

KLASIFIKASI DAUN HERBAL BERTULANG MENYIRIP MENGGUNAKAN K-NEAREST NEIGHBOR

Lintang Sari Putri Wardhani¹⁾, Anggraini Puspita Sari^{*2)}, Hendra Maulana³⁾

1. Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur, Indonesia
2. Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur, Indonesia
3. Bisnis Digital, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa Timur, Indonesia

Article Info

Kata Kunci: Klasifikasi Daun Herbal; KNN; Pengolahan Citra Digital; Pra-pemrosesan Citra

Keywords: Digital Image Recognition; Image Preprocessing; Herbal Leaf Classification; KNN

Article history:

Received 12 May 2025

Revised 16 June 2025

Accepted 21 July 2025

Available online 1 March 2026

DOI :

<https://doi.org/10.29100/jupi.v11i1.7840>

* Corresponding author.

Corresponding Author

E-mail address:

anggraini.puspita.if@upnjatim.ac.id

ABSTRAK

Indonesia memiliki kekayaan keanekaragaman hayati, termasuk berbagai jenis tanaman herbal yang dimanfaatkan untuk pengobatan tradisional. Salah satu bagian tanaman yang sering dimanfaatkan adalah daun, karena kandungan senyawa bioaktifnya dan kemudahan dalam pengolahan. Namun, proses identifikasi daun herbal sering mengalami kendala karena kemiripan bentuk antar jenis daun. Untuk mengatasi tantangan tersebut, penelitian ini mengembangkan sistem klasifikasi daun herbal bertulang menyirip menggunakan algoritma KNN yang dikombinasikan dengan ekstraksi fitur berbasis deep learning menggunakan arsitektur ResNet-50. Penelitian dilakukan dengan mengumpulkan 1200 citra daun dari dua jenis tanaman herbal. Citra mengalami tahapan pra-proses kemudian diklasifikasikan menggunakan KNN dengan berbagai parameter. Evaluasi dilakukan menggunakan *confusion matrix*, akurasi, presisi, *recall*, dan *F1-score*. Hasil menunjukkan bahwa skenario terbaik diperoleh pada kombinasi metrik *Manhattan* dengan nilai $k = 5$, dan proporsi data latih-uji 75:25, menghasilkan akurasi sebesar 98,33%. Temuan ini menunjukkan bahwa algoritma KNN mampu mengklasifikasikan daun herbal bertulang menyirip dengan akurasi yang tinggi tanpa perlu ekstraksi fitur kompleks. Hasil ini membuktikan bahwa pendekatan sederhana namun efektif seperti KNN tetap relevan dan dapat diandalkan dalam membangun sistem identifikasi tanaman berbasis citra digital, khususnya untuk mendukung pelestarian dan pemanfaatan tanaman herbal dalam bidang kesehatan, pertanian, dan edukasi.

ABSTRACT

Indonesia has a rich biodiversity, including various types of herbal plants that are used for traditional medicine. One part of the plant that is often utilized is the leaf, due to its bioactive compounds and ease of processing. However, the identification process of herbal leaves is often challenging due to the similarity in shape between types of leaves. To overcome these challenges, this research develops a classification system for pinnately bony herbal leaves using the KNN algorithm combined with deep learning-based feature extraction using the ResNet-50 architecture. The research involved collecting 1,200 leaf images from two types of herbal plants. The images underwent preprocessing stages and were then classified using KNN with various parameter configurations. Evaluation was conducted using a confusion matrix, accuracy, precision, recall, and F1-score. The best scenario was obtained using the *Manhattan* distance metric with $k = 5$ and a 75:25 training-testing data split, achieving an accuracy of 98.33%. These findings demonstrate that the KNN algorithm can classify pinnate-veined herbal leaves with high accuracy without the need for complex feature extraction. This result proves that a simple yet effective approach like KNN remains relevant and reliable for developing image-based plant identification systems, particularly to support the preservation and utilization of herbal plants in the fields of healthcare, agriculture, and education.

I. PENDAHULUAN

KEANEKARAGAMAN hayati adalah seluruh kehidupan yang ada di bumi, mencakup semua makhluk hidup mulai dari gen, spesies, dan ekosistem [1]. Indonesia, sebagai salah satu negara megabiodiversitas di dunia, memiliki tingkat keanekaragaman hayati yang tinggi. Hal ini disebabkan oleh pengaruh letak geografis Indonesia yang membentuk wilayah bioregion, sehingga menjadi pemisah antara biogeografi flora dan fauna dengan Australia [2]. Beragam tanaman pertanian seperti padi, jagung, dan kedelai berasal dari kekayaan keanekaragaman hayati. Selain tanaman pertanian yang berperan penting dalam ketahanan pangan, Indonesia juga memiliki kekayaan hayati berupa tanaman herbal yang digunakan dalam pembuatan obat-obatan tradisional yang telah menjadi bagian tak terpisahkan dari budaya serta kehidupan masyarakat sejak zaman dahulu [3]. Pemanfaatan tanaman herbal telah diwariskan secara turun-temurun dan terus berkembang seiring waktu, mencerminkan pengetahuan lokal yang mendalam terkait khasiat alami tumbuhan untuk pengobatan berbagai penyakit [4]. Ada ribuan jenis tanaman herbal tumbuh subur di berbagai wilayah Indonesia yang hingga saat ini masih dimanfaatkan secara luas, baik secara tradisional dalam bentuk jamu dan ramuan, maupun sebagai bahan baku industri farmasi modern [5].

