BAB II

LANDASAN TEORI

Untuk memudahkan dalam pemahaman dalam penyusunan laporan skripsi peneliti mengadakan studi pustaka mengenai arti dan istilah yang digunakan dalam penelitian sehingga memudahkan dalam pemecahan suatu masalah yang akan ditemui dalam perancangan sistem klasifikasi buah jambu biji merah.

2.1 Citra Digital

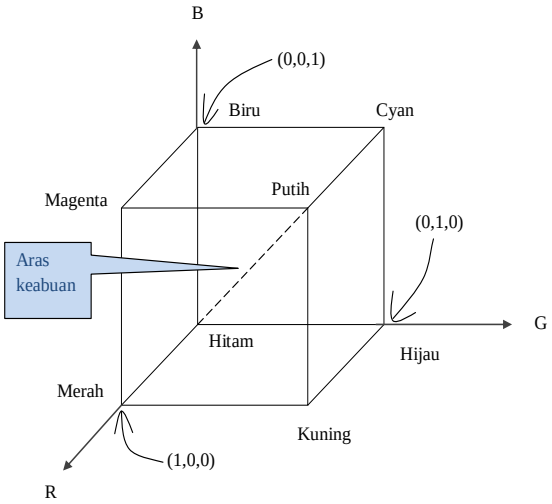
Secara umum, pengolahan citra digital menunjuk pada pemrosesan gambar 2 dimensi menggunakan komputer. Dalam konteks yang lebih luas, pengolahan Citra digital mengacu pada pemrosesan setiap data 2 dimensi. Citra digital merupakan sebuah larik (array) yang berisi nilai-nilai real maupun komplek yang direpresentasikan dengan deretan bit tertentu.

Suatu Citra dapat didefinisikan sebagai fungsi berukuran M baris dan N 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 dikata- kan bahwa citra tersebut adalah citra digital (Darma, 2010).

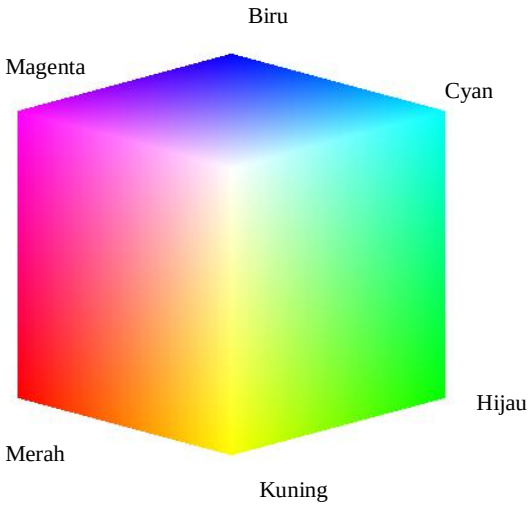
2.2 Warna RGB

Ruang warna ini menggunakan tiga komponen dasar yaitu merah (R), hijau (G), dan biru (B). Setiap piksel dibentuk oleh ketiga komponen tersebut. Model RGB biasa disajikan dalam bentuk kubus tiga dimensi, dengan warna merah, hijau, dan biru berada

pada pojok sumbu seperti gambar 2.1. Warna hitam berada pada titik asal dan warna putih berada di ujung kubus yang berseberangan. Pada gambar 2.2 memperlihatkan kubus warna secara nyata dengan resolusi 24 bit. Perlu diketahui, dengan menggunakan 24 bit, jumlah warna mencapai 16.777.216 (Mohammad, 2014).



Gambar 2. 1 Skema ruang warna RGB dalam bentuk kubus



Gambar 2. 2 Kubus warna dengan 24 bit

2.3 Warna Grayscale

Citra ini menangani gradasi warna hitam dan putih, yang tentu saja menghasilkan efek warna abu-abu. Pada jenis gambar ini, warna dinyatakan dengan intensitas. Dalam hal ini, intensitas berkisar antara 0 sampai dengan 255. Nilai 0 menyatakan hitam dan nilai 255 menyatakan putih (Mohammad, 2014).

