



Perbandingan Metode Naïve Bayes Dan K-Nearest Neighbor Dalam Mengklasifikasikan Penyakit Pada Daun Terong Berdasarkan Ekstraksi Fitur Gray Level Co-Occurrence Matrix

(Comparison Of Naïve Bayes And K-Nearest Neighbor Methods In Classifying Diseases In Eggplant Leaves Based On The Extraction Of Gray Level Co-Occurrence Matrix Features)

Merlin Puspitasari Mburu Hammu¹, Pinky Alfa Ray Leo Lede² dan Reynaldi Thimotius Abineno³

^{1,2,3} Program Studi Teknik Informatika, Fakultas Sains dan Teknologi, Universitas Kristen Wira Wacana Sumba

Jl. R. Suprato No.35, Prailiu, Kec. Kota Waingapu, Kabupaten Sumba Timur, Nusa Tenggara Timur
Email: puspitasaritesa4@gmail.com¹, pinky.leo.lede@unkriswina.ac.id², reynaldi@unkriswina.ac.id³.
Corresponding author: puspitasaritesa4@gmail.com

ABSTRACT

This study investigates the effectiveness of two popular classification algorithms, K-Nearest Neighbor (KNN) and Naïve Bayes Classifier (NBC), in identifying diseases and pests affecting eggplant leaves. To achieve this, the research utilizes images sourced from the PlantVillage dataset, a comprehensive repository of plant health images. These images are processed to extract texture features based on the Gray Level Co-occurrence Matrix (GLCM), which provides valuable information about the spatial relationship of pixel intensities. Specifically, four GLCM features—contrast, correlation, energy, and homogeneity—are calculated for each image to serve as input variables for the classifier. The goal is to determine which algorithm is more accurate in distinguishing healthy leaves from those affected by various diseases or pest attacks. The experimental results reveal that the Naïve Bayes classifier achieved an accuracy of 96,68%, showing superior performance compared to the KNN algorithm, which produced a lower accuracy. These findings suggest that Naïve Bayes, with its probabilistic approach, may be more effective for plant disease and pest classification tasks based on texture features.

Keywords: Classification, Naïve Bayes Classifier, K-Nearest Neighbor, Gray Level Co-occurrence Matrix, Eggplant Plant Disease.

ABSTRAK

Penelitian ini menyelidiki efektivitas dua algoritma klasifikasi populer, *K-Nearest Neighbor* (KNN) dan *Naïve Bayes Classifier* (NBC), dalam mengidentifikasi penyakit dan hama yang menyerang daun terong. Untuk mencapai hal ini, penelitian ini memanfaatkan citra yang bersumber dari dataset *PlantVillage*, sebuah repositori komprehensif citra kesehatan tanaman. Citra-citra tersebut diproses untuk mengekstraksi fitur tekstur berdasarkan *Gray Level Co-occurrence Matrix* (GLCM), yang menyediakan informasi berharga tentang hubungan spasial intensitas piksel. Secara khusus, empat fitur GLCM—kontras, korelasi, energi, dan homogenitas—dihitung untuk setiap citra untuk dijadikan variabel input bagi pengklasifikasi. Tujuannya adalah untuk menentukan algoritma mana yang lebih akurat membedakan daun sehat dari daun yang terkena berbagai penyakit atau serangan hama. Hasil eksperimen mengungkapkan bahwa pengklasifikasi *Naïve Bayes* mencapai akurasi 96,68%, menunjukkan kinerja yang lebih unggul dibandingkan algoritma KNN, yang menghasilkan akurasi yang lebih rendah. Temuan ini menunjukkan bahwa *Naïve Bayes*, dengan pendekatan probabilistiknya, mungkin lebih efektif untuk tugas klasifikasi penyakit tanaman dan hama berdasarkan fitur tekstur.

Kata kunci: Klasifikasi, *Naïve Bayes Classifier*, *K-Nearest Neighbor*, *Gray Level Co-occurrence Matrix*, Penyakit Tanaman Terong.



