

Klasifikasi Penyakit Daun Tomat Menggunakan Metode *K-Nearest Neighbor* dengan Segmentasi *Watershed*

<http://dx.doi.org/10.28932/jutisi.v12i2.14615>

Riwayat Artikel

Received: 27 Januari 2026 | Final Revision: 13 Agustus 2026 | Accepted: 13 Agustus 2026

Creative Commons License 4.0 (CC BY – NC)



Madadina Adilah Pamuji ^{#1}, Rinci Kembang Hapsari ^{✉#2}

[#] Teknik Informatika, Institut Teknologi Adhi Tama Surabaya
Jl. Arief Rahman Hakim No. 100, Surabaya 60111, Indonesia

¹madadnap@gmail.com

²rincikembang@itats.ac.id

✉Corresponding author: rincikembang@itats.ac.id

Abstrak — Penyakit daun tomat merupakan salah satu faktor utama yang menyebabkan penurunan produktivitas tanaman tomat dan berdampak langsung terhadap hasil pertanian. Identifikasi penyakit secara manual memiliki keterbatasan dari segi subjektivitas, waktu, dan konsistensi diagnosis. Oleh karena itu, penelitian ini mengusulkan sistem klasifikasi penyakit daun tomat berbasis pengolahan citra digital dan machine learning menggunakan kombinasi segmentasi watershed, ekstraksi fitur warna HSV dan fitur tekstur *Gray Level Co-occurrence Matrix (GLCM)*, serta algoritma *K-Nearest Neighbor (KNN)*. Dataset yang digunakan terdiri dari 8.111 citra daun tomat *single-leaf* yang terbagi ke dalam sembilan kelas penyakit dan satu kelas daun sehat, serta dataset tambahan berupa citra *multi-leaf* untuk pengujian segmentasi. Tahapan penelitian meliputi prapemrosesan citra, segmentasi daun menggunakan metode *watershed*, ekstraksi fitur HSV dan GLCM, serta proses klasifikasi menggunakan KNN dengan jarak *Euclidean*. Hasil pengujian menunjukkan bahwa sistem memperoleh akurasi sebesar 87.08% pada data latih, 82.96% pada data validasi, dan 81.73% pada data uji. Selain itu, segmentasi watershed mampu memisahkan objek daun pada citra *multi-leaf* sehingga proses klasifikasi dapat dilakukan secara individual pada setiap objek daun. Hasil penelitian menunjukkan bahwa kombinasi metode berbasis *handcrafted features* dan KNN masih efektif untuk klasifikasi penyakit daun tomat dengan kompleksitas komputasi yang relatif rendah.

Kata kunci— GLCM; HSV; *K-Nearest Neighbor*; klasifikasi citra; penyakit daun tomat.

Tomato Leaf Disease Classification Using the *K-Nearest Neighbor* Method with Watershed Segmentation

Abstract — Tomato leaf disease is one of the main factors causing a decrease in tomato plant productivity and directly affecting agricultural yields. Manual disease identification has limitations in terms of subjectivity, time efficiency, and diagnostic consistency. Therefore, this study proposes a tomato leaf disease classification system based on digital image processing and machine learning using a combination of watershed segmentation, HSV color feature extraction, Gray Level Co-occurrence Matrix (GLCM) texture features, and the *K-Nearest Neighbor (KNN)* algorithm. The dataset used consists of 8,111 single-leaf tomato leaf images divided into nine

disease classes and one healthy leaf class, along with an additional multi-leaf image dataset for segmentation testing. The research stages include image preprocessing, leaf segmentation using the watershed method, HSV and GLCM feature extraction, and classification using KNN with Euclidean distance. The experimental results show that the system achieved accuracies of 87.08% on training data, 82.96% on validation data, and 81.73% on testing data. In addition, the watershed segmentation method was able to separate leaf objects in multi-leaf images, enabling classification to be performed individually for each leaf object. The results indicate that the combination of handcrafted feature-based methods and KNN still provides effective performance for tomato leaf disease classification with relatively low computational complexity.

Keywords— GLCM; HSV; image classification; K-Nearest Neighbor; tomato leaf disease.

I. PENDAHULUAN

Penyakit pada daun tomat merupakan salah satu faktor utama yang menyebabkan penurunan produktivitas tanaman secara global dan berpotensi mengurangi hasil panen di sektor pertanian. Infeksi penyakit seperti *early blight*, *late blight*, *bacterial spot*, dan *yellow leaf curl* virus dapat menyebabkan kerusakan jaringan daun yang menghambat proses fotosintesis dan pertumbuhan tanaman. Proses identifikasi penyakit daun tomat secara manual oleh petani maupun tenaga ahli masih memiliki berbagai keterbatasan karena bersifat subjektif, membutuhkan waktu yang relatif lama, serta sangat bergantung pada pengalaman pengamat [1], [2]. Selain itu, beberapa jenis penyakit daun tomat memiliki karakteristik visual yang mirip sehingga meningkatkan risiko kesalahan diagnosis, terutama pada tahap awal infeksi [3]. Oleh karena itu, penelitian di bidang pengolahan citra digital dan machine learning berkembang pesat untuk mendukung sistem klasifikasi penyakit daun tomat secara otomatis, objektif, dan konsisten [1], [4].

