

BAB 2

TINJAUAN PUSTAKA

2.1. Citra Digital

Citra Digital adalah gambar pada bidang dwimatra (dua dimensi). Ditinjau dari sudut pandang matematis, citra adalah fungsi intensitas cahaya dua dimensi $f(x, y)$ dimana x adalah posisi baris dan y adalah posisi kolom sedangkan f adalah fungsi intensitas atau kecerahan dari citra pada koordinat (x, y) . Menurut Darma Putra [12], Citra Digital merupakan sebuah larik (array) yang berisi nilai-nilai real maupun kompleks yang direpresentasikan dengan deretan bit tertentu.

Pada umumnya citra berbentuk empat persegi panjang dan dimensi ukurannya dinyatakan sebagai (tinggi x lebar). Citra dengan tinggi N piksel, lebarnya M piksel, dan memiliki intensitas f dapat direpresentasikan sebagai suatu matriks yang berukuran N baris dan M kolom, dengan x dan y adalah koordinat spasial, dan amplitudo f di titik koordinat (x, y) dinamakan intensitas atau tingkat keabuan dari citra pada titik tersebut. apabila nilai x, y dan nilai amplitudo f secara keseluruhan berhingga (finite) dan bernilai diskrit maka dapat dikatakan bahwa citra tersebut adalah citra digital.

Citra digital dapat ditulis dalam bentuk matrik pada Persamaan (2.1).

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, N-1) \\ f(1,0) & f(1,1) & \dots & f(1, N-1) \\ \vdots & \vdots & \ddots & \vdots \\ f(M-1,0) & f(M-1,1) & \dots & f(M-1, N-1) \end{bmatrix} \quad (2.1)$$

Indeks baris (x) dan indeks kolom (y) menyatakan suatu koordinat titik pada citra. Nilai pada suatu irisan antara baris dan kolom disebut dengan picture elements, image elements, pels, atau pixels, sedangkan $f(x, y)$ merupakan intensitas (derajat keabuan) pada titik (x, y) .

2.2. Bunga

Bunga atau kembang adalah struktur reproduksi seksual pada tumbuhan berbunga. Bunga merupakan modifikasi suatu tunas (batang dan daun) yang bentuk,

warna, dan susunannya disesuaikan dengan kepentingan tumbuhan. Oleh karena itu, bunga ini berfungsi sebagai tempat berlangsungnya penyerbukan dan pembuahan yang akhirnya dapat dihasilkan alat-alat perkembangbiakan.

Pengenalan dan Klasifikasi bunga merupakan tugas yang menantang karena banyaknya spesies bunga yang memiliki bentuk dan penampilan yang serupa atau benda di sekitarnya seperti daun dan rumput [13]. Beberapa bunga dari berbagai jenis memiliki warna, bentuk, dan penampilan yang serupa. Sehingga akan cukup sulit untuk dilakukan pengenalan dan klasifikasi.

Pengenalan dan Klasifikasi bunga memiliki berbagai penerapan, karena dapat membantu dalam pencarian bunga untuk analisis pattern dan pemeliharaan bunga. Industri Florikultura terdiri dari sarana perdagangan bunga jual beli bunga, produksi umbi dan benih, pembibitan dan tanaman dalam pot serta ekstraksi minyak atsiri dari bunga [14].

2.2.1. Jenis Bunga yang digunakan

Bunga yang akan digunakan ada 5, berikut adalah gambar dari bunga yang akan digunakan :

Tabel 2.1 Jenis Bunga yang digunakan

No	Nama Bunga	Gambar Bunga
1	Buttercup	

2	Daisy	
3	Dandelion	
4	Sunflower	
5	Windflower	

2.3. Jenis Citra

Jenis citra yang digunakan pada pemrosesan citra, umumnya dibagi tiga. Ketiga citra tersebut yaitu Citra berwarna, Citra berskala keabuan dan Citra biner.

2.3.1. Citra Berwarna

Citra berwarna atau biasa disebut Citra RGB, Menyajikan tiga komponen warna, Yaitu: R(merah), G(hijau), dan B(biru). Komponen - komponen warna tersebut menggunakan 8 bit dimana nilainya berkisar antara 0 sampai dengan 255. Sehingga, Nilai yang dihasilkan mencapai $255 \times 255 \times 255$ atau 16.581.375 warna. menunjukkan contoh warna R, G dan B. [15]

Tabel 2.2 Citra Berwarna RGB

Warna	R	G	B
Merah	255	0	0
Hijau	0	255	0
Biru	0	0	255
Hitam	0	0	0
Putih	255	255	255
Kuning	0	255	255

2.3.2. Citra berskala keabuan

Citra berskala merupakan citra dengan komponen warna hitam dan putih, yang menghasilkan efek warna abu-abu. Dengan nilai komponennya 0 dan 255. Nilai 0 yang berarti hitam dan 255 adalah putih. [15]

2.3.3. Citra biner

Citra biner merupakan citra dengan piksel yang terbagi menjadi 2 Nilai yaitu nilai 0 dan 1. Nilai 0 menyatakan warna hitam dan nilai 1 menyatakan warna putih. Citra biner sendiri sering digunakan dalam pemrosesan citra dalam memperoleh tepi objek. [15]