PENDAHULUAN

Terong, yang banyak dibudidayakan di Indonesia untuk konsumsi rumah tangga sebagai sayuran serbaguna, merupakan bahan pokok dalam banyak hidangan tradisional yang dinikmati di seluruh negeri. Terong dikenal memiliki nilai gizi cukup tinggi, mengandung vitamin A, fosfor, serta senyawa bioaktif yang bermanfaat bagi kesehatan, seperti zat anti-kanker, penurunan kolesterol, dan penjaga fungsi jantung (Sulardi dkk., 2022).

Namun, meskipun permintaan pasar terhadap terong terus meningkat, produktivitasnya masih terhambat oleh berbagai tantangan budidaya, terutama serangan penyakit dan hama. Beberapa penyakit umum yang menyerang daun terong antara lain bercak daun, layu fusarium, jamur putih, virus kuning, serta gangguan dari hama serangga. Kondisi ini sering kali tidak terdeteksi secara dini oleh petani, sehingga menurunkan hasil panen secara signifikan.

Untuk mengatasi tantangan itu, teknologi pengolahan citra digital muncul sebagai solusi modern yang dapat mendeteksi penyakit tanaman dengan lebih cepat dan akurat. Dalam metode klasifikasi citra daun yang umum, atribut warna dan fitur tekstur yang diperoleh dari *Gray Level Co-occurrence Matrix* (GLCM) digunakan untuk mengenali berbagai spesies tanaman dengan tepat. Algoritma pembelajaran mesin terawasi seperti *Naïve Bayes Classifier* dan *K-Nearest Neighbors* (KNN) diterapkan, keduanya menunjukkan keefektifan dalam mengklasifikasikan gambar daun berdasarkan fitur-fitur yang diambil ini.

Penelitian sebelumnya menunjukkan efektivitas penggunaan kombinasi metode ekstraksi fitur dan algoritma klasifikasi. Misalnya, (Herdiansah dkk., 2022) menggunakan 200 citra daun herbal dengan algoritma *Backpropagation Neural Network* (BPNN) dan memperoleh akurasi 88,75%. Sementara itu, (Meiriyama dkk., 2022) menggunakan fitur HOG dan LBP dengan KNN 15 jenis tanaman herbal dan mencatat akurasi hingga 92,67%.

Melanjutkan pendekatan tersebut, penelitian ini bertujuan untuk mengimplementasikan GLCM sebagai teknik ekstraksi fitur dan membandingkan performa *Naïve Bayes Classifier* dan KNN dalam mengklasifikasikan citra daun terong yang terserang penyakit dan hama. Dataset yang digunakan diambil dari *Kaggle*.

MATERI DAN METODE

Klasifikasi

Klasifikasi adalah proses fundamental yang melibatkan pengorganisasian objek atau ide secara sistematis berdasarkan karakteristik umumnya. Metode ini membantu dalam memahami dan menganalisis informasi kompleks dengan mengelompokkan item-item yang serupa. Untuk memfasilitasi hal ini, simbol yang dikenal sebagai notasi ditetapkan untuk setiap kelompok atau item. Terdapat dua jenis notasi utama: notasi murni, yang hanya menggunakan huruf atau angka, dan notasi campuran, yang menggabungkan simbol dan karakter. Sistem ini memudahkan identifikasi, perbandingan, dan pengomunikasian klasifikasi secara efektif.

Klasifikasi K-Nearest Neighbor

Dalam pendekatan *K-Nearest Neighbor* (K-NN) standar, sistem mengklasifikasikan hama dan penyakit daun terong dengan menganalisis secara cermat ciri-ciri visual daun tanaman. Sistem membandingkan ciri-ciri ini dengan dataset pelatihan yang sudah ada, dengan fokus pada kesamaan gejala seperti bercak, perubahan warna, atau deformitas. Dengan mengukur seberapa dekat citra baru dengan contoh yang sudah diketahui, metode ini secara

akurat menentukan kategori, membantu petani dalam diagnosis tepat waktu dan pengobatan yang efektif (Saputra dkk., 2025).

