

**RISET INFORMATIKA**  
**OPTIMASI EKSTRAKSI FITUR PADA MODIFIED K-NEAREST NEIGHBORS**  
**DALAM KLASIFIKASI PENYAKIT DAUN MELON**



**Disusun oleh:**

Ananda Ayu Puspitaningrum 21081010242

**UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR**  
**FAKULTAS ILMU KOMPUTER**  
**PRODI INFORMATIKA**  
**2024**

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang

Melon merupakan tanaman semusim yang termasuk tumbuhan yang merambat, dengan daun berwarna hijau yang berbentuk lebar, dan lima sudut, dan akar tunggang yang ditumbuhi akar serabut di ujungnya. Daun memiliki ruas yang berselang – selang di seluruh batangnya[1]. Tanaman melon tidak terlepas dari penyakit virus atau bakteri dalam proses penanamannya, salah satunya bagian daun. Infeksi virus dan bakteri mengakibatkan perubahan warna pada daun tanaman yang dapat mengganggu perkembangan dan pertumbuhan tanaman[6]. Penyakit tersebut tidak hanya menurunkan kualitas hasil panen tetapi juga memengaruhi kuantitas produksi. Dalam upaya mengatasi kerugian penurunan, perlu adanya deteksi dini terhadap jenis penyakit daun melon.

Awal untuk mengidentifikasi penyakit pada daun dapat dilakukan mendeteksi data gambar dengan teknik pengolahan citra. Teknik ini memungkinkan ekstraksi seperti warna, tekstur, dan pola pada daun untuk mendeteksi dan mengklasifikasikan jenis penyakit daun[3]. Salah satu metode yang sering digunakan dalam pengolahan citra adalah *Gray Level Co-occurrence Matrix* (GLCM), yang berfungsi untuk mengekstraksi tekstur pada citra. Selain itu, ruang warna HSV (Hue, Saturation, Value) untuk ekstraksi fitur warna, karena mampu merepresentasikan bagaimana manusia memahami dan melihat warna[8].

Namun, pengolahan citra saja tidak cukup untuk mendapatkan hasil yang optimal. Hasil ekstraksi fitur ini memerlukan algoritma klasifikasi yang kuat untuk memberikan hasil diagnosis yang akurat. Algoritma *K-Nearest Neighbors* (KNN) adalah salah satu teknik sederhana dalam memproses data dan klasifikasi. Meskipun demikian, KNN memiliki kelemahan dalam menghadapi dataset yang besar dan distribusi data yang tidak merata. Untuk mengatasi masalah ini, *Modified K-Nearest Neighbors* (MKNN) telah dikembangkan dengan mengintegrasikan teknik pengoptimalan seperti penyesuaian parameter  $k$  dan pembobotan jarak[4].

*Modified K-Nearest Neighbors* (MKNN) dirancang untuk meningkatkan performansi KNN dengan cara modifikasi struktur dasar algoritma agar lebih fleksibel terhadap variasi data. Dalam algoritma *K-Nearest Neighbors* (KNN), setiap tetangga yang terdekat memiliki bobot yang sama dalam menentukan klasifikasi. Namun dalam *Modified K-Nearest Neighbors* (MKNN), bobot dari tetangga-tetangga tersebut dapat disesuaikan untuk mengakomodasi perbedaan dalam jarak antar data, sehingga memberikan klasifikasi yang lebih akurat[4].

*Modified K-Nearest Neighbors* (MKNN) juga dapat mengatasi masalah yang muncul pada data yang tidak terdistribusi merata atau memiliki data yang memiliki nilai jauh berbeda atau tidak sesuai dengan pola dalam suatu dataset, dengan cara mengubah cara pemilihan tetangga berdasarkan karakteristik data[6].

