

BAB II

LANDASAN TEORI

2.1 Tebu

2.1.1. Sejarah Tebu

Tebu (*saccharum officinarum L*) merupakan tanaman asli tropika basah. Penanaman tebu di indonesia dimulai pada saat sistem tanam paksa (Tahun 1870) yang memberikan keuntungan besar untuk khas negara pemerintahan kolonial belanda. Setelah sistem tanam paksa di hentikan, usaha perkebunan tebu di lakukan oleh pengusaha-pengusaha swasta. Perluasan perkebunan tebu tidak pernah melampaui pulau jawa karena memang jenis tanaman dan pola pertanian di pulau jawa lebih sesuai untuk penanaman tebu. Daerah jantung perkebunan yang tumbuh sejak tahun 1940-an dan berkembang sampai sekarang adalah daerah pesisir utara cirebon hingga semarang di sebelah gunung muria hingga madiun, kediri, di sepanjang probolinggo hingga ke malang melalui pasuruan. Pusat penelitian dan pengembangan gula indonesia (P3GI) pasuruan telah berperan melakukan penelitian-penelitian untuk menghasilkan varietas unggul dan berbagai produk turunannya seperti fermentasi pembuatan etanol dari tetes, pembuatan ragi roti, pakan ternak, gula pasir, karton dll.

Ada beberapa manfaat tebu diantaranya digunakan untuk di konsumsi langsung dengan cara di buat jus, di buat menjadi tetes rum dan di buat menjadi ethanol yang nantinya digunakan sebagai bahan bakar. Limbah hasil produksi dari tebu bisa dimanfaatkan menjadi listrik. Ekstrak sari tebu yang di tambah jeruk nipis dan garam bisa di konsumsi di india itu dimaksudkan untuk memberikan kekuatan gigi dan gusi. Air tebu dapat di manfaatkan sebagai penyembuh sakit tenggorokan dan mencegah sakit flu serta bisa menjaga badan kita sehat. Mengkonsumsi air tebu secara teratur dapat menjaga metabolisme tubuh kita dari kekurangan cairan karena banyak kegiatan yang sudah di lakukan sehingga dapat terhindar dari stroke. Dengan banyaknya kandungan karbohidrat sehingga dapat menambah kekuatan jantung, mata, ginjal dan otak.

2.1.2. Manfaat dan Keunggulan Tanaman Tebu

Bagian yang paling utama untuk diolah dari tebu adalah batangnya. Bagian daging tebu bisa menghasilkan banyak manfaat terutama yang di olah menjadi bahan baku gula. Bentuk pohon dari tanaman tebu yaitu batang yang berbentuk memanjang keatas, dan terdapat ruas-ruas di batangnya, daunnya terdapat pada setiap ruasnya. Tebu yang tumbuh lebih dari 200 negara, India adalah produsen gula terbesar kedua di dunia sedangkan penghasil terbesarnya adalah Brasil. Oleh karena itu, tebu terkenal pemanfaatannya sebagai bahan pokok pembuatan gula. Tanaman tebu memiliki manfaat yang sangat banyak. Baik dari segi kesehatan, segi industri, segi konsumsi rumah tangga, segi peternakan, dan segi industri rumah tangga.



Gambar 2.1. Kebun Tebu

<http://www.produknaturalnusantara.com>

2.1.3. Ciri-ciri Morfologi Tanaman Tebu

a. Batang Tebu

Batang tanaman tebu tinggi ramping, tidak mempunyai cabang dan tumbuh tegak ke atas. Tinggi badan tebu bisa mencapai 3 sampai 5 meter atau bahkan lebih. Kulit batang tebu berstruktur keras, warnanya hijau, kuning, ungu, merah tua, atau gabungannya.



