

Deteksi Infeksi pada Daun Kapas menggunakan Kombinasi Metode Ekstraksi Fitur Warna dan Tekstur

Siti Khotimatul Wildah^{*1}, Abdul Latif², Suharyanto³

^{1,2,3}Universitas Bina Sarana Informatika
Email: ¹siti.ska@bsi.ac.id, ²abdul.bl@bsi.ac.id, ³suharyanto@bsi.ac.id
^{*}Siti Khotimatul Wildah

Abstrak

Tanaman kapas merupakan salah satu tanaman penting dalam industri tekstil dan pertanian, akan tetapi sering mengalami kendala terutama dalam masalah infeksi yang disebabkan oleh patogen seperti jamur dan bakteri. Deteksi dini dan akurat infeksi pada daun kapas menjadi kunci untuk mengambil langkah pencegahan dan pengendalian yang tepat guna sehingga dapat mengurangi dampak kerugian pada hasil panen dan kualitas serat kapas. Teknologi pengolahan citra telah menunjukkan potensi untuk digunakan dalam deteksi infeksi pada tanaman, khususnya dengan menggunakan metode ekstraksi fitur warna dan tekstur. Penelitian ini bertujuan untuk melakukan kombinasi metode ekstraksi fitur warna untuk mengidentifikasi perubahan warna pada daun kapas yang terinfeksi dan fitur tekstur untuk mengenali pola atau struktur tekstur yang khas dari daun kapas yang terinfeksi serta metode pengklasifikasian *Random Forest*. Dataset yang digunakan pada penelitian ini terdiri dari 4 jenis citra dengan jumlah citra sebanyak 1786 citra. Hasil penelitian menunjukkan bahwa penggunaan kombinasi metode ekstraksi fitur warna dan tekstur meningkatkan akurasi deteksi infeksi pada daun kapas secara signifikan dibandingkan dengan penggunaan metode ekstraksi fitur tunggal. Model deteksi yang dihasilkan mampu mengenali infeksi pada daun kapas dengan tingkat akurasi yang tinggi dan meminimalkan kesalahan dalam pengklasifikasian. Nilai akurasi yang dihasilkan dari penelitian ini sebesar 0.982683.

Kata kunci: Penyakit Daun Kapas, Haralick, Color Histogram, Random Forest, Kombinasi Ekstraksi Fitur

Abstract

Cotton is one of the crucial plants in the textile and agricultural industries, but it often faces challenges, especially in terms of infection caused by pathogens like fungi and bacteria. Early and accurate detection of infections on cotton leaves is crucial to implement proper preventive and control measures, thereby reducing the impact on crop yield and cotton fiber quality. Image processing technology has shown potential for use in detecting plant infections, particularly through the utilization of color and texture feature extraction methods. This research aims to combine color feature extraction methods to identify color changes in infected cotton leaves, texture feature extraction methods to recognize distinctive texture patterns, and the Random Forest classification method. The dataset used in this study comprises four types of images, totaling 1786 images. The results reveal that the combined use of color and texture feature extraction methods significantly improves the accuracy of detecting infections on cotton leaves compared to single-feature extraction methods. The generated detection model achieves a high level of accuracy in identifying infections and minimizes misclassifications. The accuracy achieved in this research is 0.982683.

Keywords: Cotton Leaf Disease, Haralick, Color Histogram, Random Forest, Feature Extraction Combination

1. PENDAHULUAN

Kapas atau *Gossypium hirsutum* L. menjadi salah satu jenis komoditi perkebunan penghasil serat alam yang dijadikan sebagai bahan baku industri tekstil (Hamid et al., 2023), akan tetapi sering

mengalami masalah dikarenakan infeksi yang disebabkan oleh patogen seperti jamur dan bakteri. Penyakit kapas sangat mempengaruhi hasil dan kualitas kapas. Jenis hama atau penyakit yang diderita tanaman kapas dapat diketahui dari bercak penyakit pada daun kapas (Liang, 2021).

Deteksi dini penyakit atau infeksi pada tanaman merupakan cara produksi pertanian yang baik (Shah & Jain, 2019). Permasalahan penyakit atau infeksi pada tanaman merupakan masalah yang harus segera ditangani sebagai tindak pencegahan gagal panen dan peningkatan produktivitas pertanian. Pemanfaatan teknologi saat ini telah menjangkau hampir semua bidang, seperti pendidikan, kesehatan, dan pertanian. Hal ini karena teknologi dapat membantu manusia dalam bekerja lebih mudah, cepat, dan mengurangi tingkat kesalahan (Agustiani et al., 2022). Salah satu teknologi tersebut yaitu penerapan teknologi *Computer Vision* dan teknik *Machine Learning* yang telah banyak diadopsi dalam hal deteksi penyakit pada tanaman, yang berdampak pada percepatan proses identifikasi (Martina & Abdillah, 2020).