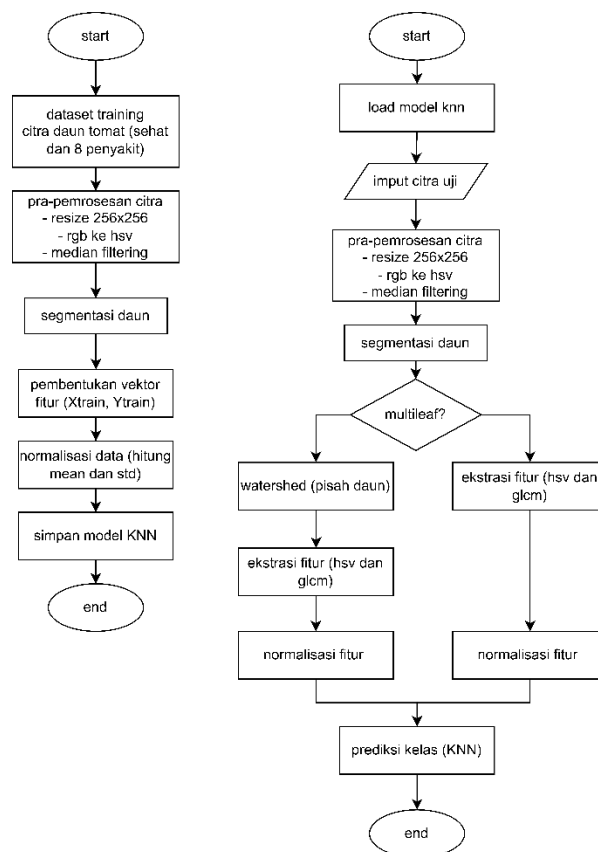
Pendekatan berbasis citra dianggap efektif karena daun merupakan bagian tanaman yang paling awal menunjukkan gejala penyakit melalui perubahan warna, pola tekstur, serta karakteristik morfologi visual [5]. Sejumlah penelitian menunjukkan bahwa fitur warna dan tekstur merupakan representasi visual yang efektif untuk membedakan daun sehat dan daun yang terinfeksi penyakit. Ruang warna *Hue Saturation Value* (HSV) terbukti mampu merepresentasikan perubahan warna daun secara lebih stabil terhadap variasi pencahayaan, sedangkan metode *Gray Level Co-occurrence Matrix* (GLCM) efektif dalam mengekstraksi pola tekstur bercak penyakit pada permukaan daun [1], [6]. Kombinasi fitur warna dan tekstur dengan algoritma klasifikasi berbasis machine learning seperti *K-Nearest Neighbor* (KNN), *Support Vector Machine* (SVM), dan *Extreme Gradient Boosting* (XGBoost) dilaporkan mampu menghasilkan performa klasifikasi yang cukup baik dengan kompleksitas komputasi yang relatif rendah [6], [7].

Di sisi lain, pendekatan berbasis *deep learning* seperti Convolutional Neural Network (CNN), Vision Transformer, dan model segmentasi otomatis menunjukkan tingkat akurasi yang sangat tinggi dalam klasifikasi penyakit daun tomat [5], [8], [9]. Meskipun demikian, metode *deep learning* umumnya membutuhkan jumlah data yang besar, proses pelatihan yang kompleks, serta sumber daya komputasi yang tinggi seperti GPU dan memori berkapasitas besar [10]. Kondisi tersebut menjadi tantangan dalam implementasi sistem deteksi penyakit tanaman pada perangkat dengan keterbatasan sumber daya komputasi, khususnya pada sistem pertanian cerdas berbasis edge computing maupun perangkat bergerak (*mobile-based agriculture system*). Oleh karena itu, pendekatan berbasis *handcrafted* features masih memiliki potensi yang relevan karena lebih ringan secara komputasi, lebih mudah diinterpretasikan, dan tidak memerlukan proses pelatihan yang kompleks.

Sebagian besar penelitian sebelumnya masih berfokus pada penggunaan citra daun tunggal (*single-leaf*) dengan latar belakang sederhana sehingga kurang merepresentasikan kondisi nyata di lapangan [11], [12]. Pada kondisi aktual, satu citra seringkali mengandung beberapa daun yang saling bersentuhan atau bertumpuk (*multi-leaf*), sehingga proses segmentasi objek menjadi lebih kompleks. Padahal, kualitas segmentasi sangat mempengaruhi proses ekstraksi fitur warna dan tekstur yang digunakan pada tahap klasifikasi [13]. Tanpa segmentasi yang baik, fitur yang dihasilkan dapat bercampur dengan objek lain maupun latar belakang sehingga menurunkan performa sistem klasifikasi secara keseluruhan, khususnya pada citra *multi-leaf* [11], [13].

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan sistem klasifikasi penyakit daun tomat berbasis pengolahan citra digital menggunakan kombinasi segmentasi *watershed*, fitur warna HSV, fitur tekstur GLCM, dan algoritma *K-Nearest Neighbor* (KNN). Metode *watershed* digunakan untuk memisahkan objek daun yang saling bersentuhan pada citra *multi-leaf* sehingga setiap daun dapat diproses dan diklasifikasikan secara individual [13]. Berbeda dengan sebagian besar penelitian sebelumnya yang berfokus pada pencapaian akurasi tertinggi menggunakan pendekatan *deep learning*, penelitian ini menitikberatkan pada pengembangan sistem klasifikasi yang efisien secara komputasi, ringan, mudah diimplementasikan, dan tetap mampu memberikan performa klasifikasi yang stabil pada kondisi citra yang lebih realistis. Sistem yang dikembangkan dievaluasi menggunakan data latih, data validasi, dan data uji pada citra daun tunggal, serta diuji lebih lanjut pada citra *multi-leaf* untuk mengevaluasi efektivitas segmentasi terhadap hasil klasifikasi. Dengan demikian, penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan sistem klasifikasi penyakit tanaman berbasis *handcrafted* features yang sesuai untuk perangkat dengan sumber daya terbatas di bidang pertanian cerdas.

II. METODE PENELITIAN



Gambar 1. Diagram alur metode penelitian klasifikasi penyakit daun tomat menggunakan HSV, GLCM, KNN, dan segmentasi *watershed*.

Penelitian ini menggunakan pendekatan eksperimental untuk mengembangkan sistem klasifikasi penyakit daun tomat berbasis citra digital. Sistem dirancang dengan memanfaatkan segmentasi *watershed* untuk memisahkan objek daun, ekstraksi fitur warna HSV dan fitur tekstur GLCM, serta algoritma *K-Nearest Neighbor* sebagai metode klasifikasi. Alur kerja sistem klasifikasi penyakit daun tomat ditunjukkan pada Gambar 1.

Dataset citra daun tomat yang digunakan dalam penelitian ini diperoleh dari sumber daring terbuka, yaitu platform Kaggle sebagai dataset utama. Dataset ini terdiri dari sembilan kelas, yaitu *bacterial spot*, *early blight*, *healthy*, *late blight*, *leaf mold*, *septoria leaf spot*, *spotted spider mite*, *target spot*, dan *yellow leaf curl virus*. Setiap citra telah diberi label sesuai kelas penyakitnya.

Seluruh citra pada dataset *single-leaf* dibagi secara terstratifikasi menjadi tiga bagian, yaitu data latih, data validasi, dan data uji. Pembagian dilakukan dengan proporsi sekitar 70% untuk data latih, 20% untuk data validasi, dan 10% untuk data uji pada setiap kelas. Distribusi jumlah citra pada setiap kelas serta pembagian data ditunjukkan pada Tabel 1.

