

Implementasi Metode *K-Nearest Neighbor* Untuk Klasifikasi Penyakit Tanaman Mentimun Pada Citra Daun

Implementation Of The K-Nearest Neighbor Method For Classification Of Cucumber Diseases On Leaf Images

Ratna Indah Juwita Harahap^{*1}, Sumi Khairani², Rismayanti³ (*correspondent author)

^{1,2,3} Prodi Teknik Informatika, Universitas Harapan Medan

E-mail: ¹ratnaahr@gmail.com, ²sumibintisyailfullah@gmail.com, ³risma.stth@gmail.com

Abstrak

Mentimun merupakan salah satu sayuran yang banyak dikonsumsi oleh masyarakat Indonesia. Akan tetapi, tanaman mentimun rentan terhadap serangan penyakit yang menyebabkan kehilangan hasil secara substansial. Contoh penyakit pada tanaman mentimun adalah embun bulu, embun tepung dan cmv (cucumber mosaic virus). Penyakit ini dapat dikenali secara visual karena memiliki ciri warna dan tekstur. Melalui sebuah citra dapat dipelajari informasi mengenai penyakit tanaman mentimun tersebut. Penelitian ini bertujuan untuk membangun sistem klasifikasi penyakit pada citra daun mentimun sehingga dapat memberikan informasi jenis penyakit. Penerapan sistem terdiri dari tahap pre-processing, ekstraksi ciri, klasifikasi dan evaluasi. Tahap pre-processing melakukan resize citra RGB lalu dikonversi ke grayscale. Tahap ekstraksi ciri dengan metode GLCM (Gray Level Co-Occurrence Matrix). Tahap klasifikasi menggunakan algoritma K-NN (K-Nearest Neighbor). Tahap evaluasi dengan confusion matrix. Hasil pengujian klasifikasi penyakit daun mentimun menggunakan algoritma K-Nearest Neighbor menghasilkan nilai akurasi terbaik dengan penggunaan nilai ketetanggaan $k = 1$ mencapai 90%.

Kata kunci : Daun mentimun, pre-processing, GLCM, KNN, Confusion Matrix

Abstract

Cucumber is a vegetable that is widely consumed by Indonesian people. However, cucumber plants are susceptible to disease attack which causes substantial yield loss. Examples of disease in cucumber plants are downy mildew, powdery mildew, and cucumber mosaic virus. This disease can be recognized visually because it has a characteristic color and texture. Through an image, information can be learned about the cucumber plant disease. This study aims to build a disease classification system on cucumber leaf images so that it can provide information on the type of disease. The application of the system consisting of pre-processing, feature extraction, classification, and evaluation stages. The pre-processing stages resizes the RGB image and then converts it to Grayscale. The feature extraction stage uses the GLCM (Gray Level Co-Occurrence) method. The classification stage uses the K-NN (K-Nearest Neighbor) algorithm. Evaluation stage is a confusion matrix. The results of the cucumber leaf disease classification test used the K-Nearest Neighbor algorithm, produced the best accuracy value by using the neighborhood value $k=1$ reaching 90%.

Keywords : Cucumber leaf, Pre-processing, GLCM, KNN, Confusion Matrix

1. PENDAHULUAN

Mentimun termasuk salah satu jenis sayuran yang memiliki banyak manfaat dalam kehidupan masyarakat sehari-hari, sehingga permintaan terhadap komoditi ini cukup besar. Buah ini disukai oleh seluruh golongan masyarakat, sehingga buah mentimun dibutuhkan dalam jumlah relatif besar dan berkesinambungan. Kebutuhan buah mentimun cenderung terus meningkat sejalan dengan pertambahan penduduk, peningkatan taraf hidup, tingkat pendidikan, dan kesadaran masyarakat tentang pentingnya nilai gizi [1]. Salah satu masalah rendahnya kualitas mentimun adalah penyakit yang menyerang daunnya. Daun yang berfungsi untuk proses fotosintesis jika terserang penyakit mengakibatkan ketidakmampuan memberikan makanan pada bagian tanaman. Pertumbuhan tanaman menjadi terhambat dan pada akhirnya tanaman menghasilkan buah dengan kualitas rendah dan dapat mengakibatkan tanaman layu atau mati.