Image processing merupakan bagian dari *computer vision* dimana diagnosis gejala penyakit pada tanaman dengan menerapkan metode *image processing* merupakan hal yang sangat diperhatikan (Tripathy, 2021). Membedakan penyakit tanaman secara dini menjadi suatu metode yang diperlukan untuk menghindari kerugian dalam suatu sistem produksi pertanian. Teknik *image processing* digunakan untuk mengamati dan menentukan penyakit pada tanaman secara cepat dan tepat (Karthika et al., 2021).

Beberapa penelitian telah dilakukan dalam deteksi infeksi atau penyakit pada daun kapas, seperti penelitian yang dilakukan oleh Kapil Prashar, Rajneesh Talwar dan Chander Kant pada tahun 2019 mengenai pengenalan penyakit daun kapas dengan menggunakan metode fitur visual. Pada penelitian ini metode *Knearest Neighbor* (kNN) dan *Support Vector Machine* (SVM) diterapkan dan berhasil memperoleh nilai akurasi sebesar 96% (Prashar et al., 2019a). Penelitian yang dilakukan oleh Shantanu Kumbhar, Amita Nilawar, Shruti Pati dan Bodireddy Mahalakshmi pada tahun 2019 mengenai deteksi penyakit pada daun kapas menggunakan CNN. Penelitian ini berhasil mengimplementasikan sebuah sistem berbasis web dengan model CNN dimana hasil akurasi yang diperoleh sebesar 80% (Kumbhar et al., 2019).

Kemudian penelitian yang dilakukan oleh Kapil Prashar, Rajneesh Talwar, Chander Kant pada tahun 2019 mengenai analisis pengenalan penyakit daun kapas. Dalam penelitian ini deskriptor fitur dinamis dirancang dengan analisis cluster yang tidak konsisten (ICA) dan peningkatan fitur penyakit (DFE), yang digabungkan sebagai deskriptor hibrida yang dikenal sebagai ICADFE untuk pengenalan penyakit tanaman kapas. ICADFE ditemukan untuk meningkatkan akurasi deteksi sekitar 80%, presisi sekitar 95% dan pengukuran f1 sekitar 88% (Prashar et al., 2019b).

Penelitian lainnya dilakukan oleh Azath M., Melese Zekiwos, dan Abey Bruck pada tahun 2021 mengenai teknologi pemrosesan citra menggunakan metode *deep learning* untuk mendiagnosis penyakit dan hama pada daun kapas. Model pada penelitian ini mencapai akurasi 96,4% untuk mengidentifikasi penyakit tersebut (Azath et al., 2021). Penelitian selanjutnya dilakukan oleh Xihuizi Liang pada tahun 2021 mengenai mengenai klasifikasi penyakit daun kapas berdasarkan metric learning. Pada penelitian ini secara eksperimental, ditunjukkan bahwa akurasi klasifikasi *DenseNet* adalah yang tertinggi, dibandingkan dengan dua jaringan lainnya, dan akurasi klasifikasi *S-DenseNet* adalah 7,7% lebih tinggi daripada rata-rata *DenseNet* untuk jumlah langkah yang berbeda. (Liang, 2021). Penelitian yang dilakukan oleh Rafael Faria Caldeira, Wesley Esdras Santiago dan Barbara Teruel pada tahun 2021 mengenai Identifikasi lesi daun kapas menggunakan teknik *deep learning*, dimana hasil model pembelajaran *GoogleNet* dan *Resnet50* menggunakan jaringan saraf *convolutional*, diperoleh presisi masing-masing sebesar 86,6% dan 89,2% (Caldeira et al., 2021).

Berdasarkan hal tersebut, penelitian ini menggunakan teknologi *computer vision* dengan teknik *image processing* untuk mendeteksi penyakit pada daun kapas. Metode identifikasi difokuskan pada analisis perbedaan warna pada bercak daun untuk menentukan jenis infeksi. Penelitian ini menggunakan kombinasi fitur ekstraksi *Color Histogram* untuk informasi warna dan fitur ekstraksi *Haralick* untuk informasi tekstur, serta menggunakan metode klasifikasi *Random Forest*.

