

Klasifikasi Penyakit Daun Pada Tanaman Terong dengan Metode K-Nearest Neighbors

Oke Hariansyah, Saprudin, Yono Cahyono, Perani Rosyani*

Fakultas Ilmu Komputer, Program Studi Teknik Informatika, Universitas Pamulang, Kota Tangerang Selatan, Indonesia

Email: ¹dosen00840@unpam.ac.id, ²dosen00845@unpam.ac.id, ³dosen00843@unpam.ac.id, ⁴dosen00837@unpam.ac.id

Email Penulis Korespondensi: dosen00837@unpam.ac.id

Submitted: 19/02/2025; Accepted: 26/03/2025; Published: 28/03/2025

Abstrak—Tanaman terong (*Solanum melongena*) merupakan salah satu komoditas pertanian yang memiliki nilai ekonomi tinggi. Namun, berbagai penyakit daun dapat menghambat pertumbuhan dan menurunkan hasil panen. Oleh karena itu, identifikasi dan klasifikasi penyakit daun secara cepat dan akurat sangat penting untuk meningkatkan produktivitas pertanian. Penelitian ini mengusulkan penggunaan metode K-Nearest Neighbors (KNN) dalam klasifikasi penyakit daun pada tanaman terong berbasis citra. Model dikembangkan dengan menggunakan fitur histogram warna dari gambar daun sebagai dasar klasifikasi. Penelitian ini melibatkan proses pengumpulan dataset gambar daun terong dengan berbagai kategori penyakit, ekstraksi fitur warna menggunakan model warna RGB dan HSV, serta implementasi model KNN dengan parameter $k=3$. Evaluasi dilakukan menggunakan metrik akurasi, precision, recall, dan F1-score. Hasil eksperimen menunjukkan bahwa model KNN mampu mencapai akurasi sekitar 87%, namun masih terdapat tantangan seperti ketidakseimbangan dataset dan kesalahan klasifikasi pada kelas dengan kemiripan pola warna. Untuk meningkatkan akurasi, penelitian ini juga mengeksplorasi metode augmentasi data serta optimalisasi parameter model KNN. Dengan adanya penelitian ini, diharapkan metode KNN dapat lebih dioptimalkan dalam mendeteksi dan mengklasifikasikan penyakit daun pada tanaman terong, sehingga dapat membantu petani dalam mengelola tanaman secara lebih efisien dan tepat guna.

Kata Kunci: Klasifikasi Penyakit Daun; K-Nearest Neighbors; Tanaman Terong, Citra Digital; Histogram Warna.

Abstract—Eggplant (*Solanum melongena*) is an important agricultural commodity with high economic value. However, various leaf diseases can hinder its growth and reduce crop yields. Therefore, rapid and accurate identification and classification of leaf diseases are crucial for improving agricultural productivity. This study proposes the use of the K-Nearest Neighbors (KNN) method for classifying eggplant leaf diseases based on image analysis. The model is developed using color histogram features extracted from leaf images as the basis for classification. This research involves collecting a dataset of eggplant leaf images with various disease categories, extracting color features using RGB and HSV color models, and implementing a KNN model with $k=3$. The model's performance is evaluated using accuracy, precision, recall, and F1-score metrics. Experimental results show that the KNN model achieves an accuracy of approximately 87%, but challenges remain, such as dataset imbalance and misclassification of disease classes with similar color patterns. To improve accuracy, this study explores data augmentation techniques and optimizes the KNN model parameters. This research aims to enhance the effectiveness of KNN in detecting and classifying eggplant leaf diseases, ultimately assisting farmers in managing their crops more efficiently and effectively.

Keywords: Leaf Disease Classification; K-Nearest Neighbors; Eggplant; Digital Image; Color Histogram.

1. PENDAHULUAN

Tanaman terong (*Solanum melongena*) merupakan salah satu komoditas pertanian yang memiliki nilai ekonomi tinggi. Namun, seperti tanaman lainnya, terong rentan terhadap berbagai penyakit daun yang dapat menghambat pertumbuhan dan mengurangi hasil panen. Identifikasi dan klasifikasi penyakit daun pada tanaman terong secara cepat dan akurat sangat penting untuk meningkatkan produktivitas pertanian serta mengurangi risiko penyebaran penyakit [1], [2], [3].

Teknologi kecerdasan buatan (AI) semakin banyak digunakan dalam bidang pertanian, terutama dalam deteksi dan klasifikasi penyakit tanaman berbasis citra. Salah satu metode yang umum digunakan dalam klasifikasi citra adalah K-Nearest Neighbors (KNN)[4], [5]. Metode ini bekerja dengan membandingkan fitur dari gambar baru dengan data pelatihan yang telah dikategorikan sebelumnya. Dengan menggunakan fitur histogram warna dari gambar daun, model KNN dapat mengklasifikasikan jenis penyakit berdasarkan pola warna yang terdeteksi [6], [7].

Penelitian yang di tuliskan oleh Eli et al pada tahun 2024 mengembangkan sistem deteksi penyakit daun anggur berbasis citra menggunakan metode K-Nearest Neighbors (KNN) yang bertujuan meningkatkan akurasi deteksi dan efisiensi klasifikasi penyakit. Dalam penelitian ini, gambar daun anggur yang sehat dan terinfeksi penyakit dikumpulkan dan dianalisis dengan metode ekstraksi fitur yang meliputi Local Binary Pattern (LBP) untuk tekstur, deteksi tepi Canny dan Hu Moments untuk morfologi, serta histogram warna HSV. Dengan rasio 80:20 untuk data pelatihan dan pengujian, nilai K yang bervariasi (1-5) diuji, dan hasil terbaik diperoleh pada $K=1$ dengan akurasi 71%. Hasil ini menunjukkan bahwa pemilihan nilai K yang optimal sangat penting untuk mencapai performa terbaik dalam klasifikasi penyakit daun anggur. Kesimpulan dari penelitian ini menunjukkan bahwa metode KNN efektif dalam mengidentifikasi penyakit daun anggur [5].

Penelitian Kiswanto et al. yang di tulis pada tahun 2024 menyajikan pendekatan berbasis pengolahan citra digital dengan algoritma K-Nearest Neighbors (KNN) untuk mengklasifikasikan penyakit pada daun jambu. Proses diawali dengan preprocessing, termasuk mengubah ukuran gambar, mengurangi noise menggunakan Gaussian Blur, meningkatkan kualitas warna dengan CLAHE (Contrast Limited Adaptive Histogram Equalization), serta melakukan restorasi gambar. Ekstraksi fitur dilakukan dengan memanfaatkan Color Moments (meliputi mean, standar deviasi, dan skewness) untuk mendapatkan informasi warna, serta analisis tekstur berbasis Gray Level Co- occurrence Matrix

(GLCM) untuk fitur seperti contrast, correlation, energy, dan homogeneity. Fitur yang dihasilkan digunakan untuk melatih model KNN, yang kemudian diuji menggunakan train-test split dengan berbagai nilai parameter kkk. Hasil menunjukkan bahwa metode ini mampu mengklasifikasikan penyakit dengan tingkat akurasi yang baik [4].

