

Identifikasi Penyakit Tanaman Anggur berdasarkan Daunnya menggunakan Naïve Bayes

Identification of Grape Plant Diseases based on the Leaves using Naïve Bayes

¹Muhammad Akbar Ramadhan*, ²Fauzan Nusyura, ³Farah Zakiyah Rahmanti

¹Teknologi Informasi, Fakultas Teknologi Informasi dan Bisnis, ITTelkom Surabaya,

²Teknik Robotika dan Kecerdasan Buatan, Fakultas Teknologi Maju dan Multidisiplin, Universitas Airlangga

³Teknologi Informasi, Fakultas Informatika, Universitas Telkom (Kampus Surabaya)

^{1,3}Jl. Ketintang No.156, Ketintang, Kec. Gayungan, Surabaya, Jawa Timur, Indonesia

²Kampus C UNAIR

*e-mail: makbarramadhan62@gmail.com

(received: 19 August 2024, revised: 29 August 2024, accepted: 4 September 2024)

Abstrak

Salah satu cara melihat ciri penyakit pada tanaman anggur adalah adanya perubahan pada warna daun. Orang awam melakukan pendeteksian penyakit pada tanaman anggur hanya berdasarkan penglihatan secara subyektif saja. Dengan dasar ini, diperlukan sebuah sistem yang dapat membantu orang awam tersebut untuk dapat mendeteksi penyakit pada tanaman anggur berdasarkan warna daunnya menggunakan metode pengklasifikasian algoritma Naïve Bayes. Algoritma ini menggunakan perhitungan yang simpel, sehingga proses yang dilakukan pun menjadi lebih cepat. Pada penelitian ini, pengujian dilakukan menggunakan model klasifikasi Naive Bayes dengan 800 data training dan 160 data validation. Hasil akurasi yang didapatkan sebesar 90% dengan menggunakan skenario color histogram pada channel RGB interval 16 dan GLCM dengan fitur dissimilarity, correleation, homogeneity, contrast jarak piksel 5. Akurasi 90% juga didapatkan pada skenario color histogram pada channel HSV dengan interval 16 dan GLCM dengan fitur dissimilarity, correlation, homogeneity jarak piksel 5. Sehingga, dapat diambil kesimpulan bahwa model klasifikasi Naive Bayes dapat memperoleh penerapan dalam identifikasi penyakit pada tanaman anggur melalui analisis warna daun.

Kata kunci: tanaman anggur, ekstraksi fitur, naïve bayes deteksi objek

Abstract

One way to see the signs of disease in grapevines is a change in leaf color. Ordinary people detect diseases in grapevines only based on subjective vision. On this basis, we need a system that can help the layman to be able to detect diseases in grapevines based on the color of the leaves using the Naïve Bayes algorithm classification method. This algorithm uses simple calculations, so the process is carried out faster. In this study, testing was carried out using the Naive Bayes classification model with 800 training data and 160 validation data. The accuracy results obtained are 90% using the color histogram scenario on channel RGB interval 16 and GLCM with features of dissimilarity, correlation, homogeneity, contrast pixel spacing 5. 90% accuracy is also obtained in the color histogram scenario on channel HSV with interval 16 and GLCM with features of dissimilarity, correlation, homogeneity at pixel spacing of 5. Thus, it can be concluded that the Naive Bayes classification model can gain application in identifying diseases in grapevines through leaf color analysis.

Keywords: grape vine, feature extraction, naïve bayes, object detection

1 Pendahuluan

Salah satu buah di Indonesia yang banyak digemari adalah buah anggur. Cita rasa buah anggur yang manis dan cara makan yang mudah karena tidak harus mengupas kulitnya menjadi salah satu

daya tarik tersendiri dari buah ini. Buah anggur memiliki sejumlah manfaat, seperti pencegahan penyakit kanker, pengurangan tanda-tanda dari insufisiensi vena kronis, peningkatan memori, perlindungan terhadap retina mata, penurunan tekanan hipertensi, dan perlambatan proses penuaan [1]. Dari sekian banyaknya manfaat yang dihasilkan oleh buah anggur, tanaman dari buah anggur sendiri memiliki kelemahan terhadap serangan hama yang dapat menimbulkan gangguan dalam pertumbuhan tanaman buah anggur [2]. Tentunya hal tersebut akan mempengaruhi kualitas dari buah yang dihasilkan. Salah satu cara melihat ciri penyakit tanaman anggur adalah perubahan pada daunnya. Umumnya, pengamatan visual dengan mata saja memiliki keterbatasan karena sulit untuk membedakan berdasarkan karakteristik tekstur dan warna pada daun tersebut [3]. Akibatnya, proses penanggulangan yang dilakukan menjadi kurang tepat sehingga menyebabkan terjadinya kerusakan pada tanaman anggur tersebut.

Di era perkembangan teknologi yang tergolong cukup pesat ini, terdapat banyak sekali penerapan teknologi yang dapat digunakan untuk dapat mempermudah manusia dalam mendapatkan solusi dari masalah yang ada termasuk pertanian. Salah satunya adalah pengolahan gambar. Pengolahan gambar dapat membantu manusia dalam mengatasi masalah yang berhubungan dengan visual seperti warna dan tekstur. Pengolahan citra biasanya digunakan secara bersamaan dengan algoritma pembelajaran mesin sebagai pendukung dalam pengambilan keputusan yang akan diinformasikan kepada *user*.