2.4. Pengenalan Pola

Pola adalah suatu entitas yang terdefinisi dan dapat diidentifikasi serta diberi nama. Pola bisa merupakan kumpulan hasil pengukuran atau pemantauan dan bisa dinyatakan dalam notasi vektor atau matriks. Secara umum pengenalan pola (pattern recognition) adalah suatu ilmu untuk mengklasifikasikan atau

menggambarkan sesuatu berdasarkan pengukuran kuantitatif fitur (ciri) atau sifat utama dari suatu objek (sitasi) [12]. Ciri-ciri tersebut digunakan untuk membedakan suatu pola dengan pola lainnya. ciri yang baik adalah ciri yang memiliki daya pembeda yang tinggi sehingga pengenalan pola berdasarkan ciri yang dimiliki dapat dilakukan dengan baik. pengenalan pola ini bersifat *conceptually driven processing* yang berarti bahwa proses dimulai dari pembentukan konsep pada objek yang dijumpai.

Menurut Theodoridis, Pengenalan pola dapat diartikan sebagai proses klasifikasi dari objek atau pola menjadi beberapa kategori atau kelas dan bertujuan untuk pengambilan keputusan [16]. Ada banyak aplikasi yang bisa dijadikan implementasi dari pengenalan pola. diantaranya pengenalan wajah manusia, pengenalan gambar mata, pengenalan penyakit berdasarkan gejala-gejala yang ditemukan pada objek dan lain-lain. Pengenalan pola bisa dibuat melalui pendekatan pemrosesan citra yang mana tujuan akhir dari pemrosesan citra tersebut digunakan untuk pengelompokan objek sehingga menghasilkan output yang diinginkan.

2.5. Resize

Resizing adalah proses pengubahan ukuran citra, baik memperbesar maupun memperkecil resolusi citra [17]. Proses ini sering dilakukan pada pengolahan citra untuk normalisasi ukuran citra digital yang akan diproses sehingga memiliki ukuran yang sama dan mempercepat proses pengolahan citra ketika ukuran citra diperkecil.

2.6. Grayscale

Grayscale merupakan proses untuk mengubah citra berwarna (RGB) menjadi citra keabuan yang berfungsi untuk mengecilkan range warna menjadi 0 sampai dengan 255. Berikut adalah rumus konversi citra berwarna (RGB) menjadi grayscale pada Persamaan (2.2).

$$I = (0.2989 * R) + (0.5879 * G) + (0.1141 * B) \quad (2.2)$$

Dimana :

I = Nilai Grayscale

R= komponen nilai merah (Red) dari suatu titik pixel

G = komponen nilai hijau (Green) dari suatu titik pixel

B = komponen nilai biru (Blue) dari suatu titik pixel

2.7. Segmentasi Citra

Segmentasi citra bertujuan mendapatkan objek-objek citra dengan cara membagi citra ke dalam beberapa daerah yang memiliki kemiripan atribut [15].

Terdapat dua teknik segmentasi citra yaitu mendeteksi diskontinuitas dan similaritas. Pendekatan diskontinuitas membagi citra berdasarkan perubahan intensitas yang tiba-tiba seperti deteksi titik, deteksi garis, dan deteksi tepi. Sedangkan pendekatan similaritas memecah citra ke dalam daerah yang sama menurut beberapa kriteria yang sudah ditentukan, seperti proses thresholding, region growing, region splitting dan merging.

2.8. Thresholding

Thresholding mengubah citra keabuan menjadi citra biner bergantung pada nilai threshold (T) sehingga dapat diketahui daerah mana yang termasuk objek dan latar belakang [18]. Jika nilai pixel lebih besar dari threshold diatur menjadi 1 sebaliknya jika kurang dari threshold diatur menjadi 0. Secara umum proses thresholding ditunjukkan pada Persamaan (2.3).

$$g(x, y) = \begin{cases} 1, & \text{jika } f(x, y) \geq T \\ 0, & \text{jika } f(x, y) < T \end{cases} \quad (2.3)$$

Persamaan (2.3) menjelaskan bahwa $g(x, y)$ adalah citra biner dari citra aras keabuan $f(x, y)$ dan T menyatakan nilai ambang.

2.8.1. Metode Otsu

Konsep thresholding *Otsu* pertama kali diperkenalkan oleh Nobuyuki Otsu (1979) untuk mengelompokkan citra biner berdasarkan bentuk histogram secara otomatis, mengasumsikan bahwa citra berisi dua kelas dasar dengan bentuk histogram bimodal (*foreground* dan *background*) [18].

Tujuan dari metode *Otsu* membagi histogram citra keabuan ke dalam dua daerah yang berbeda secara otomatis tanpa membutuhkan bantuan user. Pendekatan dengan melakukan analisis diskriminan yakni menentukan suatu variabel lalu memaksimalkan variabel tersebut agar dapat memisahkan objek dengan latar belakang [12] [19].

Berikut algoritma *Otsu* menentukan threshold (k) [12]. Nilai k berkisar antara 0 sampai 255.