TABEL 1
DISTRIBUSI DATASET SINGLE-LEAF

Kelas Penyakit	Total	Training	Validation	testing
Bacterial Spot	899	629	180	90
Early Blight	897	627	180	90
Healthy	900	630	180	90
Late Blight	898	628	180	90
Leaf Mold	908	638	180	90
Septoria Leaf Spot	899	629	180	90
Spotted Spider Mite	900	630	180	90

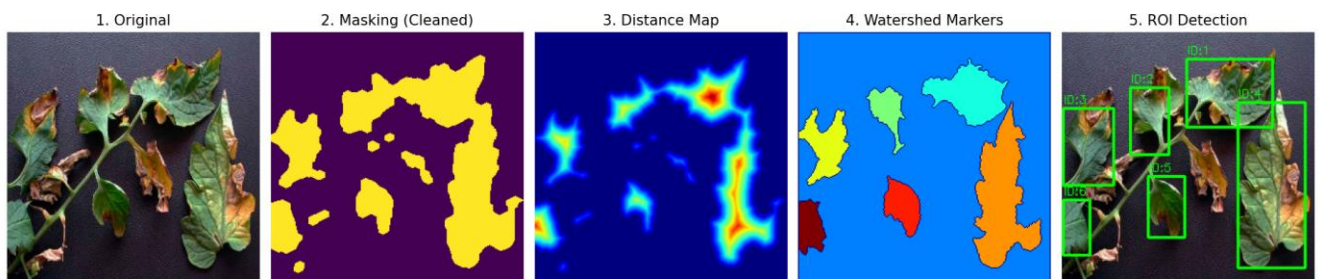
Kelas Penyakit	Total	Training	Validation	testing
Target Spot	900	630	180	90
Yellow Leaf Curl Virus	910	640	180	90
Total	8111	5681	1620	810

Pembagian dataset yang seimbang pada setiap kelas bertujuan untuk menghindari bias kelas serta memastikan proses pelatihan dan evaluasi model dilakukan secara adil dan representatif.

Selain dataset *single-leaf*, penelitian ini juga menggunakan dataset tambahan berupa citra *multi-leaf* yang diperoleh dari Kaggle dan Pinterest. Dataset *multi-leaf* digunakan khusus untuk menguji kemampuan sistem dalam memisahkan dan mengklasifikasikan beberapa daun dalam satu citra. Dataset ini terdiri dari 20 citra multi-daun dengan total 166 daun hasil segmentasi menggunakan metode *watershed*. Dataset *multi-leaf* tidak digunakan pada proses pelatihan, melainkan hanya digunakan pada tahap pengujian sistem.

Tahap pra pemrosesan citra meliputi perubahan ukuran citra, konversi ruang warna dari RGB ke HSV, serta penerapan *thresholding* berdasarkan nilai HSV untuk memisahkan objek daun dari latar belakang. Ruang warna HSV dipilih karena mampu merepresentasikan perubahan warna daun secara lebih stabil terhadap variasi pencahayaan dibandingkan ruang warna RGB. Rentang *threshold* HSV ditentukan berdasarkan distribusi warna dominan daun tomat pada dataset, yaitu *lower threshold* [25,40,40] dan *upper threshold* [85,255,255]. Rentang tersebut digunakan untuk mempertahankan area daun berwarna hijau sekaligus mengurangi pengaruh latar belakang dan *noise* pada citra. Hasil *thresholding* kemudian diproses menggunakan operasi morfologi untuk mengurangi *noise* dan memperbaiki bentuk objek daun. Tahapan ini bertujuan untuk menghasilkan citra biner yang merepresentasikan area daun secara lebih akurat.

Segmentasi objek daun dilakukan menggunakan metode *watershed* untuk memisahkan daun-daun yang saling bersentuhan atau bertumpuk dalam satu citra. Proses segmentasi diawali dengan perhitungan transformasi jarak (*distance transform*) dan dilanjutkan dengan penentuan *marker* sebagai titik awal segmentasi. Hasil segmentasi *watershed* menghasilkan beberapa region yang merepresentasikan masing-masing objek daun, sehingga setiap daun dapat diproses secara individual pada tahap berikutnya. Tahapan segmentasi *watershed* dan proses ekstraksi ROI daun ditunjukkan pada Gambar 2, yang meliputi citra asli daun tomat, hasil masking dan pembersihan morfologi, hasil *distance transform*, pembentukan marker *watershed*, serta hasil deteksi ROI daun.



Gambar 2. Tahapan segmentasi *watershed* dan ekstraksi ROI: (a) citra asli daun tomat, (b) hasil masking dan pembersihan morfologi, (c) hasil *distance transform*, (d) pembentukan marker *watershed*, dan (e) hasil deteksi ROI daun.

Ekstraksi fitur dilakukan terhadap setiap objek daun hasil segmentasi. Fitur warna diekstraksi dari ruang warna HSV dengan menghitung nilai rata-rata komponen *Hue*, *Saturation*, dan *Value* sebagai representasi karakteristik warna daun. Selain fitur warna, fitur tekstur diekstraksi menggunakan metode *Gray Level Co-occurrence Matrix (GLCM)*. Parameter tekstur yang digunakan meliputi *contrast*, *energy*, *homogeneity*, dan *correlation*. Matriks GLCM dibentuk menggunakan jarak antar piksel sebesar 1 piksel dengan empat orientasi sudut, yaitu 0°, 45°, 90°, dan 135°. Penggunaan beberapa orientasi sudut bertujuan untuk memperoleh representasi tekstur daun dari berbagai arah pola bercak penyakit. Parameter *contrast*, *energy*, *homogeneity*, dan *correlation* dipilih karena mampu merepresentasikan tingkat perbedaan intensitas, keseragaman tekstur, hubungan antar piksel, serta distribusi pola permukaan daun yang dipengaruhi oleh penyakit. Dua parameter utama GLCM yang digunakan dalam penelitian ini didefinisikan pada Persamaan (1) dan Persamaan (2).