Berbagai penelitian telah menerapkan metode ekstraksi fitur *Gray Level Co-Occurrence Matrix* (GLCM) dan *Hue, Saturation, Value* (HSV) dalam kombinasi dengan algoritma *K-Nearest Neighbors* (KNN) dan *Modified K-Nearest Neighbors* (MKNN) untuk klasifikasi penyakit tanaman. Penelitian yang dilakukan oleh Moch Aris Setyawan dkk (2022) berjudul “Klasifikasi Penyakit Daun Jagung Berdasarkan Ruang Warna HSV dan Fitur Tekstur dengan Algoritma K-NN”, yang menunjukkan bahwa penggunaan GLCM dan HSV dapat mengidentifikasi penyakit daun jagung dengan akurasi tertinggi sebesar 84% untuk nilai  $K=3$ , sedangkan akurasi menurun menjadi 80,5% untuk  $K=5$  dan 70% untuk  $K=9$ [9]. Penelitian lain oleh Yeni Minarti (2023) dalam jurnal “Klasifikasi Kesegaran Sayur Sawi Berdasarkan Citra HSV dan GLCM Menggunakan Metode *K-Nearest Neighbor*” mencatat bahwa ekstraksi fitur menggunakan GLCM dan HSV menghasilkan akurasi tertinggi sebesar 92,75% pada  $K=1$ , menunjukkan efektivitas metode ini dalam mengklasifikasikan kesegaran sayur[5]. Hasil-hasil penelitian ini menunjukkan bahwa penerapan GLCM dan HSV dalam algoritma KNN/MKNN dapat secara signifikan meningkatkan akurasi klasifikasi penyakit tanaman, memberikan kontribusi penting bagi pengelolaan kesehatan tanaman.

Penerapan kombinasi metode pengolahan citra seperti GLCM dan HSV dengan algoritma MKNN dapat meningkatkan akurasi deteksi penyakit pada daun melon secara signifikan. Melalui pendekatan ini, penelitian bertujuan untuk memberikan solusi yang lebih efektif dalam mengatasi deteksi penyakit tanaman. Dengan demikian, hasil penelitian ini diharapkan dapat dijadikan pengambilan keputusan yang lebih cepat dan tepat guna meningkatkan produktivitas dan kualitas panen.

## **1.2. Rumusan Masalah**

Rumusan masalah ini berfokus pada upaya untuk meningkatkan akurasi klasifikasi penyakit daun melon melalui pendekatan yang lebih optimal dalam ekstraksi fitur dan algoritma klasifikasi. Penelitian ini mencakup analisis terhadap metode ekstraksi fitur seperti *Gray Level Co-occurrence Matrix* (GLCM) dan ruang warna HSV, untuk mengidentifikasi sejauh mana kedua metode ini mampu meningkatkan akurasi dalam pengklasifikasian penyakit daun melon. Selain itu, penelitian ini juga berfokus pada pengembangan dan penerapan

algoritma *Modified K-Nearest Neighbors* (KNN) sebagai metode klasifikasi berbasis hasil ekstraksi fitur yang bertujuan untuk menghasilkan klasifikasi yang lebih akurat.

### **1.3. Tujuan Penelitian**

Tujuan dari penelitian ini adalah untuk mengklasifikasi penyakit pada daun melon berdasarkan hasil ekstraksi fitur *Gray Level Co-occurrence Matrix* (GLCM) dan ruang warna HSV menggunakan algoritma *Modified K-Nearest Neighbors* (KNN) sebagai metode klasifikasi berbasis hasil ekstraksi fitur untuk meningkatkan akurasi klasifikasi penyakit daun melon.

### **1.4. Manfaat Penelitian**

Penelitian ini diharapkan memberikan manfaat dalam berbagai aspek. Secara akademis, penelitian ini untuk mengembangkan pemahaman lebih dalam tentang penerapan algoritma *Modified K-Nearest Neighbors* (MKNN) dalam klasifikasi citra penyakit tanaman, yang dapat memperkaya literatur mengenai teknik klasifikasi citra. Penelitian ini juga diharapkan dapat memberikan kontribusi terhadap pengembangan algoritma MKNN, khususnya dalam meningkatkan akurasi dan ketepatan klasifikasi penyakit tanaman dengan data citra yang kompleks. Dari segi sosial, penelitian ini memiliki potensi dalam membantu petani untuk mendeteksi dan mengidentifikasi penyakit tanaman secara lebih cepat dan akurat, yang nantinya dapat meningkatkan hasil pertanian dan mengurangi kerugian yang disebabkan oleh penyakit tanaman.