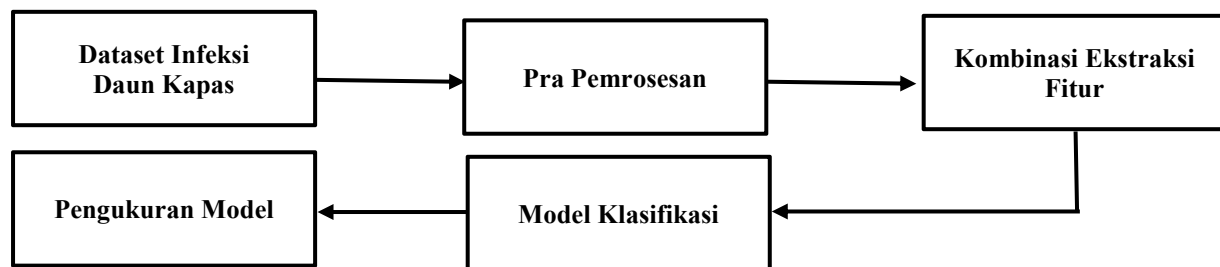
Tujuan penelitian ini adalah untuk mengevaluasi dampak penggunaan kombinasi fitur tekstur *Haralick*, fitur *Color Histogram*, dan metode klasifikasi *Random Forest* terhadap akurasi dalam klasifikasi infeksi pada daun kapas. Selain itu, penelitian ini juga membandingkan metode yang diusulkan dengan metode lain yang sudah digunakan sebelumnya dalam klasifikasi penyakit pada daun kapas untuk menilai seberapa efektif metode yang diusulkan dibandingkan dengan metode yang telah ada sebelumnya.

2. METODE PENELITIAN

Dalam penelitian ini, digunakan kombinasi analisis tekstur menggunakan metode *Haralick* dan analisis tekstur warna menggunakan metode *Color Histogram*, serta metode klasifikasi *Random Forest*

untuk mengekstrak fitur dan mengklasifikasikan infeksi pada tanaman kapas. *Dataset* yang digunakan terdiri dari 4 jenis citra dengan total 1786 citra. Analisis tekstur *Haralick* digunakan untuk mengekstrak fitur tekstur dari citra, sementara *Color Histogram* digunakan untuk mengekstrak fitur warna. Selanjutnya, *Random Forest* digunakan untuk mengklasifikasikan infeksi pada tanaman menjadi 4 jenis yang berbeda.

Dalam penelitian ini, dilakukan pengukuran menggunakan tabel *Confusion Matrix* untuk mengevaluasi hasil prediksi pada setiap kelas sehingga dapat dilihat jumlah citra yang diprediksi dengan sesuai maupun jumlah citra yang diprediksi dengan salah. Selain itu, pengukuran juga dilakukan dengan mempertimbangkan nilai akurasi, nilai presisi, dan nilai *recall* yang dihasilkan dari pengklasifikasian model tersebut. Tahapan metode penelitian dapat dilihat pada Gambar 1.

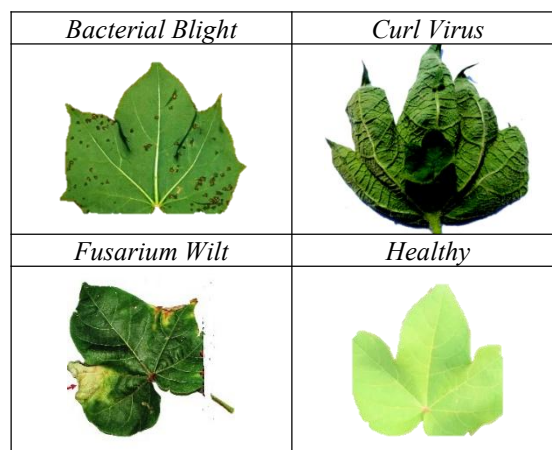


Sumber: (Hasil Penelitian (2023))

Gambar 1. Metode Penelitian

A. Dataset

Dataset yang digunakan menggunakan citra daun kapas yang terdiri dari 4 jenis yaitu, *bacterial blight* atau citra daun kapas yang terinfeksi hawar bakteri, kemudian citra *curl virus* yaitu citra daun yang keriting karena terkena virus, *fusarium wilt* atau citra daun yang layu serta citra *healthy* atau citra daun yang sehat. Keseluruhan citra yang digunakan sebanyak 1786 citra. *Sample* citra yang digunakan pada penelitian ini dapat dilihat pada Gambar 2.



Gambar 2. *Dataset*

B. Pra Pemrosesan

Kumpulan citra daun kapas diproses terlebih dahulu dengan penentuan ukuran dan dimensi citra yang akan diproses dengan tujuan untuk mempercepat proses klasifikasi model, penentuan jumlah persentase data latih dan data uji yang akan digunakan, dimana pada penelitian ini persentasi yang digunakan sebesar 80% data latih dan 20% data uji.