1. Probabilitas setiap pixel pada gray level i .

$$p_i = \frac{n_i}{N} \quad (2.4)$$

$n_i = \text{Jumlah pixel pada level ke } i$

2. Mencari nilai Zeroth Cumulative moment (zerothCM):

$$\omega(k) = \sum_{i=0}^k p_i \quad (2.5)$$

3. Mencari nilai First cumulative moment (firstCM):

$$\mu(k) = \sum_{i=0}^k i \cdot p_i \quad (2.6)$$

4. Mencari nilai Rerata intensitas Global (tMean):

$$\mu^T = \sum_{i=0}^L i \cdot p_i \quad (2.7)$$

5. Nilai ambang k ditentukan dengan memaksimalkan Persamaan (between class variance) :

$$\sigma_B^2(k^*) = \max(\sigma_B^2(k)) \quad (2.8)$$

Dengan:

$$\sigma_B^2(k) = \frac{[\mu^T \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (2.9)$$

Menurut Penelitian Mei Fang, Metode Otsu merupakan metode populer diantara semua metode thresholding dan metode terbaik dalam mendapatkan nilai threshold secara otomatis [20].

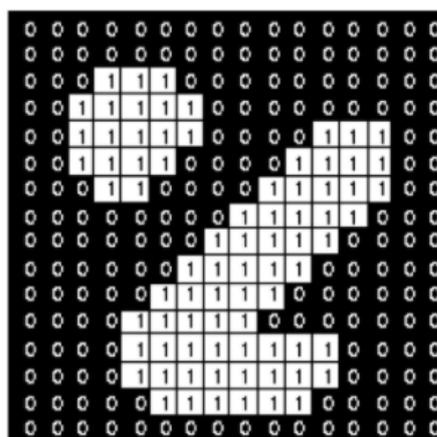
2.9. Operasi Morfologi

Operasi morfologi merupakan suatu teknik di dalam bidang pengolahan citra digital, teknik ini bekerja dengan menggunakan bentuk objek sebagai pedoman dalam pengolahan. Nilai dari setiap piksel yang terdapat dalam citra digital diperoleh pada citra digital masukan dengan piksel sekitarnya [15] jumlah piksel yang ditambahkan atau dikurangkan tergantung dari ukuran dan bentuk dari element penstrukturan yang digunakan untuk memproses citra. Elemen penstrukturan merupakan himpunan kecil yang dapat digunakan untuk meneliti citra dalam pembelajaran propertinya. Elemen penstrukturan memiliki ukuran yang lebih kecil dibandingkan dengan citra yang akan diproses [21].

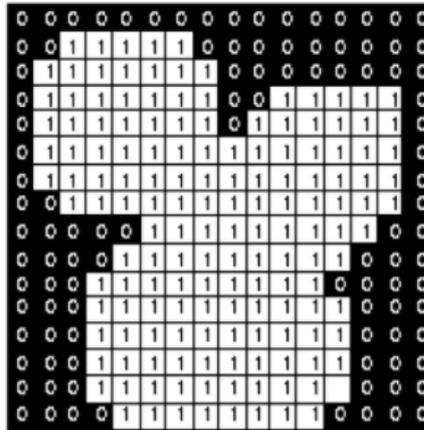
Terdapat beberapa operasi dalam operasi morfologi, beberapa diantaranya sebagai berikut :

1. Operasi Dilasi

Dilasi adalah teknik untuk memperbesar segmen objek (citra biner) dengan menambah lapisan di sekeliling objek atau dengan menjadikan titik latar (0) yang bertetangga dengan titik objek (1) menjadi titik objek (1).



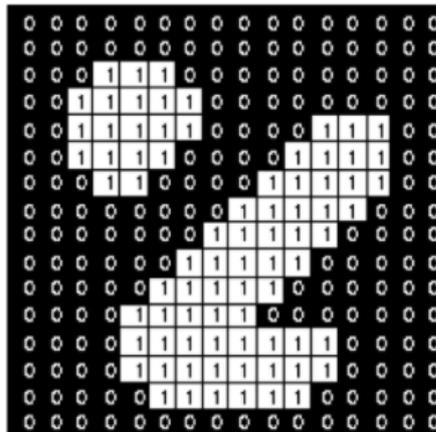
Gambar 2.1. Sebelum Dilasi



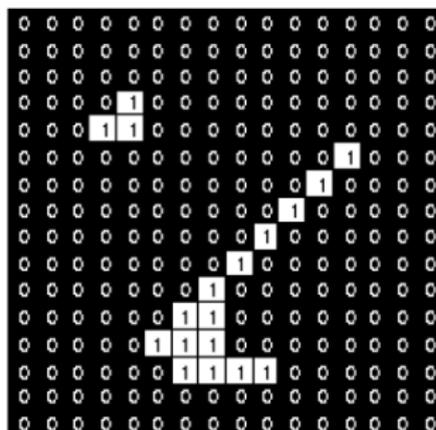
Gambar 2.2. Sesudah Dilasi

2. Operasi Erosi

Erosi atau pengikisan adalah kebalikan dari dilasi yaitu teknik yang bertujuan untuk memperkecil atau mengikis tepi objek. Atau dengan menjadikan titik objek (1) yang bertetangga dengan titik latar (0) menjadi titik latar (0).