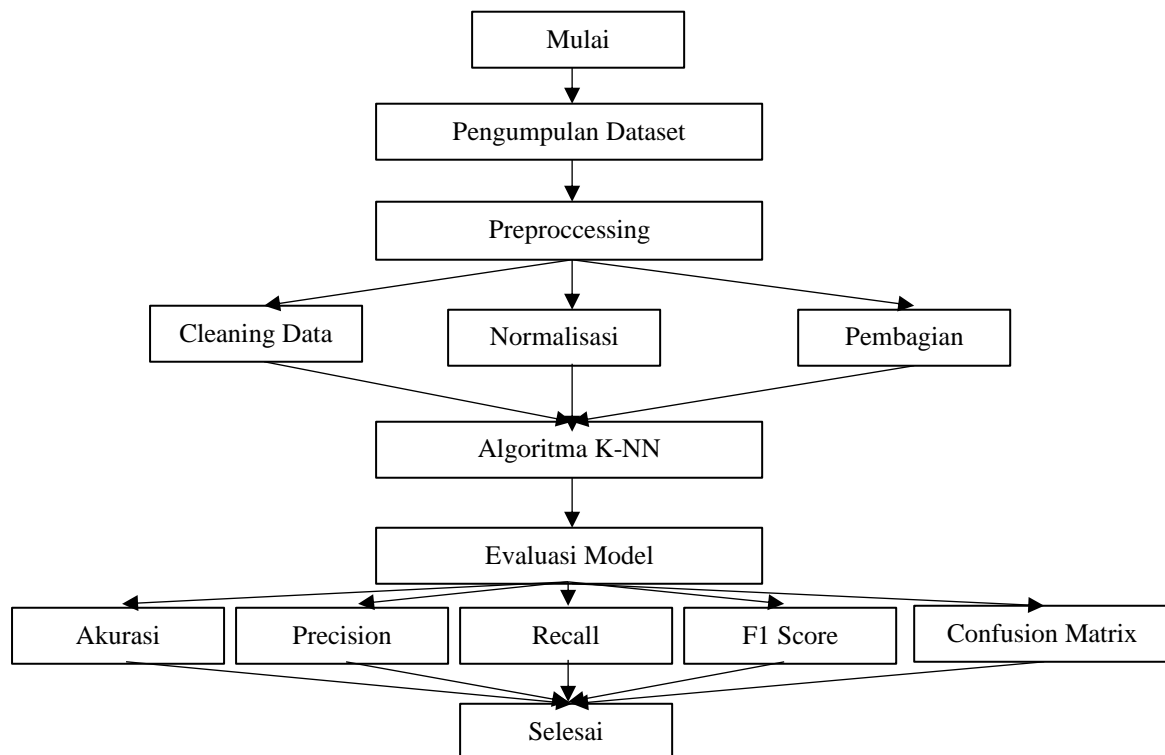
Mahrus dkk penelitian ini bertujuan untuk mencegah penyebaran penyakit pada tanaman anggur dengan mengidentifikasi penyakit pada daun menggunakan metode K-Nearest Neighbor (KNN) dan ekstraksi fitur Gray Level Co-Occurrence Matrix (GLCM) [8]. Dataset yang digunakan berasal dari PlantVillage Database, yang terdiri dari 4.063 data yang diklasifikasikan ke dalam empat kategori, yaitu Healthy, Black Rot, Esca, dan Leaf Blight. Pada penelitian ini, nilai K dalam metode KNN divariasikan antara K=1 hingga K=10 untuk menentukan parameter terbaik dalam proses klasifikasi. Sementara itu, parameter yang digunakan dalam ekstraksi fitur GLCM meliputi empat sudut derajat (0°, 45°, 90°, dan 135°) serta nilai jarak piksel tetangga sebesar 1. Ekstraksi fitur GLCM yang digunakan sebagai dasar klasifikasi mencakup Contrast, Correlation, Homogeneity, Energy, dan Entropy. Hasil evaluasi performa menggunakan Confusion Matrix menunjukkan bahwa metode ini mencapai akurasi tertinggi sebesar 88,6%, dengan presisi 50,5% dan recall 25,7%. Waktu klasifikasi terbaik diperoleh pada K = 10.

Meskipun metode KNN telah digunakan dalam berbagai penelitian untuk mendeteksi penyakit daun, belum ada penelitian yang secara spesifik mengoptimalkan metode ini untuk tanaman terong. Penelitian ini akan mengoptimalkan parameter K dalam metode K-Nearest Neighbors (KNN) untuk mendapatkan performa terbaik dalam klasifikasi penyakit daun pada tanaman terong. Dengan penelitian ini, kami menggunakan metode KNN untuk dapat lebih dioptimalkan dalam mendeteksi dan mengklasifikasikan penyakit daun pada tanaman terong, sehingga membantu petani dalam pengelolaan tanaman secara lebih efisien dan tepat guna.

2. METODOLOGI PENELITIAN

2.1 Alur Penelitian

Gambar 1 berikut merupakan alur dari penelitian yang dilakukan.



Gambar 1. Alur Penelitian

Dari Gambar 1, penelitian ini dilakukan melalui beberapa tahap utama, yaitu pengumpulan data, preprocessing data, implementasi model K-Nearest Neighbors (KNN), evaluasi model, dan analisis hasil. Berikut adalah langkah-langkah yang digunakan dalam penelitian ini:

a. Pengumpulan Data

Pengumpulan data merupakan tahap pertama yang sangat krusial dalam penelitian ini. Data yang digunakan untuk pelatihan dan pengujian model diperoleh dari berbagai sumber daring yang menyediakan dataset penyakit tanaman, khususnya daun terong. Salah satu dataset yang digunakan adalah Plant Disease Classification atau Vegetable Leaf Disease Dataset yang tersedia di platform Kaggle. Dataset ini mengandung berbagai gambar daun dengan kondisi berbeda, seperti daun yang sehat, daun terinfeksi penyakit, dan daun yang menunjukkan gejala

penyakit tertentu [9]. Setiap gambar daun dalam dataset telah diberi label berdasarkan kondisi kesehatan daun, yang dikelompokkan menjadi enam kategori utama, yaitu:

1. Healthy Leaf yaitu Daun yang memiliki kondisi normal tanpa tanda-tanda kerusakan, infeksi, atau gangguan lain, Warna daun biasanya hijau merata, bentuk daun sempurna, dan tidak terdapat perubahan tekstur atau noda.
2. Insect Pest Disease yaitu Daun yang menunjukkan tanda-tanda kerusakan akibat serangan hama serangga, seperti lubang, robekan, atau tekstur yang tidak rata. Biasanya disebabkan oleh aktivitas serangga seperti ulat atau kutu daun yang memakan jaringan daun.
3. Leaf Spot Disease Daun yang memiliki bercak-bercak kecil dengan warna coklat, kuning, atau hitam. Penyakit ini biasanya disebabkan oleh patogen jamur atau bakteri yang menyerang permukaan daun, menyebabkan perubahan warna pada area tertentu.
4. Mosaic Virus Disease Daun yang mengalami perubahan warna tidak merata, dengan pola seperti mosaik (campuran hijau terang, hijau tua, dan kuning). Penyakit ini disebabkan oleh infeksi virus yang memengaruhi pigmentasi daun, sering kali disertai deformasi bentuk daun.
5. Small Leaf Disease Daun yang berukuran lebih kecil dari normal, sering kali disebabkan oleh stres lingkungan, infeksi bakteri tertentu, atau ketidakseimbangan nutrisi. Selain ukurannya yang kecil, daun mungkin juga memiliki tekstur yang kaku atau rapuh.
6. White Mold Disease Daun yang menunjukkan pertumbuhan jamur putih berbulu halus di permukaan, sering kali pada bagian bawah daun. Penyakit ini disebabkan oleh infeksi jamur patogen dan sering terjadi di lingkungan dengan kelembapan tinggi.
7. Wilt Disease Daun yang tampak layu atau menggantung akibat gangguan pada sistem transportasi air dan nutrisi di tanaman. Penyebabnya bisa berupa infeksi jamur, bakteri, atau bahkan kondisi lingkungan seperti kekurangan air.

b. Preprocessing Data

1. Cleaning Data: tahap awal dalam pra-pemrosesan yang bertujuan untuk memastikan kualitas data yang akan digunakan dalam pelatihan. Proses ini melibatkan penghapusan gambar yang kabur atau gambar yang tidak relevan dengan penelitian. Beberapa langkah yang akan dilakukan dalam pembersihan yaitu Penghapusan Gambar Kabur maksudnya adalah gambar yang terlalu buram atau kabur akan dihapus, karena ini dapat mengurangi kemampuan model untuk mengenali fitur penting dalam gambar. Pemilihan gambar relevan yaitu hanya gambar yang menunjukkan kondisi daun yang relevan dengan penelitian yang akan dipertahankan. Gambar yang tidak sesuai, seperti gambar yang tidak mengandung gejala penyakit atau gambar yang menunjukkan daun bukan terong, akan dihapus [10].
2. Normalisasi : Setelah Cleaning data, gambar-gambar akan dinormalisasi untuk memastikan bahwa input yang diberikan kepada model berada dalam rentang yang konsisten. Proses normalisasi dengan cara Mengubah Ukuran Gambar yaitu semua gambar akan diubah ukurannya menjadi resolusi yang konsisten, misalnya 256x256 piksel, untuk memudahkan pemrosesan data. Kemudian normalisasi piksel yaitu dengan nilai piksel pada gambar akan dinormalisasi dengan membagi setiap nilai piksel dengan 255, sehingga semua nilai piksel berada dalam rentang 0 hingga 1. Normalisasi ini penting untuk mempercepat proses pelatihan dan meningkatkan konvergensi model [11], [12].
3. Pembagian dataset dengan data training 80% dan data tes 20%. Pembagian ini dilakukan menggunakan teknik stratified sampling untuk memastikan bahwa proporsi kelas yang ada pada dataset terjaga dengan baik di kedua subset, sehingga model tidak cenderung bias terhadap kelas tertentu.

c. Implementasi Algoritma K-NN

a. Pemilihan parameter K

Pada algoritma K-NN, parameter k menentukan jumlah tetangga terdekat yang digunakan untuk menentukan kelas suatu data. Eksperimen akan dilakukan dengan berbagai nilai k mulai dari 1 hingga 15 untuk menemukan nilai yang optimal. Penggunaan k-fold cross-validation (dengan k=5) akan diterapkan untuk menghindari overfitting.

b. Pelatihan Model

Cross-validation ini membagi data pelatihan menjadi 5 bagian dan menggunakan 4 bagian untuk pelatihan dan 1 bagian untuk pengujian secara bergantian, sehingga model dapat diuji dengan berbagai subset data [13], [14].

c. Penentuan bobot Euclidean Distance

Euclidean Distance digunakan dalam proses klasifikasi atau identifikasi dengan menghitung jarak antara vektor pelatihan dan vektor uji pada basis data yang ada. Perhitungan Euclidean Distance dapat dilakukan menggunakan persamaan berikut:

$$d(P, Q) = \sqrt{\sum_{i=1}^n (P_i - Q_i)^2} \quad (1)$$

d. Evaluasi Model Kinerja model dievaluasi menggunakan beberapa metrik:

Akurasi: Persentase prediksi yang benar dibandingkan dengan jumlah total prediksi.

Precision: Mengukur seberapa relevan prediksi terhadap kelas penyakit tertentu.

Recall: Mengukur sejauh mana model dapat mendeteksi penyakit yang sebenarnya ada.

F1-Score: Harmonic mean antara precision dan recall untuk memberikan gambaran keseimbangan performa model.

Confusion Matrix: Digunakan untuk menganalisis distribusi kesalahan klasifikasi antar kelas.