Klasifikasi *Naïve Bayes Classifier*

Naïve Bayes Classifier dipilih karena pendekatannya yang lugas, sehingga mudah diimplementasikan dan dipahami. Akurasinya yang tinggi dalam tugas klasifikasi, serta efisiensinya dalam menangani data terstruktur, memungkinkannya mengungguli banyak algoritma lainnya. Kombinasi kesederhanaan dan efektivitas ini menjadikan *Naïve Bayes Classifier* pilihan populer untuk berbagai aplikasi analisis data (Alfandi Safira & Hasan, 2023). Keunggulan lainnya adalah kemampuannya menangani data kategorikal maupun numerik, berkat adanya varian seperti *Gaussian*, Multinomial, dan Bernoulli. Meskipun asumsi independensi antar fitur jarang berlaku di dunia nyata, *Naïve Bayes* tetap memberikan hasil yang baik pada berbagai aplikasi klasifikasi, seperti pemfilteran spam, analisis sentimen, dan diagnosis medis. Kecepatan, skalabilitas, kebutuhan data pelatihan yang minim, serta ketahanannya terhadap data yang hilang, menjadikan *Naïve Bayes* lebih unggul dibandingkan banyak algoritma lain yang lebih kompleks secara teori.

Daun Terong

Daun terong biasanya berukuran panjang antara 12 dan 20 sentimeter. Setiap daun memiliki tangkai kokoh yang menghubungkannya dengan tanaman, helaian daun lebar yang menyediakan permukaan untuk fotosintesis, dan pelepah daun menonjol yang membentang di tengahnya, menopang struktur daun.

Tabel 1 Citra Daun Terong

No	Jenis Penyakit	Jumlah Citra
1	 Bercak Daun	15
2	 Daun Sehat	15
3	 Hama Serangga	15

4		15
		
5	Jamur Putih	15
		
6	Penyakit Layu	15
		
	Virus Mozaik	

Citra Berwarna

Citra warna, atau biasa disebut citra RGB, merupakan jenis citra yang merepresentasikan warna dengan tiga komponen utama, yaitu R (red), G (green), dan B (blue). Tiap komponen warna ini memanfaatkan 8 bit, dengan nilai yang berada di antara 0 hingga 255. Sehingga, jumlah total warna yang bisa dihasilkan adalah sebanyak 16.581.375, yang diperoleh dari $255 \times 255 \times 255$ (Wahyudi dkk., 2015).

Pengolahan Citra

Pemrosesan citra merupakan bidang vital yang berfokus pada manipulasi dan analisis data visual untuk meningkatkan kualitas dan kegunaannya. Proses ini dimulai dengan memasukkan citra, yang seringkali berkualitas buruk akibat noise, blur, atau distorsi lainnya. Langkah awal biasanya melibatkan peningkatan kualitas citra, yang bertujuan untuk memperjelas dan meningkatkan kejernihan visual citra tersebut. Teknik yang digunakan meliputi peningkatan citra untuk meningkatkan detail, perbaikan dan restorasi untuk memperbaiki kerusakan, penyempurnaan model untuk akurasi yang lebih baik, dan pemrosesan warna untuk meningkatkan fidelitas warna. Metode-metode ini secara kolektif bekerja untuk mengurangi noise yang tidak diinginkan, meminimalkan blur, dan mengatasi berbagai ketidaksempurnaan, yang pada akhirnya menghasilkan citra yang lebih jernih, lebih akurat, dan menarik secara visual yang cocok untuk dianalisis atau ditampilkan lebih lanjut (Sugiarti, 2018).

Metode Grayscale

Dalam teknik skala abu-abu yang umum, prosesnya melibatkan penyesuaian kecerahan dan kontras gambar secara bertahap dengan beralih secara bergantian antara versi asli dan versi yang dimodifikasi, menciptakan transisi halus yang meningkatkan kejelasan dan detail visual (Pramudiya dkk., 2024). Metode skala abu-abu secara signifikan meningkatkan kejernihan gambar dan daya tarik visual dengan menyesuaikan tingkat kecerahan dan kontras secara



cermat, sehingga menghasilkan tampilan yang lebih mencolok dan halus yang memikat pemirsa.