## **BAB II**

### **TINJAUAN PUSTAKA**

#### **2.1. Penelitian Terdahulu**

Penelitian-penelitian yang sudah dilakukan peneliti sebelumnya yang terkait dengan penelitian yang akan dilakukan adalah sebagai berikut.

1. Penelitian yang dilakukan oleh Moch Aris Setyawan dkk (2022) berjudul “Klasifikasi Penyakit Daun Jagung Berdasarkan Ruang Warna HSV dan Fitur Tekstur dengan Algoritma K-NN”, yang menunjukkan bahwa penggunaan GLCM dan HSV dapat mengidentifikasi penyakit daun jagung dengan akurasi tertinggi sebesar 84% untuk nilai  $K=3$ , sedangkan untuk  $K$  selanjutnya yaitu 80,5 %, 73,5 %, 72 % dan 70 % pada fitur tekstur dengan Average 76% sedangkan untuk fitur warna akurasi tertinggi yaitu 64 % dengan Average 56,2 % [9].
2. Penelitian oleh Yeni Minarti (2023) dalam jurnal “Klasifikasi Kesegaran Sayur Sawi Berdasarkan Citra HSV dan GLCM Menggunakan Metode *K-Nearest Neighbor*” mencatat bahwa ekstraksi fitur menggunakan GLCM dan HSV menghasilkan akurasi tertinggi sebesar 92,75% pada  $K=1$ , sedangkan pada  $K=3$  diperoleh akurasi sebesar 90%. Hal tersebut menunjukkan efektivitas metode ini dalam mengklasifikasikan kesegaran sayur [5].
3. Penelitian yang dilakukan oleh Hamid Parvin dkk (2019) memperkenalkan metode Modified K-Nearest Neighbors (MKNN), yang memanfaatkan konsep "robust neighbors" dalam data pelatihan untuk memberikan hasil yang lebih akurat. Dengan menerapkan pendekatan ini, MKNN dapat mencapai akurasi yang lebih tinggi dibandingkan dengan KNN tradisional, yang sering kali terpengaruh oleh noise dan outlier dalam dataset. Hasil penelitian menunjukkan bahwa MKNN tidak hanya lebih efektif dalam menangani data yang tidak bersih, tetapi juga mampu memberikan hasil klasifikasi yang lebih stabil dan dapat diandalkan, menjadikannya pilihan yang lebih baik untuk aplikasi klasifikasi penyakit tanaman. Dalam evaluasi terhadap beberapa dataset, MKNN menunjukkan peningkatan akurasi yang signifikan, dengan nilai akurasi mencapai 87.81% untuk nilai  $k=3$  dan 90.58% untuk nilai  $k=5$ , dibandingkan dengan akurasi KNN yang hanya mencapai 84.26% dan 89.35% pada dataset yang sama [7].
4. Penelitian oleh Kiman Siregar dkk (2020) berjudul “Comparative Study of Decision Tree, K-Nearest Neighbor, and Modified K-Nearest Neighbor on *Jatropha Curcas* Plant Disease Identification”, yang menunjukkan bahwa MKNN memberikan hasil yang lebih baik dengan akurasi mencapai 92%, berkat optimasi fitur yang dilakukan sebelum proses

klasifikasi. Ini menunjukkan potensi MKNN dalam aplikasi pertanian untuk deteksi dini penyakit tanaman, karena MKNN lebih tahan terhadap *noise* dan *outlier* yang umumnya terkait dengan data yang kurang akurat. Selain itu, penggunaan MKNN juga membantu meningkatkan stabilitas hasil klasifikasi[10].

## 2.2. Landasan Teori

Landasan teori yang mendasari penelitian ini meliputi konsep – konsep berikut.

