari citra utuh.
Perceptual Hashing	: Algoritma pendeteksi kemiripan struktur visual citra.
ReduceLROnPlateau	: Penurunan laju pelatihan saat performa stagnan.
Self-Attention	: Mekanisme pengukur relasi antar-piksel secara global.
Transfer Learning	: Pemanfaatan bobot model pralatih untuk kasus baru.
ViT	: Arsitektur pemroses relasi konteks citra berbasis atensi.

INTISARI

Penyakit pada tanaman pisang seperti Sigatoka, Cordana, dan Panama merupakan ancaman serius yang berdampak pada terganggunya ketahanan pangan. Diagnosis manual dinilai lambat, sehingga dibutuhkan sistem deteksi otomatis yang tangguh. Sayangnya, penggunaan model *deep learning* tunggal menghadapi keterbatasan arsitektural. *Convolutional Neural Network* (CNN) kesulitan menangkap konteks spasial jarak jauh, sedangkan Vision Transformer (ViT) rentan mengalami overfitting pada dataset berskala kecil.

Untuk menyelesaikan masalah tersebut, penelitian ini merancang arsitektur hibrida yang mengintegrasikan EfficientNetV2 (pengekstraksi fitur lokal) dan ViT (pengekstraksi konteks global). Metode ini menggunakan 4.675 citra sekunder untuk proses permodelan, serta 40 citra lapangan primer untuk pengujian validasi eksternal (mengukur pergeseran domain). Selain itu, teknik *Explainable AI* berbasis Grad-CAM diterapkan untuk memvisualisasikan landasan pengambilan keputusan dan mengatasi sifat kotak hitam.

Hasil pengujian internal menunjukkan EfficientNetV2 mencapai akurasi tertinggi 99,29%, diikuti ViT (98,58%), dan Hibrida (97,88%). Namun, pada pengujian eksternal di lapangan, ViT turun drastis menjadi 62,50%. Sementara itu, EfficientNetV2 (82,50%) dan Hibrida terbukti stabil di angka 80,00%. Analisis Grad-CAM memvalidasi bahwa model hibrida mengevaluasi geometri lesi secara holistik melalui mekanisme *self-attention*, sehingga mampu mengatasi kelemahan spesifik CNN yang hanya mengunci fitur tepi lesi secara parsial dan rentan salah klasifikasi pada kasus dengan kemiripan tekstur antar-kelas penyakit. Penelitian ini berkontribusi menghasilkan sistem deteksi andal dan transparan bagi praktisi pertanian. Untuk penelitian lanjutan, direkomendasikan menerapkan segmentasi latar belakang, pengayaan augmentasi data lapangan, kompresi arsitektur seluler, serta restrukturisasi lapisan penyandian posisi.

Kata kunci: Penyakit Daun Pisang, Arsitektur Hibrida, EfficientNetV2, Vision Transformer, Grad-CAM.

ABSTRACT

Banana plant diseases such as Sigatoka, Cordana, and Panama pose a serious threat that disrupts food security. Manual diagnosis is considered slow, thus a robust automated detection system is needed. Unfortunately, the use of single deep learning models faces architectural limitations. Convolutional Neural Networks (CNN) struggle to capture long-range spatial contexts, whereas Vision Transformers (ViT) are prone to overfitting on small-scale datasets.

To resolve these problems, this research designed a hybrid architecture integrating EfficientNetV2 (local feature extractor) and ViT (global context extractor). This method utilized 4,675 secondary images for modeling and 40 primary field images for external validation testing (measuring domain shift). Additionally, Explainable AI based on Grad-CAM was applied to visualize the decision-making foundation and overcome the black-box nature.

Internal testing results showed EfficientNetV2 achieved the highest accuracy of 99.29%, followed by ViT (98.58%), and the Hybrid model (97.88%). However, in external field testing, ViT dropped drastically to 62.50%. Meanwhile, EfficientNetV2 (82.50%) and the Hybrid model demonstrated greater stability at 80.00%. Grad-CAM analysis validated that the hybrid model evaluates lesion geometry holistically via self-attention, overcoming a specific weakness of standalone CNN architectures that fixate only on localized high-frequency edge features and are vulnerable to misclassification in cases with visually similar inter-class textures. This research contributes by producing a reliable and transparent detection system for agricultural practitioners. For future research, it is recommended to apply background segmentation, enrich field data augmentation, compress the mobile architecture, and restructure the positional encoding layer.

Keywords: Banana Leaf Disease, Hybrid Architecture, EfficientNetV2, Vision Transformer, Grad-CAM.