Metode klasifikasi algoritma *Naïve Bayes* menerapkan prinsip probabilitas bersyarat dalam mengambil keputusan, untuk menghitung kemungkinan suatu objek akan diklasifikasikan ke dalam salah satu kelas yang ada [4]. Algoritma ini akan memberikan sebuah keputusan berupa label (*target*) pada sebuah objek dengan melakukan pendekatan kriteria berdasarkan data *training* yang tersedia dengan nilai probabilitas tertinggi. Penerapan algoritma *Naïve Bayes* menggunakan data *training* dan data *validation* yang kemudian akan digunakan dalam proses *training* untuk menentukan label dari sebuah objek berdasarkan kriteria yang digunakan [5]. Keuntungan dari penggunaan algoritma *Naïve Bayes* yaitu Algoritma ini menggunakan perhitungan yang sederhana, sehingga proses komputasi yang dilakukan pun menjadi lebih cepat.

2 Tinjauan Literatur

Studi literatur yang dikumpulkan berfungsi sebagai acuan dan menambah referensi dalam hal teori serta pengaplikasian metode yang digunakan. Tinjauan Literatur yang dipakai pada penelitian ini yaitu:

2.2.1 Tanaman Anggur

Anggur merupakan tanaman yang umumnya tumbuh di dataran rendah. Proses perawatan tanaman anggur pun bisa dibilang cukup berbeda dengan tanaman lainnya. Jika tanaman lainnya membutuhkan banyak air untuk dapat membantu proses pembuahannya, sedangkan penyiraman air yang berlebihan pada tanaman anggur akan mengganggu proses pembuahannya. Kesalahan dalam proses perawatan tanaman anggur membuat tanaman ini menjadi rentan untuk terinfeksi suatu penyakit. Contoh penyakit tanaman anggur yang dapat dideteksi melalui daunnya antara lain sebagai berikut [6].

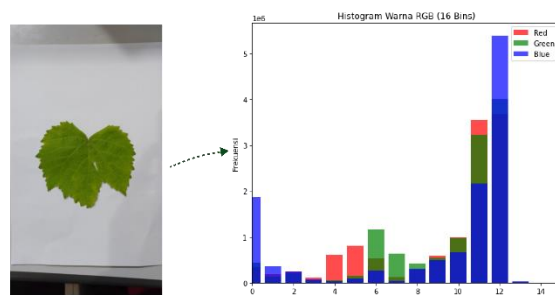
- **Bercak Daun** merupakan penyakit yang gejalanya dapat dilihat pada bagian atas daun terdapat area yang memiliki bercak-bercak berwarna hijau kekuningan.
- **Bercak Merah** memiliki gejala yaitu sebagian daun hijau yang mendadak mati dan berubah warna menjadi kecoklatan serta jaringan sekitarnya berubah menjadi kuning atau merah. Bagian daun yang kering akan menyebar hingga keseluruhan daun bisa mengerut dan rontok.
- **Embun Tepung Palsu** memiliki gejala yaitu serangan jamur yang mempengaruhi daun, tunas, dan buah muda. Serangan pada tunas akan mengakibatkan kering dan patahnya tunas. Serangan pada sisi atas daun ditandai dengan adanya bercak kuning-kehijauan yang tidak terbatas dengan jelas, dan jika buah muda diserang, buah akan membusuk.
- **Hama Tungau Merah** memiliki gejala yang terlihat yaitu serangan hama yang umumnya menghisap cairan daun tanaman anggur. Hama ini cenderung menyebabkan bercak-bercak pada daun dan dapat mengubahnya menjadi warna hitam.

2.2.2 Pengolahan Citra Digital

Pada dasarnya, pengolahan citra digital adalah manipulasi gambar dalam format dua dimensi yang dilakukan dengan bantuan komputer. Gambar yang ditangkap oleh kamera akan diproses

<http://sistemasi.ftik.unisi.ac.id>

menggunakan komputer berdasarkan piksel-piksel yang ada. Kebanyakan citra yang didapat melalui kamera menggunakan warna dasar yaitu RGB. Model warna RGB adalah warna yang biasa kita lihat sehari-hari menggunakan mata secara langsung, model warna ini juga merupakan model warna dasar



Gambar 1. Contoh histogram warna BIN 16

bagi model warna lain yang berisi saluran warna RGB. Sedangkan, model warna HSV merupakan transformasi dari kubus warna RGB [7].

2.2.3 Histogram Warna

Histogram adalah diagram vertikal yang merepresentasikan distribusi tingkat intensitas suatu piksel pada gambar, berdasarkan format warna yang digunakan [8] [9]. Histogram warna merupakan sebuah metode yang digunakan untuk melakukan pendeskripsian konten warna terhadap suatu citra dengan menghitung jumlah kemuculan atau frekuensi dari setiap warna [10]. Sebagai upaya dalam mengatasi kompleksitas perhitungan pada histogram warna channel RGB dan HSV, dilakukan penggolongan data dengan menggunakan parameter BIN. Setiap nilai BIN akan menghitung jumlah piksel dalam interval tertentu. Nilai BIN yang sering digunakan adalah 16 karena memberikan representasi yang ringkas dan mudah untuk dianalisa, dengan interval interval 0-15, 16-31, 32-47, dan seterusnya hingga 240-255, sehingga informasi dalam histogram terangkum dengan baik. Contoh dari histogram warna dengan BIN 16 dapat ditemukan pada Gambar 1.