Organisasi Kesehatan Dunia (WHO) mencatat bahwa lebih dari 80% penduduk di negara – negara berkembang masih mengandalkan pengobatan tradisional sebagai solusi utama dalam memenuhi kebutuhan kesehatan primer mereka, yang sebagian besar berbasis pada pemanfaatan tanaman herbal [6]. Hal ini menunjukkan bahwa meskipun kemajuan teknologi di bidang kesehatan terus berkembang pesat, pengobatan berbasis bahan alami tetap menjadi pilihan utama bagi masyarakat, terutama di daerah yang akses terhadap layanan medis modern masih terbatas [7]. Kepercayaan masyarakat terhadap khasiat tanaman herbal, ditambah dengan ketersediaannya yang melimpah di lingkungan sekitar, menjadikan pengobatan tradisional sebagai alternatif utama yang mudah dijangkau oleh berbagai lapisan masyarakat. Selain itu, kecenderungan masyarakat modern yang mulai kembali ke pola hidup sehat dan alami semakin memperkuat posisi tanaman herbal sebagai pilihan pengobatan ditengah gempuran obat-obatan sintesis [8]. Pengobatan tradisional yang berbasis tanaman herbal tidak hanya bertumpu pada satu bagian tanaman saja, melainkan mencakup beragam bagian tumbuhan yang masing masing memiliki kandungan senyawa aktif dengan manfaat tertentu bagi kesehatan [9]. Beberapa bagian tanaman herbal yang umum dimanfaatkan meliputi akar, rimpang, umbi, bunga, buah, biji, batang, hingga daun. Setiap bagian tersebut memiliki peranan tersendiri dalam proses pengobatan, tergantung pada jenis tanaman dan kandungan fitokimia yang dikandungnya [10]. Diantara semua bagian tanaman tersebut, daun merupakan salah satu bagian yang juga sering dimanfaatkan karena mengandung senyawa bioaktif serta lebih mudah diperoleh tanpa merusak tanaman keseluruhan [11]. Daun dapat diolah baik dengan cara diseduh, direbus, dikeringkan, atau dijadikan ekstrak, sehingga menjadikannya bagian yang praktis untuk dikembangkan dalam berbagai bentuk pengobatan tradisional maupun produk kesehatan modern [12].

Pengenalan dan identifikasi daun dari tanaman herbal menjadi aspek penting dalam mendukung pemanfaatan tanaman herbal secara optimal, baik untuk konsumsi masyarakat umum maupun sebagai bahan penelitian ilmiah dalam bidang farmasi dan kesehatan [13]. Namun, daun memiliki variasi bentuk yang seringkali memiliki kemiripan satu sama lain, sehingga proses identifikasinya menjadi rawan kesalahan. Kemiripan bentuk ini dapat menjadi tantangan besar dalam proses identifikasi manual, apalagi jika dilakukan oleh individu yang tidak memiliki latar belakang botani atau pengalaman dalam mengenali spesies tanaman [14]. Kesalahan dalam identifikasi spesies daun tidak hanya berdampak pada ketidaktepatan dalam pemanfaatan, tetapi juga dapat berisiko tinggi terhadap kesehatan pengguna [15]. Oleh sebab itu, dibutuhkan metode identifikasi yang lebih akurat dan efisien untuk membedakan daun dari berbagai tanaman herbal, agar proses pengolahan obat tradisional dapat dilakukan dengan lebih aman dan tepat sasaran.

Pada era digital yang terus berkembang pesat, menjadikan teknologi machine learning salah satu Solusi dalam menyelesaikan berbagai permasalahan identifikasi objek secara otomatis, termasuk dalam bidang botani untuk mengklasifikasikan jenis jenis daun tanaman [16]. Salah satu pendekatan yang umum digunakan adalah teknik *image classification* yang memungkinkan sistem komputer untuk mengenali dan membedakan objek visual berdasarkan pola, tekstur, bentuk, serta fitur – fitur khas lainnya dalam citra digital [17]. Teknologi ini sangat bermanfaat dalam proses identifikasi daun tanaman herbal yang memiliki kemiripan bentuk satu sama lain, karena mampu memproses data dalam jumlah besar dengan kecepatan dan tingkat akurasi yang tinggi [18]. Dengan memanfaatkan machine learning, identifikasi daun yang sebelumnya memerlukan keahlian manusia, kini dapat dilakukan secara otomatis, sehingga dapat membantu penelitian, pelestarian tanaman herbal, dan produksi obat herbal secara lebih efisien dan minim risiko kesalahan [19].

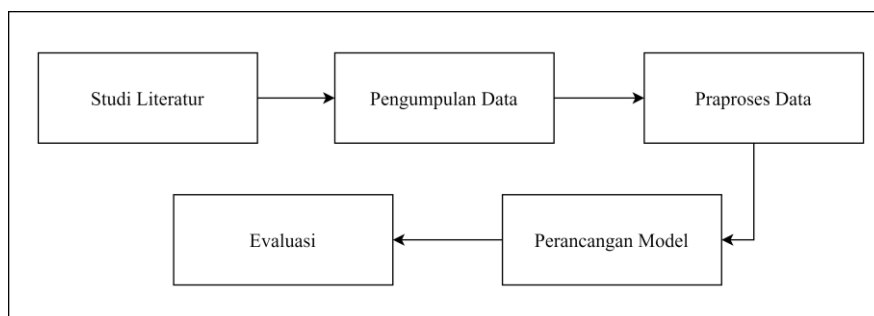
Penelitian terdahulu yang relevan dalam klasifikasi daun herbal yang dilakukan oleh Meiriyama, dkk [20] menggunakan metode K-Nearest Neighbor (KNN) dengan nilai $K = 3$ dan ekstraksi fitur Histogram of Oriented

Gradient (HOG) serta Local Binary Patterns (LBP) pada 15 jenis tanaman herbal. Penelitian tersebut melaporkan akurasi sebesar 92,67% untuk fitur HOG dengan KNN, 88,67% untuk fitur LBP dengan KNN, dan 92,67% untuk kombinasi fitur HOG dan LBP dengan KNN. Hasil ini menunjukkan bahwa kombinasi fitur HOG dan LBP tidak memberikan peningkatan signifikan terhadap akurasi klasifikasi dibandingkan penggunaan fitur tunggal HOG. Pendekatan ini memiliki keterbatasan karena bergantung pada proses ekstraksi fitur manual yang memerlukan keahlian khusus dan tidak fleksibel terhadap variasi citra yang kompleks. Di sisi lain, pendekatan berbasis deep learning seperti Convolutional Neural Network (CNN) maupun model pretrained seperti ResNet telah banyak digunakan dalam klasifikasi citra daun, seperti pada penelitian [21] yang memanfaatkan ResNet untuk mendeteksi penyakit daun gandum dan menghasilkan akurasi tinggi hingga 98%. Meskipun efektif, pendekatan ini memerlukan sumber daya komputasi yang besar dan dataset yang cukup besar agar model dapat belajar secara optimal.