### 1. *K-Nearest Neighbors (KNN)*

*K-Nearest Neighbors (KNN)* adalah salah satu algoritma klasifikasi yang paling sederhana dan banyak digunakan dalam pembelajaran mesin. Algoritma ini bekerja dengan cara mengklasifikasikan data baru berdasarkan kedekatannya dengan data yang sudah ada dalam dataset. KNN menghitung jarak antara data uji dan data latih menggunakan berbagai metrik jarak. Salah satu keunggulan KNN adalah kemampuannya untuk menangani dataset yang besar dan beragam, serta pelatihan yang cepat. Namun, KNN juga memiliki beberapa kelemahan, termasuk sensitivitas terhadap noise dan outlier, serta kebutuhan akan memori yang besar karena harus menyimpan seluruh data latih[7].

### 2. *Modified K-Nearest Neighbors (MKNN)*

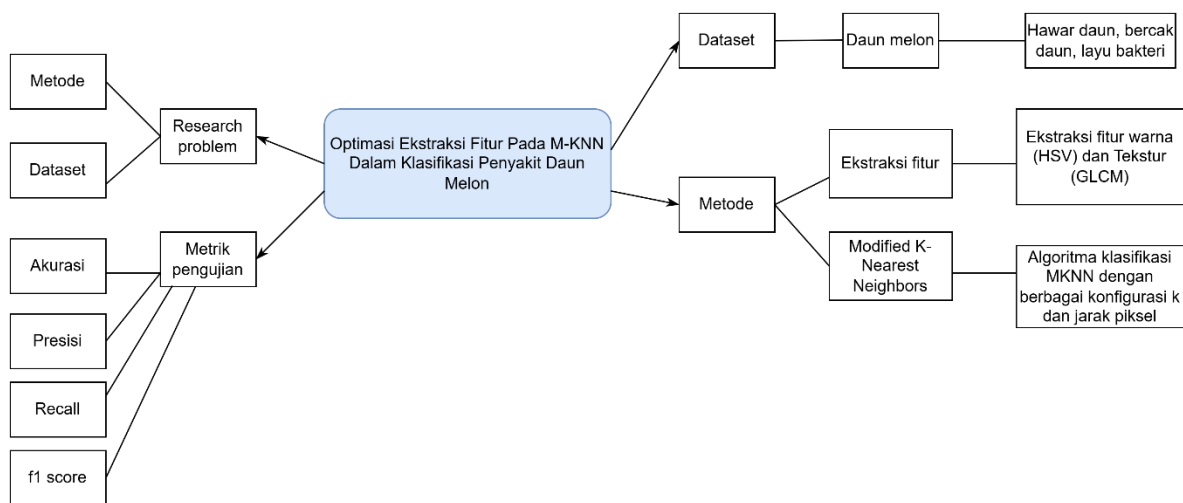
*Modified K-Nearest Neighbors (MKNN)* merupakan pengembangan dari algoritma KNN yang bertujuan untuk meningkatkan akurasi klasifikasi dengan memperhitungkan validitas tetangga terdekat. Dalam MKNN, bobot diberikan kepada setiap tetangga berdasarkan jaraknya dari titik data yang ingin diklasifikasikan, sehingga tetangga yang lebih dekat memiliki pengaruh lebih besar terhadap keputusan klasifikasi. Dalam beberapa penelitian MKNN dapat mengatasi beberapa kelemahan KNN, terutama dalam hal outlier, tetapi tetap menghadapi tantangan seperti nilai  $k$  yang bias dan kompleksitas komputasi yang tinggi[7].

### 3. Ekstraksi Fitur

Ekstraksi fitur adalah proses penting dalam pengolahan data dan pembelajaran mesin yang bertujuan untuk mengurangi dimensi dataset dengan memilih fitur-fitur yang paling relevan untuk analisis. Dalam konteks klasifikasi penyakit tanaman, ekstraksi fitur dapat mencakup analisis warna, tekstur, dan bentuk daun[10]. Proses ini tidak hanya membantu meningkatkan akurasi model klasifikasi tetapi juga mengurangi waktu komputasi. Teknik – teknik ekstraksi fitur meliputi:

- A. Fitur Warna: Merepresentasikan distribusi warna pada daun, yang dapat mencerminkan kondisi kesehatan tanaman. Model warna seperti RGB (*Red, Green, Blue*) dan HSV (*Hue, Saturation, Value*) sering digunakan dalam ekstraksi fitur warna.
- B. Fitur Tekstur: Menggambarkan pola permukaan daun, seperti kekasaran atau kehalusan, yang dapat diukur menggunakan metode seperti *Gray-Level Co-occurrence Matrix* (GLCM).