Metode Gray Level Co-occurrence Matrix

Gray Level Co-occurrence Matrix (GLCM) adalah teknik statistik yang banyak digunakan untuk menganalisis tekstur pada citra. Metode ini berfungsi dengan memeriksa pasangan piksel yang memiliki nilai intensitas tertentu, sambil mempertimbangkan hubungan spasial di antara mereka. GLCM menghitung frekuensi kemunculan pasangan intensitas abu-abu pada jarak dan orientasi tertentu, sehingga menghasilkan matriks yang mencerminkan distribusi spasial intensitas piksel. Matriks ini kemudian digunakan untuk mengekstrak fitur tekstur seperti kontras, korelasi, dan homogenitas. Secara keseluruhan, GLCM memberikan pemahaman yang lebih mendalam mengenai pola tekstur dalam citra, yang mendukung analisis dan interpretasi yang lebih akurat.

Namun, GLCM memiliki beberapa kelemahan. Salah satunya adalah ketergantungan pada ukuran jendela yang digunakan dalam analisis, yang dapat mempengaruhi hasil ekstraksi fitur tekstur. Jika ukuran jendela terlalu kecil, informasi spasial penting dapat hilang, sedangkan jendela yang terlalu besar bisa mengaburkan detail tekstur yang signifikan. Selain itu, GLCM kurang efektif dalam menangani variasi skala dan rotasi citra, sehingga mungkin tidak optimal untuk analisis citra dengan banyak variasi orientasi dan ukuran objek. GLCM juga membutuhkan daya komputasi yang cukup besar, terutama ketika diterapkan pada citra dengan resolusi tinggi atau ketika banyak fitur yang diekstrak.

1. Contrast

Contrast, yang juga dikenal sebagai inersia, adalah ukuran yang menggambarkan sejauh mana variasi intensitas piksel dalam sebuah gambar. Jika kontrasnya tinggi, itu berarti terdapat banyak perbedaan dalam intensitas piksel. Sebaliknya, kontras yang rendah menunjukkan perbedaan intensitas yang lebih sedikit. *Contrast* dapat dihitung dengan rumus sebagai berikut.

$$Contrast = \sum_i \sum_j (i - j)^2 * P(i, j)$$

2. Correlation

Correlation adalah ukuran yang menggambarkan seberapa besar hubungan antara dua piksel. *Correlation* yang tinggi menunjukkan adanya ketergantungan yang kuat antara kedua piksel tersebut. *Correlation* dapat dihitung dengan menggunakan rumus berikut.

$$Correlation = \frac{\sum_i \sum_j [ij * P(i, j)] - \mu_x * \mu_y}{\sigma_x * \sigma_y}$$

3. Energy

Energy adalah fitur GLCM yang digunakan untuk mengukur kekuatan pasangan intensitas dalam matriks GLCM, dan didefinisikan sebagai berikut.

$$Energy = \sum_i \sum_j p(i, j)^2$$

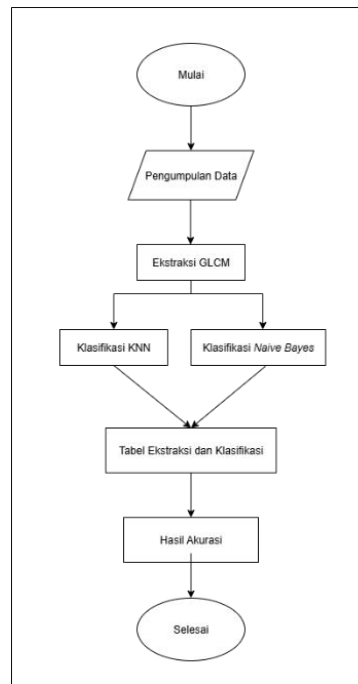
4. Homogeneity

Homogeneity adalah ukuran yang menggambarkan sejauh mana elemen-elemen dalam matriks GLCM mendekati diagonalnya. Nilai *Homogeneity* akan tinggi jika elemen-

elemen matriks dengan nilai intensitas yang serupa terletak dekat dengan diagonal. Homogeneity dapat dihitung menggunakan rumus berikut.

