

**KLASIFIKASI MULTIKELAS PENYAKIT DAUN TOMAT
MENGUNAKAN GLCM, EKSTRAKSI FITUR WARNA,
DAN BENTUK DENGAN XGBOOST YANG
DIOPTIMALKAN**

LAPORAN SCIENTIST

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



Disusun oleh :

FRANSISKO ANDRADE LAISKODAT

22.11.4665

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

**KLASIFIKASI MULTIKELAS PENYAKIT DAUN TOMAT
MENGUNAKAN GLCM, EKSTRAKSI FITUR WARNA,
DAN BENTUK DENGAN XGBOOST YANG
DIOPTIMALKAN**

LAPORAN SCIENTIST

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



Disusun oleh :

FRANSISKO ANDRADE LAISKODAT

22.11.4665

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2026

HALAMAN PERSETUJUAN

JALUR NON-REGULER

**KLASIFIKASI MULTIKELAS PENYAKIT DAUN TOMAT MENGGUNAKAN
GLCM, EKSTRAKSI FITUR WARNA, DAN BENTUK DENGAN XGBOOST
YANG DIOPTIMALKAN**

yang disusun dan diajukan oleh
Fransisko Andrade Laiskodat
22.11.4665

telah disetujui oleh Dosen Pembimbing
pada tanggal 21 Januari 2026

Dosen Pembimbing,



Mard Rahardi, S.Kom., M.Eng.
NIK, 190302393

HALAMAN PENGESAHAN

JALUR NON-REGULER

**KLASIFIKASI MULTIKELAS PENYAKIT DAUN TOMAT MENGGUNAKAN
GLCM, EKSTRAKSI FITUR WARNA, DAN BENTUK DENGAN XGBOOST
YANG DIOPTIMALKAN**

yang disusun dan diajukan oleh

Fransisko Andrade Laiskodat
22.11.4665

Telah dipertahankan di depan Dewan Penguji
pada tanggal 21 Januari 2026

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Arifivanto Hadinegoro, S.Kom., M.T.
NIK. 190302289

Ali Mustopa, S.Kom., M.Kom.
NIK. 190302192

Majid Rahardi, S.Kom., M.Eng.
NIK. 190302393



Laporan ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 21 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



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

HALAMAN PERNYATAAN KEASLIAN KARYA

Yang bertandatangan di bawah ini,

Nama mahasiswa : Fransisko Andrade Laiskodat

NIM : 22.11.4665

Menyatakan bahwa Laporan dengan judul berikut:

Klasifikasi Multikelas Penyakit Daun Tomat Menggunakan GLCM, Ekstraksi Fitur Warna, dan Bentuk dengan Xgboost yang Dioptimalkan

Dosen Pembimbing : Majid Rahardi, S.Kom, M.Eng

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 kegiatan **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 ketidak-benaran 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, 21 Januari 2026