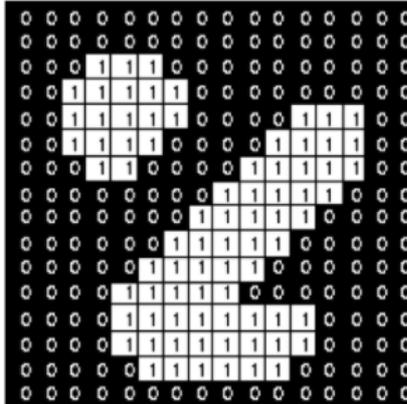
Gambar 2.3. Sebelum Erosi



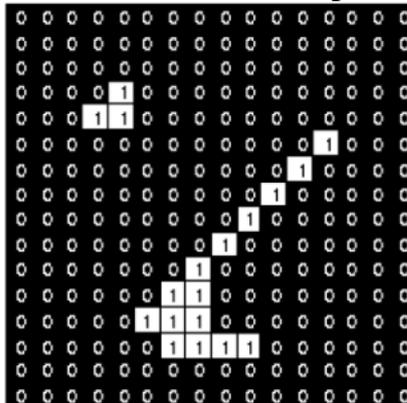
Gambar 2.4. Sesudah Erosi

3. Operasi Opening

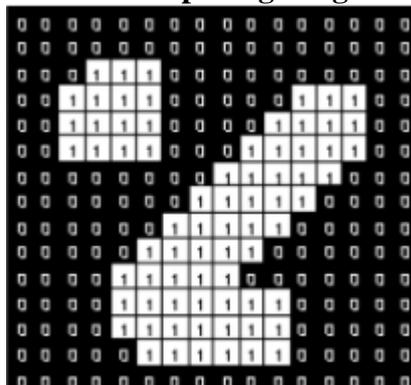
Opening adalah proses erosi yang diikuti dengan dilasi. Dimulai dengan melakukan erosi pada citra kemudian hasil tersebut kembali dilakukan erosi. Opening biasanya digunakan untuk menghilangkan objek-objek kecil dan kurus serta dapat membuat tepi citra lebih smooth.



Gambar 2.5. Sebelum Opening



Gambar 2.6. Opening dengan erosi



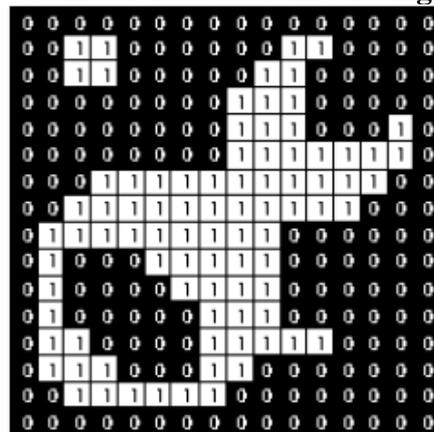
Gambar 2.7. Dilanjutkan dengan Dilasi

4. Operasi Closing

Closing merupakan kebalikan dari opening. Dimana citra terlebih dahulu dilakukan dilasi yang kemudian dilanjutkan dengan erosi. Closing bertujuan untuk mengisi lubang kecil pada objek, menggabungkan objek yang berdekatan.



Gambar 2.8. Sebelum Closing



Gambar 2.9. Setelah Closing

Kegunaan operasi morfologi adalah untuk :

1. Menghilangkan noise
2. Mengisolasi objek

Operasi Morfologi yang digunakan pada penelitian ini yaitu *area opening*, *closing*, *opening* dan *filling holes*.

2.10. Ekstraksi Fitur

Ekstraksi fitur adalah proses untuk mendapatkan fitur utama yang terdapat pada citra. Pada penelitian digunakan 2 Ekstraksi fitur, yaitu Ekstraksi fitur warna

dengan metode Model Warna HSV dan Ekstraksi fitur bentuk berupa eccentricity dan metric.

2.10.1. Model Warna HSV

HSV adalah model warna yang lebih baik untuk digunakan dalam berbagai keperluan pengolahan citra dan computer vision. Model HSV (*Hue, Saturation, Value*) menunjukkan ruang warna dalam bentuk tiga komponen utama yaitu Hue, Saturation dan Value (atau disebut juga brightness). *Hue* menunjukkan jenis warna atau corak warna yaitu tempat warna tersebut ditemukan dalam spektrum warna. *Saturation* dari suatu warna adalah ukuran seberapa besar kemurnian dari warna tersebut. *Value* atau disebut juga intensity yaitu ukuran seberapa besar kecerahan dari suatu warna atau seberapa besar cahaya datang dari suatu warna. Value dapat bernilai dari 0-100%. [12]

Perhitungan konversi RGB menjadi HSV dapat dirumuskan dengan normalisasi RGB terlebih dahulu.

$$r = \frac{R}{R + G + B}, g = \frac{G}{R + G + B}, b = \frac{B}{R + G + B} \quad (2.10)$$

Perhitungan nilai Value:

$$V = \max(r, g, b) \quad (2.11)$$

Perhitungan nilai Saturation:

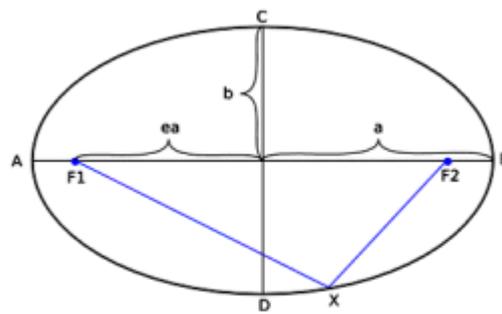
$$S = \begin{cases} 0, & V = 0 \\ 1 - \frac{\min(r, g, b)}{V}, & V > 0 \end{cases} \quad (2.12)$$

Perhitungan nilai Hue:

$$H = \begin{cases} 0, & \text{jika } S = 0 \\ \frac{60 * (g - b)}{S * V}, & \text{jika } V = r \\ 60 * \left[2 + \frac{b - r}{S * V} \right], & \text{jika } V = g \\ 60 * \left[4 + \frac{r - g}{S * V} \right], & \text{jika } V = b \end{cases} \quad (2.13)$$

2.10.2. Eccentricity

Fitur bentuk bertujuan untuk membedakan bentuk objek satu dan objek lainnya, dapat menggunakan parameter *eccentricity*. Eccentricity merupakan nilai perbandingan antara jarak foci ellips minor dengan foci ellips major suatu objek. Eccentricity memiliki rentang nilai antara 0 hingga 1. Objek yang berbentuk memanjang/mendekati bentuk garis lurus, nilai eccentricitynya mendekati angka 1, sedangkan objek yang berbentuk bulat/lingkaran, nilai eccentricitynya mendekati angka 0. Penghitungan eccentricity dapat diilustrasikan pada persamaan dibawah ini :



Gambar 2.10 Ilustrasi Parameter fitur bentuk Eccentricity

$$e = \sqrt{1 - \frac{b^2}{a^2}} \quad (2.14)$$

Dimana :

e = Eccentricity

a = Major axis

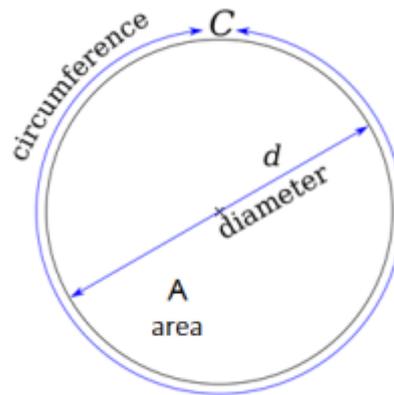
b = Minor axis

2.10.3. Metric

Parameter lainnya yang dapat digunakan untuk membedakan bentuk suatu objek yaitu *metric*. Metric merupakan nilai perbandingan antara luas dan keliling objek. Metric memiliki rentang nilai antara 0 hingga 1. Objek yang berbentuk memanjang/mendekati bentuk garis lurus, nilai metricnya mendekati angka 0,

sedangkan objek yang berbentuk bulat/lingkaran, nilai metricnya mendekati angka

1. Penghitungan metric diilustrasikan pada persamaan dibawah ini :



Gambar 2.11 Ilustrasi Parameter fitur bentuk Metric

$$M = \frac{4\pi * A}{C^2} \quad (2.15)$$

Dimana :

M = Metric

A = Area

C = Circumference

2.11. Naïve Bayes Classifier

Naive bayes merupakan pengembangan dari metode bayesian classification. Bayesian classification adalah pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Menurut Mitchell, Bayesian classification didasarkan pada teorema bayes yang memiliki kemampuan klasifikasi serupa dengan decision tree dan neural network [9]. Bayesian classification terbukti memiliki akurasi dan kecepatan yang tinggi saat diaplikasikan ke dalam database dengan data yang besar [10].

Naïve bayes adalah teknik prediksi berbasis probabilistik sederhana yang berdasar pada penerapan teorema Bayes

1. Asumsi Independensi (ketidaktergantungan) yang kuat (naif)
2. Model yang digunakan adalah “model fitur independen”

Independensi yang kuat pada fitur adalah bahwa sebuah fitur pada sebuah data tidak ada kaitannya dengan atau tidak adanya fitur yang lain dalam data yang sama.

Naïve bayes adalah suatu metode pengklasifikasian paling sederhana dengan menggunakan peluang yang ada, dimana diasumsikan bahwa setiap variabel X bersifat bebas (*independence*). Karena asumsi variabel tidak saling terikat, maka didapatkan :

$$P(X|Y = y) = \prod_{i=1}^q P(X_i|Y = y) \quad (2.16)$$

Data yang digunakan dapat bersifat kategorial diskrit maupun kontinu. Namun, pada penelitian ini akan digunakan data kontinu, karena hasil ekstraksi fitur pada citra berupa angka-angka hasil pengukuran tingkat adalah rata-rata Hue, rata-rata Saturation, rata-rata Value dan eccentricity dan metric. Maka dari itu data kontinu dapat diselesaikan dengan menggunakan langkah-langkah berikut.