$$Contrast = \sum_{i,j} (i - j)^2 P(i, j) \quad (1)$$

$$Energy = \sum_{i,j} P(i, j)^2 \quad (2)$$

Proses klasifikasi dilakukan menggunakan algoritma *K-Nearest Neighbor*. Jarak antar vektor fitur dihitung menggunakan jarak *Euclidean* sebagaimana ditunjukkan pada Persamaan (3). Nilai parameter K pada algoritma *K-Nearest Neighbor* ditetapkan sebesar 5 berdasarkan hasil *preliminary experiment* menggunakan beberapa variasi nilai K, yaitu 3, 5, 7, dan 9. Hasil pengujian menunjukkan bahwa nilai K=5 memberikan performa klasifikasi yang lebih stabil pada data validasi dibandingkan nilai lainnya.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (3)$$

Penentuan kelas dilakukan menggunakan mekanisme voting mayoritas berdasarkan K tetangga terdekat sebagaimana dinyatakan pada Persamaan (4).

$$\hat{y} = \arg \max_c \sum_{k=1}^K I(y_k = c) \quad (4)$$

Pengujian sistem dilakukan menggunakan data latih, data validasi, dan data uji. Evaluasi kinerja sistem dilakukan menggunakan metrik akurasi, presisi, *recall*, *F1-score*, dan *confusion matrix* untuk menganalisis kesalahan klasifikasi antar kelas penyakit daun tomat. Selain itu, pengujian khusus dilakukan pada citra multi-daun untuk mengevaluasi efektivitas segmentasi *watershed* dalam memisahkan objek daun serta dampaknya terhadap hasil klasifikasi.

III. HASIL DAN PEMBAHASAN

A. Hasil Segmentasi Daun Menggunakan Metode Watershed

Segmentasi merupakan tahap krusial dalam sistem klasifikasi penyakit daun tomat, khususnya ketika citra yang digunakan mengandung lebih dari satu objek daun (*multi-leaf image*). Pada penelitian ini, metode *watershed* diterapkan untuk memisahkan daun yang saling bersentuhan agar setiap daun dapat diproses secara individual pada tahap selanjutnya.

Tahapan segmentasi *watershed* ditunjukkan secara rinci pada Gambar 2. Gambar 2(a) memperlihatkan citra asli daun tomat sebagai masukan sistem. Selanjutnya, pada Gambar 2(b) dilakukan proses masking warna hijau daun serta pembersihan citra menggunakan operasi morfologi untuk menghilangkan *noise* latar belakang. Tahap ini bertujuan untuk mempertahankan area daun dan menghilangkan objek non-daun.

Pada Gambar 2(c) ditampilkan hasil *distance transform*, yang merepresentasikan jarak setiap piksel terhadap batas terdekat objek. Informasi jarak ini kemudian digunakan untuk membentuk marker *watershed* sebagaimana ditunjukkan pada Gambar 2(d). Marker digunakan sebagai titik awal proses *flooding* sehingga objek daun yang saling bersentuhan dapat dipisahkan secara lebih terkontrol. Hasil akhir segmentasi berupa deteksi *Region of Interest* (ROI) daun tomat ditunjukkan pada Gambar 2(e).

Berdasarkan hasil segmentasi, metode *watershed* mampu memisahkan sebagian besar objek daun dengan baik, termasuk daun yang saling bertumpuk pada citra *multi-leaf*. Segmentasi yang dihasilkan memiliki batas objek yang cukup jelas sehingga setiap ROI dapat digunakan secara efektif pada tahap ekstraksi fitur. Untuk mengurangi kasus *over-segmentation*, dilakukan penyaringan objek berdasarkan ukuran *minimum bounding box*. Objek dengan ukuran terlalu kecil dianggap sebagai *noise* dan tidak diproses pada tahap klasifikasi. Selain itu, penggunaan operasi morfologi sebelum proses *watershed* membantu mengurangi fragmentasi objek sehingga segmentasi menjadi lebih stabil.

Hasil pengujian menunjukkan bahwa penggunaan segmentasi *watershed* meningkatkan performa sistem klasifikasi pada citra *multi-leaf*. Sistem tanpa segmentasi menghasilkan akurasi sebesar 74.8%, sedangkan sistem dengan segmentasi *watershed* mampu mencapai akurasi sebesar 81.73%. Hal ini menunjukkan bahwa pemisahan objek daun sebelum proses ekstraksi fitur mampu menghasilkan fitur HSV dan GLCM yang lebih representatif terhadap masing-masing daun.

B. Hasil Ekstraksi Fitur Warna dan Tekstur Pada Data Train

Setelah ROI daun diperoleh dari proses segmentasi, tahap berikutnya adalah ekstraksi fitur warna dan tekstur. Pada penelitian ini, fitur warna diekstraksi menggunakan ruang warna HSV (*Hue Saturation Value*) karena lebih stabil terhadap perubahan pencahayaan dibandingkan ruang warna RGB. Fitur HSV digunakan untuk merepresentasikan perubahan warna daun akibat infeksi penyakit, seperti kekuningan, bercak kecoklatan, dan area nekrosis.

Sementara itu, fitur tekstur diekstraksi menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM). Parameter GLCM dibentuk menggunakan jarak antar piksel sebesar 1 piksel dengan empat orientasi sudut, yaitu 0°, 45°, 90°, dan 135°. Penggunaan beberapa orientasi sudut bertujuan untuk memperoleh representasi tekstur daun dari berbagai arah pola bercak penyakit. Fitur tekstur yang digunakan meliputi *contrast*, *energy*, *homogeneity*, dan *correlation*.

TABEL 2
NILAI FITUR HASIL EKSTRAKSI DATA TRAIN

No.	Label	Hue Mean	Sat Mean	Val Mean	Contrast	Energy
1.	Bacterial Spot	71.30	66.32	54.34	0.44	0.83
2.	Bacterial Spot	71.60	62.35	63.41	1.05	0.71
9	Early Blight	98.14	59.76	52.18	0.96	0.76
4.	Early Blight	117.11	67.33	59.79	2.82	0.65
5.	Healthy	100.33	56.73	50.11	7.15	0.46
6.	Healthy	94.28	55.22	66.08	7.95	0.44
7.	Late Blight	22.90	22.46	73.32	2.07	0.66
8.	Late Blight	82.72	42.83	44.77	0.54	0.84
9.	Leaf Mold	110.41	25.83	37.98	1.02	0.69
10.	Leaf Mold	27.02	40.25	50.84	1.82	0.64
11.	Septoria Leaf Spot	107.98	72.12	65.75	2.81	0.58
12.	Septoria Leaf Spot	85.90	63.93	92.46	0.66	0.80
13.	Spotted Spider Mite	17.99	13.90	93.88	2.51	0.65
14.	Spotted Spider Mite	119.32	61.58	47.03	2.19	0.61
15.	Target Spot	110.97	44.92	50.38	3.08	0.62
16.	Target Spot	109.97	62.70	50.16	2.84	0.58
17.	Yellow Leaf Curl Virus	50.84	18.41	74.61	0.19	0.91
18.	Yellow Leaf Curl Virus	38.63	13.73	90.39	0.17	0.92

Berdasarkan Tabel 2, terlihat bahwa setiap kelas penyakit memiliki karakteristik fitur yang berbeda. Daun sehat cenderung memiliki distribusi warna yang lebih seragam, sedangkan daun yang terinfeksi penyakit menunjukkan variasi warna dan tekstur yang lebih kompleks akibat munculnya bercak penyakit.

