

KLASIFIKASI PENYAKIT DAUN TOMAT BERBASIS ALGORITMA K-NEAREST NEIGHBOR

Muslih¹, Andi Danang Krismawan²

Universitas Dian Nuswantoro
Jl. Imam Bonjol 207 Semarang, 50131, Jawa Tengah, Indonesia
1muslih@dsn.dinus.ac.id, 2andi.danang@dsn.dinus.ac.id

ABSTRAK

Penelitian ini mendalami tantangan dalam pemantauan kesehatan tanaman tomat, yang memiliki peran penting dalam industri pertanian dan perekonomian petani di seluruh dunia. Penyakit-penyakit seperti busuk akar dan hawar daun merupakan ancaman serius yang dapat merusak hasil panen dan mengurangi kualitas tanaman tomat. Dalam upaya mengatasi tantangan ini, penelitian ini menggunakan algoritma K-Nearest Neighbor (KNN) sebagai metode analisis klasifikasi penyakit pada daun tomat. Evaluasi dilakukan melalui tiga percobaan dengan variasi nilai K (K=1, 2, dan 3). Hasilnya menunjukkan bahwa meskipun K=1 mencapai akurasi tertinggi, nilai K=3 memberikan keseimbangan yang baik antara akurasi, kompleksitas model, dan ketahanan terhadap overfitting. Dengan akurasi rata-rata sebesar 88%, model KNN dengan nilai K=3 menjadi pilihan yang handal dalam mengidentifikasi penyakit daun tomat dengan tingkat akurasi yang memadai, memungkinkan pemantauan yang cermat terhadap kesehatan tanaman tomat untuk pemenuhan kebutuhan pangan dunia yang berkelanjutan.

Kata Kunci: Klasifikasi Citra, KNN, Penyakit Daun, Tomat.

ABSTRACT

This research delves into the challenges of monitoring the health of tomato plants, which play a crucial role in the global agricultural industry and the livelihoods of farmers worldwide. Serious threats such as root rot and leaf blight diseases pose significant risks, damaging harvest yields, and diminishing the overall quality of tomato crops. In addressing these challenges, the study employs the K-Nearest Neighbor (KNN) algorithm as a method for classifying diseases affecting tomato leaves. The evaluation is conducted through three experiments, varying the values of K (K=1, 2, and 3). The results demonstrate that while K=1 achieves the highest accuracy, a value of K=3 strikes a balance between accuracy, model complexity, and resilience against overfitting. With an average accuracy of 88%, the KNN model with K=3 emerges as a reliable choice for identifying tomato leaf diseases with adequate accuracy, enabling meticulous monitoring of tomato plant health to meet the sustainable global food demand.

Key Word: Image Classification, KNN, Leaves Disease, Tomato.

PENDAHULUAN

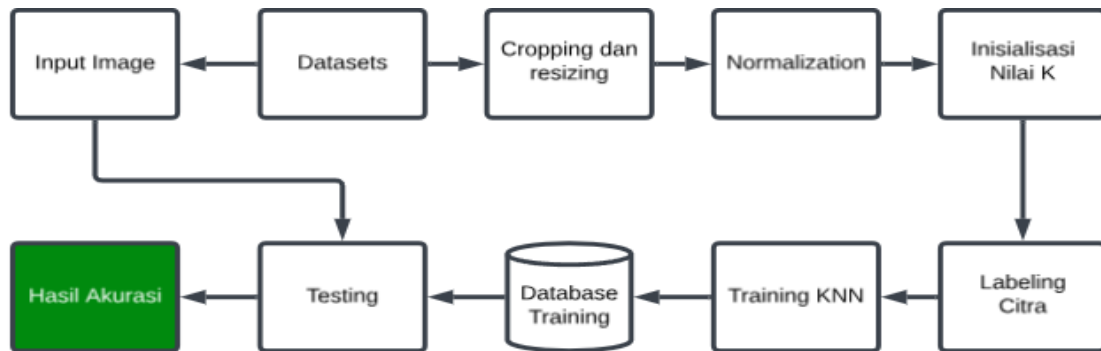
Tomat merupakan salah satu komoditas penting dalam industri pertanian dan menjadi sumber pendapatan utama bagi banyak petani di seluruh dunia (Septiadi & Mundiya, n.d.). Tingkat keberhasilan dalam pertanian tomat memegang peranan penting dalam pemenuhan kebutuhan pangan dunia. Namun, tantangan terus menerus dalam menjaga kesehatan tanaman tomat tetap optimal adalah penyakit yang dapat mengancam hasil panen. Penyakit-penyakit seperti busuk akar, hawar daun, dan berbagai penyakit daun lainnya dapat secara signifikan merusak tanaman tomat, mengurangi kualitas dan kuantitas hasil panen, serta berdampak negatif pada ekonomi petani (Huda et al., 2022). Oleh karena itu, pemantauan yang cermat terhadap kesehatan tanaman tomat menjadi sangat penting.