Training :

1. Hitung rata-rata tiap fitur dalam dataset training dengan

$$\mu = \frac{\sum x_i}{n} \quad (2.17)$$

Dimana :

$\mu = \text{mean}$

$n = \text{banyaknya data}$

$\sum x_i = \text{jumlah data}$

2. Kemudian hitung nilai standar deviasi dari dataset training tersebut pada.

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \mu)^2}{n}} \quad (2.18)$$

Dimana :

$\sigma = \text{standar deviasi}$

$\mu = \text{mean}$

$x_i = \text{nilai data}$

$n = \text{banyaknya data}$

Testing :

1. Hitung probabilitas (Prior) tiap kelas yang ada dengan cara menghitung jumlah data tiap kelas dibagi jumlah total data secara keseluruhan.
2. Selanjutnya menghitung distribusi normal. Fungsi distribusi normal yang digunakan adalah fungsi distribusi kumulatif. Rumus dari fungsi distribusi kumulatif adalah :

$$P(x_i|y) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}\sigma_y} e^{-\frac{(x-\mu)^2}{2\sigma_y^2}} \quad (2.19)$$

Atau dapat juga menggunakan rumus :

$$P(X < x) = P(X - \mu < x - \mu) = P\left(\frac{X - \mu}{\sigma} < \frac{x - \mu}{\sigma}\right) \quad (2.20)$$

Dapat ditulis :

$$P(X < x) = P\left(Z < \frac{x - \mu}{\sigma}\right) \quad (2.21)$$

Atau :

$$P(x_i|y) = P\left(Z < \frac{x - \mu}{\sigma}\right) \quad (2.22)$$

Dimana :

$x = data\ masuk$

$\pi = 3.14$

$\sigma = Standar\ deviasi$

$\mu = mean$

$y = kelas$

$Z = Z - score$

Setelah nilai Z didapatkan, cari nilai Z-score nya dengan bantuan tabel distribusi normal.

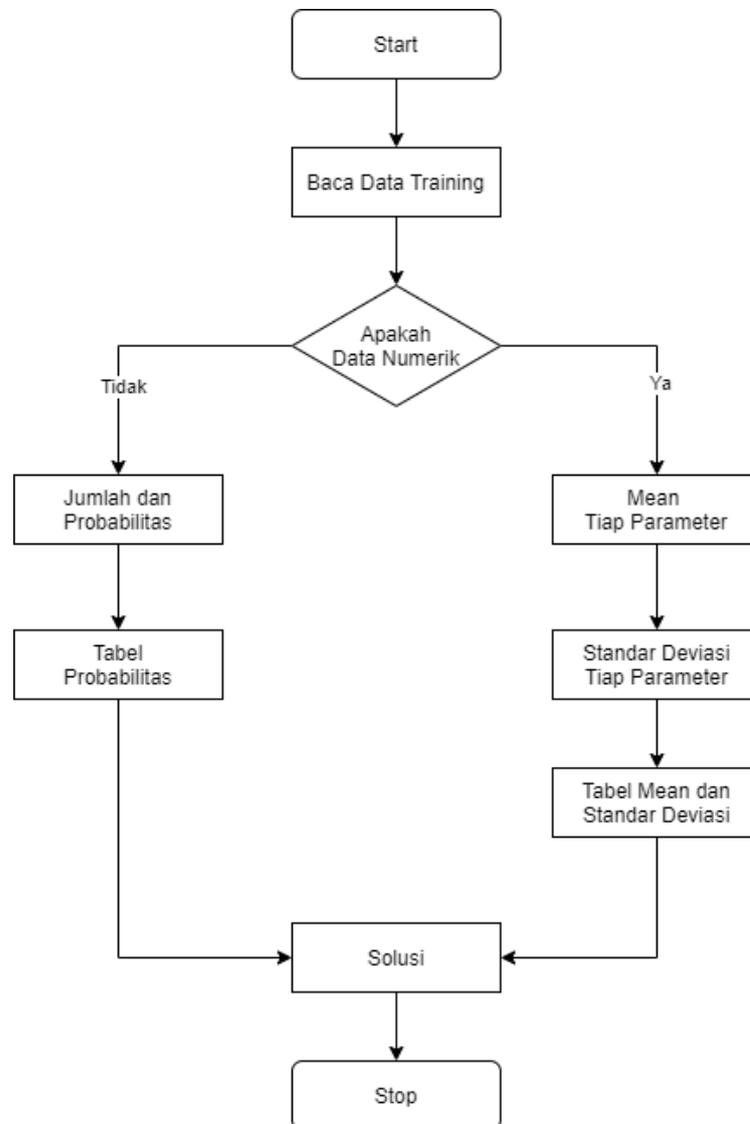
3. Setelah didapatkan nilai normal distribusinya, selanjutnya menghitung posterior masing-masing kelas dengan menggunakan persamaan (2.23) atau persamaan (2.24).

$$Posterior = \frac{P(C_i) \times P(X_1 \dots X_n | C_i)}{P(X_1 \dots X_n | C_i)} \quad (2.23)$$

Atau bisa ditulis

$$Posterior = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}} \quad (2.24)$$