Fitur HSV efektif dalam merepresentasikan perubahan warna daun akibat infeksi penyakit, sedangkan fitur tekstur berbasis GLCM mampu menangkap ketidakteraturan pola permukaan daun yang disebabkan oleh bercak penyakit. Kombinasi kedua fitur tersebut menghasilkan representasi visual yang lebih diskriminatif sehingga meningkatkan kemampuan klasifikasi menggunakan algoritma KNN.

C. Hasil Klasifikasi Penyakit Daun Tomat

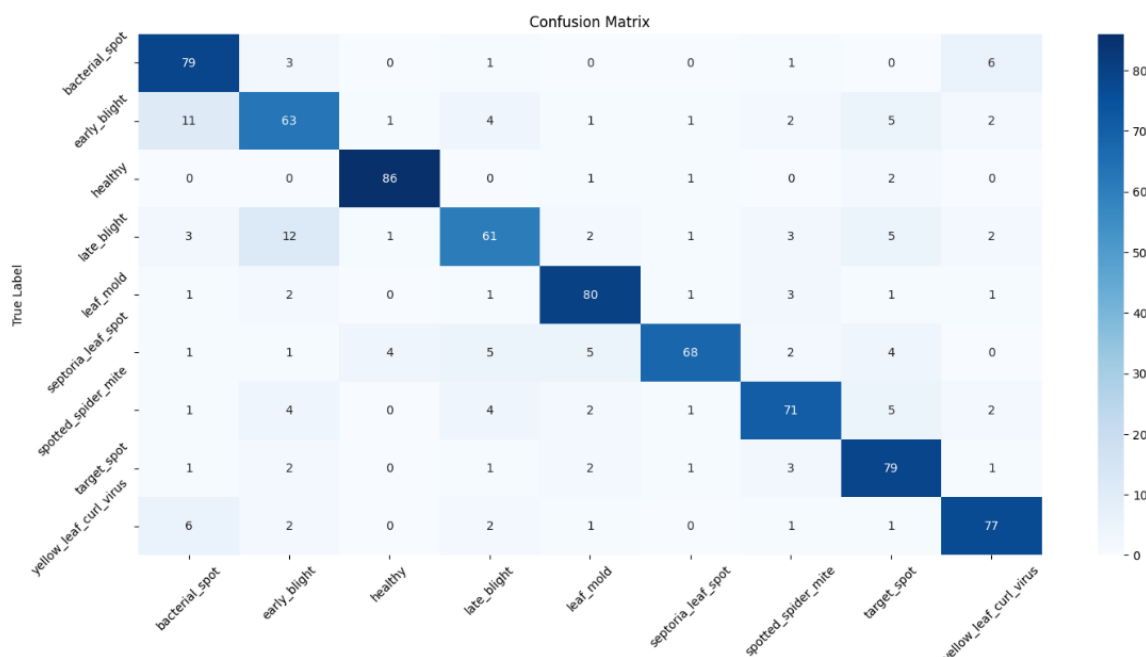
Pengujian model klasifikasi dilakukan menggunakan data test yang tidak digunakan pada tahap pelatihan. Algoritma klasifikasi yang digunakan adalah *K-Nearest Neighbor (KNN)* dengan nilai parameter $K=5$, yang dipilih berdasarkan hasil *preliminary experiment* menggunakan beberapa variasi nilai K .

TABEL 3
HASIL PENGUJIAN MODEL PADA DATA TEST

No.	Kelas Daun	TP	TN	FP	FN	Precision	Recall	F1-Score
1.	Bacterial Spot	79	696	24	11	0.7670	0.8878	0.8187
2.	Early Blight	64	691	29	26	0.6882	0.7111	0.6995
3.	Healthy	86	714	6	4	0.9384	0.9556	0.9451
4.	Late Blight	61	701	19	29	0.762	0.6778	0.7177
5.	Leaf Mold	81	706	14	9	0.8526	0.9000	0.8757
6.	Septoria Leaf Spot	68	714	6	22	0.9189	0.7556	0.8293
7.	Spotted Spider Mite	71	704	16	19	0.8161	0.7889	0.8023
8.	Target Spot	79	697	23	11	0.7745	0.8778	0.8229
9.	Yellow Leaf Curl Virus	73	709	11	17	0.8690	0.8111	0.8391
Rata-Rata						0.8204	0.8175	0.8190

Hasil pengujian menunjukkan bahwa sistem klasifikasi yang diusulkan mampu mencapai akurasi sebesar 81.73%. Nilai *precision*, *recall*, dan *F1-score* yang relatif seimbang menunjukkan bahwa model mampu bekerja secara cukup konsisten pada berbagai kelas penyakit daun tomat.

Kelas *Healthy* memiliki performa terbaik dengan nilai *F1-score* sebesar 0.9451 karena karakteristik warna hijau dan tekstur daun yang relatif seragam lebih mudah dibedakan dibandingkan kelas penyakit lainnya. Sebaliknya, kelas *Early Blight* dan *Late Blight* menunjukkan performa yang lebih rendah karena kedua penyakit tersebut memiliki pola bercak nekrosis dengan distribusi warna dan tekstur yang hampir serupa. Kemiripan karakteristik visual tersebut menyebabkan overlap fitur HSV dan GLCM pada beberapa sampel citra.