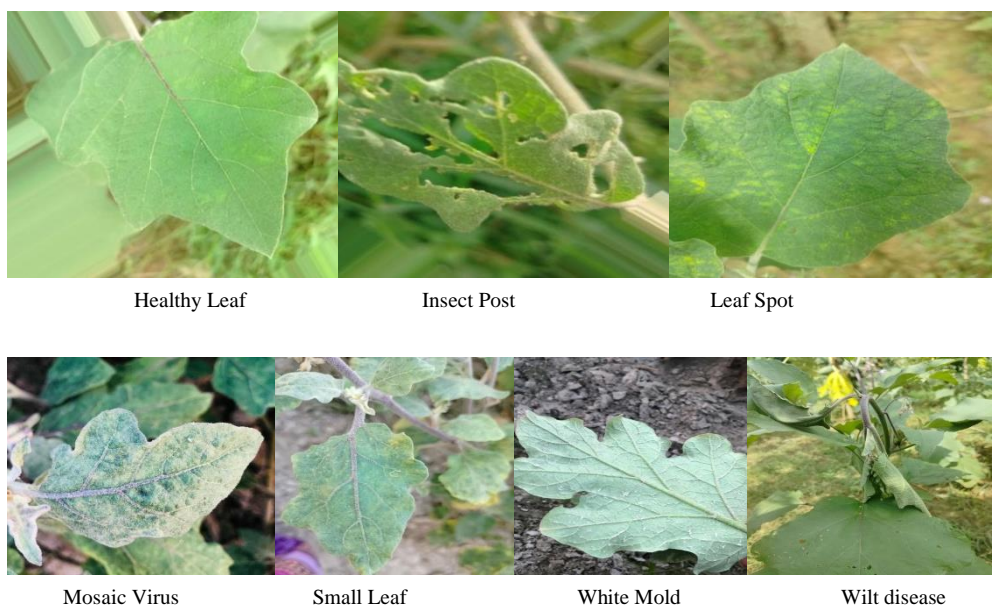
2.2 K-Nearest Neighbors (KNN)

Algoritma KNN (K-Nearest Neighbors) adalah algoritma klasifikasi yang menentukan kelas sebuah objek berdasarkan data pelatihan yang jaraknya paling dekat dengan objek tersebut. Algoritma ini merupakan metode supervised learning dan bekerja dengan cara mengukur jarak terpendek antara data uji dan data latih untuk menentukan kelas data tersebut. Setelah data dikumpulkan dalam kelompok k yang spesifik, kelas mayoritas dalam kelompok tersebut digunakan sebagai kelas prediksi untuk data uji. Algoritma ini memiliki keunggulan dalam ketahanannya terhadap data yang mengandung noise dan efektif pada data latih berjumlah besar.[5]

3. HASIL DAN PEMBAHASAN

3.1 Dataset

Dataset yang digunakan dalam penelitian ini adalah kumpulan gambar daun terong yang bersumber dari platform Kaggle dengan jumlah dataset kurang lebih 3000 yang di bagi menjadi 80% data training dan 20% data testing, dengan judul dataset "Eggplant Disease Data". Dataset ini terdiri dari file utama berformat CSV dengan nama eggplant_disease_data.csv serta folder yang menyimpan gambar-gambar daun terong. Data ini dirancang untuk membantu penelitian dalam mendeteksi dan mengklasifikasikan penyakit pada daun terong menggunakan metode machine learning, termasuk algoritma K-Nearest Neighbors (KNN) [15], [16], [17].



Gambar 2. Dataset Eggplant

Gambar 2 adalah dataset gambar yang ada didalam dataset Eggplant Disease Data, terdapat 7 jenis dataset penyakit yang di gunakan. Pada Tabel 1 dataset kemudian dibagi menjadi Training Set (80%): Digunakan untuk melatih model KNN. Testing Set (20%): Digunakan untuk mengevaluasi performa model pada data yang belum pernah dilihat.

Tabel 1. Pembagian dataset

No	Kategori	Jumlah Gambar	Data Training (80%)	Data Testing (20%)
1	Healthy Leaf	526	421	105
2	Insect Pest Disease	512	416	96
3	Leaf Spot Disease	880	704	176
4	Mosaic Virus Disease	298	238	60
5	Small Leaf Disease	114	91	23
6	White Mold Disease	174	139	35
7	Wilt Disease	496	396	100
Total		3.000	2.405	595

3.2 Preprocessing

3.2.1 Cleaning Data

Tahap ini bertujuan untuk memastikan bahwa hanya gambar yang relevan dan berkualitas tinggi yang digunakan dalam pelatihan model.

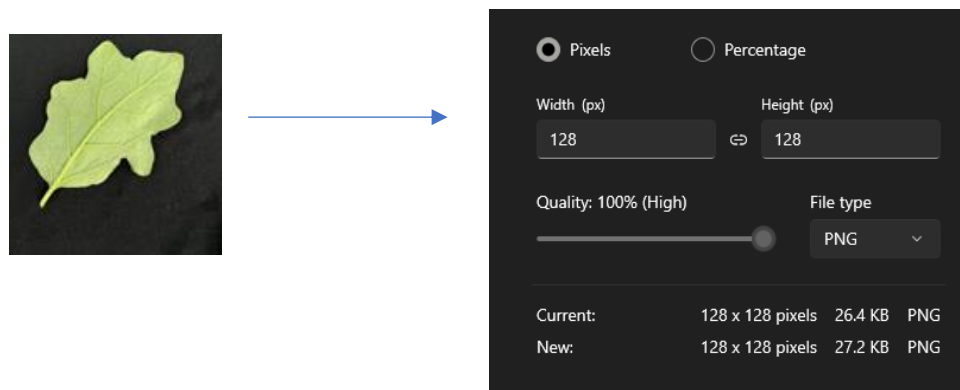


Gambar 3. Hasil Cleaning

Gambar 3 merupakan hasil cleaning data gambar yang sudah di hapus backgroundnya sehingga di dapatkan data bersih.

3.2.2 Normalisasi

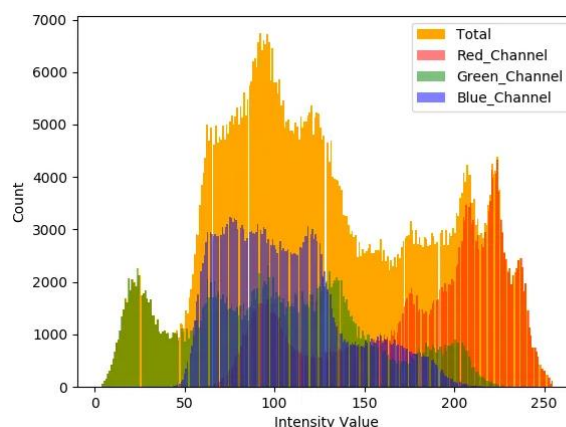
Pada tahapan ini Semua gambar diubah ukurannya menjadi 128×128 piksel (bukan 256×256 untuk efisiensi komputasi), seperti ditunjukkan pada gambar 4. Ukuran ini dipilih untuk menyeimbangkan antara kualitas informasi dan kecepatan proses pelatihan. Selain itu menormalisasikan nilai piksel



Gambar 4. Tahapan resizing dan normalisasi pixel

3.2.3 Ekstraksi Fitur

Fitur utama yang diekstrak dari gambar adalah histogram warna, baik dalam ruang warna RGB maupun HSV. Ekstraksi Histogram Warna: Histogram warna untuk setiap kanal (R, G, B) dihitung, menghasilkan vektor fitur sepanjang 768 elemen (256 nilai untuk setiap kanal).



