

BAB III

METODE PENELITIAN

3.1. Objek Penelitian

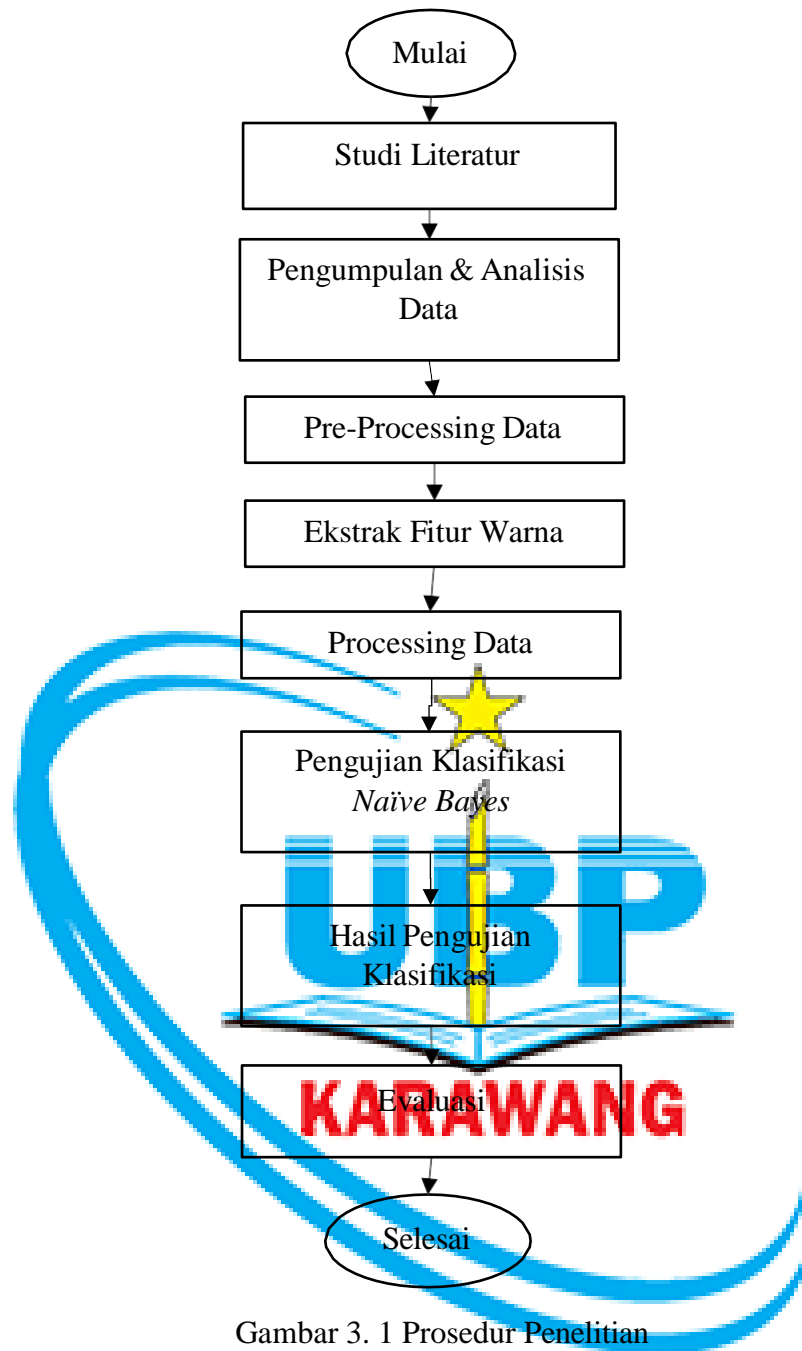
Objek dalam penelitian ini merupakan buah jambu bol dengan tingkat kematangan mentah, setengah matang dan matang. Pada tingkat kematangan buah jambu bol mentah yaitu memiliki ukuran yang lebih kecil dan berwarna hijau. Sedangkan, jambu bol yang memiliki tingkat kematangan setengah matang memiliki ukuran yang semakin besar dan berwarna merah cerah. Sementara itu, tingkat kematangan buah jambu bol yang sudah matang memiliki ukuran lebih besar dengan daging buah yang tebal dan warnanya berubah menjadi merah tua. Lokasi penelitian dilakukan di Laboratorium Riset Universitas Buana Perjuangan Karawang. Waktu Penelitian pada bulan Maret sampai dengan Agustus 2024.