Untuk meningkatkan efisiensi dalam pemantauan dan identifikasi penyakit pada daun tomat, penggunaan teknologi informatika dan algoritma kecerdasan buatan menjadi suatu solusi yang menjanjikan. Dalam konteks ini, pengembangan sebuah aplikasi klasifikasi yang dapat membantu petani dalam memantau kesehatan daun tomat menjadi suatu langkah yang sangat relevan. Aplikasi ini akan memungkinkan petani untuk dengan cepat dan akurat mengidentifikasi penyakit pada daun tomat, sehingga tindakan pengobatan dapat diambil lebih dini, mengurangi kerugian panen, dan mengoptimalkan produksi pertanian. Salah satu algoritma yang potensial untuk digunakan dalam analisis klasifikasi penyakit pada daun tomat adalah algoritma K-Nearest Neighbor (KNN) (Silvia et al., 2023).

Algoritma ini adalah metode yang berbasis pada pembelajaran mesin yang dapat digunakan untuk mengklasifikasikan data berdasarkan pola kemiripan dengan data pelatihan yang ada (Rangga et al., 2019). Dengan menerapkan algoritma KNN, aplikasi ini akan mampu memproses data gambar daun tomat dan membandingkannya dengan database penyakit yang ada untuk mengidentifikasi penyakit dengan tingkat

akurasi yang tinggi (Putra et al., 2022; Rangga et al., 2019). Oleh karena itu, penelitian ini bertujuan untuk mengembangkan aplikasi desktop berbasis algoritma K-Nearest Neighbor untuk membantu pemantauan dan klasifikasi penyakit pada daun tomat dalam upaya meningkatkan produktivitas dan keberhasilan pertanian tomat.