C. Ekstraksi fitur

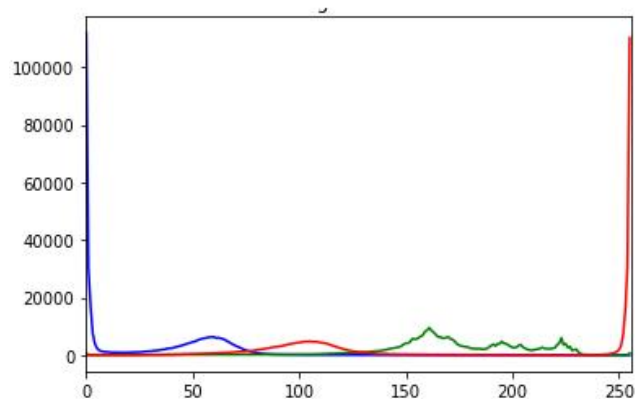
Proses ekstraksi fitur yang dilakukan pada penelitian ini berperan penting dalam pengklasifikasian karakteristik dari citra sehingga dapat mengelompokkan citra lebih baik berdasarkan ciri atau pola dari citra tersebut (Khotimatul & Latif, 2023). Ekstraksi fitur yang digunakan mengkombinasikan 2 metode ekstraksi fitur yaitu Ekstraksi Fitur Warna (*Color Histogram*) dan ekstraksi fitur tekstur (*Haralick*).

1) Fitur *Color Histogram*

Fitur Histogram Warna adalah cara untuk merepresentasikan informasi warna pada suatu citra, termasuk distribusi warna, kontras, dan kecerahan piksel. Metode ini menggunakan histogram dari hasil indeks *Vector Quantization* (VQ) untuk menggambarkan citra tersebut. Citra berwarna pertama-tama dibagi menjadi beberapa blok, dan warna yang diindeks oleh VQ dapat dihitung menggunakan dua penghitung citra, yaitu nilai minimum dan maksimum dari setiap blok citra (Khultsum & Subekti, 2021). Histogram warna menghitung jumlah piksel dari setiap jenis warna dalam suatu citra, dan dapat dibuat dengan cepat dengan hanya membaca citra sekali dan meningkatkan jumlah bin yang sesuai pada histogram (Sany, 2021), dengan persamaan sebagai berikut:

$$H_c = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \delta(f_{i,j} - c), \forall c \in C \quad (1)$$

Pada penelitian ini histogram warna pada daun kapas dengan jenis citra *Bacterial Blight* dapat dilihat pada gambar 3.



Gambar 3. Histogram Warna Bacterial Blight

Fitur tekstur *Haralick* adalah deskriptor tekstur yang sering digunakan dalam analisis citra. Perhitungan fitur Haralick juga dikenal sebagai proses kuantisasi, di mana dilakukan penghitungan reduksi tingkat keabuan dari suatu citra. (Löfstedt et al., 2019). Haralick memiliki mekanisme dengan cara menghitung 24 fitur statistik yang berbeda dari matriks GLCM (*Gray-Level Co-occurrence Matrix*) yang telah dinormalisasi. Fitur-fitur ini mengukur aspek penting dari informasi lokal dan fitur spasial dalam citra. Dari 24 fitur yang berbeda tersebut, terdapat 5 fitur yang menonjol atau memiliki pengaruh yang signifikan. Fitur tersebut diantaranya *Homogeneity, entropy, energy, correlation* dan *contrast* (Ashir et al., 2021)

2) Random Forest

Random Forest merupakan pengembangan dari algoritma *decision tree* dengan tujuan untuk mengatasi masalah yang terkait dengan tingginya potensi *overfitting* ketika menggunakan algoritma *decision tree* (Ferdita Nugraha et al., 2022). Prinsip kerja dari teknik ini adalah dengan membangun beberapa pohon keputusan yang berdasarkan pada subset acak dari data pelatihan. Prediksi akhir kemudian diambil berdasarkan agregat hasil dari pohon-pohon tersebut (Tantyo et al., 2023). Metode *Random Forest* terdiri dari *root node* yang merupakan simpul atau akar yang berada pada posisi paling atas kemudian terdiri dari internal node yang merupakan simpul cabang, dan *leaf node* atau terminal yang merupakan simpul akhir (Sandag, 2020).

Berikut ini merupakan persamaan *Random Forest*:

$$Entropy(Y) = - \sum_i P(c|Y) \log_2 P(c|Y) \quad (2)$$

Information Gain (Y, a)

$$= Entropy(Y) - \sum v \text{values}(a) \left[\frac{Y_v}{Y_a} \right] Entropy(Y_v) \quad (3)$$

3. HASIL DAN PEMBAHASAN

Pada penelitian ini, eksperimen dan pengujian dilakukan menggunakan *Software Jupyter Notebook* dengan bahasa pemrograman *Python*. Eksperimen dilakukan untuk menguji dataset, membangun model, dan mengukur hasil dari identifikasi infeksi pada daun kapas. Metode fitur ekstraksi yang digunakan adalah kombinasi dari *Color Histogram* dan *Haralick*, dan metode klasifikasi *Random Forest*.