Oleh karena itu penelitian ini membangun aplikasi klasifikasi penyakit tanaman mentimun pada citra daun sehingga dapat memberikan informasi jenis penyakit. Dataset citra daun mentimun didapatkan dari *Kaggle* dengan judul "*Cucumber plant disease dataset*" [2]. Identifikasi yang dilakukan berdasarkan fitur warna dan tekstur daun. Klasifikasi dilakukan berdasarkan 4 kelas yaitu daun normal, embun bulu, embun tepung dan cmv. Ukuran citra diubah kedalam 512x512 piksel. Manfaat dari penelitian ini adalah mengidentifikasi penyakit pada citra daun mentimun secara otomatis dan sebagai referensi penelitian klasifikasi penyakit pada citra daun mentimun dengan metode K-NN. Beberapa penelitian sejenis yang pernah dilakukan adalah sebagai berikut :

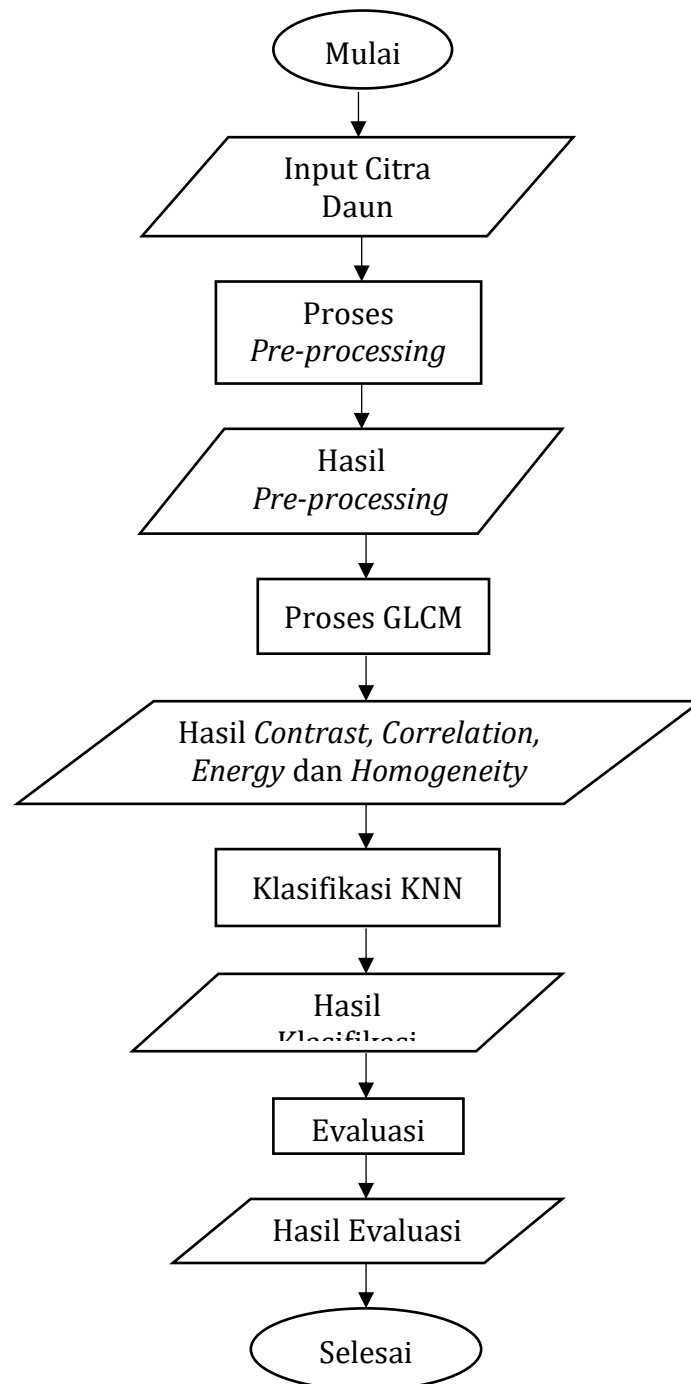
1. Penelitian yang berjudul "*The Grading of agarwood oil quality using K-Nearest Neighbor (K-NN)*". Bertujuan untuk mengklasifikasikan kualitas minyak kayu gaharu menjadi kelas, yaitu kelas satu untuk kualitas buruk dan kelas dua untuk kualitas baik. Akurasi yang dihasilkan mencapai 83% [3].
2. Penelitian yang berjudul "Identifikasi Penyakit pada Daun Tomat Berdasarkan Fitur Warna dan Tekstur". Bertujuan untuk merancang sistem identifikasi penyakit pada daun tomat dengan menggunakan RGB sebagai fitur ekstraksi warna dan GLCM sebagai fitur ekstraksi tekstur. Hasil dari sistem identifikasi penyakit pada daun tomat pada nilai $k=1$ menggunakan metode K-NN adalah penentuan label sehat, bercak bakteri, dan busuk daun. Tingkat akurasi penelitian ini adalah 92,89% [4].