Untuk mendukung pemanfaatan teknologi dalam klasifikasi daun tanaman herbal, sejumlah penelitian sebelumnya telah mengeksplorasi berbagai pendekatan klasifikasi citra digital [22] termasuk metode klasifikasi berbasis *deep learning* seperti CNN dan ResNet, pendekatan ini sering kali memerlukan sumber daya komputasi besar dan jumlah data yang cukup besar untuk pelatihan yang optimal. Berbeda dengan CNN yang memerlukan pelatihan mendalam untuk performa optimal, metode KNN langsung pada citra memungkinkan proses yang lebih ringan dan cepat tanpa risiko overfitting pada dataset kecil. Oleh karena itu, penelitian ini mengusulkan penggunaan metode KNN langsung pada citra daun herbal dengan preprocessing dasar. Metode ini dipilih untuk menyederhanakan proses klasifikasi sambil tetap mempertahankan akurasi. Pendekatan ini juga dinilai lebih praktis dan efisien untuk diterapkan dalam sistem klasifikasi ringan, terutama dalam skenario dengan keterbatasan data. Dengan menggunakan KNN secara langsung, penelitian ini berkontribusi dalam mengevaluasi sejauh mana metode klasifikasi sederhana dapat digunakan untuk membedakan jenis daun herbal tanpa ketergantungan pada proses ekstraksi fitur kompleks atau pelatihan model yang mendalam.

II. METODE PENELITIAN

Pada penelitian ini, terdapat alur kerja yang terstruktur guna menjamin bahwa setiap tahapan dilakukan secara sistematis. Alur ini berperan sebagai acuan dalam mempermudah pelaksanaan proses penelitian. Dengan perencanaan tahapan yang baik, diharapkan hasil akhir dari penelitian dapat sesuai dengan tujuan diawal. Urutan tahapan pada penelitian ini ditampilkan pada Gambar (1).



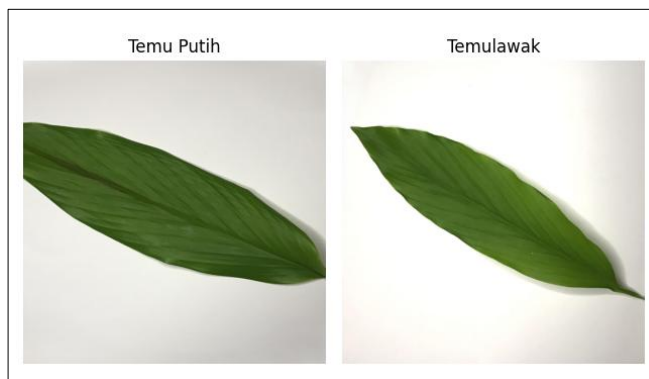
Gambar 1. Alur Penelitian

A. Studi Literatur

Pada tahap ini dilakukan literatur terhadap teori-teori yang relevan, hasil-hasil penelitian sebelumnya yang berkaitan dengan algoritma KNN. Studi literatur ini memberikan pemahaman mendalam mengenai prinsip kerja kedua pendekatan tersebut, serta mengidentifikasi celah penelitian yang belum banyak dieksplorasi. Melalui kajian pustaka ini, diperoleh dasar teori yang relevan untuk merumuskan permasalahan penelitian, membangun kerangka teori, dan merancang metodologi yang sesuai guna mencapai tujuan penelitian secara lebih terarah.

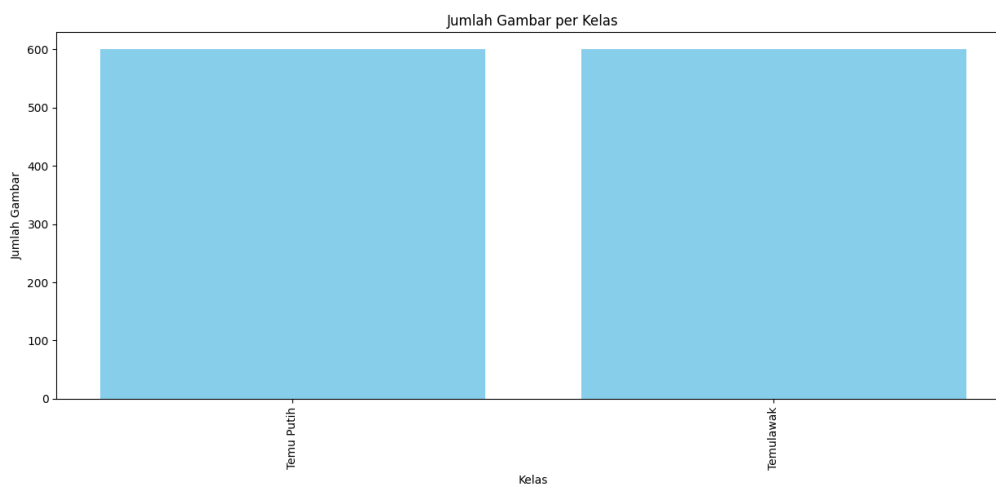
B. Pengumpulan Data

Dalam tahap pengumpulan data pada penelitian ini, data diperoleh secara langsung dari kondisi lapangan yang sebenarnya. Gambar dikumpulkan menggunakan kamera ponsel Iphone 7 dengan format JPG dan resolusi 3024x3024 piksel. Data yang diperoleh berupa citra dari dua jenis daun tanaman herbal yang memiliki karakteristik tulang daun menyirip yaitu daun temu putih, dan daun temulawak.



Gambar 2. Citra Daun

Gambar (2) merupakan citra dataset daun herbal dengan ciri daun menyirip yang digunakan. Masing-masing jenis daun diambil sebanyak 100 gambar, sehingga total citra yang dikumpulkan mencapai 200 gambar. Yang kemudian di augmentasi sehingga total data menjadi 1200 gambar, Seluruh data yang dikumpulkan adalah data primer, yang diambil langsung dari sumber asli tanpa adanya perantara.

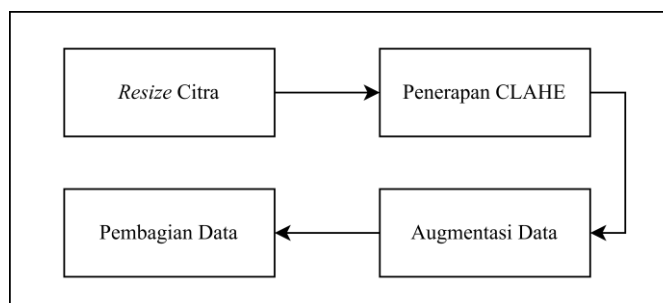


Gambar 3. Jumlah Citra

Pada Gambar (3) data yang diperoleh masih memerlukan tahap praproses untuk penyesuaian dengan kebutuhan, seperti perubahan ukuran dan augmentasi data, dan sebagainya.

C. Praproses Data

Tahapan selanjutnya dalam penelitian ini adalah pengolahan data yang mencakup proses praproses, pembagian data, serta ekstraksi fitur. Tahap praproses citra melibatkan sejumlah langkah untuk mempersiapkan data agar siap digunakan dalam pemodelan. Gambar (4) merupakan tahapan praproses data pada penelitian ini.