HASIL DAN PEMBAHASAN

Alur penelitian



Gambar 1 Alur Penelitian

Studi ini dimulai dengan mengumpulkan data gambar daun terong dari dataset publik, selanjutnya dilanjutkan dengan tahap pra-pemrosesan, yaitu mengubah gambar RGB menjadi *grayscale*. Selanjutnya, fitur tekstur diekstraksi dengan menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM) untuk mendapatkan empat fitur utama, yaitu kontras, korelasi, energi, dan homogenitas. Nilai-nilai fitur itu selanjutnya digunakan sebagai input dalam proses klasifikasi dengan algoritma *K-Nearest Neighbor* dan *Naïve Bayes Classifier*. Hasil klasifikasi dievaluasi menggunakan confusion matrix untuk memperoleh metrik akurasi, presisi, recall, dan F1-score.

Analisis Klasifikasi *Naïve Bayes Classifier*

Klasifikasi *Naif Bayes Classifier* mengevaluasi kinerjanya dengan menganalisis berbagai fitur statistik seperti kontras, korelasi, energi, dan homogenitas. Dengan membandingkan berbagai kelas dan hasil target, metode ini menentukan seberapa baik metode ini membedakan antarkategori. Keberhasilan diukur melalui identifikasi prediksi yang benar dan meminimalkan prediksi yang salah, sehingga memastikan keandalan dan efektivitas model.

Berikut ini merupakan hasil klasifikasinya.

Tabel 2 Hasil Klasifikasi *Naïve Bayes Classifier*

No	Contrast	Correlation	Energy	Homogeneity	Probabilitas	Target	Kelas	Hasil
1	860.321	0.747	0.318	0.679	0.997	1	1	Benar
2	226.750	0.642	0.190	0.629	0.436	1	1	Benar
3	1469.75	0.617	0.174	0.587	0.374	1	1	Benar
4	763.178	0.495	0.174	0.580	0.374	1	1	Benar



5	1065.50	0.661	0.182	0.589	0.273	1	1	Benar
6	1692.357	0.688	0.167	0.575	0.367	2	2	Benar
7	870.142	0.745	0.174	0.586	0.514	2	2	Benar
8	308.678	0.842	0.174	0.577	0.546	2	2	Benar
9	625.392	0.731	0.167	0.584	0.491	2	2	Benar
10	399.285	0.527	0.174	0.583	0.367	2	2	Benar
11	290.285	0.802	0.182	0.597	0.328	3	3	Benar
12	1481.035	0.729	0.174	0.589	0.517	3	3	Benar
13	1442.821	0.659	0.174	0.572	0.296	3	3	Benar
14	533.214	0.754	0.167	0.576	0.514	3	3	Benar
15	560.5	0.728	0.174	0.573	0.497	3	3	Benar
16	2748.571	0.570	0.188	0.579	0.999	4	4	Benar
17	1967.071	0.688	0.174	0.571	0.585	4	4	Benar
18	823.928	0.709	0.174	0.573	0.481	4	4	Benar
19	1501.107	0.591	0.235	0.614	0.439	4	4	Benar
20	2817.214	0.518	0.188	0.576	0.999	4	4	Benar
21	952.25	0.576	0.182	0.579	0.343	5	5	Benar
22	352.071	0.631	0.217	0.580	0.646	5	5	Benar
23	368.678	0.567	0.214	0.583	0.665	5	5	Benar
24	924.714	0.669	0.197	0.608	0.390	5	5	Benar
25	261.821	0.837	0.266	0.610	0.891	5	5	Benar
26	1501.821	0.562	0.174	0.607	0.446	6	4	Salah
27	1445.428	0.776	0.203	0.609	0.484	6	6	Benar
28	464.607	0.665	0.192	0.574	0.294	6	6	Benar
29	606.964	0.664	0.174	0.615	0.351	6	6	Benar
30	339.607	0.813	0.192	0.614	0.393	6	6	Benar
Akurasi						96.68%		