2.4 Ekstrasi Ciri Statistik

Ekstraksi ciri merupakan metode pengambilan ciri yang didasarkan pada karakteristik histogram citra.Histogram menunjukkan probabilitas kemunculan nilai derajat keabuan piksel pada suatu citra. Dari nilai-nilai pada histogram yang dihasilkan,

dapat dihitung beberapa parameter ciri, antara lain adalah *Contrast, Correlation, Energy, Homogeneity, Mean, Variance, Standard deviation, Skewness, Kurtosis, Entropy* dan *inverse difference moment (IDM)* (Abdul Fadlil, 2012).

1. *Contrast*

   Contrast adalah ukuran variasi antar derajat keabuan suatu daerah citra. Rumus untuk mencari *contrast* dari suatu gambar sebagai berikut :

$$C = \sum_{i,j}(i-j)^2 p(i,j)$$ Rumus……....(1)

   i,j menyatakan baris ke-, kolom ke-*j*
   P(i,j) menyatakan piksel di lokasi baris ke-*i* kolom ke-*j*.

2. *Correlation*

   Correlation adalah ukuran tingkat abu-abu ketergantungan linier antara piksel pada posisi tertentu terhadap piksel lain. Rumus untuk menentukan *correlation* dari sutu gambar sebagai berikut :

$$Cr = \frac{\sum_i \sum_j (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$ Rumus……....(2)

   μ_x μ_y menyatakan rata-rata dari p_x p_y
   σ_x σ_y menyatakan standar deviasi p_x p_y

3. *Energy*

   Energy disebut juga *angular second moment* (ASM) menunjukkan nilai yang tertinggi saat piksel-piksel gambar homogen. Rumus untuk mencari *energy* dari suatu gambar sebagai berikut :

$$H = \sum_i \sum_j \frac{p(i,j)}{1+|i-j|}$$ Rumus……....(3)

4. *Homogeneity*

   Homogeneity menunjukkan nilai distribusi antara elemen. Rumus untuk mencari *homogeneity* dari suatu gambar sebagai berikut :

$$Cr = \frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \qquad \text{Rumus……....(4)}$$

5. *Mean*

   Mean merupakan rata-rata dari nilai-nilai piksel pada gambar. Rumus untuk mencari *mean* pada suatu gambar sebagai berikut :

$$\mu = \frac{1}{MN} \sum_{i=1}^{M} \sum_{i=1}^{N} P(i,j) \qquad \text{Rumus……....(5)}$$

   P(i,j) adalah nilai piksel pada titik *(i,j)* dari gambar dengan ukuran *M x N* (*M,N* menyatakan banyaknya tingkat keabuan)

6. *Variance*

   Variance menunjukan seberapa banyak tingkat keabu-abuan yang beragam dari rata-rata. Rumus untuk menghitung *variance* dari suatu gambar sebagai berikut :

$$var = \sum_i \sum_j (i-j)^2 p(i,j) \qquad \text{Rumus……....(6)}$$

7. *Standard deviation*

   Standard deviation adalah perhitungan dari akar rata-rata atau *mean* dari nilai piksel keabuan . *Standard deviation* dapat dirumuskan sebagai berikut :

$$\sigma = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{i=1}^{N} (P(i,j) - \mu)^2} \qquad \text{Rumus……....(7)}$$

8. *Skewness*

   Skewness adalah tingkat asimmetri dari distribusi piksel disekitar *mean*. *Skewness* adalah nilai dari karakteristik distribusi ukuran. Rumus umum untuk menghitung *skewness* sebagai berikut :

$$S = \frac{1}{MN}\sum_{i=1}^{M}\sum_{i=1}^{N}\left[\frac{(P(i,j)-\mu)}{\sigma}\right]^3$$ Rumus……....(8)

Dimana *P(i,j)* adalah nilai piksel pada titik *(i,j)*, μ dan σ adalah rata-rata (*mean*) dan *standard deviation*.

9. *Kurtosis*

Kurtosis merupakan ukuran dari tingkat distribusi normal dari yang tertinggi atau terendah. Rumus umum untuk menghitung *kurtosis* dalam ekstraksi gambar sebagai berikut :

$$K = \frac{1}{MN}\sum_{i=1}^{M}\sum_{i=1}^{N}\left[\frac{(P(i,j)-\mu)}{\sigma}\right]^4 - 3$$ Rumus……....(9)

Dimana *P(i,j)* adalah nilai piksel pada titik *(i,j)*, μ dan σ adalah rata-rata (*mean*) dan *standard deviation*.