Penelitian ini melibatkan beberapa tahapan, dimulai dari persiapan dataset berupa citra daun kapas yang terdiri dari 4 jenis, meliputi daun sehat hingga daun kapas yang terinfeksi. Selanjutnya, dilakukan prapemrosesan citra untuk membersihkan data dari noise dan meningkatkan kualitas citra. Setelah itu, dilakukan penerapan metode ekstraksi fitur untuk mengambil informasi penting dari setiap citra. Pengambilan metode fitur ekstraksi yang optimal ditentukan berdasarkan hasil pengujian, mulai dari penerapan 1 metode ekstraksi fitur hingga penerapan dari kombinasi beberapa metode ekstraksi fitur.

Penerapan metode ekstraksi yang terpilih digunakan untuk menguji model pada dataset daun kapas tersebut. Metode ekstraksi fitur yang diujikan meliputi ekstraksi fitur *color histogram*, *haralick*, *hu-moment*, dan *zernike*. Dari berbagai metode ekstraksi fitur yang diuji, metode yang memberikan hasil baik dengan akurasi tinggi dan nilai standar deviasi yang rendah akan dipilih. Hasil pengujian yang telah dilakukan dapat dilihat melalui tabel percobaan yang disajikan sebagai bentuk evaluasi dari penelitian ini.

Tabel 1. Penerapan Metode Ekstraksi Fitur

| No | Fitur Ekstraksi | Metode Klasifikasi | Nilai Akurasi | Nilai Standar Deviasi |
|----|-----------------|---|-----------------|-----------------------|
| 1 | Color Histogram | Logistic Regression (LR) | 0.930832 | 0.019485 |
| | | Linear Discriminant Analysis (LDA) | 0.926077 | 0.026908 |
| | | k-Nearest Neighbor (k-NN) | 0.910396 | 0.019549 |
| | | Classification and Regression Tree (CART) | 0.930026 | 0.021248 |
| | | Random Forest (RF) | 0.974852 | 0.011538 |
| | | Naïve Bayes (NB) | 0.011538 | 0.036795 |
| | | Support Vectore Machine (SVM) | 0.958354 | 0.015270 |
| 2 | Haralick | Logistic Regression (LR) | 0.606078 | 0.049720 |
| | | Linear Discriminant Analysis (LDA) | 0.644617 | 0.051917 |
| | | k-Nearest Neighbor (k-NN) | 0.799502 | 0.029817 |
| | | Classification and Regression Tree (CART) | 0.944181 | 0.012432 |
| | | Random Forest (RF) | 0.969328 | 0.013398 |
| | | Naïve Bayes (NB) | 0.567538 | 0.054496 |
| | | Support Vectore Machine (SVM) | 0.620208 | 0.054870 |
| 3 | Hu-Moment | Logistic Regression (LR) | 0.420639 | 0.024074 |
| | | Linear Discriminant Analysis (LDA) | 0.426932 | 0.026066 |
| | | k-Nearest Neighbor (k-NN) | 0.742169 | 0.029498 |
| | | Classification and Regression Tree (CART) | 0.943387 | 0.026533 |
| | | Random Forest (RF) | 0.948905 | 0.023140 |
| | | Naïve Bayes (NB) | 0.362408 | 0.025753 |
| | | Support Vectore Machine (SVM) | 0.502411 | 0.026879 |
| 4 | Zernike | Logistic Regression (LR) | 0.413589 | 0.046571 |
| | | Linear Discriminant Analysis (LDA) | 0.430044 | 0.034272 |
| | | k-Nearest Neighbor (k-NN) | 0.723284 | 0.041371 |
| | | Classification and Regression Tree (CART) | 0.742926 | 0.052785 |

| | | | |
|--|-------------------------------|-----------------|-----------------|
| | Random Forest (RF) | 0.893123 | 0.028015 |
| | Naïve Bayes (NB) | 0.360089 | 0.038251 |
| | Support Vectore Machine (SVM) | 0.599096 | 0.034498 |