Dalam proses klasifikasi, nilai kelas dibandingkan dengan target yang diinginkan. Jika nilai kelas sama dengan target, maka hasilnya dianggap benar. Tapi, jika kelas yang terbaca tidak sesuai dengan target, maka hasilnya dianggap salah. Setelah dijalankan di VSCode, ditemukan satu kesalahan, yaitu: pada target 6, yaitu citra daun terong virus mozaik yang terbaca sebagai kelas 4 (citra daun terong jamur putih). Klasifikasi dianggap benar jika kelas yang terbaca sesuai dengan target. Setelah proses klasifikasi untuk menentukan hasil benar atau salah selesai, langkah berikutnya adalah menghitung menggunakan *Naïve Bayes Classifier* untuk menentukan akurasi berdasarkan nilai probabilitas.

Analisis Klasifikasi *K-Nearest Neighbor*

Dalam proses klasifikasi menggunakan KNN, akurasi ditentukan dengan menghitung jarak Euclidean antara setiap data uji dan seluruh data pelatihan. Data pelatihan yang memiliki jarak terdekat dianggap sebagai tetangga terdekat, dan klasifikasi data uji dilakukan berdasarkan label yang paling sering muncul di antara tetangga tersebut.

Berikut ini merupakan hasil klasifikasinya.

Tabel 3 Klasifikasi *K-Nearest Neighbor*

No	Contrast	Correlation	Energy	Homogeneity	Euclidean	Target	Kelas	Hasil
1	860.321	0.747	0.318	0.679	0.953	1	1	Benar
2	226.750	0.642	0.190	0.629	0.538	1	1	Benar
3	1469.75	0.617	0.174	0.587	0.614	1	1	Benar
4	763.178	0.495	0.174	0.580	0.247	1	1	Benar
5	1065.50	0.661	0.182	0.589	0.499	1	1	Benar



6	1692.357	0.688	0.167	0.575	0.742	2	3	Salah
7	870.142	0.745	0.174	0.586	0.731	2	3	Salah
8	308.678	0.842	0.174	0.577	1.009	2	3	Salah
9	625.392	0.731	0.167	0.584	0.737	2	3	Salah
10	399.285	0.527	0.174	0.583	0.191	2	3	Salah
11	290.285	0.802	0.182	0.597	0.929	3	3	Benar
12	1481.035	0.729	0.174	0.589	0.850	3	3	Benar
13	1442.821	0.659	0.174	0.572	0.569	3	3	Benar
14	533.214	0.754	0.167	0.576	0.741	3	3	Benar
15	560.5	0.728	0.174	0.573	0.676	3	3	Benar
16	2748.571	0.570	0.188	0.579	1.002	4	4	Benar
17	1967.071	0.688	0.174	0.571	0.668	4	4	Benar
18	823.928	0.709	0.174	0.573	0.771	4	4	Benar
19	1501.107	0.591	0.235	0.614	0.664	4	4	Benar
20	2817.214	0.518	0.188	0.576	1.007	4	4	Benar
21	952.25	0.576	0.182	0.579	0.661	5	5	Benar
22	352.071	0.631	0.217	0.580	0.382	5	5	Benar
23	368.678	0.567	0.214	0.583	0.367	5	5	Benar
24	924.714	0.669	0.197	0.608	0.635	5	5	Benar
25	261.821	0.837	0.266	0.610	0.978	5	5	Benar
26	1501.821	0.562	0.174	0.607	0.657	6	6	Benar
27	1445.428	0.776	0.203	0.609	0.847	6	6	Benar
28	464.607	0.665	0.192	0.574	0.615	6	5	Salah
29	606.964	0.664	0.174	0.615	0.650	6	6	Benar
30	339.607	0.813	0.192	0.614	0.897	6	6	Benar
Akurasi						83.33%		