Berdasarkan uraian beberapa penelitian tersebut, maka penelitian ini menggunakan beberapa metode penyelesaian yaitu fitur warna menggunakan citra RGB ke *Grayscale* dan ekstraksi tekstur menggunakan *Gray Level Co-occurrence matrix* (GLCM). Penelitian ini menggunakan metode K-NN untuk tahap klasifikasi.

2. METODOLOGI PENELITIAN

Metode yang digunakan penelitian terdiri dari tahapan *pre-processing*, ekstraksi ciri, klasifikasi dan evaluasi. Sistem klasifikasi dirancang melalui aplikasi MATLAB R2015a melalui menu *user interface* (GUI).

Flowchart adalah gambar dengan simbol bagan tertentu yang mendeskripsikan urutan dari sebuah proses secara terperinci dan hubungan antar proses. Berikut tahapan *Flowchart* yang dapat dilihat pada Gambar 1.



Gambar 1. *Flowchart Sistem*

2.1. *Pre-processing*

Pre-processing adalah salah satu tahapan yang digunakan untuk menghilangkan *noise* pada data citra yang akan digunakan. Pada tahap *pre-processing* ini terdiri dari *resize*, *grayscale* dan *thresholding*. Proses ini menghasilkan data citra dengan warna biner yaitu hitam (0) dan putih (1) [5]. Proses awal dilakukannya perbaikan suatu citra untuk menghilangkan *noise* atau bagian-bagian

yang tidak diperlukan pada gambar input untuk proses selanjutnya. Pada tahap ini dikenal proses meningkatkan kualitas citra yang bertujuan untuk meningkatkan keberhasilan pada tahap pengolahan citra digital.

2.1.1 *Resize*

Resize adalah proses mengubah ukuran dari data citra menjadi lebih kecil maupun lebih besar. Pada proses memperkecil ukuran citra, dilakukan dengan cara mengurangi jumlah *pixel* yang terdapat pada citra. *Resize* dilakukan untuk mengurangi kinerja dari komputer sehingga proses berjalan lebih cepat dari semula, kemudian *resize* juga memakan lebih sedikit ruang memori [6]. *Resize* merupakan proses mengubah ukuran besar citra dalam satuan piksel. Tahapan *resize* pada normalisasi dilakukan dengan tujuan menyesuaikan ukuran citra latih dan citra uji. Perubahan ukuran citra dapat menghasilkan citra yang lebih besar maupun lebih kecil dari citra asli.

2.1.2 *Grayscale*

Grayscale merupakan konversi antara citra berwarna ke dalam citra berskala keabuan. *Grayscale image* direpresentasikan dalam bentuk array dua dimensi. Setiap bagian dalam array memperlihatkan intensitas (*greylevel*) dari *image* pada posisi yang bersesuaian.

Warna abu-abu merupakan hasil kombinasi antara warna merah, hijau dan biru dengan nilai intensitas yang sama dalam ruang RGB [7]. Metode yang biasanya digunakan terdapat pada persamaan berikut ini :

$$R + G + B/3$$

Keterangan :

R : Unsur warna merah

G : Unsur warna hijau

B : Unsur warna biru

2.2. Ekstraksi Ciri

Tahap kedua dalam penelitian ini yaitu melakukan ekstraksi ciri menggunakan *Grey Level Co-occurrence Matrix* (GLCM). *Grey Level Co-occurrence Matrix* (GLCM) adalah matriks yang mempresentasikan hubungan ketetanggaan antarpiksel dalam citra pada berbagai arah orientasi 0 dan jarak spasial d . Pengukuran tekstur pada orde pertama menggunakan perhitungan statistik didasarkan pada nilai piksel citra asli semata, seperti varians dan tidak memperhatikan hubungan ketetanggaan piksel. Pada orde kedua, hubungan antar pasangan dua piksel citra asli diperhitungkan [8].