Berdasarkan tabel 1 tersebut menunjukkan bahwa penerapan metode *Color Histogram*, memberikan hasil dengan nilai akurasi yang lebih unggul dibandingkan dengan metode ekstraksi fitur lainnya dengan nilai akurasi yang dihasilkan sebesar 0.974852 dan nilai standar deviasi sebesar 0.011538. Dimana urutan performa metode ekstraksi fitur tersebut dimulai dari *Color Histogram*, *Haralick*, *Hu-Moment* dan *Zernike*. Metode ekstraksi fitur *Haralick* memiliki nilai yang cukup unggul dengan nilai akurasi yang dihasilkan metode ini sebesar 0.969328 dan nilai standar deviasi sebesar 0.013398, sehingga metode ekstraksi fitur *haralick* dapat diperhitungkan untuk kombinasi metode ekstraksi fitur yang akan diujikan lainnya. Berdasarkan hasil dari pengujian penerapan 1 metode ekstraksi fitur dilanjutkan dengan eksperimen penerapan metode ekstraksi fitur gabungan atau kombinasi, yang dapat dilihat pada tabel 2.

Tabel 2. Penerapan Metode Ekstraksi Fitur Kombinasi

| No | Fitur Ekstraksi | Metode Klasifikasi | Nilai Akurasi | Nilai Standar Deviasi |
|----|-------------------------------|---|---------------|-----------------------|
| 1 | Color Histogram dan Haralick | Logistic Regression (LR) | 0.953135 | 0.021021 |
| | | Linear Discriminant Analysis (LDA) | 0.943399 | 0.021059 |
| | | k-Nearest Neighbor (k-NN) | 0.907247 | 0.019468 |
| | | Classification and Regression Tree (CART) | 0.944956 | 0.029058 |
| | | Random Forest (RF) | 0.983501 | 0.014257 |
| | | Naïve Bayes (NB) | 0.760218 | 0.044120 |
| | | Support Vectore Machine (SVM) | 0.954398 | 0.018238 |
| 2 | Hu-Moment dan Haralick | Logistic Regression (LR) | 0.598179 | 0.040523 |
| | | Linear Discriminant Analysis (LDA) | 0.622607 | 0.046153 |
| | | k-Nearest Neighbor (k-NN) | 0.844316 | 0.024160 |
| | | Classification and Regression Tree (CART) | 0.948105 | 0.011815 |
| | | Random Forest (RF) | 0.970109 | 0.013089 |
| | | Naïve Bayes (NB) | 0.547078 | 0.043247 |
| | | Support Vectore Machine (SVM) | 0.643799 | 0.037171 |
| 3 | Color Histogram dan Hu-Moment | Logistic Regression (LR) | 0.935538 | 0.016415 |
| | | Linear Discriminant Analysis (LDA) | 0.922964 | 0.027839 |
| | | k-Nearest Neighbor (k-NN) | 0.917434 | 0.013841 |
| | | Classification and Regression Tree (CART) | 0.932382 | 0.023919 |
| | | Random Forest (RF) | 0.980346 | 0.016219 |
| | | Naïve Bayes (NB) | 0.756213 | 0.048073 |
| | | Support Vectore Machine (SVM) | 0.963054 | 0.027279 |
| 4 | Color Histogram dan Zernike | Logistic Regression (LR) | 0.929995 | 0.030411 |
| | | Linear Discriminant Analysis (LDA) | 0.937857 | 0.027865 |
| | | k-Nearest Neighbor (k-NN) | 0.912685 | 0.043968 |
| | | Classification and Regression Tree (CART) | 0.918215 | 0.021260 |
| | | Random Forest (RF) | 0.977990 | 0.016432 |
| | | Naïve Bayes (NB) | 0.759449 | 0.028292 |
| | | Support Vectore Machine (SVM) | 0.960667 | 0.018644 |

Berdasarkan tabel 2 tersebut menunjukkan bahwa kombinasi metode *Color Histogram dan Haralick* memberikan hasil yang lebih baik dibandingkan dengan metode kombinasi lainnya maupun

dengan penggunaan 1 metode ekstraksi fitur. Hasil yang diberikan dari menggunakan kombinasi metode ekstraksi fitur *Color Histogram* dan *Haralick* memberikan nilai akurasi sebesar 0.983501 dan nilai standar deviasi sebesar 0.017540. Hasil eksperimen menggunakan kombinasi metode ekstraksi fitur menggunakan 2 metode yaitu *Color Histogram* dan *Haralick* meningkat sebesar 0,008649. Kombinasi metode *Color Histogram* dan *Hu-Moment* memberikan hasil yang hampir menyamai kombinasi *Color Histogram* dan *Haralick*. Pemilihan metode klasifikasi pada penelitian ini ditentukan berdasarkan hasil pengujian pada tabel 1 dan tabel 2, dimana metode klasifikasi *Random Forest* selalu unggul dibanding dengan metode klasifikasi lainnya. Berdasarkan hal tersebut maka pada penelitian ini kombinasi metode yang diambil menggunakan metode ekstraksi fitur *Color Histogram*, *Haralick* dan *Random Forest*.

