

**KLASIFIKASI PENYAKIT PADA DAUN JAMBU BERBASIS  
PENGOLAHAN CITRA DENGAN ANALISIS TEKSTUR  
MENGUNAKAN ALGORITMA K-NEAREST NEIGHBORS (KNN)**

**Salsabila Zahra<sup>1</sup>, Silvana Oyasi Ayu<sup>2</sup>, Dedi Kiswanto<sup>3</sup>**

Universitas Negeri Medan

E-mail: [salsabilazahra139@gmail.com](mailto:salsabilazahra139@gmail.com)<sup>1</sup>, [oyasisilvana@gmail.com](mailto:oyasisilvana@gmail.com)<sup>2</sup>, [dedykiswanto@unimed.ac.id](mailto:dedykiswanto@unimed.ac.id)<sup>3</sup>

**Abstrak**

Identifikasi penyakit pada daun merupakan langkah penting dalam mendukung deteksi dini dan pengelolaan tanaman secara efektif. Penelitian ini 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 k. Hasil menunjukkan bahwa metode ini mampu mengklasifikasikan penyakit dengan tingkat akurasi yang baik, sehingga berpotensi menjadi solusi bagi sistem monitoring pertanian berbasis teknologi.

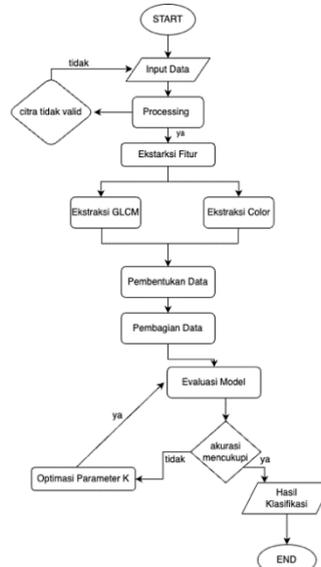
**Kata Kunci** — Pengolahan Citra, Klasifikasi Penyakit, Daun Jambu, GLCM, Color Moments, K-Nearest Neighbors (KNN).

**1. PENDAHULUAN**

Pengolahan citra digital telah berkembang pesat dalam berbagai bidang, termasuk dalam pertanian, khususnya untuk mendeteksi dan mengidentifikasi penyakit pada tanaman. Identifikasi penyakit pada daun tanaman merupakan langkah penting untuk mencegah kerugian yang signifikan pada hasil panen. Dalam hal ini, teknologi pengolahan citra memainkan peran penting dalam menghasilkan analisis yang akurat dan efisien (Rustanto, 2024). Metode ekstraksi fitur menggunakan Gray Level Co-occurrence Matrix (GLCM) merupakan salah satu pendekatan yang sering digunakan karena kemampuannya mengungkapkan informasi tekstur pada gambar. Fitur seperti energi, kontras, homogenitas, dan korelasi dari GLCM memberikan gambaran mendalam tentang pola tekstur daun yang berhubungan dengan kondisi penyakit (Huda, A. A., Setiaji, B., & Hidayat, F. R, 2022). Setelah proses ekstraksi fitur, algoritma K-Nearest Neighbors (KNN) digunakan untuk klasifikasi. Algoritma ini dipilih karena kesederhanaannya dan kemampuannya untuk bekerja baik pada dataset kecil maupun besar, serta kinerjanya yang konsisten dalam berbagai studi klasifikasi penyakit daun (Sulaiman, J. K., Pongdatu, G. A., Rusman, ., 2023). Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi penyakit pada daun jambu berbasis pengolahan citra menggunakan kombinasi GLCM dan KNN. Sistem ini diharapkan dapat menjadi solusi yang efektif untuk membantu petani dalam mendeteksi penyakit sejak dini, sehingga pengelolaan tanaman dapat dilakukan dengan lebih tepat.

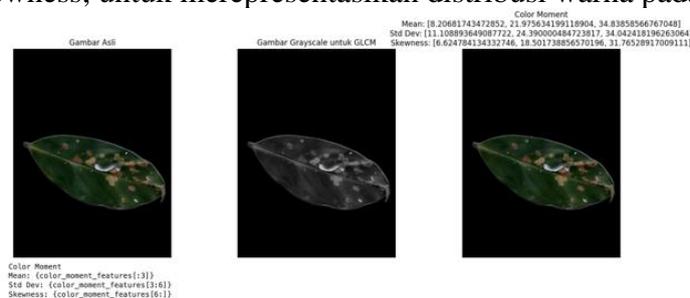
## 2. METODE PENELITIAN

Penelitian ini dilakukan untuk merancang sistem klasifikasi penyakit pada daun jambu menggunakan pengolahan citra digital dengan pendekatan analisis tekstur dan warna. Proses penelitian disusun secara sistematis melalui beberapa tahap utama: pengumpulan data, pengolahan awal citra (preprocessing), ekstraksi fitur, pengembangan model klasifikasi, dan evaluasi performa.