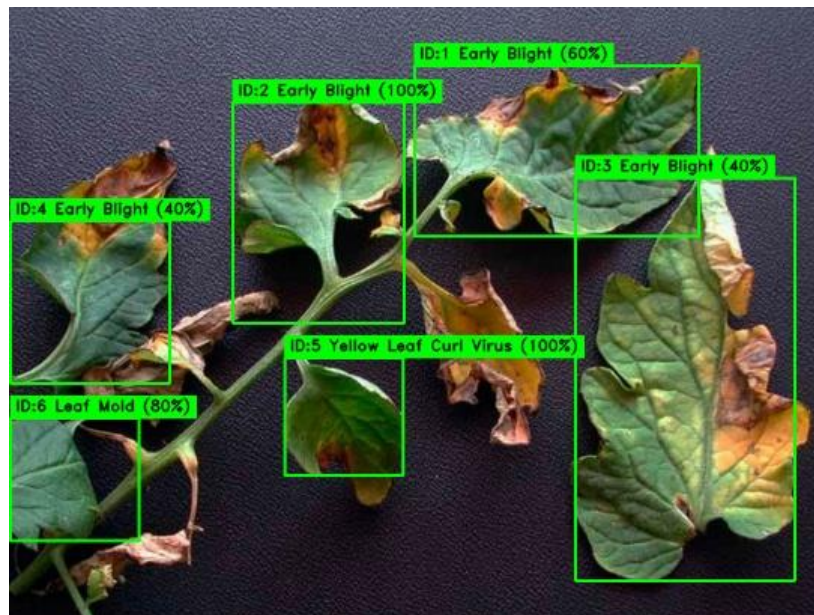
Gambar 3. Confusion Matrix Hasil Klasifikasi

Berdasarkan *confusion matrix*, pada Gambar 3 sebagian besar kesalahan klasifikasi terjadi pada kelas penyakit yang memiliki kemiripan pola bercak dan distribusi warna. Hal ini menunjukkan bahwa pendekatan berbasis *handcrafted features* masih memiliki keterbatasan dalam membedakan penyakit dengan karakteristik visual yang sangat mirip.

D. Visualisasi Hasil Deteksi dan Prediksi Penyakit Daun

Selain evaluasi numerik, penelitian ini juga menyajikan visualisasi hasil deteksi dan prediksi penyakit daun tomat. Gambar hasil prediksi menampilkan citra uji beserta ROI daun hasil segmentasi *watershed* yang telah diberi label kelas penyakit berdasarkan hasil prediksi model KNN.

Visualisasi ini diletakkan setelah penyajian Tabel 3 karena berfungsi sebagai bukti visual yang mendukung hasil evaluasi numerik. Dari hasil visualisasi tersebut, terlihat bahwa sistem mampu memberikan label penyakit yang sesuai pada sebagian besar objek daun, baik pada citra *single-leaf* maupun *multi-leaf*. Hal ini menunjukkan bahwa proses segmentasi *watershed* berperan penting dalam memastikan setiap daun diproses dan diklasifikasikan secara *individual*.



Gambar 4. Visualisasi hasil prediksi penyakit daun tomat pada citra uji

Visualisasi hasil prediksi penyakit daun tomat ditunjukkan pada Gambar 4. Gambar tersebut memperlihatkan hasil segmentasi *watershed* yang berhasil memisahkan setiap objek daun secara individual, diikuti dengan pemberian label kelas penyakit berdasarkan hasil klasifikasi menggunakan algoritma KNN. Visualisasi ini menunjukkan bahwa sistem mampu mendeteksi dan mengklasifikasikan setiap daun secara terpisah, termasuk pada citra yang mengandung lebih dari satu objek daun.

E. Analisis Pengaruh Segmentasi Watershed terhadap Kinerja Sistem

Segmentasi *watershed* memberikan kontribusi signifikan terhadap peningkatan kinerja sistem klasifikasi. Dengan memisahkan objek daun sebelum tahap ekstraksi fitur, fitur HSV dan GLCM yang dihasilkan menjadi lebih spesifik terhadap masing-masing daun dan tidak terpengaruh oleh objek lain atau latar belakang.

Perbandingan hasil klasifikasi antara sistem dengan dan tanpa segmentasi menunjukkan bahwa penggunaan *watershed* mampu meningkatkan akurasi dan stabilitas prediksi, terutama pada citra *multi-leaf*. Hal ini menegaskan bahwa tahap segmentasi merupakan komponen penting dalam sistem klasifikasi penyakit daun tomat berbasis citra digital.

F. Perbandingan dengan Penelitian Lain yang Menekankan Segmentasi

Dalam beberapa tahun terakhir, penelitian mengenai deteksi dan klasifikasi penyakit daun tanaman banyak menekankan pentingnya tahap segmentasi sebelum proses klasifikasi dilakukan. Segmentasi yang baik mampu memisahkan area daun dari latar belakang maupun objek lain sehingga fitur visual yang diekstraksi menjadi lebih representatif terhadap karakteristik penyakit daun. Selain itu, segmentasi juga berpengaruh terhadap peningkatan performa sistem klasifikasi, khususnya pada citra dengan objek daun yang saling bertumpuk atau bersentuhan.

Pendekatan segmentasi modern berbasis *deep learning* seperti *U-Net* dan variannya banyak digunakan pada penelitian klasifikasi penyakit tanaman. Penelitian *MC-UNet* yang dikembangkan oleh Deng et al. [12] menunjukkan bahwa segmentasi berbasis *deep learning* mampu menghasilkan akurasi segmentasi yang tinggi dengan kemampuan menangkap detail tepi objek daun secara lebih presisi. Pendekatan serupa juga digunakan oleh Shoaib et al. [11], yang mengkombinasikan segmentasi berbasis *U-Net* dengan model klasifikasi lanjutan dan menghasilkan tingkat akurasi klasifikasi di atas 98%.

Untuk mengevaluasi posisi penelitian yang diusulkan terhadap penelitian sebelumnya, dilakukan perbandingan dengan beberapa metode klasifikasi penyakit daun tomat berbasis machine learning dan *deep learning* sebagaimana ditunjukkan pada Tabel 4.