4. Setelah didapatkan nilai posterior, kelas yang sesuai adalah nilai posterior terbesar.



Gambar 2.12. Skema Naive Bayes

2.12. K-fold Cross Validation

K-Fold Cross Validation adalah teknik untuk melakukan validasi pada dataset untuk menemukan akurasi yang baik. Teknik ini membagi dataset sebanyak k subset. Satu dari subset ini akan dijadikan sebagai data uji dan $k-1$ subset sisanya digunakan untuk proses data latih. Proses ini dilakukan sebanyak k kali sehingga setiap subset akan menjadi data uji dari model. Proses ini akan mendapatkan k buah nilai performa dari proses pembelajaran. Semua nilai performa ini akan dicari rata-ratanya dan nilai dengan rata-rata tertinggi akan dipilih sebagai model. k-fold cross validation memiliki kelebihan dapat mengklasifikasi dataset lebih efisien, namun metode ini memiliki kelemahan dalam proses komputasi yang digunakan akan lebih besar karena akan melakukan proses sebanyak k kali [22].

2.13. Akurasi

Sebuah sistem yang melakukan identifikasi atau klasifikasi diharapkan mampu melakukan klasifikasi semua data set dengan benar. Namun tidak dapat dipungkiri bahwa kinerja suatu sistem yang melakukan klasifikasi tidak akan selalu bisa 100% benar. Oleh karena itu, sistem harus diukur kinerjanya. Umumnya cara mengukur kinerja klasifikasi menggunakan *confusion matrix* [21].

Pengukuran terhadap kinerja suatu sistem klasifikasi merupakan hal yang penting. Kinerja sistem klasifikasi menggambarkan seberapa baik sistem dalam mengklasifikasikan data. Confusion matrix merupakan salah satu metode yang dapat digunakan untuk mengukur kinerja suatu metode klasifikasi. Pada dasarnya confusion matrix mengandung informasi yang membandingkan hasil klasifikasi yang dilakukan oleh sistem dengan hasil klasifikasi yang seharusnya. Pada pengukuran kinerja menggunakan confusion matrix, terdapat 4 (empat) istilah sebagai representasi hasil proses klasifikasi. Keempat istilah tersebut adalah True Positive (TP), True Negative (TN), False Positive (FP) dan False Negative (FN). Nilai True Negative (TN) merupakan jumlah data negatif yang terdeteksi dengan benar, sedangkan False Positive (FP) merupakan data negatif namun terdeteksi sebagai data positif. Sementara itu, True Positive (TP) merupakan data positif yang terdeteksi benar. False Negative (FN) merupakan kebalikan dari True Positive,

sehingga data positif, namun terdeteksi sebagai data negatif. Pada jenis klasifikasi biner yang hanya memiliki 2 keluaran kelas, confusion matrix dapat disajikan seperti pada Tabel 2.3.

Tabel 2.3 Confusion Matrix

Kelas	Positif	Negatif
Positif	TP	FN
Negatif	FP	TN

Berdasarkan nilai True Negative (TN), False Positive (FP), False Negative (FN), dan True Positive (TP) dapat diperoleh nilai akurasi. Nilai akurasi menggambarkan seberapa akurat sistem dapat mengklasifikasikan data secara benar. Dengan kata lain, nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Nilai akurasi dapat diperoleh dengan persamaan 2.25 atau persamaan 2.26.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} * 100\% \quad (2.25)$$

Atau

$$Akurasi = \frac{Jumlah\ prediksi\ benar}{Jumlah\ seluruh\ prediski\ yang\ dilakukan} * 100\% \quad (2.26)$$

2.14. Penelitian Terkait

Beberapa penelitian terdahulu yang berhubungan dengan penelitian ini, diantaranya :

Dalam sebuah penelitian tentang pengenalan bunga dengan berdasarkan image processing, gambar bunga diubah ukurannya untuk pemrosesan lebih cepat. Untuk menghilangkan background pada bunga sehingga pada gambar hanya terdapat bunganya saja. Digunakan algoritma pemotongan grafik dan konversi RGB ke grayscale. Setelah sistem menerima karakteristik tepi dengan algoritma seven-moment Hu dan karakteristik warna, *K-Nearest Neighbour* digunakan untuk klasifikasi, didapatkan nilai akurasi sebesar 80% [4].

Dalam sebuah penelitian tentang pengklasifikasian bunga menggunakan Fusion descriptor dan SVM. Klasifikasi bunga memiliki kesulitannya sendiri karena memiliki variasi besar dalam skala dan sudut pandang dalam bunga yang khas. Penelitian ini bertujuan untuk mengembangkan pendekatan klasifikasi bunga yang efektif menggunakan teknologi ekstraksi fitur. Metode yang diusulkan memberikan hasil tingkat akurasi 86,17%, dengan menggunakan dataset Oxford 17 [6].

Dalam sebuah penelitian mengenai pengenalan bunga dengan menggunakan metode segmentasi untuk memisahkan gambar bunga dari latar belakangnya. Gambar bunga diwakili menggunakan kombinasi warna dan fitur GIST, yang kemudian diklasifikasikan dengan SVM. Pendekatan yang diusulkan mendapatkan nilai akurasi 85,93% [7].