## 3. HASIL DAN PEMBAHASAN

Pengolahan citra digital dilakukan untuk menganalisis penyakit pada daun jambu dengan memanfaatkan fitur-fitur tekstur dan warna. Proses ini diawali dengan ekstraksi fitur tekstur menggunakan metode GLCM (Gray-Level Co-occurrence Matrix), yang mengukur atribut seperti kontras, energi, homogenitas, dan korelasi pada citra grayscale. Selain itu, fitur warna diekstraksi menggunakan Color Moments, yaitu Mean, Standard Deviation, dan Skewness, untuk merepresentasikan distribusi warna pada daun.



Gambar 1.

### Visualisasi Output

1. Gambar Asli: Menunjukkan daun jambu yang memiliki bercak-bercak penyakit, sebagai hasil citra input.
2. Gambar Grayscale: Citra grayscale dihasilkan untuk mendukung ekstraksi fitur tekstur menggunakan metode GLCM.
3. Fitur dan Statistik Warna: Dihitung berdasarkan kanal RGB untuk memberikan deskripsi distribusi warna pada daun, yang mencerminkan kondisi kesehatan atau adanya penyakit.

Bisa dilihat bahwa kombinasi fitur warna (Color Moments) dan tekstur (GLCM) mampu memberikan representasi yang baik terhadap karakteristik penyakit pada daun jambu. Fitur-fitur ini dapat membedakan pola daun yang sehat dan yang terinfeksi dengan bercak penyakit. Algoritma KNN memberikan hasil klasifikasi yang cukup baik karena

sifatnya yang sederhana dan efektif. Namun, kinerjanya dapat ditingkatkan lebih lanjut dengan optimasi parameter nilai k dan pemilihan metrik jarak yang sesuai.

### **Klasifikasi Menggunakan KNN**

Hasil pengujian algoritma K-Nearest Neighbors (KNN) pada dataset, seperti yang ditunjukkan pada table 3. memberikan gambaran kinerja model pada berbagai nilai parameter k (jumlah tetangga terdekat). Dataset uji terdiri dari 45 sampel, dengan akurasi model yang bervariasi dari 88,88% hingga 100% tergantung pada nilai k yang digunakan. Secara rinci, nilai k = 3 menghasilkan akurasi tertinggi sebesar 100%, menunjukkan bahwa model mampu menangkap pola lokal dengan sangat baik pada pengaturan ini. Untuk nilai k kecil lainnya, seperti k = 1 dan k = 5, akurasi tetap tinggi, masing-masing sebesar 97,77% dan 95,55%, meskipun sedikit lebih rendah dibandingkan k = 3.

Tabel 1. Hasil pengujian

	K	Jumlah Citra Uji	Jumlah Benar	Akurasi (%)
0	1	45	44	97,77
1	3	45	45	100
2	5	45	43	95,55
3	7	45	43	95,55
4	9	45	44	97,55
5	11	45	43	95,55
6	13	45	43	95,55
7	15	45	42	93,33
8	17	45	40	88,88

Namun, ketika nilai k meningkat, seperti pada k = 15 dan k = 17, akurasi cenderung menurun, masing-masing menjadi 93,33% dan 88,88%. Penurunan ini dapat disebabkan oleh model yang menjadi terlalu sederhana (underfitting) karena mempertimbangkan terlalu banyak tetangga, sehingga kehilangan sensitivitas terhadap pola lokal. Ketika nilai k terlalu besar, model kurang mampu menangkap detail pada data, sehingga menurunkan akurasi klasifikasi dalam memisahkan kelas yang berdekatan.

Secara keseluruhan, nilai k kecil (seperti 1 hingga 5) memberikan performa yang lebih baik, karena model lebih responsif terhadap pola lokal. Namun, nilai k yang terlalu kecil, seperti k = 1, berisiko menyebabkan overfitting, di mana keputusan klasifikasi hanya bergantung pada satu tetangga terdekat, yang berpotensi terpengaruh oleh noise. Sebaliknya, nilai k besar menghasilkan model yang lebih stabil tetapi dapat mengabaikan informasi penting pada distribusi data lokal.