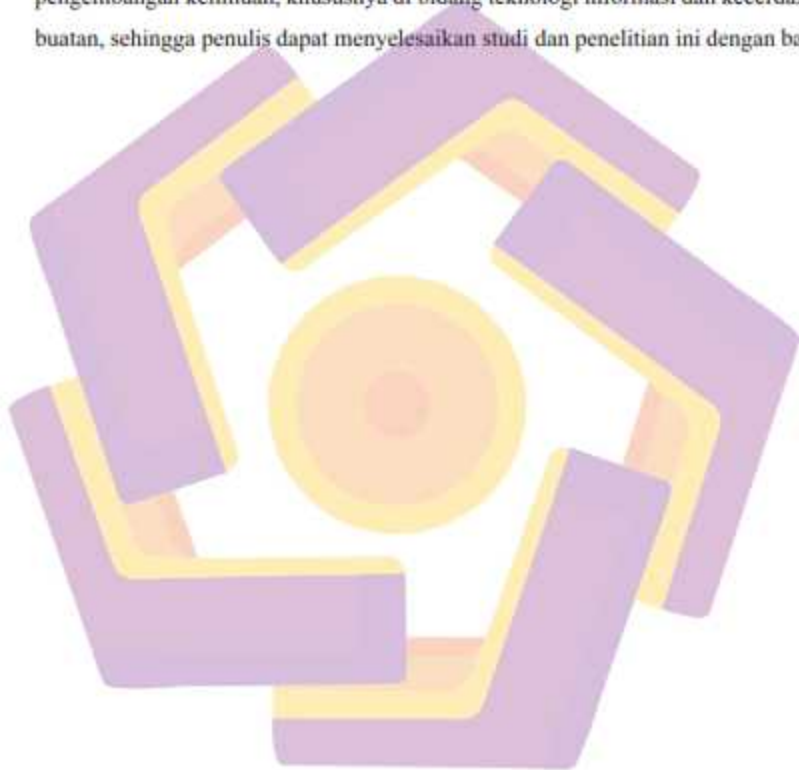
Yang Menyatakan,



Fransisko Andrade Laiskodat

HALAMAN PERSEMBAHAN

Laporan penelitian ini penulis persembahkan kepada Universitas AMIKOM Yogyakarta sebagai institusi pendidikan yang telah memberikan lingkungan akademik, fasilitas, serta dukungan yang kondusif dalam proses pembelajaran dan pengembangan keilmuan, khususnya di bidang teknologi informasi dan kecerdasan buatan, sehingga penulis dapat menyelesaikan studi dan penelitian ini dengan baik.



KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan laporan penelitian ini dengan baik. Laporan ini disusun sebagai salah satu syarat dalam menyelesaikan studi pada Program Studi Informatika.

Dalam proses penyusunan laporan ini, penulis menyadari bahwa keberhasilan tidak terlepas dari bantuan, bimbingan, serta dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Majid Rahardi, S.Kom., M.Eng., selaku Dosen Pembimbing, yang telah meluangkan waktu, tenaga, dan pikiran untuk memberikan bimbingan, arahan, serta masukan yang sangat berharga selama proses penyusunan laporan penelitian ini.
2. Bapak Paulus Laiskodat dan Ibu Dewi Indarti Laiskodat, selaku orang tua penulis, yang senantiasa memberikan doa, dukungan moral, serta dukungan finansial sehingga penulis dapat menyelesaikan penelitian ini dengan baik.
3. Nadia Martha Lefi, yang telah memberikan dukungan, kerja sama, serta bantuan kepada penulis selama proses penyusunan laporan ini.

Penulis menyadari bahwa laporan penelitian ini masih memiliki keterbatasan dan kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun demi penyempurnaan laporan ini. Semoga laporan penelitian ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan, khususnya di bidang pengolahan citra digital dan kecerdasan buatan.