Gambar 5. Hasil Histogram



Histogram pada Gambar 5 menunjukkan distribusi intensitas warna pada kanal merah, hijau, dan biru. Terlihat bahwa kanal merah (Red_Channel) dan kanal hijau (Green_Channel) memiliki intensitas tinggi di beberapa titik, menandakan dominasi warna pada gambar. Warna ini mengindikasikan area klorosis, nekrosis, atau bercak yang sering muncul sebagai gejala penyakit.

3.3 Implementasi Algoritma K-NN

Algoritma K-Nearest Neighbors (KNN) digunakan untuk klasifikasi. Detail parameter yaitu $n_neighbors=3$: KNN menggunakan 3 tetangga terdekat untuk melakukan klasifikasi. distance metric: Secara default, KNN menggunakan jarak Euclidean untuk menghitung kesamaan antar vektor fitur.

3.3.1 Pemilihan Parameter K

Pada algoritma K-NN, parameter k menentukan jumlah tetangga terdekat yang digunakan untuk menentukan kelas suatu data. Eksperimen akan dilakukan dengan berbagai nilai k mulai dari 1 hingga 15 untuk menemukan nilai yang optimal. Penggunaan k-fold cross-validation (dengan $k=5$) akan diterapkan untuk menghindari overfitting. Cross-validation ini membagi data pelatihan menjadi 5 bagian dan menggunakan 4 bagian untuk pelatihan dan 1 bagian untuk pengujian secara bergantian, sehingga model dapat diuji dengan berbagai subset data.

3.3.2 Pelatihan Model

Model K-NN akan diimplementasikan menggunakan pustaka scikit-learn yang sudah menyediakan berbagai metode untuk pelatihan dan pengujian model K-NN. Beberapa parameter yang perlu diperhatikan dalam pelatihan adalah:

- Metode Pengukuran Jarak: Dua metrik jarak yang umum digunakan dalam K-NN adalah Euclidean Distance dan Manhattan Distance. Euclidean distance mengukur jarak langsung antara dua titik, sedangkan Manhattan distance mengukur jarak berdasarkan jalur horizontal dan vertikal. Pilihan metrik jarak ini akan dievaluasi selama eksperimen.
- Bobot Jarak: Pada beberapa kasus, bobot jarak diterapkan untuk memberi pengaruh lebih besar pada tetangga yang lebih dekat dibandingkan yang lebih jauh, sehingga model dapat memberikan prediksi yang lebih akurat.

3.3.3 Penentuan Bobot Jarak Euclidean Distance

Dalam eksperimen ini, akan diuji Euclidean Distance sebagai penentu bobot jarak Euclidean Distance digunakan dalam proses klasifikasi atau identifikasi dengan menghitung jarak antara vektor pelatihan dan vektor uji pada basis data yang ada.

3.4 Evaluasi Model

Evaluasi hasil klasifikasi menggunakan K-Nearest Neighbors (KNN) melibatkan beberapa aspek penting untuk menilai akurasi dan performa model. Salah satunya adalah confusion matrix, yang memberikan gambaran jelas mengenai prediksi yang benar dan salah untuk setiap kelas, memungkinkan perhitungan metrik seperti akurasi, presisi, recall, dan F1-score [14], [18]. Selain itu, class distribution membantu mengidentifikasi apakah ada ketidakseimbangan kelas dalam dataset, yang dapat mempengaruhi prediksi KNN, karena model ini bergantung pada distribusi data dalam menentukan tetangga terdekat. Color histogram, di sisi lain, dapat digunakan dalam analisis citra, dimana distribusi warna dalam gambar menjadi fitur penting dalam mengklasifikasikan objek atau pola. Ketiga elemen ini memberikan wawasan yang komprehensif untuk mengevaluasi dan mengoptimalkan model KNN dalam berbagai aplikasi klasifikasi.

a. Confusion matrix

Matriks ini adalah alat yang digunakan untuk mengevaluasi kualitas dari model klasifikasi, termasuk K-Nearest Neighbors (KNN). Ia membandingkan hasil prediksi yang dihasilkan oleh model dengan label asli atau yang sebenarnya. Dalam confusion matrix, terdapat empat elemen utama:

True Positives (TP): Prediksi benar untuk kelas yang benar.

False Positives (FP): Prediksi salah, dimana model memprediksi kelas yang salah sebagai benar.

True Negatives (TN): Prediksi benar untuk kelas yang salah.

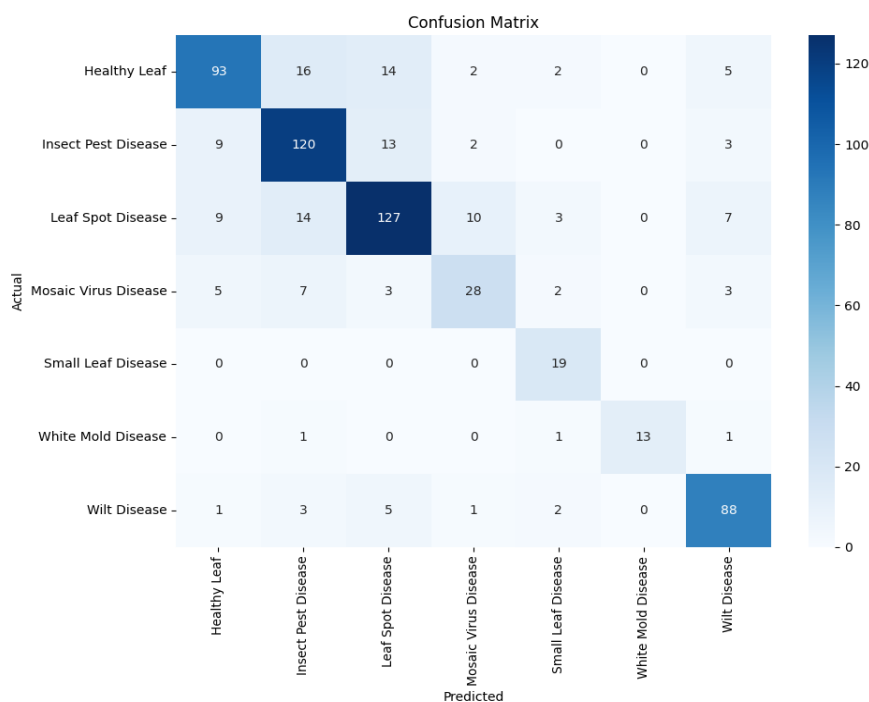
False Negatives (FN): Prediksi salah, dimana model tidak mengenali kelas yang benar.