METODE PENELITIAN



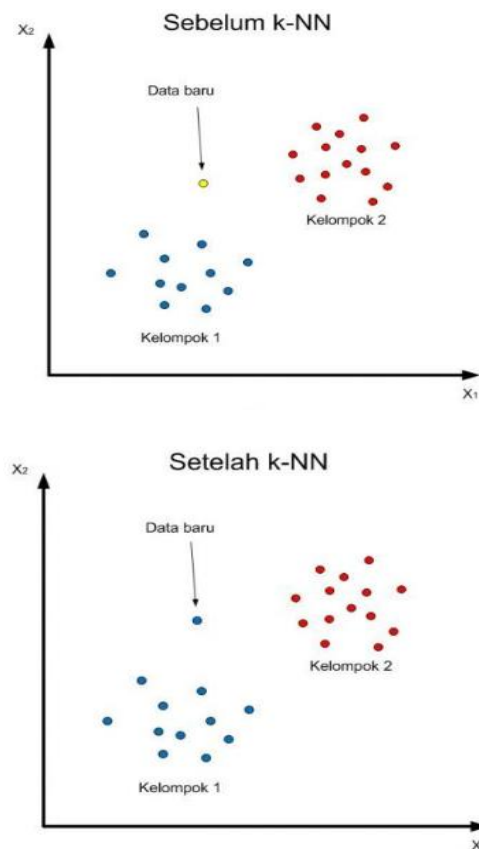
Gambar 1. Alur metode penelitian

Dalam metodologi penelitian ini, untuk mencapai hasil yang akurat dalam klasifikasi citra. Pertama-tama, mengumpulkan dataset citra yang akan digunakan sebagai input untuk penelitian ini. Selanjutnya, dilakukan proses cropping dan resizing terhadap citra-citra tersebut agar memiliki ukuran yang konsisten. Setelah itu, dilakukan normalisasi terhadap data untuk memastikan bahwa nilai-nilai piksel berada dalam rentang yang standar. Kemudian, dilakukan inisialisasi nilai K untuk algoritma K-Nearest Neighbors (KNN) yang akan diterapkan pada penelitian ini. Setelah proses pelatihan selesai, dilakukan pengujian performa model menggunakan dataset uji terpisah dan mencatat hasil akurasi untuk evaluasi. Berdasarkan metode yang diusulkan, alur secara visual dapat dilihat pada gambar 1.

K-Nearest Neighbor (KNN)

KNN adalah teknik pembelajaran mesin yang sangat relevan dan efektif (Rofiq et al., n.d.). Dalam konteks ini, KNN bekerja dengan cara mengukur jarak antara sampel data yang belum diketahui (citra penyakit daun tomat) dengan tetangga terdekatnya yang sudah diketahui kelasnya (penyakit daun tomat yang telah diidentifikasi sebelumnya). Berdasarkan jarak tersebut, model KNN memprediksi kelas penyakit daun tomat dari sampel yang belum diketahui dengan melihat mayoritas kelas dari

tetangga terdekatnya (Rachmawanto et al., 2020).

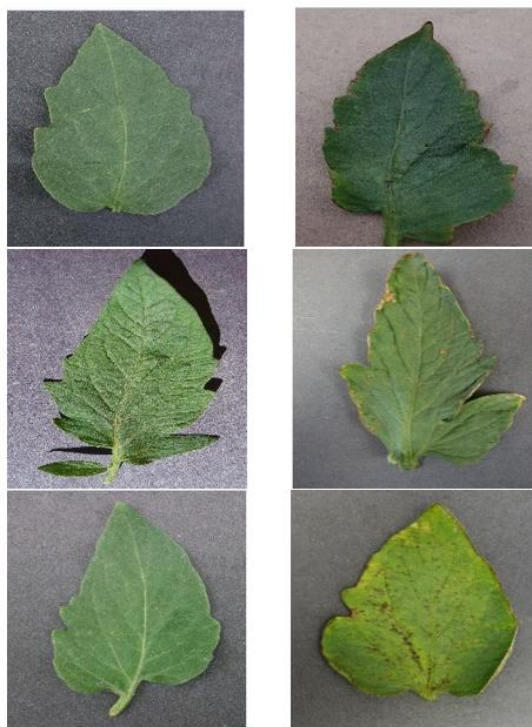


Gambar 2. Ilustrasi penggunaan KNN

Proses ini memungkinkan KNN untuk mengidentifikasi dan mengklasifikasikan penyakit daun tomat dengan membandingkan

pola-pola visual pada daun yang terinfeksi dengan data yang telah ada dalam dataset pelatihan. Dengan menggunakan algoritma ini, para peneliti dapat mengembangkan sistem yang otomatis dan cepat dalam mengenali jenis penyakit daun tomat, membantu para petani mengambil tindakan pencegahan atau pengobatan yang sesuai untuk menjaga kesehatan tanaman tomat mereka. Berdasarkan ilustrasi dari algoritma KNN, dapat dilihat pada Gambar 2.