Gambar 4. Alur Praproses

1. Resize

Langkah pertama dalam tahap praproses citra adalah mengubah ukuran gambar yang telah dikumpulkan menjadi 224x224 piksel. Penyeragaman ukuran ini dilakukan untuk memastikan bahwa seluruh gambar memiliki dimensi yang konsisten dan kompatibel dengan format input yang dibutuhkan oleh model klasifikasi yang digunakan

dalam penelitian ini. Keceragaman dimensi sangat penting untuk menjaga keteraturan dalam proses pelatihan dan pengujian model yang ditunjukkan pada Gambar (5).



Gambar 5. *Resize* Citra

2. Penerapan CLAHE

Setelah ukuran citra seragam, tahap berikutnya menerapkan teknik Contrast limited adaptive histogram equalization (CLAHE) yang ditunjukkan pada Gambar (6). Metode ini diterapkan untuk meningkatkan kontras pada gambar agar memungkinkan detail-detail penting dalam citra menjadi lebih terlihat dan mudah dikenali oleh model.



Gambar 6. Penerapan CLAHE

3. Augmentasi Data

Augmentasi data merupakan teknik dalam pemrosesan citra yang digunakan untuk memperbanyak jumlah data latih dengan memodifikasi gambar asli secara sistematis tanpa mengubah label atau maknanya. Tujuannya adalah untuk meningkatkan performa model pembelajaran mesin dengan memberikan variasi yang lebih kaya dari data yang tersedia, sehingga model menjadi lebih belajar terhadap perbedaan orientasi atau posisi objek.



Gambar 7. Augmentasi Citra

Pada Gambar (7), augmentasi dilakukan melalui berbagai transformasi sederhana seperti rotasi gambar sebesar 90 derajat searah jarum jam, 180 derajat, dan 90 derajat berlawanan arah jarum jam, serta flipping horizontal dan vertikal. Teknik-teknik tersebut membantu memperluas keragaman dataset tanpa perlu mengumpulkan gambar baru secara manual.

4. Pembagian Data

Setelah semua citra melewati proses praproses, dataset kemudian dibagi kedalam dua bagian utama yaitu data pelatihan dan data pengujian. Data pelatihan berfungsi untuk melatih model dalam mengenali pola-pola yang ada, dengan proporsi data yang lebih besar agar proses pembelajaran optimal. Sementara itu, data pengujian digunakan untuk mengevaluasi performa model secara objektif menggunakan data yang belum pernah dilihat sebelumnya.

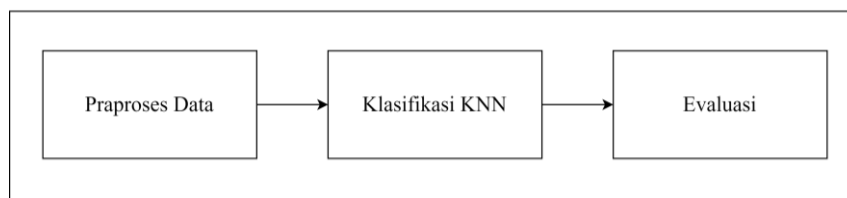
TABEL I
DISTRIBUSI DATASET

Skenario	Proporsi	%	Jumlah
1.	Training	75	900
2.	Test	25	300

Pada penelitian ini, dilakukan skenario pembagian data yaitu 75% data latih dan 25% data uji dari total keseluruhan gambar sebanyak 1200 setelah dilakukan augmentasi. Proporsi 75:25 memberikan cukup banyak data untuk proses pembelajaran model, sehingga pola-pola penting dalam citra daun dapat dikenali dengan lebih baik.

D. Pelatihan Model KNN

K-Nearest Neighbor merupakan algoritma klasifikasi yang bekerja dengan prinsip pengukuran jarak, di mana data dikategorikan berdasarkan kedekatannya terhadap data lain dalam ruang fitur [24]. Untuk mencapai performa yang optimal, penting untuk menentukan nilai k yang tepat, yaitu jumlah tetangga terdekat yang digunakan dalam proses pengambilan keputusan klasifikasi.



Gambar 9. Alur Model KNN

Gambar (9) menunjukkan alur model KNN pada penelitian ini. Dalam penerapan algoritma KNN untuk tugas klasifikasi, langkah awal yang dilakukan adalah menetapkan nilai parameter K yaitu jumlah tetangga terdekat yang akan dijadikan acuan dalam proses prediksi. Setelah nilai K ditentukan, algoritma akan mengidentifikasi sejumlah data yang paling dekat dengan data baru dengan menghitung jarak antar data. Kemudian, klasifikasi dilakukan berdasarkan mayoritas kelas dari K tetangga tersebut. Pada penelitian ini, metode pengukuran jarak yang digunakan ditunjukkan pada Tabel (2).

TABEL II
METRIC PENGUJIAN MODEL K-NN

Skenario	Keterangan
K	a. 5
	b. 11
Distance	a. <i>Euclidean</i>
	b. <i>Manhattan</i>

Pada penelitian ini menggunakan dua metrik yang umum digunakan pada KNN. Pengukuran jarak memiliki peran penting dalam menilai keterkaitan antar data. Nilai K yang lebih kecil seperti 5 memungkinkan model untuk lebih responsif terhadap pola lokal di data, sehingga cocok ketika data memiliki cluster yang cukup terpisah. Sementara itu, nilai K yang lebih besar seperti 11 memberikan efek smoothing yang lebih besar, karena keputusan klasifikasi didasarkan pada lebih banyak tetangga, yang dapat membantu mengurangi noise dan membuat prediksi lebih stabil. Metode ini digunakan untuk menilai tingkat kemiripan atau perbedaan antara satu data dengan yang lain, sehingga dapat membantu menentukan teknik pengukuran jarak yang paling sesuai. Berikut merupakan rumus jarak *Euclidean*, dan *Manhattan* yang ditunjukkan pada Persamaan (1), dan (2).