Melalui matriks ini, kita dapat menghitung berbagai metrik evaluasi seperti akurasi, presisi, recall, dan F1-score, yang membantu menilai sejauh mana model KNN dapat membuat prediksi yang tepat dan seberapa baik performa model secara keseluruhan [19], [20], [21]. Gambar 2 merupakan hasil perhitungan dari confusion matrix, sedangkan Tabel 2 merupakan nilai dari hasil confusion matrix untuk setiap label yang dibentuk.

Tabel 2. Confusion Matrix

	Precision	Recall	F1-Score
Healthy Leaf	0.92	0.86	0.89
Insect Pest Disease	0.85	0.87	0.86
Leaf Spot Disease	0.86	0.89	0.88
Mosaic Virus Disease	0.86	0.79	0.83

Small Leaf Disease	0.67	0.84	0.74
White Mold Disease	1.00	0.88	0.93
Wilt Disease	0.90	0.89	0.89
Accuracy			0.87
Macro Avg	0.87	0.86	0.86
Weight Avg	0.87	0.87	0.87

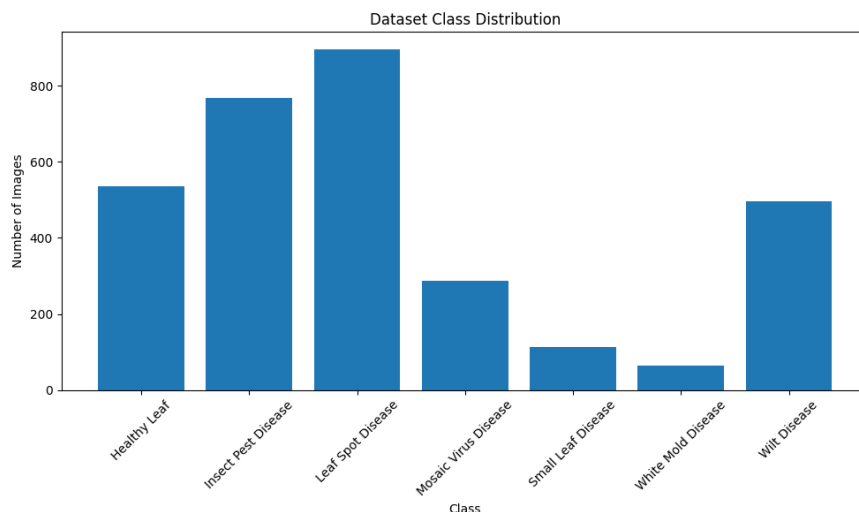


Gambar 6. Confusion matrix

Gambar 6 merupakan hasil dari evaluasi kualitas dari model klasifikasi, termasuk K-Nearest Neighbors (KNN). Kita juga dapat melihat jenis penyakit tanamannya.

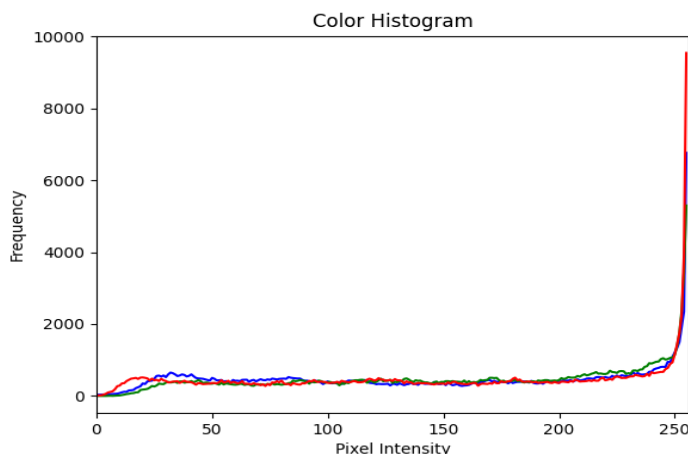
b. Class distribution

Class distribution menunjukkan bagaimana data dalam dataset dibagi ke dalam berbagai kelas. Dalam klasifikasi dengan KNN, distribusi kelas sangat penting karena KNN bekerja dengan mencari tetangga terdekat dari data yang diuji. Jika terdapat ketidakseimbangan kelas (misalnya, lebih banyak data untuk kelas tertentu), model KNN bisa saja cenderung memprediksi kelas yang lebih banyak, karena prediksi model dipengaruhi oleh tetangga yang lebih dominan. Distribusi kelas yang tidak seimbang ini dapat menyebabkan hasil yang bias atau kurang akurat, sehingga penting untuk memeriksa distribusi kelas terlebih dahulu [22], [23].



Gambar 7. class retribution

Gambar 7 adalah hasil *class Retribution* dari pembagian *class* didalam dataset. Color Histogram (Histogram Warna). Histogram warna sering digunakan dalam pengolahan citra untuk menggambarkan distribusi intensitas warna dalam gambar. Dalam konteks KNN untuk pengklasifikasian gambar, histogram warna dapat digunakan sebagai fitur untuk mengklasifikasikan gambar. KNN akan membandingkan histogram warna antara gambar yang diuji dengan gambar lainnya dalam dataset, dan menentukan kelas berdasarkan kesamaan pola warna tersebut. Ini sangat berguna dalam tugas-tugas seperti pengenalan objek atau pengklasifikasian gambar berdasarkan warna dominannya [8], [24].