2.2.4 Gray Level Co-occurrence Matrix (GLCM)

Penggunaan metode GLCM (*Gray Level Co-occurrence Matrix*) dengan tujuan memberikan informasi tambahan kepada model algoritma klasifikasi, sehingga model dapat memperoleh detail lebih dari citra yang sedang diolah. Tekstur adalah atribut yang digunakan untuk mengidentifikasi area yang dimaksud dalam suatu citra [11]. Fitur GLCM akan melakukan perhitungan terhadap hubungan spasial antara dua buah piksel dengan tingkat keabuan tertentu, dalam arah dan jarak yang ditentukan, pada suatu citra [12]. Pada dasarnya, metode algoritma GLCM digunakan untuk menghitung sejauh mana kemiripan antara dua piksel, seperti kemiripan nilai piksel antara dua piksel, dalam jarak (d) dan arah sudut yang ditentukan [13]. Arah sudut yang biasa digunakan dalam metode GLCM yaitu 0°, 45°, 90°, dan 135°. Sementara itu, jarak antara piksel yang ditentukan biasanya adalah 1 piksel. Tahapan ekstraksi fitur menggunakan metode GLCM mencakup konversi citra ke grayscale, penentuan sudut dan jarak piksel untuk matriks kookurensi, lalu perhitungan nilai fitur dari matriks tersebut [14]. Beberapa fitur yang dimiliki oleh GLCM adalah sebagai berikut [15].

a. Dissimilarity

Dissimilarity digunakan untuk mengukur kemiripan dari suatu teksur, nilai dari dissimilarity akan menjadi besar teksur tidak seragam dan akan menjadi kecil bila teksur seragam. Berikut adalah rumus dari dissimilarity.

b. Correlation

Correlation digunakan untuk mengukur keterhubungan linear satu pixel terhadap pixel lainnya. Berikut adalah rumus dari correlation.

c. Homogeneity

Homogeneity digunakan untuk mengukur tingkat kesamaan sebuah citra pada tingkat keabuan. Berikut adalah rumus dari homogeneity.

d. Contrast

Contrast digunakan untuk mengukur perbedaan intensitas antara terang dan gelap pada pixel yang saling berdekatan. Berikut adalah rumus dari contrast.

2.2.5 Algoritma Naïve Bayes

Algoritma Naïve Bayes adalah teknik klasifikasi dengan konsep statistik. Algoritma ini menggunakan konsep probabilitas untuk mendistribusikan data ke dalam label atau kategori tertentu. Algoritma ini juga mengacu pada kemampuan untuk meramalkan kejadian di masa mendatang berdasarkan informasi dan pengetahuan dari kejadian sebelumnya. Oleh karena itu, algoritma ini dikenal sebagai Teorema Bayes [16]. Salah satu keunggulan dari penggunaan *Naive Bayes* adalah hanya memerlukan sedikit data untuk melakukan proses *training* dalam mengestimasi nilai yang diperlukan pada proses klasifikasi [17]. Selain itu, algoritma ini juga memiliki kelemahan yaitu ketika probabilitas kondisional suatu data bernilai nol, maka hasil kemungkinan prediksi pun akan nol juga. Persamaan dari algoritma *naïve bayes* terhadap data kontinu memiliki parameter utama yaitu rata-rata / mean (μ) dan varian (σ^2).

2.2.6 Confussion Matrix

Confussion Matrix digunakan sebagai metode dalam melihat kinerja model klasifikasi dengan membandingkan benar atau salah dari hasil prediksi dengan label yang sebenarnya [18]. Beberapa acuan pengukuran yang biasa digunakan yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-Score* [19]. *Accuracy* adalah penggambaran seberapa dekat suatu pengukuran dengan nilai yang diterima, yang dapat dipengaruhi oleh kesalahan sistematis yang disebabkan oleh instrumen pengukuran dan prosedur pengukuran. *Precision* adalah hasil pengukuran yang berulang kali dilakukan dalam kondisi yang tidak berubah, yang menunjukkan hasil yang konsisten dan seragam. *Recall* adalah ukuran tingkat kesuksesan sistem saat mengenali kembali suatu informasi dengan benar, baik dalam hal data negatif maupun positif [20]. Terakhir, *F1-Score* atau sering juga disebut sebagai *f-measure* menyajikan hasil prediksi dengan mengukur *Accuracy*, *Precision*, dan *Recall* [21]. Berikut adalah rumus yang digunakan pada penelitian ini sebagai acuan dalam menghitung *accuracy*, *precision*, *recall*, dan *F1-score*.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1 - Score = \frac{2*(Recall*Precision)}{Recall + Precision} \quad (4)$$

3 Metode Penelitian

Tahapan pada metode penelitian ini adalah sebagai berikut:

1. Pengumpulan Data

Penelitian ini menggunakan data yang didapat pada saat melakukan survey pada balai KP Banjarsari. Terdapat beberapa ketentuan dalam pengambilan citra yaitu, dilakukan pada siang hari agar meminimalisir *noise*. Selain itu, pengambilan citra dilakukan menggunakan *smartphone* android dengan jarak kurang lebih 25 cm dan *background* berwarna putih agar kedalaman gambar yang didapat masih tajam dan tidak pecah pecah ketika dilakukan proses *pre-processing*. *Background* berwarna putih digunakan agar fitur yang nantinya akan diesktraksi tidak banyak terkontaminasi oleh *background* yang ikut terambil.