### 2.3. Mind Mapping



Proses diatas dijelaskan bahwa optimasi ekstraksi fitur pada algoritma Modified K-Nearest Neighbors (M-KNN) untuk klasifikasi penyakit daun melon dimulai dengan pengambilan data berupa gambar daun melon yang terkena penyakit seperti hawar daun, bercak daun, atau layu bakteri. Selanjutnya, dilakukan ekstraksi fitur menggunakan dua metode utama, yaitu ekstraksi fitur warna dengan model HSV (Hue, Saturation, Value) dan ekstraksi tekstur menggunakan metode GLCM (Gray Level Co-occurrence Matrix). Fitur yang telah diekstraksi kemudian digunakan dalam algoritma M-KNN, yang dioptimasi dengan berbagai konfigurasi, termasuk pengaturan nilai  $k$  (jumlah tetangga terdekat) dan jarak piksel.

## **BAB III**

### **DESAIN DAN IMPLEMENTASI SISTEM**

#### **3.1. Metode Penelitian**

##### **3.1.1. Pendekatan Penelitian**

Penelitian ini dilakukan untuk mengklasifikasi penyakit pada daun melon berdasarkan hasil ekstraksi fitur *Gray Level Co-occurrence Matrix* (GLCM) dan ruang warna HSV menggunakan algoritma *Modified K-Nearest Neighbors* (KNN). Ekstraksi fitur dilakukan dengan mempertimbangkan fitur tekstur dan warna. Algoritma MKNN akan diterapkan untuk mengoptimalkan proses klasifikasi dengan memanfaatkan hasil ekstraksi fitur tersebut, sehingga diharapkan dapat meningkatkan akurasi dan efisiensi dalam identifikasi penyakit pada daun melon.

##### **3.1.2. Pengumpulan Data**

Pengumpulan data dilakukan dengan pemotretan sample daun melon yang dibudidayakan di Greenhouse Melon Ngrowo Bening Edu Park Madiun. Pengambilan data citra dilakukan menggunakan kamera handphone resolusi HD 48 MP. Daun yang difoto didasarkan pada tiga jenis penyakit daun melon seperti hawar daun, bercak daun, dan layu bakteri.

##### **3.1.3. Preprocessing Data**

Tahap preprocessing dilakukan dengan mengubah data citra daun yang telah dikumpulkan. Langkah awal tahapan preprocessing adalah cropping, yaitu: memotong sebagian objek di area gambar tertentu agar dapat memisahkan objek yang satu dengan lainnya, kemudian menghilangkan background di sekitar citra daun dengan latar belakang putih agar objek citra dapat dikenali. Setelah itu, dilanjutkan dengan resize yaitu mengubah ukuran pixel citra menjadi lebih kecil agar citra lebih cepat diproses dan tidak banyak menghabiskan memori penyimpanan.

##### **3.1.4. Ekstraksi Fitur**

Setelah tahapan preprocessing, ekstraksi fitur akan dilakukan terhadap citra daun melon. Ekstraksi fitur merupakan proses untuk melihat nilai-nilai ciri yang terdapat pada suatu citra. Ekstraksi fitur menghasilkan nilai yang telah diekstrak akan digunakan untuk proses pelatihan. Fitur yang digunakan pada ekstraksi fitur yaitu