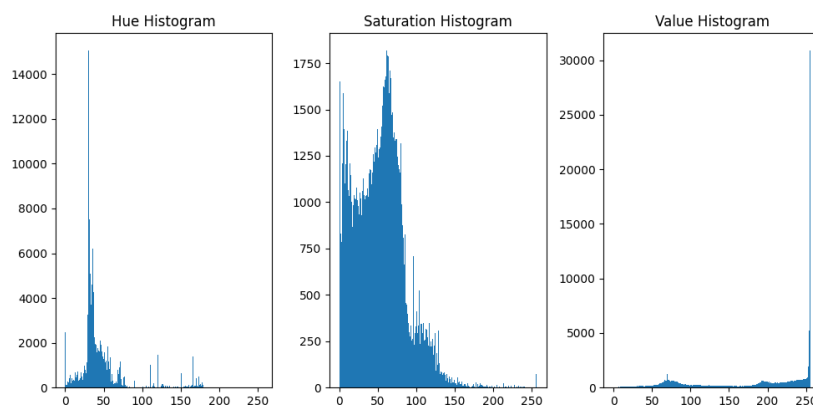


Gambar 8. Color histogram

Gambar 8 merupakan hasil perhitungan warna dan intensitas warna dari setiap dataset. Fitur ini untuk mengklasifikasikan gambar.

c. HSV histogram

Representasi distribusi warna dalam suatu gambar berdasarkan ruang warna Hue, Saturation, dan Value (HSV). Histogram ini digunakan untuk menganalisis komponen warna (Hue), intensitas warna (Saturation), dan kecerahan (Value) dari gambar. Dalam konteks K-Nearest Neighbors (KNN), HSV histogram dapat digunakan sebagai fitur utama dalam proses klasifikasi citra. Gambar-gambar dalam dataset akan dikonversi ke ruang warna HSV, dan histogramnya dihitung untuk masing-masing komponen. KNN kemudian membandingkan histogram gambar yang diuji dengan histogram gambar lain dalam dataset berdasarkan jarak (misalnya, jarak Euclidean) untuk menentukan kelasnya. Pendekatan ini sering digunakan dalam klasifikasi berbasis visual, seperti pengenalan objek atau deteksi penyakit pada daun berdasarkan warna dan pola [25].



Gambar 9. hsv histogram

Gambar 9 menunjukkan histogram distribusi warna dari sebuah citra daun dalam ruang warna Hue, Saturation, dan Value (HSV). Histogram ini digunakan untuk merepresentasikan karakteristik visual dari gambar, yang kemudian digunakan sebagai fitur dalam proses klasifikasi berbasis K-Nearest Neighbors (KNN). Penjabaran masing-masing komponen adalah sebagai berikut:

1. Hue Histogram (kiri)

Histogram ini memperlihatkan distribusi *jenis warna* pada gambar. Terlihat bahwa mayoritas piksel terkonsentrasi pada rentang nilai Hue rendah (sekitar 0–50), yang biasanya merepresentasikan warna-warna seperti merah, jingga, dan kuning. Ini mengindikasikan bahwa citra tersebut memiliki dominasi warna tertentu yang bisa menjadi penanda kondisi daun, misalnya warna kekuningan akibat infeksi atau kekeringan.

2. Saturation Histogram (tengah)
Histogram ini menunjukkan *kejenuhan warna*. Sebagian besar piksel memiliki nilai Saturation antara 30–100, yang berarti warna dalam gambar cukup jenuh dan tidak terlalu abu-abu. Ini menandakan bahwa gambar masih mempertahankan kualitas warna yang baik untuk keperluan analisis visual.
3. Value Histogram (kanan)
Histogram Value mencerminkan tingkat *kecerahan*. Puncak tertinggi berada di sekitar nilai 250, menunjukkan banyak piksel yang sangat terang. Hal ini dapat disebabkan oleh pencahayaan kuat saat pengambilan gambar atau permukaan daun yang memantulkan cahaya. Distribusi nilai Value ini juga penting untuk membedakan antara daun sehat dan terinfeksi, karena perubahan kecerahan bisa menjadi indikator visual dari penyakit.

4. KESIMPULAN

Berdasarkan latar belakang yang telah disampaikan, dapat disimpulkan bahwa klasifikasi penyakit daun pada tanaman terong merupakan langkah penting dalam meningkatkan produktivitas pertanian. Teknologi kecerdasan buatan, khususnya metode K-Nearest Neighbors (KNN), telah banyak digunakan dalam analisis citra daun untuk mendeteksi penyakit dengan lebih cepat dan akurat. Meskipun metode KNN menunjukkan hasil yang cukup baik dengan akurasi sekitar 87%, masih terdapat beberapa tantangan yang perlu diatasi, seperti ketidakseimbangan dataset dan kesalahan klasifikasi akibat kemiripan pola warna antar penyakit. Oleh karena itu, penelitian ini bertujuan untuk mengoptimalkan model klasifikasi dengan meningkatkan teknik preprocessing data dan eksplorasi fitur citra yang lebih kompleks. Dengan adanya penelitian ini, diharapkan solusi yang lebih efektif dapat diterapkan untuk membantu petani dalam mengelola dan mencegah penyebaran penyakit pada tanaman terong secara lebih efisien.