2. Pemrosesan Data

Pada tahap ini, dilakukan seleksi citra dari hasil pengumpulan data agar hanya citra berkualitas yang masuk ke proses *preprocessing*, dengan validasi sesuai label dari KP Banjarsari untuk memastikan kesesuaian dengan ciri penyakit sebenarnya. Ciri penyakit yang dilihat dan disesuaikan pada saat proses validasi ada 5 (lima) yaitu bercak daun, bercak merah, daun sehat, embun tepung palsu dan hama tungau merah.

3. Preprocessing Data

Proses *preprocessing* dilakukan agar tiap data citra memiliki karakteristik yang sama. Sehingga model dapat mempelajari karakteristik masing masing data dengan lebih mudah. Tahapan pada proses

preprocessing data diawali dari data citra asli yang dinormalisasi cahaya, kemudian *cropping* dan *resize* menjadi 1080 x 1080px.

4. Ekstraksi Fitur

Metode ekstraksi fitur yang diimplementasikan adalah metode histogram warna pada channel RGB dan HSV untuk fitur warna dan metode *Gray Level Co-occurrence Matrix* (GLCM) untuk mengambil fitur tekstur dengan parameter *dissimilarity*, *correlation*, *homogeneity* dan *contrast* jarak 1 hingga 5 dan arah derajat 0°, 45°, 90°, dan 135°.

5. Implementasi Naïve Bayes

Metode klasifikasi yang diterapkan dalam penelitian ini adalah Naïve Bayes, salah satu contoh algoritma *supervised learning*. Algoritma ini dipilih karena menggunakan perhitungan yang sederhana, sehingga proses komputasi yang dilakukan menjadi lebih cepat.

6. Evaluasi

Metode evaluasi yang diterapkan adalah *confussion matrix*. Metode ini digunakan untuk mengukur performa dari model yang dibuat. Parameter yang digunakan untuk melihat kinerja dari model klasifikasi dengan metode *confussion matrix* adalah akurasi, presisi, *recall*, dan *f1-score*.

4 Hasil dan Pembahasan

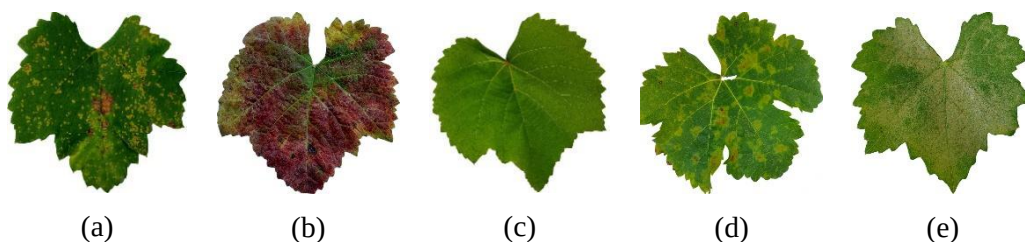
Hasil dari klasifikasi penyakit tanaman anggur menggunakan ekstraksi fitur warna menggunakan histogram warna dan tekstur menggunakan GLCM didapatkan dengan tahapan sebagai berikut.

1. Hasil Pengumpulan Data

Data yang didapatkan pada proses pengumpulan data sebanyak 218 data. Data yang didapat kemudian akan melalui tahap pemrosesan data sebelum memasuki proses *preprocessing*.

2. Pemrosesan data

Pada proses ini terdapat tahap validasi terhadap data citra yang akan di seleksi oleh pihak KP Banjarsari sesuai dengan label yang digunakan agar dataset dengan labelnya sesuai dengan ciri penyakit yang sebenarnya. Contoh gambar pada masing masing label dapat dilihat pada Gambar 2. Jumlah citra yang telah proses validasi berjumlah 150. Sebanyak 68 data tidak masuk validasi karena data tersebut tidak sesuai dengan ciri penyakit pada label yang digunakan. Karena data yang tersisa masih kurang untuk melakukan proses training pada model klasifikasi, maka beberapa data yang ada tersebut akan melalui proses augmentasi untuk membuat data citra baru dengan cara melakukan rotasi sebanyak tujuh derajat yaitu 45°, 90°, 135°, 180°, 225°, 270°, 315°. Setelah proses augmentasi tersebut, didapatkan total dataset akhir yang berjumlah 960 data.