Dalam tugas klasifikasi, sistem mengevaluasi nilai kelas yang diprediksi terhadap label target yang sebenarnya. Jika kedua nilai cocok, prediksi dianggap benar, yang menandakan klasifikasi berhasil. Sebaliknya, jika berbeda, sistem menganggap prediksi salah, yang menunjukkan perlunya penyesuaian atau pembelajaran lebih lanjut untuk meningkatkan akurasi. Setelah dijalankan di VSCode, ditemukan kesalahan sebanyak 5 kali, antara lain: target 2 (citra daun terong sehat) terbaca sebagai kelas 3 (citra daun terong hama serangga). Dalam proses klasifikasi standar, keakuratan klasifikasi dikonfirmasi ketika kelas yang diprediksi cocok dengan kelas target yang sebenarnya. Setelah verifikasi bahwa data gambar telah dibaca dan diproses secara akurat, akurasi keseluruhan kemudian dihitung menggunakan algoritma *K-Nearest Neighbors* (K-NN). Hal ini memastikan evaluasi kinerja model yang andal.

Pembagian Data

Makalah ini meneliti kinerja dua algoritma klasifikasi yang sering digunakan, yaitu *Naive Bayes Classifier* (NBC) dan *K-Nearest Neighbors* (KNN), dalam mengidentifikasi tiga jenis penyakit pada daun terong secara tepat. Proses ekstraksi fitur dilakukan menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM) untuk memperoleh ciri-ciri tekstur dari citra daun. Tujuan dari penelitian ini adalah untuk membandingkan keefektifan kedua metode tersebut dalam menghasilkan hasil klasifikasi yang terbaik. Dataset yang digunakan terdiri dari 120 gambar daun yang terbagi ke dalam enam folder sebagai data pelatihan dan data pengujian.



Gambar 2 Pembagian Data Latih



Gambar 3 Pembagian Data Uji

Sebanyak 90 citra digunakan sebagai data latih, yang dikelompokkan ke dalam enam folder, masing-masing berisi 15 gambar. Sementara itu, data uji terdiri dari 30 citra yang juga dibagi ke dalam enam folder, dengan setiap folder memuat 5 gambar. Data uji ini bersifat independen dan tidak diambil dari data latih, untuk mencegah duplikasi yang dapat memengaruhi hasil klasifikasi.

Kedua metode klasifikasi, yakni *Naïve Bayes Classifier* (NBC) dan *K-Nearest Neighbor* (KNN), diuji secara terpisah untuk memungkinkan perbandingan tingkat akurasi dalam mengklasifikasikan daun terong. Proses klasifikasi didasarkan pada fitur tekstur yang diekstraksi menggunakan *Gray Level Co-occurrence Matrix* (GLCM), yang menghasilkan empat parameter utama: *Contrast*, *Correlation*, *Energy*, dan *Homogeneity*. Nilai-nilai ini selanjutnya digunakan sebagai input pada tahap klasifikasi.

Hasil pengujian menunjukkan bahwa metode NBC mampu mengklasifikasikan 120 citra daun terong dengan tingkat akurasi sebesar 96,68%, sedangkan metode KNN menghasilkan akurasi sebesar 83,33%. Dari hasil ini dapat disimpulkan bahwa NBC memiliki performa yang lebih baik dibandingkan KNN dalam mengidentifikasi jenis penyakit pada daun terong.

KESIMPULAN

Dari hasil analisis dan pembahasan terhadap dua metode klasifikasi, yaitu *Naïve Bayes Classifier* (NBC) dan *K-Nearest Neighbor* (KNN), yang diterapkan pada enam set data citra daun terong masing-masing terdiri dari 15 gambar untuk pelatihan dan 5 gambar untuk pengujian diperoleh bahwa NBC menunjukkan performa yang lebih unggul dibandingkan KNN. NBC berhasil mencapai akurasi sebesar 96,68%, sedangkan KNN hanya mencapai 83,33%. Berdasarkan perbandingan tersebut, dapat disimpulkan bahwa metode NBC lebih efektif dalam mengklasifikasikan citra daun terong dalam penelitian ini dibandingkan metode KNN.