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

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

Keterangan:

- d : jarak
- x : data pusat pada kluster
- y : data atribut
- i : tiap-tiap data
- n : jumlah data
- x_i : data pusat kluster ke i
- y_i : data pusat kluster ke i

E. Evaluasi Pengujian Model

Evaluasi kinerja model klasifikasi umumnya dilakukan dengan menggunakan *confusion matrix*, yang menyajikan informasi mengenai prediksi benar dan salah dari model. *Confusion matrix* terdiri dari empat komponen utama, yaitu *True Positive* (TP), ketika model dengan tepat memprediksi kelas positif, *True Negative* (TN), saat model benar dalam memprediksi kelas negative, *False Positive* (FP), yaitu ketika model secara keliru memprediksi data negatif sebagai positif, dan *False Negative* (FN), di mana model salah memprediksi data positif sebagai negatif. Dari nilai-nilai ini, dapat dihitung metrik evaluasi lain seperti akurasi, presisi, recall, dan F1-score, yang memberikan gambaran lebih komprehensif mengenai performa model, khususnya dalam mengatasi ketidakseimbangan data dan keakuratan prediksi di masing-masing kelas. Nilai akurasi serta *confusion matrix* dapat dihitung menggunakan rumus pada Persamaan (3), (4), (5), dan (6).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

III. HASIL DAN PEMBAHASAN

Pada penelitian ini, dilakukan pengujian sebanyak 4 skenario klasifikasi menggunakan model KNN. Setiap skenario disusun dengan menggunakan dua jenis metrik jarak yaitu *Euclidean* dan *Manhattan*. Selain itu, pembagian data menggunakan variasi proporsi yaitu 75% data latih dan 25% data uji. Seluruh skenario dievaluasi berdasarkan performa klasifikasi terhadap data uji, dan hasil terbaik dari masing-masing kombinasi dipilih untuk dianalisis lebih lanjut.

A. Klasifikasi KNN Metrik Euclidean

Pada metrik *Euclidean*, dilakukan dua skenario pengujian dengan variasi proporsi data, yaitu 75:25 dengan 900 data latih dan 300 data uji. Pengujian dilakukan dengan menggunakan beberapa kombinasi nilai parameter, yaitu jumlah tetangga (k) sebanyak 5 dan 11 dan *weights* diatur default. Hasil akurasi dari setiap pengujian ditampilkan pada Tabel (3).

TABEL III
HASIL PENGUJIAN METRIK *EUCLIDEAN*

Skenario	K	Akurasi
1.	5	0,96333
2.	11	0,95667

Berdasarkan hasil yang ditampilkan pada Tabel (3), akurasi terbaik pada metrik *Euclidean* dengan proporsi data 75:25 adalah sebesar 0,96333 pada jumlah tetangga $K = 5$. Hasil ini menunjukkan bahwa metode KNN dengan metrik *Euclidean* memiliki kemampuan klasifikasi yang cukup baik dalam membedakan jenis daun herbal, meskipun tanpa menggunakan proses ekstraksi fitur yang kompleks.

B. Klasifikasi KNN Metrik Manhattan

Pada metrik *Manhattan*, dilakukan dua skenario pengujian dengan variasi proporsi data, yaitu 75:25 dengan 900 data latih dan 300 data uji. Pengujian dilakukan dengan menggunakan beberapa kombinasi nilai parameter, yaitu jumlah tetangga (k) sebanyak 5 dan 11 dan *weights* diatur default. Hasil akurasi dari setiap pengujian ditampilkan pada Tabel (4).

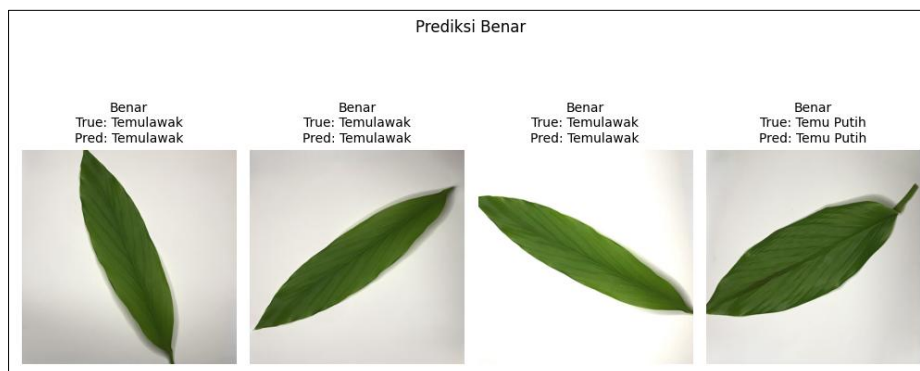
TABEL IV
HASIL PENGUJIAN METRIK *MANHATTAN*

Skenario	K	Akurasi
1.	5	0,98333
2.	11	0,96333

Berdasarkan hasil yang ditampilkan pada Tabel (4), akurasi terbaik pada metrik *Manhattan* dengan proporsi data 75:25 adalah sebesar 0,98333 pada jumlah tetangga $K = 5$. Hasil ini menunjukkan performa KNN yang sangat optimal dalam klasifikasi citra daun herbal. Tingginya akurasi ini mengindikasikan bahwa metode sederhana seperti KNN masih sangat relevan dan kompetitif untuk digunakan, terutama dalam sistem dengan keterbatasan komputasi.

C. Hasil Skenario Terbaik

Berdasarkan hasil keseluruhan skenario yang telah diuji, diperoleh satu skenario terbaik yang menghasilkan performa klasifikasi paling optimal. Skenario ini dipilih dari kombinasi parameter yang memberikan akurasi tertinggi, setelah melalui serangkaian proses pengujian dengan berbagai konfigurasi metrik jarak serta proporsi data latih dan uji. Skenario terbaik diperoleh pada metrik *Manhattan* dengan proporsi data 75:25 dan jumlah tetangga $K = 5$, yang menghasilkan akurasi sebesar 98,33%. Hasil ini menunjukkan bahwa pendekatan klasifikasi berbasis KNN dengan konfigurasi yang tepat mampu memberikan kinerja yang sangat baik meskipun tanpa menggunakan teknik deep learning. Temuan ini berpotensi memberikan dampak nyata dalam penerapan praktis, khususnya di sektor farmasi dan studi botani. Di bidang farmasi, sistem klasifikasi otomatis daun herbal dapat membantu mempercepat proses identifikasi bahan baku tanaman herbal, terutama dalam produksi berskala besar yang menuntut efisiensi dan konsistensi. Sementara itu, dalam konteks botani, teknologi ini mendukung digitalisasi dokumentasi dan konservasi keanekaragaman tanaman herbal, serta membantu peneliti dalam mengenali spesies tumbuhan berdasarkan ciri visualnya.