TABEL 4
PERBANDINGAN DENGAN PENELITIAN SEBELUMNYA

Penelitian	Metode	Dataset	Akurasi	Kompleksitas Komputasi
Nyasulu et al. (2023) [1]	GLCM + Machine Learning	Tomato Fungal Disease Dataset	84%	Rendah
Agusta & Kaswidjanti (2023) [6]	HSV + GLCM + KNN	PlantVillage	89%	Rendah
Wajid et al. (2024) [10]	CNN + Machine Learning	PlantVillage	96%	Tinggi
Sun et al. (2025) [8]	Deep learning Feature Fusion	Tomato Leaf Dataset	98%	Sangat Tinggi
Deng et al. (2025) [12]	FCMNet Multimodal Fusion	Tomato Leaf Dataset	98.7%	Sangat Tinggi
Sahu et al. (2025) [14]	Mask R-CNN	Multi-leaf Tomato Dataset	97%	Sangat Tinggi
Gong, X. & Zhang, S (2023) [15]	YOLOv5	Natural Scene Tomato Leaves	95%	Tinggi
Penelitian ini	Watershed + HSV + GLCM + KNN	PlantVillage + Multi-leaf	81.73%	Rendah

Berdasarkan Tabel 4, metode berbasis *deep learning* umumnya menghasilkan tingkat akurasi yang lebih tinggi dibandingkan pendekatan berbasis *handcrafted features*. Penelitian oleh Sun et al. [8] dan Deng et al. [12] mampu mencapai akurasi di atas 98% menggunakan arsitektur *deep learning* modern yang memanfaatkan proses ekstraksi fitur otomatis dan segmentasi berbasis jaringan saraf.

Meskipun demikian, metode berbasis *deep learning* umumnya membutuhkan sumber daya komputasi yang besar, proses pelatihan yang kompleks, serta ketergantungan terhadap GPU dan dataset dalam jumlah besar. Sebaliknya, penelitian ini menggunakan pendekatan berbasis *handcrafted features* yang lebih ringan secara komputasi dengan memanfaatkan segmentasi *watershed*, fitur warna HSV, fitur tekstur GLCM, dan algoritma KNN.

Penggunaan kombinasi HSV dan GLCM pada penelitian ini terbukti cukup efektif dalam merepresentasikan karakteristik visual penyakit daun tomat. Fitur HSV mampu menangkap perubahan warna daun akibat infeksi penyakit, seperti kekuningan, bercak kecoklatan, dan nekrosis daun. Sementara itu, fitur tekstur berbasis GLCM mampu merepresentasikan pola ketidakrataan permukaan daun akibat munculnya bercak penyakit. Kombinasi kedua fitur tersebut menghasilkan representasi visual yang lebih diskriminatif sehingga mampu meningkatkan kemampuan klasifikasi KNN pada data multikelas.

Selain itu, segmentasi *watershed* juga memberikan pengaruh terhadap distribusi fitur yang dihasilkan. Dengan memisahkan objek daun secara individual sebelum tahap ekstraksi fitur, nilai HSV dan GLCM yang diperoleh menjadi lebih spesifik terhadap masing-masing daun serta tidak tercampur dengan latar belakang maupun objek daun lainnya. Hal tersebut menyebabkan distribusi fitur antar kelas menjadi lebih terpisah sehingga membantu meningkatkan stabilitas proses klasifikasi.

Berdasarkan hasil pengujian, penggunaan segmentasi *watershed* mampu meningkatkan akurasi sistem dari 74.8% menjadi 81.73% pada citra *multi-leaf*. Hasil tersebut menunjukkan bahwa proses segmentasi memiliki kontribusi yang signifikan terhadap peningkatan performa sistem klasifikasi, khususnya pada kondisi citra yang lebih kompleks.

Meskipun akurasi penelitian ini masih berada di bawah pendekatan *deep learning*, metode yang diusulkan memiliki keunggulan pada efisiensi komputasi, kemudahan implementasi, serta tidak memerlukan GPU maupun proses pelatihan model yang kompleks. Oleh karena itu, pendekatan berbasis *handcrafted features* masih relevan untuk diterapkan pada perangkat dengan sumber daya terbatas, seperti sistem pertanian berbasis *edge computing* dan perangkat bergerak.

Namun demikian, sistem yang diusulkan masih memiliki beberapa keterbatasan. Pada kondisi lapangan dengan pencahayaan ekstrem, bayangan objek, dan *occlusion* antara daun yang terlalu kompleks, proses segmentasi dan ekstraksi fitur dapat mengalami penurunan performa. Selain itu, beberapa kelas penyakit, seperti *Early Blight* dan *Late Blight*, memiliki karakteristik visual yang sangat mirip sehingga menyebabkan terjadinya *overlap* fitur HSV dan GLCM pada beberapa sampel citra.

G. Pembahasan Umum

Hasil pengujian menunjukkan bahwa kombinasi segmentasi *watershed*, fitur HSV, fitur tekstur GLCM, dan algoritma KNN mampu melakukan klasifikasi penyakit daun tomat pada citra *multi-leaf* dengan akurasi sebesar 81.73%. Penggunaan segmentasi *watershed* memberikan pengaruh terhadap peningkatan performa sistem karena mampu memisahkan objek daun sebelum proses ekstraksi fitur dilakukan.

Pada citra *multi-leaf*, objek daun sering saling berdekatan maupun bertumpuk sehingga proses segmentasi menjadi tahap yang penting. Dengan memisahkan objek daun secara individual, fitur warna HSV dan fitur tekstur GLCM yang dihasilkan menjadi lebih spesifik terhadap masing-masing objek daun dan tidak tercampur dengan latar belakang maupun objek lainnya. Kondisi tersebut menyebabkan distribusi fitur antar kelas menjadi lebih terpisah sehingga membantu meningkatkan stabilitas proses klasifikasi menggunakan KNN.

Fitur HSV mampu merepresentasikan perubahan warna daun akibat infeksi penyakit, sedangkan fitur GLCM mampu menangkap karakteristik tekstur permukaan daun yang mengalami perubahan akibat bercak penyakit. Kombinasi kedua fitur tersebut menghasilkan representasi visual yang lebih informatif sehingga cukup efektif digunakan pada proses klasifikasi data multikelas.

Berdasarkan hasil pengujian, penggunaan segmentasi *watershed* meningkatkan akurasi sistem dari 74.8% menjadi 81.73%. Hasil tersebut menunjukkan bahwa proses segmentasi memiliki kontribusi yang signifikan terhadap performa sistem klasifikasi, khususnya pada citra dengan kompleksitas objek yang lebih tinggi.

Meskipun akurasi yang diperoleh masih berada di bawah pendekatan berbasis *deep learning*, metode yang diusulkan memiliki keunggulan pada efisiensi komputasi dan kemudahan implementasi. Sistem tidak memerlukan proses pelatihan model yang kompleks maupun penggunaan GPU sehingga lebih ringan untuk diterapkan pada perangkat dengan sumber daya terbatas. Namun demikian, sistem yang diusulkan masih memiliki keterbatasan pada kondisi pencahayaan yang tidak seragam, keberadaan bayangan objek, serta occlusion antar daun yang kompleks. Selain itu, beberapa kelas penyakit memiliki karakteristik visual yang mirip sehingga menyebabkan terjadinya overlap fitur pada proses klasifikasi.