Yogyakarta, 25 Desember 2025

Penulis

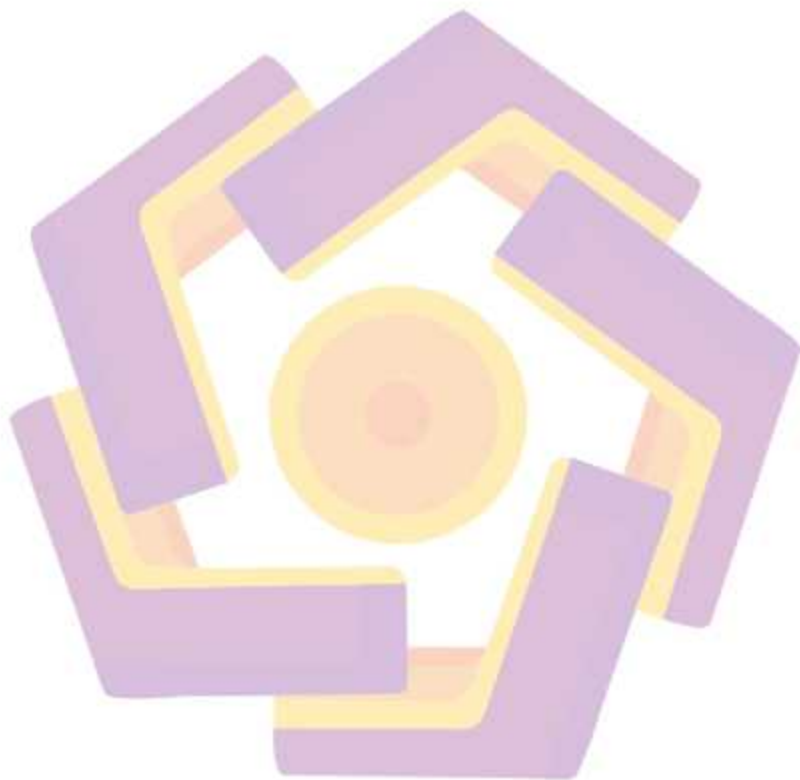
DAFTAR ISI

Halaman Judul.....	i
Halaman Persetujuan.....	ii
Halaman Pengesahan	iii
Halaman Pernyataan Keaslian Karya.....	iv
Halaman Persembahan	v
Kata Pengantar	vi
Daftar Isi.....	vii
Daftar Tabel.....	ix
Daftar Gambar.....	x
Daftar Lampiran.....	xi
Daftar Lambang dan Singkatan.....	xii
Daftar Istilah.....	xiii
Intisari	xv
<i>Abstract</i>	xvi
Bab I Pendahuluan	1
1.1 Gambaran Umum	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan.....	3
Bab II Tinjauan Pustaka	4
2.1 Studi Literatur.....	4
2.2 Landasan Teori	6
2.2.1 Penyakit Daun Tomat.....	6
2.2.2 Pengolahan Citra Digital	7
2.2.3 Ekstraksi Fitur Citra	7
2.2.4 Normalisasi Data	7
2.2.5 Principal Component Analysis (PCA)	8
2.2.6 Synthetic Minority Over-sampling Technique (SMOTE).....	8
2.2.7 XGBoost (Extreme Gradient Boosting).....	8
2.2.8 Hyperparameter Tuning	9

2.2.9	Evaluasi Model.....	9
BAB III Metode Penelitian		10
3.1	Metode.....	10
3.1.1.	Riset Studi Literatur.....	11
3.1.2.	Akuisisi Data.....	11
3.1.3.	Preprocessing Citra	12
3.1.4.	Ekstraksi Fitur.....	12
3.1.5.	Normalisasi Dan Reduksi Dimensi	14
3.1.6.	Penyeimbangan Data.....	17
3.1.7.	Pembagian Data	18
3.1.8.	Pemodelan.....	18
3.1.9.	Evaluasi Model.....	19
3.1.10.	Dokumentasi.....	21
BAB IV Pembahasan		22
4.1.	Sub Pembahasan.....	22
4.1.1.	Ekstraksi Fitur.....	22
4.1.2.	Metrik Evaluasi	24
4.1.3.	Confusion Matrix	26
4.1.4.	Perbandingan Model Lain.....	27
4.1.5.	Kurva Precision-Recal	28
BAB V Kesimpulan		30
5.1.	Kesimpulan.....	30
5.2.	Saran.....	30
Referensi		32
Curriculum Vitae		35
Lampiran dan Bukti Pendukung.....		36
a.	Letter of Acceptance (LOA).....	36
b.	Lembar Review	37
c.	Bukti Terbit/Terindex	39
d.	Bukti pembayaran	39

DAFTAR TABEL

Tabel 4. 1 Hasil Ekstraksi Fitur Citra.....	22
Tabel 4. 2 Hasil Evaluasi Model XGBoost	25