Pengukuran model pada penelitian ini dapat dilihat pula berdasarkan tabel *confusion matrix* untuk melihat seberapa akurat model yang diujikan.

Tabel 3. Confusion Matrix

| | | | | | |
|-------------------------|----------------------|------------------|------------|---------------|---------|
| Predicted Values | Bacterial Blight | 57 | 0 | 0 | 0 |
| | Curl Virus | 2 | 67 | 1 | 0 |
| | Fusarium Wilt | 0 | 0 | 52 | 1 |
| | Healthy | 1 | 0 | 0 | 74 |
| | | Bacterial Blight | Curl Virus | Fusarium Wilt | Healthy |
| | Actual Values | | | | |

Pada tabel *confusion matrix* menunjukkan bahwa hasil yang diperoleh pada penelitian ini termasuk kedalam kategori yang baik dimana setiap jenis kelas yang dapat diprediksi dengan tepat.

Tabel 4. Hasil Pengukuran Model

| Class | Precision | Recall | F1-Score | Support |
|------------------|-----------|--------|----------|---------|
| Bacterial Blight | 1.00 | 1.00 | 1.00 | 67 |
| Curl Virus | 0.97 | 1.00 | 0.98 | 63 |
| Fusarium Wilt | 1.00 | 1.00 | 1.00 | 54 |
| Healthy | 1.00 | 0.97 | 0.99 | 71 |
| Accuracy | | | 0.99 | 255 |
| Macro Avg | 0.99 | 0.99 | 0.99 | 255 |
| Weighted Avg | 0.99 | 0.99 | 0.99 | 255 |

Tabel 4 menunjukkan hasil pengukuran, validasi dan analisis model dalam penelitian infeksi daun kapas. Yang diukur berdasarkan tabel laporan klasifikasi yaitu nilai *precision*, *recall*, *f1-score* dan *support*.

4. KESIMPULAN

Hasil dari penelitian yang dilakukan untuk deteksi infeksi pada daun kapas dengan menggabungkan metode ekstraksi fitur *Haralick* dan *Color Histogram*, serta mengimplementasikan klasifikasi *Random Forest*, menunjukkan bahwa pendekatan ini berhasil meningkatkan akurasi deteksi infeksi pada daun kapas secara signifikan dibandingkan dengan menggunakan metode ekstraksi fitur tunggal. Penggunaan metode ekstraksi fitur *Haralick* memungkinkan pengenalan pola tekstur yang khas pada daun kapas yang terinfeksi, sementara *Color Histogram* memungkinkan identifikasi perubahan warna yang terjadi akibat infeksi. Kombinasi kedua metode ini memungkinkan penanganan yang lebih komprehensif dalam deteksi dan identifikasi penyakit pada daun kapas. Model klasifikasi *Random Forest* yang diimplementasikan dalam penelitian ini berhasil mengenali infeksi pada daun kapas dengan tingkat akurasi yang tinggi. Hasil ini menunjukkan bahwa penggunaan metode *Random Forest* dalam kombinasi dengan ekstraksi fitur *Haralick* dan *Color Histogram* dapat memberikan hasil deteksi yang handal dan akurat. Kesimpulannya, pendekatan yang menggabungkan metode ekstraksi fitur *Haralick* dan *Color Histogram*, serta klasifikasi *Random Forest*, mampu meningkatkan efektivitas deteksi penyakit pada daun kapas. Hasil penelitian memberikan nilai akurasi 0.983501 dan nilai standar deviasi sebesar 0.017540.