Pengumpulan Data



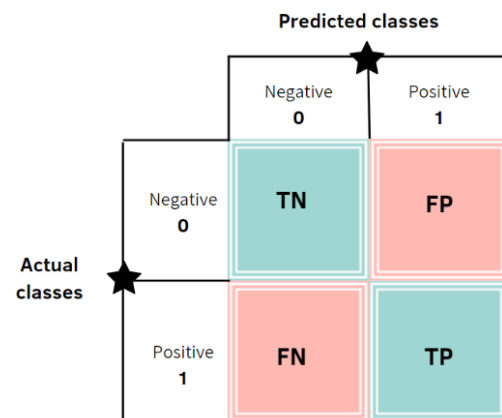
(a) Sehat (b) Infeksi
 Gambar 3. Sample Citra Datasets

Dataset yang digunakan dalam penelitian ini telah melalui serangkaian proses pengolahan data yang meliputi *cropping*, *resizing*, dan normalisasi. Proses *cropping* dilakukan untuk memotong citra-citra tersebut sehingga fokus pada area daun tomat yang penting untuk analisis penyakit. Selanjutnya, citra-citra tersebut diubah ukurannya (*resizing*) agar memiliki dimensi yang konsisten, memudahkan proses perbandingan dan analisis.

Evaluasi Dengan Confusion Matrix

Dalam evaluasi kinerja model klasifikasi penyakit daun tomat, kami menggunakan matriks kebingungan (*confusion matrix*) untuk mengukur sejauh mana model kami

dapat mengklasifikasikan penyakit dengan akurat (Darwis, 2023; Kim & Cho, 2022). Matriks ini memberikan gambaran yang jelas tentang jumlah prediksi yang benar (*true positive*) dan prediksi yang salah (*false positive*) untuk setiap kelas penyakit. Dengan memperhatikan nilai-nilai diagonal utama matriks kebingungan, kami dapat menilai tingkat akurasi dari model kami. Selain itu, melalui matriks ini, kami juga dapat mengidentifikasi tingkat kesalahan yang terjadi, seperti kesalahan dalam mengklasifikasikan penyakit tertentu sebagai penyakit lain (*false negative*).



Gambar 4. Evaluasi confusion matrix

Berdasarkan hasil dari tahap evaluasi confusion matrix, Rumus/Formula berbasis confusion matrix dapat dilihat pada tabel dibawah ini.

Tabel 1. Rumus confusion matrix

Rumus Confusion Matrix	
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1 - Score	$\frac{2 \times Precision \times Recall}{Precision + Recall}$

HASIL DAN PEMBAHASAN

Dalam hasil dan pembahasan penelitian ini, kami melakukan tiga percobaan yang berbeda dengan variasi nilai K pada algoritma K-Nearest Neighbors (KNN). Percobaan ini dilakukan dengan memilih nilai K sebesar 1, 2, dan 3. Hasil dari ketiga percobaan tersebut menunjukkan variasi yang signifikan dalam kinerja model klasifikasi pada penelitian ini.

Tabel 2. Akurasi berdasarkan nilai K

Matrik Evaluasi	K = 1	K = 2	K = 3	Average
Accuracy	96%	90%	82%	88%
Precision	99%	92%	90%	94%
Recall	100%	100%	100%	100%
F1-Score	99.5%	96%	95%	97%

Tabel 3. Testing data

No	Testing	Sample Image	Hasil
1.	K=1	Segar 1.jpg	TP
		Infeksi 1.jpg	TP
2.	K=2	Segar 1.jpg	TP
		Infeksi 1.jpg	TP
3.	K=3	Segar 1.jpg	TP
		Infeksi 1.jpg	TN
4.	K=1	Segar 2.jpg	TP
		Infeksi 2.jpg	TP
5.	K=2	Segar 2.jpg	TP
		Infeksi 2.jpg	TP
6.	K=3	Segar 2.jpg	TP
		Infeksi 2.jpg	TN
7.	K=1	Segar 3.jpg	TP
		Infeksi 3.jpg	TP
8.	K=2	Segar 3.jpg	TP
		Infeksi 3.jpg	TP
9.	K=3	Segar 3.jpg	TP
		Infeksi 3.jpg	TP

Hasil dari 9 percobaan dari Tabel 3 menunjukkan bahwa algoritma KNN dalam klasifikasi ini cukup bagus untuk digunakan, dimana dari keseluruhan percobaan terdapat 2 hasil yang menunjukkan TN (Kesalahan Identifikasi).