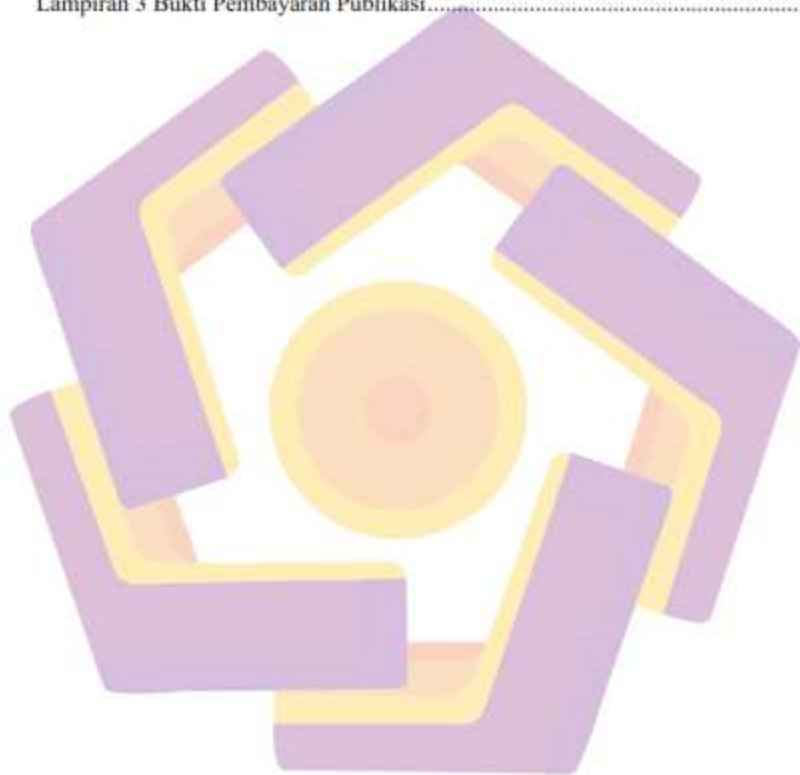
DAFTAR GAMBAR

Gambar 3. 1 Alur Penelitian.....	10
Gambar 3. 2 Contoh Data Mentah	12
Gambar 3. 3 Visualisasi Dataset Setelah Ekstraksi Fitur PCA 3D.....	16
Gambar 3. 4 Sebelum SMOTE	17
Gambar 3. 5 Setelah SMOTE.....	18
Gambar 4. 1 Metrik Evaluasi	25
Gambar 4. 2 Confusion Matrix	27
Gambar 4. 3 Perbandingan Akurasi Model Lain.....	28
Gambar 4. 4 Kurva Precision-Recall	29



DAFTAR LAMPIRAN

Lampiran 1 Bukti Letter of Acceptance	36
Lampiran 2 Bukti Terbit Jurnal pada journal of Applied Informates and Computing	39
Lampiran 3 Bukti Pembayaran Publikasi.....	39