Gambar 2. Contoh citra yang terprediksi dengan benar (a) label bercak daun, (b) label bercak merah, (c) label daun sehat, (d) label embun tepung palsu dan (e) label hama tungau merah



Gambar 3. Sebelum *preprocessing*



Gambar 4. Setelah *preprocessing*

Tabel 1. Hasil ekstraksi fitur pada CSV

Label	hist_r1	hist_r2	hist_r3	hist_r4	hist_r5	hist_r6		contrast 0	contrast 45	contrast 90	contrast 135
1	0.0016	0.390	0.2077	0.1466	0.0639	0.357	...	0.0005	0.0005	0.0005	0.0006
1	0.0158	0.2711	0.2122	0.0501	0.0183	0.0093	...	0.0003	0.0004	0.0003	0.0004
1	0.1033	0.1990	0.1743	0.0806	0.0392	0.0197	...	0.0004	0.0005	0.0004	0.0004
1	0.1032	0.1990	0.1744	0.0806	0.0392	0.0198	...	0.0004	0.0004	0.00004	0.0005
1	0.1165	0.2086	0.1782	0.0861	0.0430	0.0218	...	0.0004	0.0005	0.0005	0.0005
...	...	...	...	...	...	...	...	...	...	...	...
5	0.0008	0.0250	0.1139	0.1782	0.1809	0.1271	...	0.0006	0.0006	0.0005	0.0006
5	0.0000	0.0007	0.0128	0.0611	0.1329	0.1547	...	0.0005	0.0006	0.0006	0.0006
5	0.0052	0.1275	0.2259	0.1611	0.0886	0.0571	...	0.0006	0.0006	0.0005	0.0006
5	0.0034	0.1063	0.1985	0.1426	0.0781	0.0493	...	0.0004	0.0005	0.0004	0.0005
5	0.0034	0.1063	0.1982	0.1430	0.0780	0.0949	...	0.0004	0.0005	0.0004	0.0005

3. Preprocessing Data

Tahapan *preprocessing* data dilakukan agar data yang diolah memiliki karakteristik yang sama. Citra daun yang telah terseleksi di tahap pemrosesan data akan melalui proses normalisasi pada bagian kecerahan, agar tidak ada citra yang terlalu gelap atau terlalu terang sebelum memasuki proses *cropping*. Setelah itu, citra akan melalui proses *cropping* menggunakan metode ROI (*Region of interest*) agar objek daun dapat terseleksi dengan lebih baik dan meminimalisir adanya background atau objek lain yang mengganggu. Setelah itu, hasil dari *cropping* tersebut akan melewati proses perubahan pada bagian *background* menjadi warna putih untuk menyeragamkan background data baru sesuai dengan yang ada pada data *training*. Hasil dari proses perubahan *background* menjadi putih tersebut akan melalui proses *resize* menjadi 1080 x 1080px agar memiliki ukuran pixel yang sama. Visualisasi gambar sebelum menjalani proses *preprocessing* dapat ditemukan pada Gambar 3, sedangkan visualisasi gambar yang sudah dilakukan proses *preprocessing* dapat ditemukan pada Gambar 4.

4. Ekstraksi Fitur

Pemrosesan fitur melalui metode histogram warna pada channel RGB dan HSV untuk fitur warna dan metode GLCM untuk fitur tekstur sebagai pembanding. Dalam proses ekstraksi fitur, citra yang telah melalui *preprocessing* mengalami tiga tahapan ekstraksi yang berbeda sesuai dengan skenario pengujian. Pada ekstraksi warna RGB, citra dikonversi dari BGR ke RGB, kemudian histogram warna dihitung. Pada ekstraksi warna HSV, citra diubah dari RGB menjadi HSV, kemudian histogram warna dihitung. Pada ekstraksi fitur GLCM, citra RGB dikonversi menjadi citra Gray atau abu-abu sebelum ekstraksi fitur GLCM. kemudian, data fitur akan melalui proses normalisasi agar memiliki skala yang sama, dan disimpan dalam file CSV sesuai skenario yang ditentukan. Hasil ekstraksi fitur yang telah di simpan di csv dapat ditemukan pada Tabel 1.

5. Klasifikasi menggunakan Naïve Bayes

Skenario pengujian dilakukan agar mendapatkan hasil yang terbaik. Skenario yang digunakan dalam pengujian yaitu metode *Color Histogram*, metode GLCM dan penggabungan metode *Color Histogram* dengan metode GLCM. Seluruh skenario pada penelitian ini diuji dengan algoritma klasifikasi *Naïve Bayes* menggunakan 800 data *training* (160 data pada masing masing label) dan 160 data *validation* (32 data pada masing masing label).

1. Hasil Pengujian Parameter Tuning Metode *Color Histogram*

Pada skenario pertama, dilakukan percobaan yang bertujuan untuk mendapatkan parameter terbaik dari metode *Color Histogram* dengan dua channel warna yaitu RGB dan HSV. Terdapat dua parameter yang digunakan pada pengujian metode ini yaitu *channel* dan *interval*. Parameter *channel* yang digunakan adalah RGB dan HSV dengan nilai *interval* yang digunakan adalah 16. Hasil dari parameter *tuning color histogram* dilihat pada Tabel 2.