###### **A. Hue Saturation Value**

Merupakan ekstraksi fitur warna yaitu nilai variasi merah (R), hijau (G), dan biru (B). Ruang warna HSV akan digunakan untuk menghasilkan area daun yang terinfeksi penyakit. Untuk



memperoleh nilai HSV, harus mengkonversi citra RGB menjadi HSV dengan persamaan – persamaan dibawah ini

$$r = \frac{R}{(R + G + B)}$$

$$g = \frac{G}{(R + G + B)}$$

$$b = \frac{B}{(R + G + B)}$$

$$V = \max (r,g,b)$$

$$S = \begin{cases} 0, & \text{jika } v = 0 \\ 1, & -\frac{\min(r, g, b)}{v} v > 0 \end{cases}$$

$$f(x) = \begin{cases} 0, & \text{jika } S = 0 \\ \frac{60 * (g - b)}{S * V}, & \text{jika } V = r \\ 60 * \left[ 2 + \frac{b - r}{S * V} \right], & \text{jika } V = g \\ 60 * \left[ 4 + \frac{r - g}{S * V} \right], & \text{jika } V = b \end{cases}$$

$$H = H + 360 \text{ jika } H < 0$$

## B. Gray Level Co-occurrence Matrix (GLCM)

Merupakan ekstraksi fitur tekstur berupa kontras, energy, homogeneity, correlation, dan dissimilarity. Konsep sederhana metode algoritma GLCM digunakan untuk menghitung kesamaan piksel misalnya kesamaan nilai piksel antara piksel I dan piksel J, dengan jarak (d) dan sudut tertentu. Fitur yang digunakan dipaparkan dalam persamaan – persamaan berikut.

1. Contrast: keberadaan variasi atas ke abuan piksel citra

$$Contrast = \sum_i \sum_j (i - j)^2 p_{(i,j)}$$

2. Energy: representasi ukuran keseragaman citra

$$Energy = \sum_i \sum_j p_{(i,j)}^2$$

3. Homogeneity: representasi ukuran keseragaman distribusi intensitas piksel citra

$$Homogeneity = \sum_i \sum_j \frac{p(i,j)}{1+|i-j|}$$

4. Correlation: representasi fitur tekstur terhadap ketergantungan linier derajat keabuan dari setiap piksel yang bertetangga dengan obyek citra

$$Correlation = \sum_i \sum_j \frac{(i-\mu_i)(j-\mu_j)p(i,j)}{\sigma_i \sigma_j}$$

### 3.1.5. Modified K-Nearest Neighbors

Modified K-Nearest Neighbors (MKNN) adalah varian dari algoritma *K-Nearest Neighbors* (KNN) yang dirancang untuk meningkatkan akurasi klasifikasi dengan memperhitungkan kedekatan (jarak) masing-masing tetangga dalam menentukan label kelas.

Perhitungan diawali dengan perhitungan validasi data pada semua data di data latih. Lalu, dilakukan perhitungan untuk mencari Weight Voting pada semua data uji menggunakan validitas data. Weight masing – masing tetangga dihitung dengan persamaan.

$$\frac{1}{(de + 1)}$$

Kemudian, validitas dari setiap data pada data latih dikalikan dengan weight berdasarkan pada jarak Euclidean. Sehingga metode MKNN, didapatkan persamaan weight voting tiap tetangga sebagai berikut.

$$W(i) = Validitas(i) \times \frac{1}{de+0,5}$$

Keterangan:

W(i): Perhitungan Weight Voting

Validasi (x): Nilai Validasi

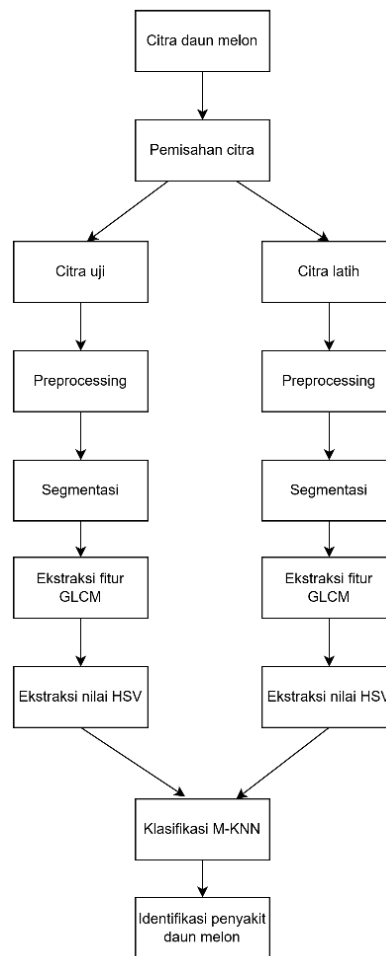
de: Jarak Euclidean

### 3.1.6. Akurasi Sistem

Akurasi perhitungan ini dilakukan untuk mengetahui tingkat akurasi metode MKNN dalam hasil klasifikasi. Akurasi dapat diperoleh dari presentase kebenaran, yaitu perbandingan antara jumlah data uji dengan jumlah data keseluruhan dikalikan 100%. Akurasi bisa didapat melalui persamaan berikut.