REFERENSI

- Agustiani, S., Arifin, Y. T., Junaidi, A., Wildah, S. K., & Mustopa, A. (2022). *Klasifikasi Penyakit Daun Padi menggunakan Random Forest dan Color Histogram*. 10(1).
- Ashir, A. M., Ibrahim, S., Abdulghani, M., Ibrahim, A. A., & Anwar, M. S. (2021). Diabetic Retinopathy Detection Using Local Extrema Quantized Haralick Features with Long Short-Term Memory Network. *International Journal of Biomedical Imaging*, 2021. <https://doi.org/10.1155/2021/6618666>
- Azath, M., Zekiwas, M., & Bruck, A. (2021). Deep Learning-Based Image Processing for Cotton Leaf Disease and Pest Diagnosis. *Journal of Electrical and Computer Engineering*, 2021. <https://doi.org/10.1155/2021/9981437>
- Caldeira, R. F., Santiago, W. E., & Teruel, B. (2021). Identification of cotton leaf lesions using deep learning techniques. *Sensors*, 21(9). <https://doi.org/10.3390/s21093169>
- Ferdita Nugraha, A., Aziza, R. F. A., & Pristyanto, Y. (2022). Penerapan metode Stacking dan Random Forest untuk Meningkatkan Kinerja Klasifikasi pada Proses Deteksi Web Phishing. *Jurnal Infomedia*, 7(1), 39. <https://doi.org/10.30811/jim.v7i1.2959>
- Hamid, H., Abubakar, H., & Jahuddin, R. (2023). Deteksi Dini Cendawan Terbawa Benih Kapas Impor Di Sulawesi Selatan. *Journal Agro Ecotech Indonesia*, 86–92.
- Karthika, J., Mathan Kumar, K., Santhos, M., Sharan, T., & Sri Hariharan, S. (2021). Retraction: Disease Detection in Cotton Leaf Spot Using Image Processing. *Journal of Physics: Conference Series*, 1916(1). <https://doi.org/10.1088/1742-6596/1916/1/012224>
- Khotimatul, S., & Latif, A. (2023). *Kombinasi Metode Fitur Ekstraksi untuk Identifikasi Penyakit pada Daun Teh Combination of Feature Extraction for Identification Disease in Tea Leaves*. 11(3), 447–452. <https://doi.org/10.26418/justin.v11i3>.
- Khultsum, U., & Subekti, A. (2021). Penerapan Algoritma Random Forest dengan Kombinasi Ekstraksi Fitur Untuk Klasifikasi Penyakit Daun Tomat. *Jurnal Media Informatika Budidarma*, 5(1), 186. <https://doi.org/10.30865/mib.v5i1.2624>
- Kumbhar, S., Patil, S., Nilawar, A., Mahalakshmi, B., & Nipane, M. (2019). Farmer Buddy-Web Based Cotton Leaf Disease Detection Using CNN. *International Journal of Applied Engineering Research*, 14(11), 2662–2666. <http://www.ripublication.com>
- Liang, X. (2021). Few-shot cotton leaf spots disease classification based on metric learning. *Plant Methods*, 17(1), 1–11. <https://doi.org/10.1186/s13007-021-00813-7>
- Löfstedt, T., Brynolfsson, P., Asklund, T., Nyholm, T., & Garpebring, A. (2019). Gray-level invariant Haralick texture features. *PLoS ONE*, 14(2), 1–18. <https://doi.org/10.1371/journal.pone.0212110>
- Martina, S., & Abdillah, F. (2020). Pola Komersialisasi Teh Hijau Pada Industri Hospitality : Pandangan Pelaku Usaha. *Sadar Wisata: Jurnal Pariwisata*, 3(1), 1–12. <https://doi.org/10.32528/sw.v3i1.3368>
- Prashar, K., Talwar, R., & Kant, C. (2019a). CNN based on Overlapping Pooling Method and Multi-layered Learning with SVM KNN for American Cotton Leaf Disease Recognition. *2019 International Conference on Automation, Computational and Technology Management, ICACTM 2019*, 330–333. <https://doi.org/10.1109/ICACTM.2019.8776730>
- Prashar, K., Talwar, R., & Kant, C. (2019b). Inconsistent cluster analysis with disease feature enhancement (ICADFE) for American cotton leaf disease recognition. *International Journal of Engineering and Advanced Technology*, 8(5), 1497–1505. <https://doi.org/10.35940/ijeat.E1220.0585C19>
- Sandag, G. A. (2020). Prediksi Rating Aplikasi App Store Menggunakan Algoritma Random Forest. *CogITo Smart Journal*, 6(2), 167–178. <https://doi.org/10.31154/cogito.v6i2.270.167-178>
- Sany, D. S. (2021). Implementasi dan Pengujian Algoritma Fitur Warna pada Image Retrieval. *Media Jurnal Informatika*, 12(1), 1. <https://doi.org/10.35194/mji.v12i1.1187>
- Shah, N., & Jain, S. (2019). Detection of Disease in Cotton Leaf using Artificial Neural Network. *Proceedings - 2019 Amity International Conference on Artificial Intelligence, AICAI 2019*, 473–476. <https://doi.org/10.1109/AICAI.2019.8701311>
- Tantyo, H., Sari, D. K., & Wijaya, A. R. (2023). *Prediksi Potensial Gempa Bumi Indonesia Menggunakan Metode Random Forest Dan Feature Selection*. 6, 83–89.
- Tripathy, S. (2021). Detection of cotton leaf disease using image processing techniques. *Journal of Physics: Conference Series*, 2062(1). <https://doi.org/10.1088/1742-6596/2062/1/012009>