Hasil dari parameter *tuning* dengan akurasi tertinggi yaitu 90% didapatkan pada *channel* HSV dengan menggunakan *interval* 16. Sedangkan pada *channel* RGB, nilai akurasi tertinggi yang didapatkan hanya sebesar 88%. Hal ini juga didukung dengan atribut evaluasi model yaitu *precision*, *recall*, dan *F1-core* pada *channel* HSV yaitu sebesar 92% untuk *precision*, 90% untuk *recall*, dan 90% untuk *F1-Score*. Sedangkan pada *channel* RGB mendapatkan nilai sebesar 90% untuk *precision*, 88% untuk *recall*, dan 88% untuk *F1-Score*. Berdasarkan hasil tersebut, dapat dilihat bahwa pada *channel* HSV mendapatkan hasil cukup baik pada identifikasi terhadap data yang digunakan pada penelitian ini daripada *channel* RGB.

2. Hasil Pengujian Parameter Tuning Metode GLCM

Pada skenario kedua, dilakukan percobaan yang bertujuan untuk mendapatkan parameter terbaik dari metode GLCM. Terdapat dua parameter yang digunakan pada pengujian metode ini yaitu jarak dan fitur dengan menggunakan arah sudut 0°, 45°, 90°, 135° pada seluruh percobaan. Parameter fitur yang digunakan adalah *dissimilarity*, *correlation*, *homogeneity*, dan *contrast* dengan menggunakan jarak piksel 1 sampai 5. Hasil dari parameter *tuning* GLCM dapat dilihat pada

Tabel 3.

Hasil akurasi tertinggi yang didapatkan pada percobaan metode GLCM ke 5 dengan nilai akurasi sebesar 51%, *precision* senilai 57%, *recall* 51%, dan *F1-Score* 50%. Berdasarkan analisis tersebut, dapat disimpulkan bahwa skenario metode GLCM yang memiliki kinerja cukup baik dengan menggunakan 4 atribut GLCM *Dissimilarity*, *Correlation*, *Homogeneity*, *Contrast* dan jarak piksel 5 karena memiliki nilai *precision* dan *F1-Score* yang lebih tinggi, walaupun memiliki selisih 2% lebih sedikit pada nilai akurasi dan *recall*.

3. Hasil Pengujian Penggabungan Metode *Color Histogram* Dengan Metode GLCM

Pada skenario ketiga, dilakukan penggabungan dari kedua metode yaitu *Color Histogram* dan GLCM. Parameter yang digunakan pada penggabungan metode ini adalah parameter terbaik pada metode *Color Histogram* yang telah didapat pada skenario pengujian sebelumnya dan parameter pada pengujian metode GLCM. Hasil dari percobaan penggabungan kedua metode dapat dilihat pada Tabel 4 untuk skenario *Color Histogram* HSV dan Tabel 5 untuk skenario *Color Histogram* RGB.

Tabel 2. Hasil paramter tuning color histogram

Channel	Interval	Akurasi	Precision	Recall	F1-Score
H, S, V	16	90%	92%	90%	90%
R, G, B	16	88%	90%	88%	88%

Tabel 3. Hasil parameter tuning GLCM

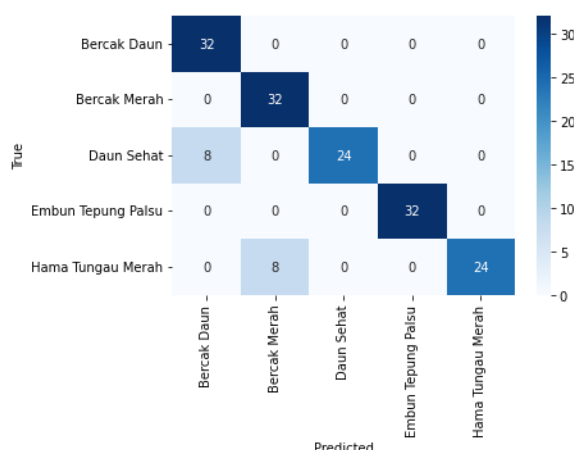
No	Fitur GLCM	Jarak	Akurasi	Precision	Recall	F1-Score
1	Dissimilarity, Correlation, Homogeneity, Contrast	1	37%	54%	37%	36%
2		2	39%	46%	39%	38%
3		3	40%	44%	40%	40%
4		4	50%	53%	50%	49%
5		5	51%	57%	51%	50%

Tabel 4. Hasil penggabungan metode color histogram HSV dan GLCM

No	Fitur GLCM	Jarak	Akurasi	Precision	Recall	F1-Score
1	Dissimilarity, Correlation, Homogeneity, Contrast	1	88%	90%	88%	87%
2		2	88%	90%	88%	87%
3		3	88%	90%	88%	87%
4		4	85%	88%	85%	84%
5		5	85%	88%	85%	84%

Tabel 5. Hasil penggabungan metode color histogram RGB dan GLCM

No	Fitur GLCM	Jarak	Akurasi	Precision	Recall	F1-Score
1	Dissimilarity, Correlation, Homogeneity, Contrast	1	89%	92%	89%	89%
2		2	89%	92%	89%	89%
3		3	90%	92%	90%	90%
4		4	90%	92%	90%	90%
5		5	90%	92%	90%	90%



Gambar 5. Confussion matrix

Tabel 6. Hasil evaluasi tiap label

Label	Jumlah Data Uji	Precision	Recall	F1-Score
Bercak Daun	32	80%	100%	89%
Bercak Merah	32	80%	100%	89%
Daun Sehat	32	100%	75%	86%
Embun Tepung Palsu	32	100%	100%	100%
Hama Tungau Merah	32	100%	75%	86%