Fitur ekstraksi GLCM pada penelitian ini hanya mengambil 4 fitur yaitu *correlation*, *energy*, *homogeneity* dan *contrast*.

a. *Correlation*

$$\text{Correlation} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{\sigma_i \sigma_j} \quad (1)$$

Atau untuk sudut ganda $\mu_i = \mu_j = \mu$ dan $\sigma_i = \sigma_j = \sigma$ maka :

$$co = \frac{1}{\sigma^2} ((\sum_{i=0}^{G-1} \sum_{j=0}^{G-1} (i * j)P(i, j)) - \mu^2) \quad (2)$$

b. Energy

$$\text{Energy} = \sum_{i,j} P(i, j)^2 \quad (3)$$

c. Homogeneity

$$\text{Homogeneity} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{P(i, j)}{1 + (i - j)^2} \quad (4)$$

Atau

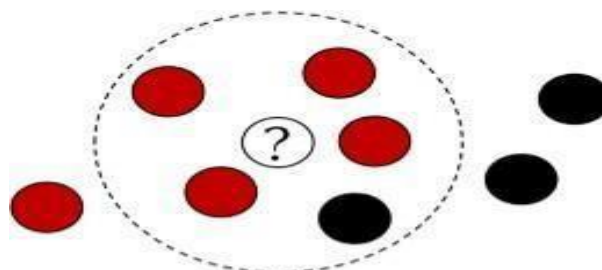
$$\text{Homogeneity} = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \frac{P(i, j)}{1 + |i - j|} \quad (5)$$

d. Contrast

$$\text{Contrast} = \sum_{i,j} (i - j^2) P_{i,j} \quad (6)$$

2.3. Classification

Klasifikasi dalam penelitian ini menggunakan *K-Nearest Neighbor* (K-NN). *K-Nearest Neighbor* (K-NN) adalah algoritma klasifikasi yang berdasarkan K tetangga terdekat akan mencari nilai K data latih yang jaraknya paling dekat dengan data uji dikalkulasi dengan menghitung nilai jarak antar titik yang mempresentasikan data latih menggunakan rumus jarak *Euclidean* [9].



Gambar 2. Ilustrasi Klasifikasi K-NN

Adapun rumus jarak *Euclidean Distance* terdapat dalam persamaan berikut.

$$d_{x,y} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (7)$$

Berikut langkah-langkah perhitungan metode KNN adalah sebagai berikut :

- Langkah 1 : Menentukan parameter K
- Langkah 2 : Menghitung jarak antara data uji dengan data latih menggunakan rumus *Euclidean Distance*
- Langkah 3 : Mengurutkan hasil jarak tersebut secara ascending dan menetapkan tetangga terdekat berdasarkan jarak minimum ke-K
- Langkah 4 : Dengan menggunakan klasifikasi K-Nearest Neighbor yang paling mayoritas, maka dapat diprediksikan klasifikasi dari data uji.

2.4. Evaluasi

Pada penelitian ini akan dilakukan evaluasi menggunakan metode *Confusion Matrix*. *Confusion Matrix* adalah matriks untuk mengukur kinerja suatu klasifikasi yang memiliki 4 variabel yaitu TP (True Positive) berarti ketika sistem memprediksi positif dan hasilnya benar, TN (True Negative) berarti ketika sistem memprediksi negatif dan hasilnya benar, FP (False Positive) berarti ketika sistem memprediksi positif dan hasilnya salah, FN (False Negative) berarti ketika sistem memprediksi negatif dan hasilnya salah [10]. Berikut salah satu perhitungan yang digunakan sebagai pengujian pada sistem adalah sebagai berikut.

$$akurasi = \frac{\sum \text{data yang tepat diklasifikasi}}{\sum \text{data yang digunakan}} \times 100\% \quad (8)$$

Berikut metode *confusion matrix* dapat dilihat pada Tabel 1 berikut.

Tabel 1. Confusion Matrix

		True Values	
		True	False
Prediction	True	TP	FP
		Correct Result	Unexpected result
	False	FN	TN
		Missing result	Correct absence of result

2.5 Penyakit Mentimun Pada Citra Daun

2.5.1 Embun Bulu (*Downy Mildew*)

Embun bulu merupakan penyakit penting pada daun tanaman mentimun. Penyakit embun bulu merupakan penyakit yang sangat merusak pada tanaman cucurbitaceae. Embun bulu disebabkan oleh *Pseudoperonospora cubensis*. *P. cubensis*

termasuk kingdom Straminipila, filum Oomycota, kelas Oomycetes, ordo Peronosporales, family peronosporaceae dan genus Pseudoperonospora [11].