Penelitian ini memiliki beberapa kelemahan yang perlu diperhatikan dan dapat dijadikan bahan evaluasi untuk penelitian selanjutnya. Pertama, penelitian ini menggunakan dataset yang terbatas, yaitu hanya terdiri dari 120 gambar daun terong dengan enam kelas penyakit. Keterbatasan jumlah dataset ini berpotensi mempengaruhi kemampuan model dalam melakukan generalisasi. Oleh karena itu, disarankan untuk meningkatkan jumlah dataset secara signifikan dengan menambahkan lebih banyak gambar yang mencakup berbagai kondisi tanaman, baik yang sehat maupun yang terinfeksi penyakit atau hama. Dataset yang lebih besar



dan beragam akan menghasilkan model yang lebih kuat dan dapat mengidentifikasi lebih banyak variasi dalam citra daun. Kedua, keragaman kelas penyakit dalam penelitian ini terbatas pada lima jenis penyakit. Untuk meningkatkan akurasi dan validitas model, sebaiknya penelitian selanjutnya mencakup lebih banyak kelas penyakit, sehingga model dapat dilatih untuk membedakan berbagai jenis penyakit yang dapat menyerang tanaman terong. Terakhir, terkait validasi model, penelitian ini hanya menggunakan data pelatihan dan pengujian yang terbatas. Oleh karena itu, penggunaan teknik cross-validation atau pembagian data yang lebih bervariasi dapat membantu mengurangi risiko overfitting dan memberikan gambaran yang lebih akurat tentang kinerja model pada data yang tidak pernah terlihat sebelumnya. Dengan mengikuti saran-saran tersebut, diharapkan penelitian berikutnya dapat meningkatkan akurasi dan validitas model klasifikasi yang dikembangkan, serta memberikan hasil yang lebih efektif dalam mendeteksi penyakit pada tanaman.

DAFTAR PUSTAKA

- Alfandi Safira, & Hasan, F. N. (2023). Analisis Sentimen Masyarakat Terhadap Paylater Menggunakan Metode Naive Bayes Classifier. *Zonasi: Jurnal Sistem Informasi*, 5(1), 59–70. <https://doi.org/10.31849/Zn.V5i1.12856>
- Herdiansah, A., Borman, R. I., Nurnaningsih, D., Sinlae, A. A. J., & Al Hakim, R. R. (2022). Klasifikasi Citra Daun Herbal Dengan Menggunakan Backpropagation Neural Networks Berdasarkan Ekstraksi Ciri Bentuk. *Jurikom (Jurnal Riset Komputer)*, 9(2), 388. <https://doi.org/10.30865/Jurikom.V9i2.4066>
- Meiriyama, M., Devella, S., & Adelfi, S. M. (2022). Klasifikasi Daun Herbal Berdasarkan Fitur Bentuk Dan Tekstur Menggunakan Knn. *Jatissi (Jurnal Teknik Informatika Dan Sistem Informasi)*, 9(3), 2573–2584. <https://doi.org/10.35957/Jatissi.V9i3.2974>
- Pramudiya, R., Asyraq, C., & Kadafi, A. (2024). Analisis Gambar Menggunakan Metode Grayscale Dan Hsv (Hue, Saturation, Value). 14(3).
- Saputra, G. A., Ariyanto, D. A., Febriansyah, R., & Anugrah, R. (2025). Studi Literatur: Penggunaan K-Nearest Neighbors Untuk Klasifikasi Penyakit Daun Pada Tanaman Terong. 2(8).
- Sugiarti, S. (2018). Peningkatan Kualitas Citra Dengan Metode Fuzzy Possibility Distribution. *Ilkom Jurnal Ilmiah*, 10(1), 100–104. <https://doi.org/10.33096/Ilkom.V10i1.226.100-104>
- Sulardi, Hakim, T., Wasito, M., & Najla, L. (2022). *Agribisnis Budidaya Terong Unggu*. Pt Dewangga Energi Internasional Anggota Ikapi (403/Jba/2021).
- Wahyudi, E., Triyanto, D., & Ruslianto, I. (2015). *Identifikasi Teks Dokumen Menggunakan Metode Profile Projection Dan Template Matching*. 03.