Pada hasil percobaan penggabungan metode *Color Histogram* HSV dan metode GLCM, dapat dilihat bahwa nilai akurasi tertinggi yang didapatkan adalah 88% dengan *precision* 90%, *recall* 88%, dan *F1-Score* 87%. Sedangkan, pada percobaan penggabungan metode *Color Histogram* RGB dan metode GLCM, dapat dilihat bahwa nilai akurasi tertinggi yang didapatkan adalah 90% dengan *precision* 92%, *recall* 90%, dan *F1-Score* 90%. Berdasarkan analisa dari hasil penggabungan metode *Color Histogram* dan metode GLCM menunjukkan bahwa penggabungan fitur warna RGB memiliki kinerja yang lebih baik jika digabungkan dengan metode GLCM daripada fitur warna HSV pada penelitian ini.

6. Evaluasi Model

Berdasarkan hasil pengujian yang telah dilakukan, skenario terbaik yang didapatkan adalah *color histogram* RGB dengan interval 16 dan GLCM dengan fitur *dissimilarity*, *correlation*, *homogeneity*,

<http://sistemasi.ftik.unisi.ac.id>

dan *contrast* dengan jarak piksel 5. Hal ini karena skenario tersebut memiliki jumlah vektor yang paling panjang yaitu 64 sehingga pola ekstraksi fitur yang dihasilkan dapat memberikan informasi lebih yang dibutuhkan agar performa klasifikasi menjadi lebih baik. Nilai akurasi yang didapatkan sebesar 90%, *precision* 92%, *recall* 90% dan *F1-Score* 90%. *Confussion matrix* dari hasil proses pengujian pada skenario ini dapat dilihat pada Gambar 5.

Berdasarkan *confussion matrix* tersebut, citra yang teridentifikasi benar sebanyak 144 citra. Citra yang teridentifikasi dengan benar adalah 32 citra label bercak daun, 32 citra label bercak merah, 24 citra label daun sehat, 32 citra label embun tepung palsu dan 24 citra label hama tungau merah. Sedangkan, citra yang teridentifikasi salah adalah 8 citra label daun sehat yang teridentifikasi label bercak daun dan 8 citra label hama tungau merah yang teridentifikasi label bercak merah. Berdasarkan pada Tabel 6 dapat dilihat bahwa label yang memiliki nilai evaluasi terbaik adalah label embun tepung palsu dengan nilai *precision*, *recall* dan *F1-Score* sebesar 100%.

5 Kesimpulan

Penelitian ini menggunakan algoritma Naïve bayes dengan 800 data *training* dan 160 data *validation*. Pengujian pada histogram warna *channel* HSV mendapatkan nilai akurasi sebesar 90%, *precision* senilai 92%, *recall* senilai 90% dan *F1-Score* senilai 90%. Sedangkan, pada *channel* RGB menghasilkan akurasi senilai 88%, *precision* senilai 90%, *recall* senilai 88% dan *F1-Score* senilai 88%. Pengujian terhadap metode GLCM mendapatkan hasil terbaik nilai akurasi sebesar 51%, *precision* 57%, *recall* 51%, dan *F1-Score* 50%. Pengujian selanjutnya dilakukan dengan menggabungkan ekstraksi fitur histogram warna dan GLCM mendapatkan akurasi sebesar 90%, *precision* 92%, *recall* 90%, dan *F1-Score* 90% pada *channel* warna RGB. Saran untuk penelitian kedepannya adalah objek atau label pada sistem ini dapat ditambahkan menjadi lebih banyak lagi, optimalisasi *preprocessing* agar dapat menghilangkan *background random* dengan optimal selain satu warna saja.

Referensi

- [1] W. Saputro dan D. B. Sumantri, "Implementasi Citra Digital Dalam Klasifikasi Jenis Buah Anggur Dengan Algoritma K-Nearest Neighbors (KNN) Dan Data Augmentasi," *Journal of Information Technology and Computer Science (INTECOMS)*, vol. 5, no. 2, pp. 248-253, 2022.
- [2] R. Apriyanto dan M. Ahsan, "Sistem Analisis Diagnosa Penyakit Tanaman Anggur dengan Pendekatan Certainty Factor Berbasis Android," *Jurnal Teknologi Informasi dan Industri*, vol. 2, 2019.
- [3] S. S. Simanjuntak, H. Sinaga, K. Telaumbanua dan Andri, "Klasifikasi Penyakit Daun Anggur Menggunakan Metode GLCM, Color Moment dan K*Tree," *Jurnal SIFO Mikroskil*, vol. 21, no. 2, pp. 93-104, 2020.
- [4] R. R. Waliyansyah dan C. Fitriyah, "Perbandingan Akurasi Klasifikasi Citra Kayu Jati Menggunakan Metode Naive Bayes dan K-Nearest Neighbor (KNN)," *Jurnal Edukasi dan Penelitian Informatika*, vol. 5, no. 2, pp. 157-163, 2019.
- [5] I. S. Manuel dan I. Ernawati, "Implementasi GLCM dan Algoritma Naive Bayes dalam Klasifikasi Jenis Bunga Anggrek," *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, pp. 99-109, 2020.
- [6] "cybex.pertanian.go.id," Kementrian Pertanian, 08 2019. [Online]. Available: <http://cybex.pertanian.go.id/mobile/artikel/70375/Pengendalian-Hama-Dan-Penyakit--Pada-Tanaman-Anggur/>. [Diakses 12 06 2023].
- [7] M. Z. Andrekha dan Y. Huda, "Deteksi Warna Manggis Menggunakan Pengolahan Citra dengan