Gambar 2.2. Batang Tebu

b. Daun Tebu

Daun tanaman tebu merupakan daun tidak lengkap karena hanya terdiri dari pelepah dan helaian daun, tanpa tangkai daun. Daun berpangkul langsung pada buku batang dengan pola selang seling. Pelepah daun memeluk batang, makin ke atas makin sempit. Pada pelepah daun terdapat bulu-bulu dan telinga daun. Bentuk tulang daun tanaman tebu sejajar.



Gambar 2.3. Daun Tebu

<https://pixabay.com/en/leaf-nerve-american-cane-2690705/>

c. Bunga Tebu

Bunga tebu sering di katakan bunga majemuk yang tersusun atas oomalai dengan pertumbuhan terbatas. Panjang bunga majemuk yaitu sekitar 70-90cm. Setiap bunga bunga mempunyai tiga daun kelopak, satu daun mahkota, tiga benang sari, dan dua kepala p



Gambar 2.4. Bunga Tebu

<http://gendhispethak.wordpress.com>

d. Akar Tebu

Tebu memiliki akar serabut dengan panjang yang bisa mencapai satu meter. Sewaktu tanaman tebu masih muda atau masih berbentuk bibit, ada 2 macam akar, yaitu akar stek dan akar tunas. Akar stek berasal dari stek batangnya, tidak berumur panjang dan hanya berguna saat tanaman masih berumur muda. Akar tunas berasal dari tunasnya, berumur panjang dan akar tetap ada selama tanaman masih hidup. Pada tanah yang sesuai, Akar tebu dapat tumbuh panjang mencapai 1 meter.



Gambar 2.5. Akar Tebu

<https://www.google.co.id/search?q=akar+tebu&tbn>

2.2. Jenis Citra

Nilai suatu pixel memiliki nilai dalam rentang tertentu, dari nilai minimum sampai nilai maksimum. Jangkauan yang berbeda-beda tergantung dari jenis warnanya. Namun secara umum jangkauannya adalah 0-255. Citra dengan penggambaran seperti ini digolongkan kedalam citra integer. Berikut Jenis-jenis citra berdasarkan nilai pixelnya [Putra, D. 2010].

2.2.1 Segmentasi Manual

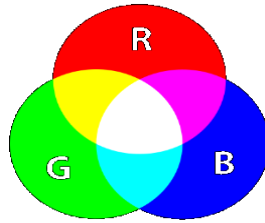
Segmentasi manual ini dilakukan dengan cara manual dikarenakan citra asli tidak dapat di segmentasi secara otomatis dengan matlab. Citra awal akan di segmentasi manual dengan *phothosop* untuk mendapatkan hasil objek yang lebih bagus, Citra RGB kemudian dikonversi menjadi *grayscale* untuk mendapatkan citra *gray* (abu-abu). Dengan proses *grayscale* ini dapat mempermudah untuk memproses gambar lebih lanjut, karena Citra *gray* hanya memiliki satu nilai kanal pada setiap pixelnya yang bernilai 0-255.

2.2.2. Citra RGB

RGB sering disebut sebagai warna *additive*. Hal ini karena warna dihasilkan oleh cahaya yang ada. Beberapa alat yang menggunakan color model RGB antara lain; mata manusia, projector, TV, kamera video, kamera digital, dan alat-alat yang menghasilkan cahaya. Proses pembentukan cahayanya adalah dengan mencampur ketiga warna tadi. Skala intensitas tiap warnanya dinyatakan dengan rentang 0 sampai 255.

Ketika warna Red memiliki intensitas sebanyak 255, begitu juga dengan Green dan Blue, maka terjadilah warna putih. Sementara ketika ketiga warna tersebut mencapai intensitas 0, maka terjadilah warna hitam, sama seperti ketika berada diruang gelap tanpa cahaya, yang tampak hanya warna hitam. Hal ini bisa dilihat ketika menonton bioskop tua di mana proyektor yang digunakan masih menggunakan proyektor dengan 3 warna dari lubang yang terpisah, bisa terlihat

ketika film menunjukkan ruangan gelap, cahaya yang keluar dari ketiga celah proyektor tersebut berkurang [Novi. DE. 2012].