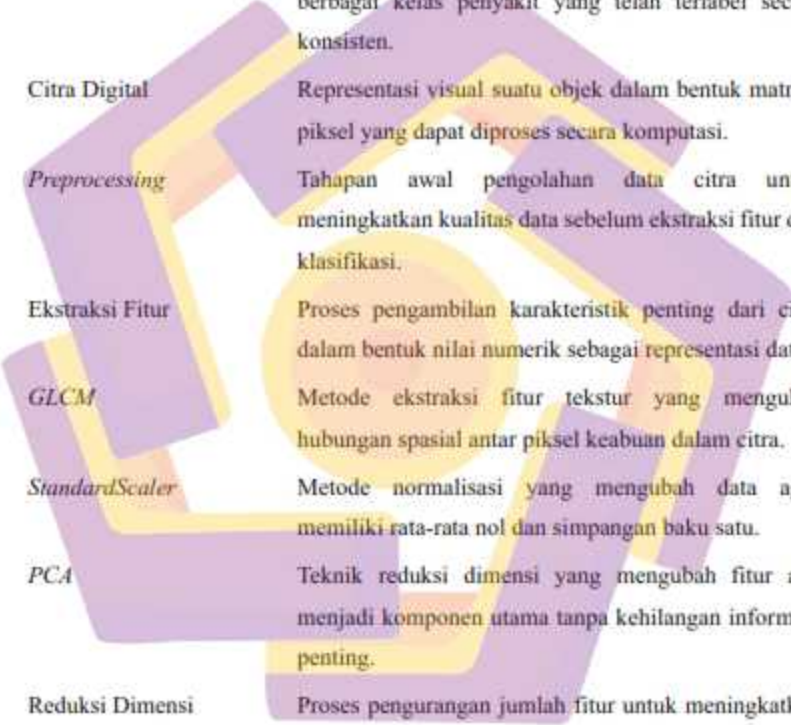


DAFTAR LAMBANG DAN SINGKATAN



AI	<i>Artificial Intelligence</i>
ASM	<i>Angular Second Moment</i>
CNN	<i>Convolutional Neural Network</i>
CV	<i>Cross Validation</i>
F1-score	<i>Harmonic mean antara precision dan recall</i>
GLCM	<i>Gray Level Co-occurrence Matrix</i>
IoT	<i>Internet of Things</i>
Log Loss	<i>Logarithmic Loss</i>
ML	<i>Machine Learning</i>
PCA	<i>Principal Component Analysis</i>
PR Curve	<i>Precision-Recall Curve</i>
RGB	<i>Red, Green, Blue</i>
SMOTE	<i>Synthetic Minority Over-sampling Technique</i>
VIT	<i>Vision Transformer</i>
XGBoost	<i>Extreme Gradient Boosting</i>

DAFTAR ISTILAH



Dataset	Kumpulan data citra daun tomat yang digunakan sebagai bahan pelatihan dan pengujian model klasifikasi.
<i>PlantVillage</i>	Dataset publik berisi citra daun tanaman dengan berbagai kelas penyakit yang telah terlabel secara konsisten.
Citra Digital	Representasi visual suatu objek dalam bentuk matriks piksel yang dapat diproses secara komputasi.
<i>Preprocessing</i>	Tahapan awal pengolahan data citra untuk meningkatkan kualitas data sebelum ekstraksi fitur dan klasifikasi.
Ekstraksi Fitur	Proses pengambilan karakteristik penting dari citra dalam bentuk nilai numerik sebagai representasi data.
<i>GLCM</i>	Metode ekstraksi fitur tekstur yang mengukur hubungan spasial antar piksel keabuan dalam citra.
<i>StandardScaler</i>	Metode normalisasi yang mengubah data agar memiliki rata-rata nol dan simpangan baku satu.
<i>PCA</i>	Teknik reduksi dimensi yang mengubah fitur asli menjadi komponen utama tanpa kehilangan informasi penting.
Reduksi Dimensi	Proses pengurangan jumlah fitur untuk meningkatkan efisiensi komputasi dan kinerja model.
<i>SMOTE</i>	Metode penyeimbangan data dengan menghasilkan sampel sintetis pada kelas minoritas.
<i>XGBoost</i>	Algoritma ensemble berbasis boosting yang menggabungkan banyak decision tree untuk menghasilkan prediksi yang kuat.



<i>Hyperparameter</i>	Parameter yang ditentukan sebelum pelatihan model dan memengaruhi proses pembelajaran.
<i>Hyperparameter Tuning</i>	Proses pencarian kombinasi hyperparameter terbaik untuk meningkatkan performa model.
<i>RandomizedSearchCV</i>	Metode pencarian hyperparameter secara acak dengan validasi silang untuk efisiensi komputasi.
Evaluasi Model	Tahapan pengukuran kinerja model menggunakan metrik statistik tertentu.
<i>Confusion Matrix</i>	Matriks evaluasi yang menunjukkan perbandingan antara kelas aktual dan hasil prediksi model.
<i>Accuracy</i>	Ukuran proporsi prediksi yang benar terhadap seluruh data pengujian.
<i>Precision</i>	Ukuran ketepatan model dalam memprediksi suatu kelas tertentu.
<i>Recall</i>	Ukuran kemampuan model dalam mengenali seluruh data yang benar dari suatu kelas.
<i>F1-Score</i>	Rata-rata harmonis antara precision dan recall yang menggambarkan keseimbangan kinerja model.
<i>Log Loss</i>	Metrik evaluasi yang mengukur tingkat kesalahan prediksi probabilistik model.
<i>Precision-Recall Curve</i>	Kurva evaluasi yang menunjukkan hubungan antara precision dan recall pada berbagai ambang keputusan.
<i>Edge Device</i>	Perangkat dengan sumber daya terbatas seperti smartphone atau IoT yang digunakan untuk pemrosesan lokal.