Hasil evaluasi model K-Nearest Neighbors (KNN) dengan parameter k = 9 menunjukkan performa yang sangat baik dalam mengklasifikasikan data uji. Akurasi keseluruhan model tercatat mencapai 98%, yang berarti 44 dari 45 sampel uji berhasil diklasifikasikan dengan benar. Hasil ini mencerminkan kemampuan model dalam membuat prediksi yang sangat tepat, meskipun terdapat sedikit kesalahan pada satu sampel.

Analisis lebih lanjut melalui laporan klasifikasi memberikan informasi yang lebih rinci mengenai kinerja model pada setiap kelas yang diuji. Model menunjukkan nilai precision dan recall yang sangat baik untuk semua kelas yang diuji seperti pada table 4.

Untuk kelas LB (Label 1), model memperoleh precision 1.00, yang menunjukkan bahwa setiap prediksi untuk kelas ini akurat tanpa adanya kesalahan klasifikasi ke kelas lain. Namun, recall untuk kelas LB sedikit lebih rendah, yaitu 0.95, yang berarti ada satu sampel yang salah diprediksi sebagai kelas LG. Ini menunjukkan bahwa meskipun model sangat tepat dalam memprediksi kelas LB, ada sedikit kekeliruan dalam menangkap seluruh sampel pada kelas ini.

Pada kelas LG (Label 3), model menunjukkan precision yang sangat baik, yaitu 0.93, yang artinya 93% dari seluruh prediksi untuk kelas ini benar. Di sisi lain, recall untuk kelas LG adalah 1.00, menunjukkan bahwa seluruh sampel yang seharusnya berada di kelas LG diprediksi dengan benar, tanpa adanya kesalahan. Hal ini menunjukkan bahwa model mampu menangani kelas LG dengan sangat baik, bahkan dalam situasi dengan ketidakseimbangan kelas yang cukup besar.

Untuk kelas NR (Label 3), model memperoleh hasil sempurna dengan precision, recall, dan F1-score masing-masing sebesar 1.00, yang berarti model berhasil mengklasifikasikan seluruh sampel dalam kelas ini dengan akurat. Ini menunjukkan bahwa model sangat efisien dalam memisahkan kelas NR dari kelas lainnya, tanpa ada kesalahan dalam prediksi.

Tabel 2. Klasifikasi

Kelas	Precision	Recall	f-1 score	Support
LB	1.00	0.95	0.97	19
LG	0.93	1.00	0.96	13
NR	1.00	1.00	1.00	13
Accuracy			0.98	45
Macro avg	0.98	0.98	0.98	45
Weighted avg	0.98	0.98	0.98	45

Average untuk precision, recall, dan F1-score menunjukkan performa yang konsisten dan seimbang di seluruh kelas. Macro average yang sebesar 0.98 menandakan bahwa model memiliki kinerja yang relatif merata di seluruh kelas meskipun ada ketidakseimbangan jumlah sampel antar kelas. Weighted average yang juga mencapai 0.98 menunjukkan bahwa model tidak hanya akurat dalam melakukan prediksi, tetapi juga stabil dalam menangani ketidakseimbangan data.

Hasil confusion matrix mempertegas temuan ini, dengan hanya satu kesalahan klasifikasi yang terjadi, yaitu satu sampel kelas LB yang diprediksi sebagai kelas LG. Ini menunjukkan bahwa model cukup efektif dalam memisahkan kelas-kelas yang ada. Kelas LG dan NR diprediksi dengan sempurna tanpa kesalahan, yang menunjukkan bahwa model mampu menangani kelas-kelas ini dengan baik, meskipun ada variasi dalam data.

Confusion Matrix :

$$\begin{bmatrix} 18 & 1 & 0 \\ 0 & 13 & 0 \\ 0 & 0 & 13 \end{bmatrix}$$

Namun, ada sedikit penurunan kinerja pada kelas LB, di mana satu kesalahan klasifikasi tercatat. Hal ini bisa disebabkan oleh beberapa faktor, seperti kedekatan antara kelas LB dan LG dalam ruang fitur atau adanya noise dalam data yang memengaruhi keputusan klasifikasi model. Meskipun demikian, dengan  $k = 9$ , model menunjukkan akurasi yang cukup tinggi secara keseluruhan, menandakan bahwa model mampu menyeimbangkan kompleksitas dan kemampuan untuk menggeneralisasi data.