Gambar 10. Prediksi Benar



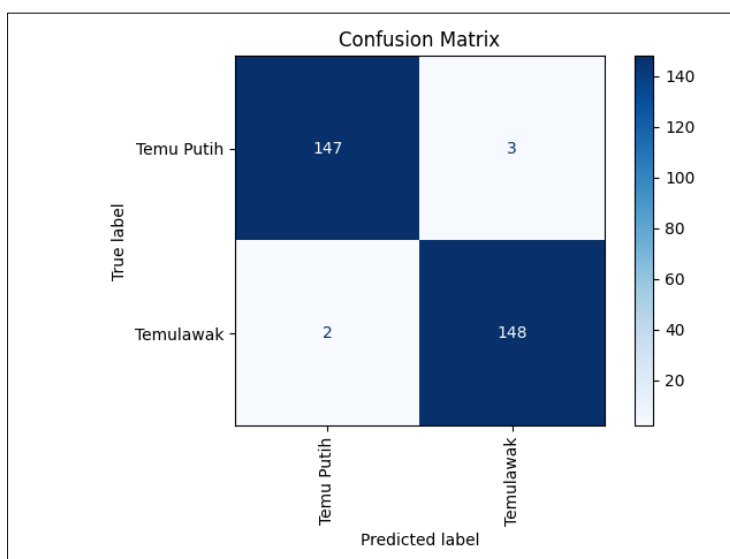
Gambar 11. Prediksi Salah

Prediksi hasil klasifikasi menunjukkan bahwa sebagian besar citra daun herbal berhasil dikenali dengan sangat baik oleh model, dengan jumlah kesalahan prediksi yang relatif sedikit. Gambar (10) menampilkan contoh hasil prediksi benar, di mana objek yang dikenali sesuai dengan label aslinya. Hal ini mencerminkan kemampuan model dalam mengenali pola visual dan ciri khas dari masing-masing jenis daun herbal. Sementara itu, Gambar (11) menunjukkan contoh prediksi salah, yang mengindikasikan area di mana model masih mengalami kesulitan dalam membedakan jenis daun tertentu.

Daftar Kesalahan Prediksi:		
File	True	Pred
Temulawak509.jpg	Temulawak	Temu Putih
Temulawak37.jpg	Temulawak	Temu Putih
TemuPutih528.jpg	Temu Putih	Temulawak
TemuPutih157.jpg	Temu Putih	Temulawak
TemuPutih509.jpg	Temu Putih	Temulawak

Gambar 12. Daftar Prediksi Kesalahan

Selain visualisasi prediksi benar dan salah, Gambar (12) menyajikan daftar lengkap citra yang mengalami kesalahan klasifikasi. Daftar ini mencakup nama file dan label prediksi yang diberikan model, sehingga memudahkan peneliti untuk melakukan evaluasi lebih mendalam terhadap karakteristik citra yang keliru dikenali. Informasi ini penting untuk mengidentifikasi pola kesalahan serta menjadi dasar dalam pengembangan model lebih lanjut. Kesalahan ini dapat disebabkan oleh kemiripan visual antar kelas atau variasi citra yang memengaruhi akurasi klasifikasi. Untuk memahami performa model secara keseluruhan, analisis lebih lanjut dapat dilakukan menggunakan *confusion matrix* pada Gambar (13) yang menggambarkan distribusi klasifikasi benar dan salah di setiap kelas. *Confusion matrix* memberikan wawasan penting mengenai kelas mana yang paling akurat diklasifikasikan serta kelas mana yang cenderung mengalami kesalahan.



Gambar 11. *Confusion matrix* Metric Terbaik

Gambar (13) menampilkan *confusion matrix* hasil klasifikasi untuk dua jenis tanaman, yaitu temu putih dan temulawak. Pada *matrix* tersebut, baris merepresentasikan label sebenarnya (*true label*), sedangkan kolom menunjukkan label hasil prediksi (*predicted label*). Model menunjukkan performa klasifikasi yang sangat baik, yang ditandai oleh nilai tinggi pada diagonal utama yaitu 147 sampel temu putih dan 148 sampel temulawak berhasil diklasifikasikan dengan benar. Terdapat sedikit kesalahan klasifikasi, dengan 3 sampel temu putih yang diprediksi sebagai temulawak, dan 2 sampel temulawak yang salah diklasifikasikan sebagai temu putih. Hasil ini menunjukkan bahwa model memiliki tingkat akurasi yang tinggi dan cukup andal dalam membedakan kedua jenis tanaman tersebut secara visual.

```
Classification Report:
      precision    recall  f1-score   support

 Temu Putih      0.99      0.98      0.98       150
 Temulawak       0.98      0.99      0.98       150

 accuracy              0.98       300
 macro avg              0.98       300
 weighted avg           0.98       300

 Accuracy : 0.98333
 Precision : 0.98335
 Recall    : 0.98333
 F1-Score  : 0.98333

=== Hasil Perhitungan TP, FP, FN, TN ===

Kelas: Temu Putih
True Positive (TP): 147
False Positive (FP): 2
False Negative (FN): 3
True Negative (TN): 148

Kelas: Temulawak
True Positive (TP): 148
False Positive (FP): 3
False Negative (FN): 2
True Negative (TN): 147
```

Gambar 14. *Classification report*

Gambar (14) menunjukkan *classification report* dari hasil klasifikasi dua jenis tanaman herbal, yaitu temu putih dan temulawak. Secara keseluruhan, model menghasilkan akurasi tinggi sebesar 98,33%, dengan nilai precision, recall, dan F1-score yang seimbang di angka 98% untuk masing-masing kelas. Untuk kelas temu putih, model memiliki precision sebesar 0.99, recall 0.98, dan F1-score 0.98 dari total 150 sampel, di mana 147 sampel berhasil diklasifikasikan dengan benar TP, 3 sampel salah diklasifikasikan sebagai temulawak FN, dan terdapat 2 kesalahan klasifikasi dari kelas temulawak yang dikira sebagai temu putih FP. Sementara itu, untuk kelas temulawak, model menunjukkan hasil yang serupa dengan *precision* 0.98, *recall* 0.99, dan *F1-score* 0.98. Sebanyak 148 sampel berhasil diklasifikasikan dengan benar, 2 sampel salah diklasifikasikan sebagai temu putih, dan terdapat 3 prediksi keliru dari kelas temu putih yang dikira sebagai temulawak. Nilai-nilai metrik tersebut menunjukkan bahwa model memiliki performa yang sangat baik, dengan kemampuan yang konsisten dalam membedakan kedua jenis tanaman secara akurat. Hal ini menjadikan model ini potensial untuk digunakan dalam sistem klasifikasi otomatis tanaman herbal di dunia nyata.