INTISARI

Klasifikasi otomatis penyakit daun tomat merupakan komponen penting dalam mendukung pengembangan pertanian presisi berbasis kecerdasan buatan. Penyakit daun tomat dapat menyebabkan penurunan produktivitas hingga 60–80%, sehingga diperlukan sistem diagnosis yang cepat, akurat, dan efisien. Penelitian ini bertujuan untuk membangun model klasifikasi multikelas penyakit daun tomat dengan mengombinasikan fitur tekstur, warna, dan bentuk yang diekstraksi dari citra daun, serta menerapkan algoritma XGBoost yang telah dioptimalkan. Dataset PlantVillage digunakan sebagai sumber data, dengan tahapan praproses meliputi ekstraksi fitur menggunakan GLCM, histogram warna, dan ciri bentuk dasar, normalisasi data, reduksi dimensi menggunakan Principal Component Analysis (PCA), serta penyeimbangan kelas menggunakan Synthetic Minority Over-Sampling Technique (SMOTE). Model kemudian dilatih menggunakan XGBoost yang dioptimalkan melalui RandomizedSearchCV untuk mendapatkan kombinasi hiperparameter terbaik. Hasil eksperimen menunjukkan bahwa model mampu mengklasifikasikan sepuluh kelas penyakit dengan akurasi tinggi sebesar 97,63%, serta macro dan weighted F1-score sebesar 0,98, menandakan performa stabil pada seluruh kelas. Temuan ini membuktikan bahwa kombinasi fitur buatan (handcrafted features) dengan XGBoost teroptimasi mampu menghasilkan model klasifikasi yang ringan, efisien, dan kompetitif dibandingkan metode berbasis deep learning yang membutuhkan komputasi lebih besar. Pendekatan ini berpotensi diterapkan pada sistem pertanian cerdas, aplikasi mobile, maupun perangkat lapangan untuk deteksi dini penyakit tanaman secara real-time.

Kata kunci: *Feature Extraction, Gray-Level Co-occurrence Matrix (GLCM), Principal Component Analysis (PCA), Smote, XGBoost*

ABSTRACT

Automatic classification of tomato leaf diseases plays a vital role in advancing artificial intelligence-driven precision agriculture. Tomato leaf diseases can significantly reduce crop productivity by up to 60–80%, making fast, accurate, and efficient diagnostic systems crucial for agricultural sustainability. This study aims to develop a multiclass classification model for tomato leaf diseases by integrating handcrafted texture, color, and shape features extracted from leaf images, and employing an optimized XGBoost algorithm. The publicly available PlantVillage dataset was utilized, involving preprocessing stages such as feature extraction using GLCM, color histograms, and basic shape descriptors, data normalization, dimensionality reduction using Principal Component Analysis (PCA), and class balancing with the Synthetic Minority Over-Sampling Technique (SMOTE). The model was trained using XGBoost optimized through RandomizedSearchCV to identify the best hyperparameter configuration. Experimental results show that the proposed model achieves a high classification accuracy of 97.63%, along with macro and weighted F1-scores of 0.98, demonstrating stable performance across all ten disease classes. These findings indicate that combining handcrafted features with an optimized XGBoost classifier provides a lightweight, efficient, and competitive alternative to deep learning approaches that often require higher computational resources. The proposed method shows strong potential for implementation in smart farming systems, mobile-based diagnostic tools, and field-deployable solutions for real-time plant disease detection.

Keyword: Feature Extraction, Gray-Level Co-occurrence Matrix (GLCM), Principal Component Analysis (PCA), Smote, XGBoost