Tabel 3. 1 Tabel Waktu Penelitian

No	Item	Bulan 1				Bulan 2				Bulan 3				Bulan 4				Bulan 5			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1.	Studi literatur																				
2.	Pengumpulan Data dan Analisis Data																				
3.	Pra-Processing Data																				
4.	Ekstra Fitur Warna																				
5.	Processing Data																				
6.	Pengujian Klasifikasi Naïve Bayes																				
7.	Pengujian Klasifikasi HSV																				
8.	Evaluasi																				

3.2. Prosedur Penelitian

Terdapat 8 (Delapan) tahapan pada prosedur percobaan dalam penelitian ini. Langkah pertama adalah membaca dan mengamati, diikuti dengan pengumpulan dan analisis data. Selanjutnya lakukan proses *pre-processing* data, ekstrak fitur warna dan *processing* data. Langkah selanjutnya adalah pengujian dengan mengklasifikasi data menggunakan algoritma *Naive Bayes* dan HSV. Setelah dilakukan klasifikasi data, hasilnya akan di evaluasi dan digunakan dalam penelitian selanjutnya. Prosedur penelitian dijelaskan pada Gambar 3.1.



Gambar 3. 1 Prosedur Penelitian

3.2.1. Studi Literatur

Pada tahap ini melakukan *review* pengumpulan data dengan mencari referensi data dari buku, jurnal *paper* yang berkaitan dengan pokok bahasan penelitian implementasi algoritma *Naïve Bayes* dan *Hue Saturation Value* (HSV) untuk klasifikasi kematangan buah jambu bol berdasarkan warna buah. Tahap selanjutnya yaitu penulis akan melakukan pengumpulan dan analisis data.

3.2.2. Analisis dan Pengumpulan Data

Tahapan ini dilakukan dengan menganalisis data dengan cara melakukan pengumpulan data pada hasil *review* studi literatur dan observasi. Data tersebut berhubungan dengan implementasi algoritma *Naïve Bayes* dan *Hue Saturation Value* (HSV) untuk klasifikasi kematangan buah jambu bol berdasarkan warna buah. Pemilihan 3 jenis data citra buah jambu bol tersebut bertujuan untuk mempermudah proses klasifikasi tingkat kematangan. Data yang digunakan pada penelitian adalah sebuah keputusan yang dihasilkan dari pengujian sistem. Contoh data yang diambil seperti tabel dibawah ini.



Gambar 3. 2 Buah Jambu Bol Mentah
(Sumber : Roring. dkk, 2022)



Gambar 3. 3 Buah Jambu Bol Setengah Matang
(Sumber : Roring. dkk, 2022)



Gambar 3. 4 Buah Jambu Bol Matang
(Sumber : Roring. dkk, 2022)

3.2.3. *Preprocessing Data*

Preprocessing data merupakan langkah penting dalam *machine learning* untuk mempersiapkan data agar dapat digunakan secara efektif oleh model pembelajaran. Berikut adalah tahapan umum dalam *preprocessing* data:

A. *Preprocessing Data*:

1. Mengimpor dataset.
2. Mengkategorikan nilai target menjadi kelas.
3. Memisahkan fitur dan label.
4. Membagi data menjadi set pelatihan dan pengujian.
5. Melakukan normalisasi (jika diperlukan).

B. *Training Model*:

1. Melatih model *Naïve Bayes*.
2. Melatih model HSV.

C. *Evaluasi Model*:

1. Mengukur akurasi model pada data pengujian.
2. Membandingkan hasil akurasi kedua model.

3.2.4. *Ekstrak Fitur Warna*

Fungsi ekstrak fitur warna digunakan untuk mengonversi gambar dari ruang warna BGR (*Blue, Green, Red*) ke ruang warna HSV (*Hue, Saturation, Value*) dan kemudian mengekstraksi nilai rata-rata dari setiap channel (H, S, V). Fitur ini akan digunakan sebagai input untuk model klasifikasi.

3.2.5. *Processing Data*

Pemrosesan data (*data processing*) merupakan elemen kunci di era digital, di mana sejumlah besar data dihasilkan setiap hari. Dengan mengumpulkan, membersihkan, menganalisis dan memvisualisasikan data dapat memperoleh wawasan berharga yang mendukung banyak aspek operasi dan pengambilan keputusan. Pada tahap *Processing* metode yang digunakan yaitu Algoritma *Naïve Bayes* dan *Hue Saturation Value* (HSV).

A. Tahap Algoritma Algoritma *Naïve Bayes* :

1. Pengumpulan Data:
 - a. Kumpulkan dataset gambar buah jambu bol dari tiga kategori: matang, setengah matang dan mentah.
 - b. Gambar disimpan dalam tiga folder terpisah.
 2. Pembersihan dan Pra-pemrosesan Data:
 - a. Baca gambar dari folder menggunakan OpenCV.
 - b. Ubah ukuran gambar menjadi ukuran yang konsisten (misalnya, 100x100 piksel).
 - c. Konversi gambar dari ruang warna BGR ke ruang warna HSV.
 - d. Ekstrak fitur dari gambar dengan menghitung nilai rata-rata dari channel H, S dan V.
 3. Pembagian Dataset:
 - a. Bagi dataset menjadi data latih dan data uji menggunakan metode train-test split (misalnya, 90% untuk data latih dan 10% untuk data uji).
 4. Pelatihan Model:
 - a. Latih model *Naive Bayes* menggunakan data latih.
 5. Prediksi dan Evaluasi:
 - b. Gunakan model yang dilatih untuk memprediksi data uji.
 - c. Hitung metrik evaluasi seperti akurasi, precision, recall dan F1-score.
 - d. Visualisasikan hasil menggunakan confusion matrix.
- B. Tahap *Hue Saturation Value* (HSV):
1. Pengumpulan Data:
 - a. Sama seperti pada algoritma *Naive Bayes*, kumpulkan dataset gambar buah jambu bol dari tiga kategori: matang, setengah matang dan mentah.
 2. Pembersihan dan Pra-pemrosesan Data:
 - a. Sama seperti pada algoritma *Naive Bayes*, baca gambar, ubah ukurannya dan konversi ke ruang warna HSV.
 - b. Ekstrak fitur dengan menghitung nilai rata-rata dari channel H, S dan V.