Gejala pertama biasanya terlihat berwarna hijau keputihan pada bagian atas permukaan daun. Gejala lanjutan akan membuat daun menjadi berwarna kuning dan bentuknya menjadi tidak rata dan dibatasi oleh tulang daun [12], kemudian dapat membuat daun tanaman kering, gugur hingga mati [13]. Berikut citra daun mentimun yang terkena penyakit embun bulu dapat dilihat pada Gambar 3.



Gambar 3. Citra daun Embun Bulu

2.5.2 Embun Tepung (*Powdery Mildew*)

Embun tepung (*Powdery Mildew*) disebabkan oleh jamur *Erysiphe cichoracearum* sering menyerang tanaman labu-labuan, bahkan kadang-kadang cukup berat, khususnya pada mentimun (*Cucumis sativus* L), labu (*Cucurbita moschata*) dan labu air (*Lagenaria leucantha*). Disamping terdapat pada labu-labuan, penyakit tepung juga terdapat pada berbagai tanaman seperti anggur, apel, mangga, pepaya, tembakau bahkan tanaman hias [14].

Gejala yang ditunjukkan oleh *E. cichoracearum* adalah permukaan daun dan batang muda ditutupi tepung putih dan kemudian berubah menjadi kuning dan mengering. Lapisan putih bertepung tersebut merupakan miselium, konidiofor dan konidium jamur penyebab penyakit. Berikut citra daun mentimun yang terkena penyakit embun tepung dapat dilihat pada Gambar 4.



Gambar 4. Citra daun Embun Tepung

2.5.3 CMV (*Cucumber Mozaic Virus*)

CMV merupakan virus mozaik yang memiliki inang lebih dari 1200 spesies tanaman dalam 100 famili monokotil dan dikotil, termasuk tanaman hortikultura dan tanaman perkebunan. CMV merupakan virus yang sangat penting pada tanaman mentimun, karena selalu terdapat di antara virus yang lainnya, dan mengakibatkan kerugian yang cukup besar. Penurunan produksi akibat virus mozaik ini dapat dengan cepat tersebar ke pertanaman di sekitar sumber virus sesuai dengan aktivitas kutu daun (aphids) yang berfungsi sebagai vektornya. Sampai saat ini beberapa usaha yang dilakukan untuk pengendalian CMV pada tanaman mentimun belum memberikan hasil seperti yang diharapkan.

Pertumbuhan tanaman yang diserang virus realtif lebih kerdil. Mula-mula tulang daun menguning atau terjadi jalur kuning sepanjang tulang daun. Daun menjadi belang hijau tua dan hijau muda, ukuran daun lebih kecil dan lebih sempit dari ukuran daun yang normal, atau menjadi seperti tali sepatu karena lembaran daun menghilang yang tinggal hanya tulang saja. Virus mozaik sering menyebabkan gejala bisul atau kutil pada buah [15]. Berikut citra daun mentimun yang terkena penyakit cmv dapat dilihat pada Gambar 5.