Dalam Penelitian ini masih memiliki beberapa keterbatasan, khususnya dalam membedakan spesies tanaman herbal yang memiliki kemiripan visual tinggi, seperti temu putih dan temulawak. Meskipun model menunjukkan akurasi yang tinggi sebesar 98,33%, masih terdapat kesalahan klasifikasi yang disebabkan oleh kemiripan bentuk, tekstur, dan warna daun antar spesies. Hal ini menunjukkan bahwa model masih rentan terhadap fitur visual yang saling tumpang tindih. Selain itu, proses *preprocessing* yang dilakukan tidak mencakup peningkatan kontras atau penyesuaian pencahayaan secara eksplisit, maka variasi kondisi pencahayaan ekstrem atau sudut pengambilan gambar yang tidak standar masih menjadi tantangan bagi model. Kondisi pencahayaan yang tidak konsisten pada beberapa citra juga turut memengaruhi performa model. Variasi intensitas cahaya dan bayangan dapat menyebabkan fitur visual yang terekstraksi menjadi kurang representatif, sehingga model kesulitan mengenali pola yang sebenarnya ada. Faktor lain yang perlu diperhatikan adalah sudut pandang objek yang tidak umum atau sebagian daun yang terpotong, yang dapat menurunkan akurasi prediksi. Oleh karena itu, untuk meningkatkan performa, disarankan untuk menambahkan teknik augmentasi yang lebih adaptif, seperti rotasi acak, perubahan pencahayaan, atau distorsi geometri ringan guna meningkatkan keragaman data latih. Oleh karena itu, untuk penelitian selanjutnya, disarankan agar dataset diperluas dengan jumlah citra yang lebih beragam, mencakup berbagai kondisi pencahayaan, orientasi, dan latar belakang. Rekomendasi lainnya adalah penerapan teknik augmentasi gambar lanjutan, seperti *Generative Adversarial Networks (GAN)* atau *adaptive augmentation*, yang dapat menghasilkan citra sintesis dengan karakteristik visual yang lebih bervariasi. Di samping itu, penggunaan metode ekstraksi fitur yang lebih kompleks atau arsitektur model yang lebih dalam dan canggih misalnya CNN dengan lapisan lebih dalam atau pendekatan berbasis transformer dapat meningkatkan sensitivitas model terhadap detail halus, sehingga akurasi klasifikasi antar spesies dengan kemiripan tinggi dapat lebih ditingkatkan.

Hasil penelitian ini menunjukkan akurasi klasifikasi sebesar 98,33% menggunakan metode KNN dengan metrik *Manhattan* dan jumlah tetangga ($k = 5$), yang sedikit lebih tinggi dibandingkan dengan penelitian terdahulu oleh Isman, dkk [12] yang menguji klasifikasi daun herbal di Sulawesi Tenggara dengan akurasi sebesar 97,5%. Perbedaan ini kemungkinan disebabkan oleh penerapan teknik *preprocessing* yang lebih efektif, seperti CLAHE, yang meningkatkan kualitas citra dan kontras fitur visual, serta optimasi parameter KNN yang lebih tepat pada penelitian ini, salah satunya melalui pemilihan metrik *Manhattan* yang terbukti memberikan hasil klasifikasi yang

lebih akurat dibandingkan *Euclidean* pada dataset ini. Hal ini dikarenakan *Manhattan* lebih stabil terhadap dimensi tinggi dan variasi kecil antar fitur yang umum ditemukan dalam citra daun, serta lebih tahan terhadap pengaruh nilai ekstrem dibandingkan *Euclidean*. Karakteristik ini membuat *Manhattan* lebih sesuai untuk menangani data citra daun herbal yang memiliki dimensi fitur tinggi dan variasi visual yang kompleks.

Kontribusi utama penelitian ini terletak pada peningkatan performa klasifikasi daun herbal melalui pemilihan metrik jarak dan konfigurasi parameter yang optimal, serta penggunaan preprocessing yang membantu model mengenali pola visual dengan lebih akurat. Dengan demikian, penelitian ini memberikan nilai tambah yang signifikan dalam pengembangan sistem klasifikasi daun herbal yang dapat diaplikasikan secara praktis contohnya, sistem ini dapat digunakan dalam identifikasi cepat spesies herbal di lapangan untuk keperluan produksi obat tradisional dan pengelolaan sumber daya tanaman di laboratorium botani. Penelitian ini tidak hanya menunjukkan peningkatan performa klasifikasi, tetapi juga membuka peluang pengembangan aplikasi nyata dalam bidang farmasi dan botani, sekaligus menjadi landasan bagi penelitian selanjutnya untuk mengatasi tantangan klasifikasi pada spesies dengan kemiripan visual tinggi.

IV. KESIMPULAN

Penelitian ini membuktikan bahwa algoritma klasifikasi KNN dengan praproses citra berupa resizing dan augmentasi data mampu menghasilkan performa klasifikasi yang tinggi pada citra daun herbal bertulang menyirip. Konfigurasi terbaik diperoleh dengan menggunakan metrik *Manhattan*, nilai $K = 5$, dan proporsi data latih-uji 75:25, yang menghasilkan akurasi sebesar 98,33%. Hasil evaluasi lain seperti *confusion matrix* menunjukkan kemampuan model dalam membedakan empat jenis daun herbal, yaitu temu putih, dan temulawak. Penelitian ini membuka peluang penerapan sistem klasifikasi daun herbal berbasis citra digital untuk bidang farmasi, botani, dan edukasi. Penelitian selanjutnya disarankan untuk memperluas dataset serta menerapkan teknik augmentasi dan metode lain guna meningkatkan performa model dalam mengatasi variasi dan kemiripan visual antar spesies.