10. *Entropy*

Entropy menghitung keacakan intensitas gambar (cline, dkk, 1990:1037- 1045). Rumus *entropy* (E) dari suatu gambar sebagai berikut :

$$E = \sum_i \sum_j p(i,j) log\big(p(i,j)\big)$$ Rumus…….... (10)

11. *Inverse Difference Moment (IDM)*

Inverse Difference Moment (IDM) menunjukan kesamaan piksel. Rumus *Inverse Difference Moment* (IDM) suatu gambar sebagai berikut :

$$IDM = \sum_{i=0}\sum_{j=0}\frac{p(i,j)}{1+(i-j)^2}$$ Rumus…….... (11)

2.5 Matlab

MATLAB (Matrix Laboratory) adalah sebuah program untuk analisis dan komputer numeric, merupakan satu bahasa pemrograman matematika lanjutan yang

dibentuk dengan dasar pemikiran menggunakan sifat dan bentuk matriks. Pada awalnya, program ini merupakan interface untuk koleksi rutin-rutin numeric proyek LINPACK dan EISPACK, dikembangkan menggunakan bahasa FORTRAN. Namun sekarang, program ini merupakan produk komersial dari perusahaan Mathwork, Inc. yang dalam perkembangan selanjutnya dikembangkan menggunakan bahasa dan assembler (Mohammad, 2014).

2.6 Jambu Biji Merah

Jambu biji merah (*Psidium guajava*) merupakan jenis buah tropis yang keberadaannya sulit digantikan dengan buah-buah lainya karena jambu biji merah memiliki kandungan yang sangat dibutuhkan oleh manusia. Buah jambu biji merah merupakan keluarga *mytaceae* yang memiliki bentuk bulat mempunyai warna hijau jika belum matang, kuning muda jika sudah matang dan kuning kemerahan apabila telah busuk, daging buah berwarna merah. Buah jambu biji merah memiliki kulit tipis dan permukaannya halus sampai kasar (Bambang, 2010).

2.7 Metode K-NN

a. Definisi *K - Nearest Neighbour* (K-NN)

K-Nearest Neighbour (K-NN) merupakan sebuah metode untuk melakukan klasifikasi terhadap objek baru berdasarkan (k) tetangga terdekatnya (Gourunescu, 2011). K-NN termasuk algoritma supervised learning, dimana hasil dari query instance yang baru, diklasifikasikan berdasarkan mayoritas dari kategori pada K-NN. Kelas yang paling banyak muncul yang akan menjadi kelas hasil klasifikasi.

Kelebihan K-NN :

- Lebih efektif di data training yang besar
- Dapat menghasilkan data yang lebih akurat

Kekurangan K-NN :

- Perlu ditentukan nilai k yang paling optimal yang menyatakan jumlah tetangga terdekat
- Biaya komputasi cukup tinggi karena perhitungan jarak harus dilakukan pada setiap query instance bersama-sama dengan seluruh instan dari training sample.

b. Prinsip Kerja *K-Nearcst Neighbour* (K-NN)

Prinsip kerja *K-Nearest Neighbour* (K-NN) adalah mencari jarak terdekat antara data yang akan dievaluasi dengan k tetangga (*Neighbour*) terdekatnya dalam data pelatihan. Berikut urutan proses kerja K-NN :

1. Menentukan parameter k (jumlah tetangga paling dekat).

   Parameter kepada testing ditentukan berdasarkan nilai k optimum pada saat training. Nilai k optimum diperoleh dengan mencoba-coba.

2. Menghitung jarak *euclidean* (*euclidean distance*) Obyek terhadap data sampel yang diberikan.

$$d_i = \sqrt{\sum_{i=1}^{p} (x_{2i} - x_{1i})^{\ 2}} \qquad \text{Rumus …… (12)}$$

Keterangan :

x1 = Sampel Data

x2 = Data Uji / Testing

i = Variabel Data

d = Jarak

p = Dimensi Data

3. Mengurutkan objek-objek tersebut ke dalam kelompok yang mempunyai jarak *euclidean* terkecil.
4. Mengumpulkan kategori y (klasifikasi *nearest neighbor*).