Gambar 2.6 Warna RGB

2.2.3. Citra Gray

Citra grayscale merupakan citra digital yang hanya memiliki satu nilai kanal pada setiap pixelnya, dengan kata lain nilai bagian RED=GREEN=BLUE. Nilai tersebut digunakan untuk menunjukkan tingkat intensitas. Warna yang memiliki adalah warna dari hitam, keabuan dan putih. Tingkat keabuan disini merupakan warna abu dengan berbagai tingkatn dari hitam hingga mendekati putih. Citra grayscale berikut memiliki kedalaman warna 8 bit (256 kombinasi warna keabuan) [Putra, D. 2010].



Gambar 2.7. Citra Grayscale

2.3. Pemrosesan Data Awal (*Pre-Processing*)

2.3.1. Konversi Citra RGB ke Grayscale

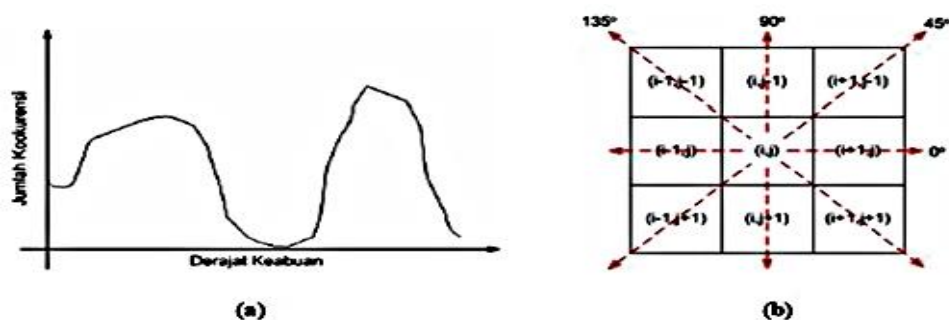
Merubah citra Rgb menjadi citra grayscale adalah salah satu contoh proses pengolahan citra grayscale adalah dengan menghitung rata-rata nilai intenstas RGB setiap pixel penyusun tersebut Rumusan matematis yang di gunakan adalah:
Citra Abu-Abu = *rgb2gray* (*a*).....(2.1)

Dimana : *a* = gambar yang di konversi ke *gray*

2.3.2 Analisis Tekstur

Tekstur merupakan karakteristik intrinsik dari suatu citra yang terkait dengan tingkat kekasaran (*roughness*), granularitas (*granulation*) dan keteraturan (*regularity*) susunan struktural piksel. Aspek tekstural dari sebuah citra dapat dimanfaatkan sebagai dasar dari segmentasi, klasifikasi, maupun interpretasi citra.

Analisis tekstur lazim dimanfaatkan sebagai proses antara untuk melakukan klasifikasi dan interpretasi citra. Suatu proses klasifikasi citra berbasis analisis tekstur pada umumnya membutuhkan tahapan ekstraksi ciri, yang terdiri dari tiga macam metode yaitu metode statistik, metode spektral, dan metode struktural. Metode *First Order* dan *Co-occurrence Matrix*, termasuk dalam metode statistik dimana dalam perhitungannya menggunakan distribusi derajat keabuan (histogram) dengan mengukur tingkat kekontrasan, granularitas, dan kekasaran suatu daerah dari hubungan ketetanggaan antar piksel didalam citra. Paradigma statistik ini penggunaannya tidak terbatas, sehingga sesuai untuk tekstur-tekstur alami yang tidak terstruktur dari sub pola dan himpunan aturan (mikrostruktur). Metode statistik terdiri dari ekstraksi ciri order pertama, ekstraksi ciri order kedua dan ekstraksi ciri order ketiga. Ekstraksi ciri statistik order kedua dilakukan dengan matriks kookurensi, yaitu suatu matriks antara yang merepresentasikan hubungan ketetanggaan antar piksel dalam citra pada berbagai arah orientasi dan jarak spasial. Ilustrasi ekstraksi ciri statistik di tunjukkan pada gambar 2.3.