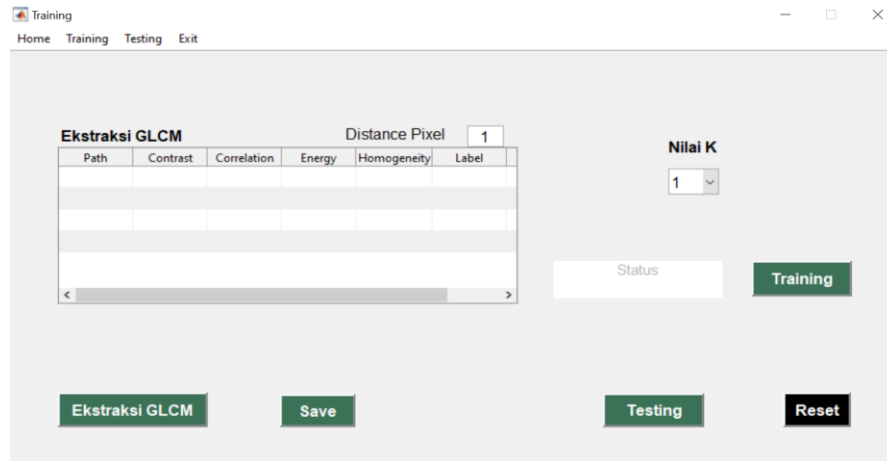


Gambar 5. Citra daun CMV

3 HASIL DAN PEMBAHASAN

3.1. Tampilan Pelatihan

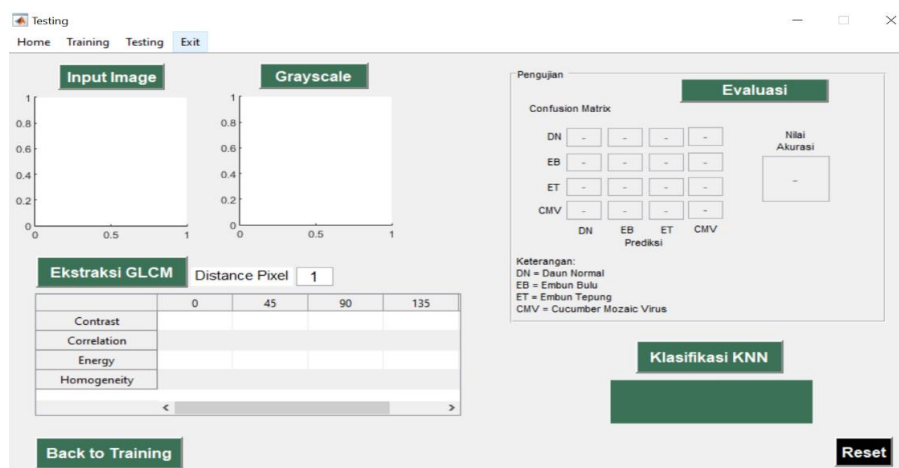
Tampilan Pelatihan (*Training*) adalah halaman yang berisikan informasi tentang ekstraksi *Grey Level Co-occurrence Matrix* (GLCM), nilai K yang akan menghasilkan nilai data. Informasi data terdiri dari 140 citra daun yang terdiri dari 35 citra pada setiap jenis penyakit daun mentimun yaitu Daun Normal, *Embun Bulu* (*Downy Mildew*), *Embun Tepung* (*Powdery Mildew*) dan CMV (*Cucumber Mozaic Virus*). Pada halaman ini terdapat tombol untuk memilih citra latih dan melakukan pelatihan yang hasilnya akan disimpan ke dalam *database* sistem. Adapun tampilan pelatihan dapat dilihat pada Gambar 6.



Gambar 6. Tampilan Pelatihan

3.2. Tampilan Pengujian

Tampilan Pengujian adalah halaman yang berisikan fitur untuk melakukan klasifikasi yang dimulai dari memilih citra daun yang akan diuji sampai dengan menampilkan hasil klasifikasi KNN dan nilai akurasi penyakit daun mentimun. Dari tampilan halaman *testing* terdapat banyak tombol yang digunakan untuk mendukung pengguna dalam melakukan pengujian. Tombol pertama yaitu “*Input Image*” yang berguna untuk membuka dialog explorer untuk memilih citra yang akan diuji pada sistem. Tombol kedua yaitu “*Grayscale*” untuk *resize* citra dan mengubah kedalam bentuk *grayscale*. Tombol ketiga yaitu “*Ekstraksi GLCM*” yang berguna untuk menampilkan nilai dari ekstraksi ciri. Tombol keempat yaitu “*Hasil Klasifikasi*” dimana akan menghitung jarak terdekat dengan metode *K-Nearest Neighbor* yang akan menampilkan jenis penyakit daun mentimun. Kemudian tombol kelima yaitu “*Evaluasi*” yang dapat memberikan nilai akurasi pada setiap citra yang di uji pada sistem. Adapun tampilan pengujian dapat dilihat pada Gambar 7 berikut.



Gambar 7. Tampilan Pengujian

3.3. Pembahasan

Pada penelitian ini digunakan 160 citra dari 4 jenis citra daun mentimun yaitu 3 penyakit daun yaitu embun bulu, embun tepung, cmv dan 1 daun normal. Dimana masing-masing jenis citra daun mentimun terdiri dari 140 citra sebagai data latih dan 20 citra sebagai data uji. Berikut pembagian data terdapat dalam Tabel 2.