Dengan hasil ini, dapat disimpulkan bahwa model KNN dengan  $k = 9$  memberikan kinerja yang sangat baik dalam memisahkan kelas-kelas pada dataset ini. Model ini sangat efektif dalam mengklasifikasikan data dan berpotensi digunakan dalam masalah klasifikasi serupa. Meskipun model ini sudah menunjukkan hasil yang baik, masih ada peluang untuk peningkatan lebih lanjut, seperti eksperimen dengan parameter  $k$  yang lebih bervariasi atau penggunaan teknik pemilihan fitur yang lebih canggih untuk meningkatkan akurasi klasifikasi dan meminimalkan potensi kesalahan pada kelas tertentu.

#### 4. KESIMPULAN

Penelitian ini mengembangkan sistem klasifikasi penyakit pada daun jambu menggunakan pengolahan citra digital dengan analisis tekstur (GLCM) dan warna (Color Moments), serta algoritma K-Nearest Neighbors (KNN). Model ini berhasil mengklasifikasikan tiga kelas penyakit—Leaf Blight (LB), Leaf Gall (LG), dan Normal (NR)—dengan akurasi tinggi, terutama pada nilai  $k = 3$  yang mencapai 100%. Hasil evaluasi menunjukkan precision, recall, dan F1-score yang sangat baik, dengan model mencapai akurasi keseluruhan 98%. Sistem ini efektif untuk mendeteksi penyakit daun jambu secara dini, meskipun masih ada potensi untuk peningkatan melalui optimasi parameter  $k$  dan teknik pemilihan fitur lanjutan.

#### DAFTAR PUSTAKA

- Astiningrum, M., Arhandi, P. P., & Ariditya, N. A. (2020). Identifikasi penyakit pada daun tomat berdasarkan fitur warna dan tekstur. *Jurnal Informatika Polinema*, 6(2), 47-50.
- Hasdimeyra, E. (2023). *Klasifikasi Penyakit Pada Citra Daun Sayuran Kangkung Berdasarkan Fitur Warna Menggunakan Algoritma KNN* (Doctoral dissertation, Universitas Islam Riau).
- Huda, A. A., Setiaji, B., & Hidayat, F. R. (2022). Implementasi Gray Level Co-occurrence Matrix (GLCM) untuk Klasifikasi Penyakit Daun Padi. *Pseudocode Journal*.
- Ilhamy, R. S., & Sanjaya, U. P. (2023). Algoritma K-Nearest Neighbors (KNN) untuk Klasifikasi Citra Buah Pisang dengan Ekstraksi Ciri Gray Level Co-Occurrence. *Jurnal Telematika*, 17(2).
- Letik, M. E. F., & Bisilisin, F. Y. (2024). Klasifikasi Tanaman Herbal Berdasarkan Tekstur Daun Menggunakan Backpropagation Berbasis Citra. *KETIK: Jurnal Informatika*, 1(6), 1-10.
- Manek, P. G., Baso, B., & Meidyani, B. (2022). Identifikasi Tingkat Kematangan Buah Pinang Menggunakan K-Nearest Neighbor Berdasarkan Fitur Tekstur dan Warna. *Jurnal Teknologi Informasi dan Terapan Universitas Timor (JITU)*, 2(2).
- Muzakir, A. (2021). Perangkat Lunak Mobile untuk Mendeteksi Daun pada Tanaman Menggunakan Algoritma K-Nearest Neighbor (K-NN). *Journal of Information Technology Ampera*, 2(1), 10-18.
- Rachmawanto, E., & Hadi, H. (2021). OPTIMASI EKSTRAKSI FITUR PADA KNN DALAM KLASIFIKASI PENYAKIT DAUN JAGUNG. *Dinamik*, 26(2), 58-67.
- Rustanto, D. W. (2024). Identifikasi Penyakit Daun pada Tanaman Padi Menggunakan Ekstraksi Fitur GLCM dan Metode K-Nearest Neighbor (KNN). *JUSTIN (Jurnal Sistem dan Teknologi Informasi)*, 12(1).
- Sulaiman, J. K., Pongdatu, G. A., Rusman, J., & Artikel, I. (2023). Klasifikasi Penyakit pada Tanaman Kopi Arabika Menggunakan Metode K-Nearest Neighbor (KNN) Berbasis Citra. *Indonesian Journal of Computer Science*