- Opencv Python,” *Jurnal Vocational Teknik Elektronika dan Informatika*, vol. 9, no. 4, pp. 27-33, 2021.
- [8] H. Syahputra, F. Arnia dan K. Munadi, “Karakterisasi Kematangan Buah Kopi Berdasarkan Warna Kulit Kopi Menggunakan Histogram dan Momen Warna,” *Jurnal Nasional Teknik Elektro*, vol. 8, no. 1, pp. 42-50, 2019.
- [9] S. Ratna, “Pengolahan Citra Digital dan Histogram dengan Python dan Text Editor Phycharm,” *Technologia*, vol. 11, no. 3, pp. 181-186, 2020.
- [10] S. Y. E. Simarmata, Y. A. Sari dan S. Adinugroho, “Klasifikasi Citra Makanan Menggunakan Algoritme Learning Vector Quantization Berdasarkan Ekstraksi Fitur Color Histogram dan Gray Level Co-occurrence Matrix,” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 3, pp. 2369 - 2378, 2019.
- [11] F. K. Fikriah, M. B. Sulthan, N. Mujahidah dan M. K. Roziqin, “Naïve Bayes untuk Klasifikasi Penyakit Daun Bawang Merah Berdasarkan Ekstraksi Fitur Gray Level Co-occurrence Matrix (GLCM),” *Jurnal Komtika (Komputasi dan Informatika)*, vol. 6, no. 2, pp. 133-141, 2022.
- [12] Z. Y. Lamasgi, Serwin, Y. Lasena dan Husdi, “Identifikasi Tingkat Kesegaran Ikan Tuna menggunakan Metode GLCM dan KNN,” *Jambura Journal of Electrical and Electronics Engineering*, vol. 4, no. 1, pp. 70-76, 2022.
- [13] E. H. Rachmawanto dan . H. P. Hadi, “Optimasi Ekstraksi Fitur pada Kini dalam Klasifikasi Penyakit Daun Jagung,” *DINAMIK*, vol. 22, no. 2, pp. 58-67, 2021.
- [14] H. Sulaiman, D. Riana dan A. Rifai, “Perbandingan Algoritma Decision Tree C4.5 dan Naive Bayes pada Analisis Tekstur Gray Level Co-occurrence Matrix Menggunakan Citra Wajah,” *Jurnal Sistem Informasi*, vol. 10, no. 2, pp. 470-479, 2021.
- [15] L. Hakim, S. P. Kristanto, D. Yusuf dan F. N. Afia, “Pengenalan Motif Batik Banyuwangi berdasarkan Fitur Gray Level Co-Occurence Matrix,” *Jurnal TEKNOINFO*, vol. 16, no. 1, pp. 1-7, 2022.
- [16] D. Alita, I. Sari, A. R. Isnain dan Styawati, “Penerapan Naive Bayes Classifier untuk Pendukung Keputusan Penerima Beasiswa,” *JDMSI*, vol. 2, no. 1, pp. 17-23, 2021.
- [17] M. F. Rifai, H. Jatnika dan B. Valentino, “Penerapan Algoritma Naïve Bayes Pada Sistem Prediksi Tingkat Kelulusan Peserta Sertifikasi Microsoft Office Specialist (MOS),” *PETIR: Jurnal Pengkajian dan Penerapan Teknik Informatika*, vol. 12, no. 2, pp. 131 - 144, 2019.
- [18] A. S. Sastrawan, I. G. A. Gunadi dan I. N. Sukajaya, “Perbandingan Kinerja Algoritma Dempster Shafer dan Fuzzy Naive Bayes dalam Klasifikasi Penyakit Demam Berdarah dan Tifus,” *Jurnal Ilmu Komputer Indonesia (JIK)*, vol. 4, no. 2, pp. 24-32, 2019.
- [19] E. Fitriani, “Perbandingan Algoritma C4.5 dan Naive Bayes untuk Menentukan Kelayakan Penerima Bantuan Program Keluarga Harapan,” *Sistemasi: Jurnal Sistem Informasi*, vol. 9, pp. 103-115, 2020.
- [20] H. Azis, F. T. Admojo dan E. Susanti, “Analisis Perbandingan Performa Metode Klasifikasi pada Dataset Multiclass Citra Busur Panah,” *Techno.COM*, vol. 19, no. 3, pp. 286-294, 2020.

- [21] I. M. Erwin, Risnandar, E. Prakasa dan B. Sugiarto, "Identifikasi dan Evaluasi F-Measure Citra Kayu Berbasis Deep Convolutional Neural Network (DCNN)," *Jurnal Teknologi Informasi dan Ilmu Komputer (JTIIK)*, vol. 7, no. 6, pp. 1089 - 1098, 2019.