Tabel 2. Pembagian Data

Klasifikasi Daun	Jumlah Dataset	Dataset Latih	Dataset Uji
Daun Normal	40	35	5
Embun Bulu	40	35	5
Embun Tepung	40	35	5
CMV	40	35	5

Pengujian ini menentukan akurasi dari hasil klasifikasi dengan menggunakan nilai parameter $k = 1, 3$ dan 5 . Berikut hasil pengujian klasifikasi dapat dilihat pada Tabel 3.

Tabel 3. Hasil Pengujian Klasifikasi

Nilai k	Kelas	Jenis Penyakit	Hasil		Akurasi
			Benar	Salah	
K = 1	1	Daun Normal	5	0	90%
	2	Embun Bulu	5	0	
	3	Embun Tepung	5	0	
	4	CMV	3	2	
K = 3	1	Daun Normal	4	1	75%
	2	Embun Bulu	4	1	
	3	Embun Tepung	4	1	
	4	CMV	3	2	
K = 5	1	Daun Normal	5	0	75%
	2	Embun Bulu	3	2	
	3	Embun Tepung	4	1	
	4	CMV	3	2	

4. KESIMPULAN

Berdasarkan hasil dari pengujian yang telah dilakukan maka dapat disimpulkan bahwa implementasi metode K-Nearest Neighbor dengan ekstraksi fitur warna dan tekstur daun dapat mengklasifikasi jenis penyakit tanaman mentimun pada citra daun dengan nilai $k=1$ sehingga menghasilkan nilai akurasi terbaik mencapai 90% menggunakan metode *confusion matrix*.

DAFTAR PUSTAKA

- [1] Cahyono B. "Teknik dan Strategi Budidaya Mentimun", Yogyakarta, 2003.
- [2] Cucumber plant diseases dataset | Kaggle, 2022.
- [3] N. Ismail, M. Rahiman, M. T. I. C. "The grading of agarwood oil quality using k-Nearest Neighbor (k-NN)," 2013.
- [4] M. Astiningrum, P. A.-S. I., "Identifikasi Penyakit pada Daun Tomat Berdasarkan Fitur Warna dan Tekstur," 2019.
- [5] A. Sonita and K. Khairunnisyah, "Aplikasi Pendeteksi Obat dan Makanan Menggunakan OCR (Optical Character Recognition)," *J. Inform. Upgris*, vol. 4, 2018.
- [6] R. Bahri, I. M.-T. INFORMATIKA, "Perbandingan algoritma template matching dan feature extraction pada optical character recognition," 2012.
- [7] T. Aruan, "Identifikasi Jenis Tanaman Jamur Beracun Menggunakan Pendekatan K-Nearest Neighbor," *Skripsi*, pp. 1-61, 2017.
- [8] R. Widodo *et al.*, "Pemanfaatan Ciri Gray Level Co-Occurrence Matrix (GLCM) Citra Buah Jeruk Keprok (*Citrus reticulata* Blanco) untuk Klasifikasi Mutu," vol. 2, no. 11, pp. 5769-5776, 2018.
- [9] A. Anwar, S. I.-J. A. I. "Perbenihan sayuran di Indonesia: Kondisi terkini dan prospek bisnis benih sayuran," 2005.
- [10] F. Felix, S. Faisal, T. F. M. Butarbutar, and P. Sirait, "Implementasi CNN dan SVM untuk Identifikasi Penyakit Tomat via Daun," *J. SIFO Mikroskil*, vol. 20, no. 2, pp. 117-134, 2019.
- [11] E. A. Savory, L. L. Granke, L. M. Quesada-Ocampo, M. Varbanova, M. K. Hausbeck, and B. Day, "The cucurbit downy mildew pathogen *Pseudoperonospora cubensis*," *Mol. Plant Pathol.*, vol. 12, no. 3, pp. 217-226, Apr. 2011.
- [12] Babadoost, M. "Bacterial spot of cucurbits", Report on plant disease, RPD no. 949, 2012.
- [14] Semangun H, "Penyakit tanaman hortikultura di indonesia", Yogyakarta, 2000.
- [15] Semangun H, "Pengantar ilmu penyakit tumbuhan", Yogyakarta, 2006.