IV. SIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa kombinasi segmentasi *watershed*, ekstraksi fitur warna HSV, fitur tekstur GLCM, dan algoritma *K-Nearest Neighbor* (KNN) mampu digunakan untuk membangun sistem klasifikasi penyakit daun tomat secara otomatis dengan performa yang cukup stabil dan efisien secara komputasi. Sistem yang diusulkan mampu mencapai akurasi sebesar 81.73% pada data uji dengan nilai *precision*, *recall*, dan *F1-score* yang relatif seimbang pada sebagian besar kelas penyakit.

Hasil penelitian juga menunjukkan bahwa segmentasi *watershed* memberikan kontribusi signifikan terhadap peningkatan performa sistem, khususnya pada citra *multi-leaf*. Penggunaan segmentasi sebelum tahap ekstraksi fitur mampu meningkatkan akurasi sistem dari 74.8% menjadi 81.73%, karena fitur HSV dan GLCM yang dihasilkan menjadi lebih spesifik terhadap masing-masing objek daun.

Selain itu, penelitian ini membuktikan bahwa pendekatan berbasis *handcrafted features* masih memiliki potensi yang relevan sebagai alternatif metode klasifikasi penyakit tanaman dengan kebutuhan komputasi rendah. Kombinasi HSV dan GLCM mampu merepresentasikan karakteristik visual penyakit daun tomat secara cukup efektif, sementara algoritma KNN memberikan proses klasifikasi yang sederhana dan mudah diimplementasikan tanpa memerlukan GPU maupun proses pelatihan yang kompleks.

Meskipun demikian, sistem yang diusulkan masih memiliki keterbatasan pada kondisi pencahayaan ekstrem, *occlusion* antara daun, serta kemiripan visual antar beberapa kelas penyakit daun tomat. Oleh karena itu, penelitian selanjutnya dapat difokuskan pada pengembangan metode segmentasi yang lebih adaptif, penambahan fitur bentuk dan morfologi, optimasi parameter klasifikasi, maupun integrasi pendekatan *deep learning* ringan (*lightweight deep learning*) untuk meningkatkan performa sistem pada kondisi lapangan yang lebih kompleks.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Program Studi Teknik Informatika, Fakultas Teknik Elektro dan Teknologi Informasi, Institut Teknologi Adhi Tama Surabaya atas dukungan akademik dan fasilitas yang diberikan sehingga penelitian ini dapat dilaksanakan dengan baik. Apresiasi juga disampaikan kepada pihak-pihak yang telah memberikan arahan dan dukungan selama proses penelitian dan penyusunan makalah ini.

DAFTAR PUSTAKA

- [1] C. Nyasulu *et al.*, "A comparative study of Machine Learning-based classification of Tomato fungal diseases: Application of GLCM texture features," *Heliyon*, vol. 9, no. 11, 2023.
- [2] A. R. Al-Shamasneh and R. W. Ibrahim, "Classification of tomato leaf images for detection of plant disease using conformable polynomials image features," *MethodsX*, vol. 13, 2024.
- [3] A. Ahmad, D. Saraswat, and A. El Gamal, "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools," *Smart Agric. Technol.*, vol. 3, 2023.
- [4] B. Yang *et al.*, "A novel plant type, leaf disease and severity identification framework using CNN and transformer with multi-label method," *Sci. Rep.*, vol. 14, no. 1, pp. 1–15, 2024.
- [5] A. Das, F. Pathan, J. R. Jim, M. M. Kabir, and M. F. Mridha, "Deep learning-based classification, detection, and segmentation of tomato leaf diseases: A state-of-the-art review," *Artif. Intell. Agric.*, vol. 15, no. 2, pp. 192–220, 2025.
- [6] S. W. Agusta and W. Kaswidjanti, "The Implementation of Color Feature Extraction and Gray Level Co-occurrence Matrix Combination in K-

- Nearest Neighbor Classification Method for Tomato Leaf Disease Identification,” *Telematika*, vol. 20, no. 2, p. 250, 2023.
- [7] F. A. Laiskodat and M. Rahardi, “Multiclass Classification of Tomato Leaf Diseases Using GLCM , Color , and Shape Feature Extraction with Optimized XGBoost,” *J. Appl. Informatics Comput.*, vol. 9, no. 6, pp. 3594–3600, 2025.
- [8] H. Sun *et al.*, “Efficient deep learning-based tomato leaf disease detection through global and local feature fusion,” *BMC Plant Biol.*, vol. 25, no. 1, Dec. 2025.
- [9] S. Deng, J. Zhu, Y. Hu, M. He, and Y. Xia, “Tomato Leaf Disease Identification Framework FCMNet Based on Multimodal Fusion,” *Plants*, vol. 14, no. 15, Aug. 2025.
- [10] A. H. Wajid, N. Saher, S. A. Nawaz, M. Arshad, and M. Nasir, “Tomato Leaf Disease Detection and Classification Using Convolutional Neural Network and Machine Learning,” *J. Comput. Biomed. Informatics*, 2024.
- [11] M. Shoaib *et al.*, “Deep learning-based segmentation and classification of leaf images for detection of tomato plant disease,” *Front. Plant Sci.*, vol. 13, Oct. 2022.
- [12] Y. Deng *et al.*, “An Effective Image-Based Tomato Leaf Disease Segmentation Method Using MC-UNet,” *Plant Phenomics*, vol. 5, p. 49, 2023.
- [13] Z. Bi *et al.*, “Apple leaf disease severity grading based on deep learning and the DRL-Watershed algorithm,” *Sci. Rep.*, vol. 15, no. 1, 2025.
- [14] K. Sahu, S. Tiwari, M. K. Singh, and J. Pahareeya, “Plant disease detection using a hybrid dilated CNN with attention mechanisms and optimized mask RCNN segmentation,” *Sci. Rep.*, vol. 15, no. 4, pp. 1–26, 2025.
- [15] X. Gong and S. Zhang, “A High-Precision Detection Method of Apple Leaf Diseases Using Improved Faster R-CNN,” *Agriculture*, vol. 13, 2023.