Dalam sebuah penelitian mengenai pengenalan bunga dengan menggunakan metode jarak manhattan dengan ekstraksi ciri histogram dan metode klasifikasi Euclidean dengan statistik orde 1. Ditemukan bahwa metode klasifikasi manhattan memiliki tingkat akurasi yang lebih akurat dalam mengklasifikasi citra jenis bunga dengan nilai akurasi 85%. Sedangkan metode klasifikasi Euclidean mendapatkan nilai akurasi 77%. Penelitian ini juga menjelaskan bahwa cahaya ruangan dapat mempengaruhi kualitas hasil citra [23].

Dalam sebuah penelitian mengenai pengenalan bunga dengan menggunakan metode Euclidean distance diperoleh hasil akurasi sebesar 83,3% dari 12 gambar uji yang dikenali hanya 10 gambar. Pada hasil pengujian, diketahui bahwa untuk pengenalan bunga, bunga yang berhasil dikenali sebanyak 10 gambar. Sehingga tingkat keberhasilannya sebesar 51,8% [24].

Dalam sebuah penelitian tentang pengenalan pola bentuk bunga menggunakan ekstraksi fitur principal component analysis dan klasifikasi *K-Nearest Neighbor*. Dengan banyaknya jenis dan bentuk bunga maka dapat diklasifikasikan berdasarkan kesamaan cirinya melalui proses pengenalan pola. Hasil ekstraksi ciri dengan principal component analysis dipengaruhi oleh pola bentuk citra bunga. Makin seragam pola bentuk citra bunga maka memiliki vektor

ciri yang hampir sama sehingga lebih mudah dikenali pada saat proses pengenalan pola. Metode yang diusulkan menghasilkan akurasi sebesar 97,70% [25].

Dalam sebuah penelitian tentang klasifikasi pengenalan buah dengan menggunakan algoritma *Naïve Bayes* menunjukkan bahwa nilai rata-rata estimasi akurasi, sensitivitas, presisi dan spesifisitas masing-masing adalah 81%, 73%, 100% dan 70%. Ini menunjukkan bahwa *naïve bayes* memiliki potensi yang baik untuk mengidentifikasi varietas apel non destruktif dan akurat [26].

Dalam sebuah penelitian mengenai klasifikasi belimbing dengan menggunakan klasifikasi *naïve bayes* dengan metode ekstraksi ciri warna RGB, menunjukkan akurasi sebesar 80%. Ekstraksi fitur warna RGB dapat digunakan sebagai input bagi *naïve bayes* untuk pengenalan pola citra dan mengklasifikasikan buah belimbing [27].

Dalam sebuah penelitian mengenai klasifikasi bunga dengan metode LBP dan SURF sebagai fitur dan SVM sebagai classifier, berdasarkan hasil uji coba diperoleh hasil terbaik mencapai akurasi 97,2%. Menggunakan kombinasi fitur LBP dan SURF dengan classifier SVM menunjukkan hasil yang memuaskan [28].

Dalam sebuah penelitian mengenai klasifikasi bunga dengan metode Isomap untuk pencarian fiturnya menggunakan *Naïve Bayes* sebagai klasifikasinya, berdasarkan hasil uji coba diperoleh hasil terbaik mencapai akurasi 80%. Menggunakan kombinasi fitur dimensi 2 dan 3 dengan classifier *Naïve Bayes* menunjukkan hasil yang memuaskan [29].

Dalam sebuah penelitian pengenalan ekspresi wajah yang dilakukan oleh Irfan Maliki dan rekannya [30]. Digunakan metode Markov Stationary Feature - Vector Quantization untuk ekstraksi fiturnya dengan Support Vector Machine sebagai klasifikasinya. Dengan jumlah dataset yang digunakan adalah 1440 data dengan 6 klasifikasi dari ekspresi wajah. Didapatkan hasil akurasinya adalah 97.41% yang menyatakan bahwa metode ini dapat diterapkan pada pengenalan ekspresi wajah.

Dari berbagai hasil penelitian tersebut sebagian besar peneliti yang meneliti tentang pengenalan bunga memilih menggunakan metode Support Vector Machine atau SVM dan K-Nearest Neighbor atau KNN. Metode SVM dan KNN termasuk

kedalam supervised learning yang merupakan proses pembelajaran yang terawasi. Karena telah banyak peneliti sebelumnya yang telah membuktikan bahwa metode SVM dan KNN dapat digunakan untuk pengenalan bunga, maka pada penelitian ini akan digunakan metode lain yaitu metode *Naïve Bayes Classifier* yang merupakan pengklasifikasian statistic yang dapat digunakan untuk memprediksi probabilitas keanggotaan suatu kelas. Dan beberapa peneliti sebelumnya juga telah membuktikan bahwa metode ini baik dalam klasifikasi. Oleh karena itu untuk proses pengenalan bunga dalam penelitian ini akan digunakan metode *Naïve Bayes Classifier* untuk proses pengenalan dan untuk mengukur akurasi metode *Naïve Bayes Classifier* dalam mengenali bunga.