DAFTAR PUSTAKA

- [1] Rizaldi Mokodompit, Novri Youla Kandowanko, and Marini Susanti Hamidun, "Keanekaragaman Tumbuhan di Kampus Universitas Negeri Gorontalo Kecamatan Tilong Kabila Kabupaten Bone Bolango," *BIOSFER*, vol. 7, pp. 75–80, 2021.
- [2] C. Kusmana and A. Hikmat, "The Biodiversity of Flora in Indonesia," *Journal of Natural Resources and Environmental Management*, vol. 5, no. 2, pp. 187–198, Dec. 2015, doi: 10.19081/jpsl.5.2.187.
- [3] I. Wulandari and M. Dahlan, "Pemanfaatan Tanaman Herbal Sebagai Obat Tradisional Pada Masyarakat Tompobulu Kabupaten Bantaeng," 2021.
- [4] Ravico, Anggi Desviana Siregar, and Nuzulul Ramadhona, "Pendidikan Karakter Melalui Kearifan Lokal: Kajian Nilai-nilai Tradisi Pengobatan Menta Gumeng dalam Kehidupan Masyarakat Kerinci," *Journal Of General and Character Education*, vol. 3, pp. 1–16, 2024.
- [5] Mar'atus Sholikhathul Ummah and Nuning Lisdiana, "Pengembangan UMKM Jamu Tradisional di Desa Sumberagung Kecamatan Klego Kabupaten Boyolali," *SENYUM Boyolali*, vol. 2, pp. 68–71, 2021.
- [6] R. Riyanti Kusumadewi, E. Uliaswati, K. Kamidah, and A. Adela Devada, "MINUMAN JAHE MADU UPAYA MEREDAKAN BATUK PADA BALITA," *Avicenna : Journal of Health Research*, vol. 7, no. 1, Mar. 2024, doi: 10.36419/avicenna.v7i1.1031.
- [7] T. Rahma Aizaniva, M. Fitri Yeni, and F. Yulina Ade, "Tanaman yang Dijadikan Obat Di Kota Solok dan Kabupaten Solok," 2024.
- [8] K. Nisa, "Pemanfaatan Tanaman Kunyit (*Curcuma longa*) Untuk Dijadikan Jamu Tradisional Sebagai Obat Penyakit Maag Didaerah Sumber, Kabupaten Cirebon," *SCIENTIAE EDUCATIA: JURNAL PENDIDIKAN SAINS*, 2020, doi: 10.24235/sc.educatia.vxix.xxxx.
- [9] P. R. Prayoga, P. Purnawansyah, T. Hasanuddin, and H. Darwis, "Klasifikasi Daun Herbal Menggunakan K-Nearest Neighbor dan Support Vector Machine dengan Fitur Fourier Descriptor," *Edumatic: Jurnal Pendidikan Informatika*, vol. 7, no. 1, pp. 160–168, Jun. 2023, doi: 10.29408/edumatic.v7i1.17521.
- [10] M. Samir, H. Darwis, F. Umar, and P. Korespondensi, "FOURIER DESCRIPTOR PADA KLASIFIKASI DAUN HERBAL MENGGUNAKAN SUPPORT VECTOR MACHINE DAN NAIVE BAYES FOURIER DESCRIPTOR ON CLASSIFICATION OF HERBAL LEAVES USING SUPPORT VECTOR MACHINE AND NAIVE BAYES," vol. 10, no. 6, pp. 1205–1212, 2023, doi: 10.25126/jtiik.2023107309.
- [11] Meiriyama and Sudiadi, "Penerapan Algoritma Random Forest Untuk Klasifikasi Jenis Daun Herbal," 2022.
- [12] Isman, Andani Ahmad, and Abdul Latief, "Perbandingan Metode KNN Dan LBPH Pada Klasifikasi Daun Herbal," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 5, no. 3, pp. 557–564, Jun. 2021, doi: 10.29207/resti.v5i3.3006.
- [13] M. H. Ahmad, M. Hana, T. Ghazi Pratama, and H. Aulida, "KLASIFIKASI EMPAT JENIS DAUN HERBAL MENGGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK," 2023.
- [14] L. P. R. Noviana and I. N. B. S. Nugraha, "Perbandingan Klasifikasi Citra Daun Herbal Menggunakan Metode Logistic Regression dan Decision Tree Classifier Berdasarkan Fitur (Warna, GLCM, Bentuk)," *JITU : Journal Informatic Technology And Communication*, vol. 7, no. 2, pp. 126–133, Nov. 2023, doi: 10.36596/jitu.v7i2.1241.
- [15] K. Tanaman Herbal Berdasarkan Tekstur Daun Menggunakan, M. Ezperanza Florentiana Letik, F. Yusuf Bisilisin, and I. Artikel Abstrak, "KETIK : Jurnal Informatika Faatuatua Media Karya," 2024.
- [16] F. Y. Meilan E, "KETIK : Jurnal Informatika Faatuatua Media Karya," 2024.
- [17] A. F. Novianti, M. Atthariq, J. N. Dini, A. B. Kaswar, and J. C. Lapendy, "KLASIFIKASI TINGKAT KESEGARAN DAUN BAWANG MENGGUNAKAN JARINGAN SYARAF TIRUAN BERBASIS PENGOLAHAN CITRA DIGITAL," *Jurnal Sistem Informasi dan Informatika (Simika)*, vol. 7, no. 2, pp. 223–237, Jul. 2024, doi: 10.47080/simika.v7i2.3378.
- [18] N. Nurdiansyah, M. Muliadi, R. Herteno, D. Kartini, and I. Budiman, "IMPLEMENTASI METODE PRINCIPAL COMPONENT ANALYSIS (PCA) DAN MODIFIED K-NEAREST NEIGHBOR PADA KLASIFIKASI CITRA DAUN TANAMAN HERBAL," *Jurnal Mnemonic*, vol. 7, no. 1, pp. 1–9, Feb. 2024, doi: 10.36040/mnemonic.v7i1.6664.
- [19] 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.
- [20] M. Meiriyama, S. Devella, and S. M. Adelfi, "Klasifikasi Daun Herbal Berdasarkan Fitur Bentuk dan Tekstur Menggunakan KNN," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 9, no. 3, pp. 2573–2584, Sep. 2022, doi: 10.35957/jatisi.v9i3.2974.
- [21] A. Ridhovan *et al.*, "PENERAPAN METODE RESIDUAL NETWORK (RESNET) DALAM KLASIFIKASI PENYAKIT PADA DAUN GANDUM."

- [22] E. Safitri, R. Heppy Ria Sibarani, Y. SM Sidabutar, and D. Kiswanto, “KLASIFIKASI PENYAKIT DAUN ANGGUR BERBASIS CITRA MENGGUNAKAN METODE K-NEAREST NEIGHBORS (KNN),” *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 6, pp. 12633–12642, Nov. 2024, doi: 10.36040/jati.v8i6.12004.
- [23] K. Mochammad *et al.*, “IMPLEMENTASI ARSITEKTUR ALEXNET DAN RESNET34 PADA KLASIFIKASI CITRA PENYAKIT DAUN KENTANG MENGGUNAKAN TRANSFER LEARNING,” 2023.
- [24] A. P. Sari, A. N. Sihananto, and A. Prasetya, “Implementasi Metode K-NN dalam Klasterisasi Kasus Kesehatan Jantung,” 2022. [Online]. Available: www.elektro.itn.ac.id