3. Pembagian Dataset:

- a. Bagi *dataset* menjadi data latih dan data uji menggunakan metode train-test split (misalnya, 90% untuk data latih dan 10% untuk data uji).

4. Pelatihan Model Sederhana:

Tentukan ambang batas nilai channel V untuk memisahkan antara buah matang, setengah matang dan mentah (misalnya, buah dianggap matang jika nilai rata-rata V lebih dari ambang batas tertentu).

5. Prediksi dan Evaluasi:

- a. Gunakan aturan ambang batas untuk memprediksi kematangan buah pada data uji.
- b. Hitung metrik evaluasi seperti akurasi.
- c. Visualisasikan hasil menggunakan confusion matrix.

3.2.6. Pengujian Klasifikasi

Pada pengujian klasifikasi pada penelitian ini dibagi menjadi 2 yaitu klasifikasi Algoritma *Naïve Bayes* dan *Hue Saturation Value* (HSV).

A. Algoritma *Naïve Bayes*

Naive Bayes menghitung probabilitas posterior $P(H|X)$ dari kelas H yang diberikan fitur X menggunakan *Teorema Bayes*. Ini dilakukan dengan mengalikan probabilitas prior $P(H)$ dengan likelihood $P(X|H)$ dan membagi dengan bukti $P(X)$:

$$P(H | X) = \frac{P(X|H) \cdot P(H)}{P(X)} \quad (2)$$

1. Menghitung Probabilitas Prior:

$P(H)$ Probabilitas dari setiap kelas dalam dataset.

2. Menghitung Likelihood:

$P(X|H)$ Probabilitas fitur X yang diberikan kelas H.

3. Menghitung Probabilitas Posterior:

$P(H|X)$ Probabilitas kelas H yang diberikan fitur X.

4. Menghitung Probabilitas Bukti:

$P(X)$ adalah probabilitas bukti dari fitur X.

B. HSV

Menggunakan nilai rata-rata dari *channel V* (Value) untuk klasifikasi dengan menentukan nilai ambang batas untuk *channel V*. Misalnya, jika rata-rata nilai *V* lebih besar dari ambang batas, maka gambar dianggap matang, sebaliknya dianggap mentah.

3.2.7. Evaluasi

Berdasarkan prosedur pengujian yang telah dilakukan, akan mendapatkan hasil pengujian klasifikasi Algoritma *Naive Bayes* dan HSV. Hasil pengujian klasifikasi berupa *Accuracy*, *precision*, *recall* dan *F1-score*.

1. *Accuracy* :

Accuracy adalah untuk mengetahui jumlah data yang telah di klasifikasi secara benar. Adapun untuk rumus perhitungan *Accuracy* sebagai berikut :

$$Accuracy = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Predictions}} \quad (4)$$

2. *Precision* :

Precision faktor penting dalam menentukan keakuratan prediksi sistem. Ini melibatkan perhitungan rasio prediksi yang benar dengan jumlah total prediksi yang dibuat oleh sistem. Adapun untuk rumus perhitungan *precision* sebagai berikut :

$$Precision = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positive}} \quad (5)$$

3. *Recall* :

Recall adalah untuk nilai tingkat keberhasilan dalam mengenali suatu kelas yang perlu dikenali. Adapun untuk rumus perhitungan *Recall* adalah sebagai berikut :

$$Recall = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negative}} \quad (6)$$

Keterangan :

TP = *True Positif* (memprediksi *positif* dan itu benar).

TN = *True Negatif* (memprediksi *negatif* dan itu benar).

FP = *False Positif* (memprediksi *positif* dan itu salah).

FN = *False Negatif* (memprediksi *negatif* dan itu salah).

4. *F1-Score* :

F1-Score adalah metrik evaluasi untuk klasifikasi yang menggabungkan presisi dan recall. Ini memberikan gambaran tentang seberapa baik model memprediksi kelas positif dengan nilai maksimum 1 dan nilai minimum 0.

Adapun untuk rumus perhitungan *F1-Score* sebagai berikut :

$$F1 - Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (7)$$