REFERENCES

- [1] A. Sanusi Mashuri and A. Sunyoto, “Klasifikasi Penyakit Pada Daun Cabai Menggunakan Arsitektur VGG16,” *Journal homepage: Journal of Electrical Engineering and Computer (JEECOM)*, vol. 6, no. 2, 2024, doi: 10.33650/jeeecom.v4i2.
- [2] Q. N. Azizah and) Andreyestha, “Klasifikasi Penyakit Daun Apel Menggunakan Convolutional Neural Network,” *Jurnal Ilmu Komputer dan Informatika*, vol. 3, no. 1, 2022
- [3] Q. N. Azizah, “Klasifikasi Penyakit Daun Jagung Menggunakan Metode Convolutional Neural Network AlexNet,” *sudo Jurnal Teknik Informatika*, vol. 2, no. 1, pp. 28–33, Feb. 2023, doi: 10.56211/sudo.v2i1.227.
- [4] S. Zahra, S. O. Ayu, and D. Kiswanto, “Klasifikasi Penyakit Pada Daun Jambu Berbasis Pengolahan Citra Dengan Analisis Tekstur Menggunakan Algoritma K-Nearest Neighbors (KNN),” *Jurnal Komputer Multidisipliner*, Vol 7, No 12, 2024.
- [5] E. Safitri *et al.*, “Klasifikasi Penyakit Daun Anggur Berbasis Citra Menggunakan Metode K-Nearest Neighbors (KNN),” *Jurnal Jati*, 2024.
- [6] H. Mubarak, S. Murni, and M. M. Santoni, “Penerapan Algoritma K-Nearest Neighbor untuk Klasifikasi Tingkat Kematangan Buah Tomat Berdasarkan Fitur Warna” *Prosiding Senamika*, 2021.
- [7] R. Amalia, I. H. Ikasari, and P. Rosyani, “Deteksi Objek dengan Model Warna Ycber dan Similarity Distance Object Detection with YCbCr Color model and similarity distance,” vol. 09, no. 2, pp. 98–100, 2021, doi: 10.26418/justin.v9i2.44230.
- [8] F. Mahrus Fathoni, C. Aji Putra, A. Lina Nurlaili, “Klasifikasi Penyakit Daun Anggur Menggunakan Metode K-Nearest Neighbor Berdasarkan Gray Level Co-Occurrence Matrix,” vol. 3, no. 1, 2024, [Online]. Available: <https://ojs.unsiq.ac.id/index.php/biner>
- [9] Teti Desyani, M. Mahromi, F. A. Ramadhan, M. Alfiansyah, M. I. Maulana, and P. Rosyani, “Classification of Plant Leaf Diseases Using Convolutional Neural Networks,” *International Journal of Integrative Sciences*, vol. 4, no. 1, pp. 195–206, Feb. 2025, doi: 10.55927/ijis.v4i1.13478.
- [10] I. H. Ikasari, R. Amalia, and P. Rosyani, “Segmentasi Citra Bunga Menggunakan Blob Analisis,” *Building of Informatics, Technology and Science (BITS)*, vol. 3, no. 3, pp. 228–234, 2021, doi: 10.47065/bits.v3i3.1050.
- [11] R. Amalia, A. F. Zaidan, S. Ramadhan, F. Septian, A. M. Aqsha, and P. Rosyani, “Classification of Autoimmune Diseases Using the K-Nearest Neighbors Algorithm,” *Formosa Journal of Science and Technology*, vol. 4, no. 1, pp. 337–348, Jan. 2025, doi: 10.55927/fjst.v4i1.13443.
- [12] A. Venkataramana, K. Suresh Kumar, N. Suganthi, and R. Rajeswari, “Prediction of Brinjal Plant Disease Using Support Vector Machine and Convolutional Neural Network Algorithm Based on Deep Learning,” *Journal of Mobile Multimedia*, vol. 18, no. 3, pp. 771–788, 2022, doi: 10.13052/jmm1550-4646.18315.
- [13] R. Indah Juwita Harahap, S. Khairani, “Implementasi Metode K-Nearest Neighbor Untuk Klasifikasi Penyakit Tanaman Mentimun Pada Citra Daun Implementation Of The K-Nearest Neighbor Method For Classification Of Cucumber Diseases On Leaf Images,” *Jurnal Ilmu Komputer dan Sistem Informasi*, Vol 3, No 2, 2024, doi.org/10.70340/jirsi.v3i2.123
- [14] H. Salsabila, E. Rachmawati, and F. Sthevanie, “Klasifikasi Gender Berdasarkan Citra Wajah Menggunakan Metode Local Binary Pattern dan K-Nearest Neighbor,” *e-Proceeding of Engineering*, vol. 8, no. 2, pp. 3137–3146, 2021.
- [15] P. Rosyani, A. M. Lutfi, E. Purwadi, Kamaluddin, Y. A. Hanaan, and I. H. Ikasari, “Application of Random Forest for Rice Plant Disease Classification,” *International Journal of Integrative Sciences*, vol. 4, no. 1, pp. 141–150, Feb. 2025, doi: 10.55927/ijis.v4i1.13477.
- [16] S. Apandi *et al.*, “Classification of Lung Diseases Using the Decision Tree Method,” *Formosa Journal of Science and Technology*, vol. 4, no. 1, pp. 393–412, Jan. 2025, doi: 10.55927/fjst.v4i1.13442.
- [17] Ines Heidiani Ikasari, R. Y. Saputra, S. Prasdio, M. F. Kurniagis, P. Rosyani, and Z. Janariandana, “Classification of Pneumonia Medical Images with Convolutional Neural Networks,” *International Journal of Integrative Sciences*, vol. 4, no. 1, pp. 127–134, Feb. 2025, doi: 10.55927/ijis.v4i1.13511.



- [18] K. Sulastri, “Klasifikasi Naïve Bayes pada Analisis Sentimen atas Penolakan Dibukanya Larangan Ekspor Benih Lobster,” *KERNEL: Jurnal Riset Inovasi Bidang Informatika dan Pendidikan Informatika*, vol. 1, no. 2, pp. 68–75, 2020, doi: 10.31284/j.kernel.2020.v1i2.1501.
- [19] A. C. Milano, “Klasifikasi Penyakit Daun Padi Menggunakan Model Deep Learning Efficientnet-B6,” *Jurnal Informatika dan Teknik Elektro Terapan*, vol. 12, no. 1, Jan. 2024, doi: 10.23960/jitet.v12i1.3855.
- [20] N. Istiqomah and M. Murinto, “Klasifikasi Penyakit Tanaman Padi Berbasis Citra Daun Menggunakan Convolutional Neural Network (CNN),” *JSTIE (Jurnal Sarjana Teknik Informatika) (E-Journal)*, vol. 12, no. 1, p. 18, Feb. 2024, doi: 10.12928/jstie.v12i1.27314.
- [21] E. Anggiratih, S. Siswanti, S. K. Octaviani, and A. Sari, “Klasifikasi Penyakit Tanaman Padi Menggunakan Model Deep Learning Efficientnet B3 dengan Transfer Learning,” *Jurnal Ilmiah SINUS*, vol. 19, no. 1, p. 75, Jan. 2021, doi: 10.30646/sinus.v19i1.526.
- [22] S. Dwi, Y. Kusuma, H. Al Islami, and D. P. Rosyani, “Penerapan Naive Bayes Untuk Klasifikasi Penyakit Endokrin Pada Pasien Lansia,” *KERNEL: Jurnal Riset Inovasi Bidang Informatika dan Pendidikan Informatika*, vol. 5, no. 2, 2024, doi: 10.31284/j.kernel.2024.
- [23] Santi Rahayu, P. Rosyani, R. Y. Saputra, R. A. Umar, S. Prasdio, and W. A. Syach, “Application of Expert System in Rice Seedling Selection Based on Smart Data With Methods: Knowledge-Based System and Decision Tree,” *International Journal of Integrative Sciences*, vol. 4, no. 1, pp. 217–224, Feb. 2025, doi: 10.55927/ijis.v4i1.13510.
- [24] Fuad Mahrus Fathoni, “Klasifikasi Penyakit Daun Tomat Menggunakan Algoritma K-NN Berdasarkan Ekstraksi Fitur GLCM dan LBP,” *Jurnal Teknik Informatika dan Teknologi Informasi*, vol. 4, no. 1, pp. 39–50, Jan. 2024, doi: 10.55606/jutiti.v4i1.3417.
- [25] A. Putra Pranjaya, F. Rizki, R. Kurniawan, and N. K. Daulay, “Klasifikasi Penyakit Pada Daun Tanaman Padi Berbasis YoloV5 (You Only Look Once),” *KLIK: Kajian Ilmiah Informatika dan Komputer*, vol. 4, no. 6, pp. 3127–3136, 2024, doi: 10.30865/klik.v4i6.1916.