$$Akurasi = \frac{jumlah\ data\ uji\ benar}{jumlah\ seluruh\ data\ uji} \times 100\%$$

### 3.2. Desain Sistem



Gambar diatas merupakan alur penelitian. Proses dimulai dengan membagi dataset menjadi data uji dan data latih. Kemudian, dilakukan preprocessing yaitu dengan mengubah background citra menjadi putih dan mengubah ukuran pikselnya. Citra dilakukan segmentasi warna untuk mendapatkan area daun yang terinfeksi. Setelah dilakukan segmentasi citra diubah ke grayscale untuk mendapatkan nilai matriks kookurensi dilanjutkan mencari nilai energi, kontras, korelasi, dan homogenitas untuk ekstraksi fitur GLCM. Menghitung nilai HSV yang digunakan sebagai fitur untuk klasifikasi. Ketika kedua ekstraksi fitur sudah dilakukan nilai tersebut akan dijadikan acuan untuk klasifikasi dengan menggunakan algoritma klasifikasi M-KNN. Setelah mendapatkan hasil klasifikasi, kemudian hitung akurasi untuk mendapatkan pembuktian tingkat akurasi dari proses klasifikasi yang telah didapatkan.

## DAFTAR PUSTAKA

- [1] budi daya melon agromedia ebook”. Jakarta, PT Agromedia Pustaka. hlm 6. 2007. 1
- [2] Banupriya .N., & Rajaneni .D.(2022). Integration of Image Processing and Machine Learning for Plant Disease Identification. *Journal of Agriculture and Food Research*, 10, 100456. 6
- [3] Khalid, M. M., & Karan , O. (2023). Deep Learning for Plant Disease Detection. *International Journal of Mathematics, Statistics, and Computer Science*, 2, 75–84. <https://doi.org/10.59543/ijmscs.v2i.8343> 3
- [4] Huang, L., Xu, W., & Sun, Y. (2021). Improved K-Nearest Neighbors Algorithm for Crop Disease Classification. *Computers and Electronics in Agriculture*, 187, 106290. <https://doi.org/10.1016/j.compag.2021.106290> 5
- [5] Minarti, Y. (2023). Klasifikasi Kesegaran Sayur Sawi Berdasarkan Citra HSV dan GLCM Menggunakan Metode K-Nearest Neighbor. *Jurnal Teknologi Pertanian*, Universitas Mataram. 8
- [6] M. Fajar et al., “JIP (Jurnal Informatika Polinema) Foreground Extraction pada Citra Daun Melon dengan Bantuan Deep Neural Network”. 2021 2
- [7] Parvin, H., Alizadeh, H., & Minaei-Bidgoli, B. (2019). MKNN: Modified K-Nearest Neighbor. *Global Journal of Computer Science and Technology*. 9
- [8] Patel, H., Shah, M., & Jani, R. (2020). Feature Extraction Techniques in Image Processing: A Study. *International Journal of Advanced Research in Computer Science*, 11(1), 1–8. 4
- [9] Setyawan, M. A., Kasih, P., & Widyadara, M. A. D. (2022). Klasifikasi Penyakit Daun Jagung Berdasarkan Ruang Warna HSV dan Fitur Tekstur dengan Algoritma K-NN. *Jurnal Teknik Informatika dan Teknologi Informasi*, 4(1), 86–95. 7
- [10] Siregar, K. (2020). Comparative Study of Decision Tree, K-Nearest Neighbor, and Modified K-Nearest Neighbor on *Jatropha Curcas* Plant Disease Identification. 10