SIMPULAN DAN SARAN

Dari hasil evaluasi yang dilakukan pada tiga percobaan dengan nilai K yang berbeda, dapat disimpulkan bahwa performa terbaik tercapai pada nilai K=1 dengan akurasi sebesar 96%, presisi 99%, recall 100%, dan F1-Score 99.5%. Namun, perlu diperhatikan bahwa nilai K yang terlalu kecil (K=1) cenderung membuat model terlalu sensitif terhadap noise dalam data dan dapat menyebabkan overfitting. Oleh karena itu, untuk mencari keseimbangan antara kompleksitas dan akurasi, nilai K=3 dapat dianggap sebagai pilihan yang lebih baik dalam konteks ini, meskipun dengan tingkat akurasi yang sedikit lebih rendah (82%). Namun, penting untuk mempertimbangkan aspek-aspek lain seperti kecepatan komputasi dan penggunaan memori saat memilih nilai K yang tepat. Dengan rata-rata akurasi sebesar 88%, model dengan nilai K=3 memberikan keseimbangan yang baik antara akurasi dan kompleksitas, sehingga merupakan pilihan yang dapat diandalkan

dalam mengklasifikasikan penyakit daun tomat.

UCAPAN TERIMAKASIH

Terima kasih kepada Universitas Dian Nuswanotero atas dana penelitian internal tahun anggaran 2023.

DAFTAR PUSTAKA

- Darwis, H. (2023). K-Nearest Neighbor dan Convolutional Neural Network pada Klasifikasi Penyakit Tanaman Bawang Merah K-Nearest Neighbor And Convolutional Neural Network in Shallot Plant Diseases Classification. In *Agustus* (Vol. 22, Issue 3).
- Huda, A. A., Setiaji, B., & Hidayat, F. R. (2022). Implementasi Gray Level Co-Occurrence Matrix (GlcM) Untuk Klasifikasi Penyakit Daun Padi. In *Jurnal Pseudocode* (Vol. 9). www.ejournal.unib.ac.id/index.php/pseudocode
- Kim, S. C., & Cho, Y. S. (2022). Predictive System Implementation to Improve the Accuracy of Urine Self-Diagnosis with Smartphones: Application of a Confusion Matrix-Based Learning Model through RGB Semiquantitative Analysis. *Sensors*, 22(14). <https://doi.org/10.3390/s22145445>
- Putra, P., H Pardede, A. M., & Syahputra, S. (2022). Analisis Metode K-Nearest Neighbour (Knn) Dalam Klasifikasi Data Iris Bunga. *Jurnal Teknik Informatika Kaputama (JTIK)*, 6(1).
- Rangga, M., Nasution, A., & Hayaty, M. (2019). Perbandingan Akurasi dan Waktu Proses Algoritma K-NN dan SVM dalam Analisis Sentimen Twitter. *JURNAL INFORMATIKA*, 6(2), 212–218. <http://ejournal.bsi.ac.id/ejournal/index.php/ji>
- Rofiq, M., Huda, K., & Nafi'iyah, N. (n.d.). Identifikasi Penyakit Daun Kentang Berdasarkan Fitur Warna, Tekstur, dan Bentuk dengan SVM dan KNN Identification of Potato Leaf Disease Based on Color, Texture, and Shape Features with SVM and KNN. In *Seminar Nasional Multimedia & Artificial Intelligence SMAI* (Vol. 2020).

Septiadi, D., & Mundiya, A. I. (n.d.).
*Karakteristik Dan Analisis Finansial
Usahatani Tomat Di Kabupaten
Lombok Timur Characteristics And
Financial Analysis Of Tomato Farming
In East Lombok District* (Vol. 31, Issue
3). Desember.

Silvia, N., Cahyani, T., Hidayat, N., &
Santoso, E. (2023). *Klasifikasi Penyakit
Tanaman Kacang Tanah menggunakan
Metode MKNN (Modified K-Nearest
Neighbor)* (Vol. 7, Issue 3). [http://j-
ptiik.ub.ac.id](http://j-ptiik.ub.ac.id)