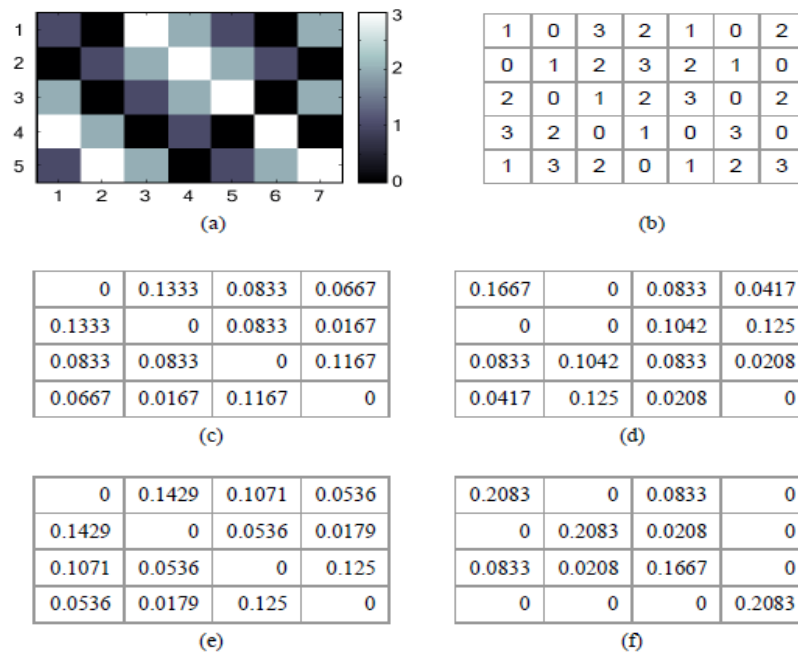


Gambar 2.8. Ilustrasi ekstraksi ciri statistik, (a) Histogram citra sebagai fungsi probabilitas kemunculan nilai intensitas pada citra, (b) Hubungan ketetanggaan antar piksel sebagai fungsi orientasi dan jarak spasial.

http://i1211.photobucket.com/albums/cc432/ardite/2011-01-23_060324.jpg

Co-occurrence berarti kejadian bersama, yaitu jumlah kejadian satu level nilai piksel bertetangga dengan satu level nilai piksel lain dalam jarak (d) dan orientasi sudut (θ) tertentu. Jarak dinyatakan dalam piksel dan orientasi dinyatakan dalam derajat. Orientasi dibentuk dalam empat arah sudut dengan interval sudut 45° , yaitu 0° , 45° , 90° , dan 135° . Sedangkan jarak antar piksel biasanya ditetapkan sebesar 1 piksel.

Co-occurrence Matrix merupakan matriks bujursangkar dengan jumlah elemen sebanyak kuadrat jumlah level intensitas piksel pada citra. Setiap titik (p, q) pada co-occurrence matrix berorientasi θ berisi peluang kejadian piksel bernilai p bertetangga dengan piksel bernilai q pada jarak d serta orientasi θ dan $(180-\theta)$.



Gambar 2.9 Ilustrasi pembuatan matriks kookurensi

- | | |
|---|--|
| (a) Citra Masukan | (b) Nilai Intensitas Citra masukan |
| (c) Hasil Matriks kookurensi 0° | (d) Hasil Matriks kookurensi 45° |
| (e) Hasil Matriks kookurensi 90° | (f) Hasil Matriks kookurensi 135° |

Setelah memperoleh matriks kookurensi tersebut, Langkah selanjutnya yakni mencari nilai rata-rata dari sudut 0° , 45° , 90° , dan 135° , hasil perhitungan dapat dilihat pada gambar 2.8

0.0938	0.0690	0.0893	0.0405
0.0690	0.0521	0.0655	0.0399
0.0893	0.0655	0.0625	0.0656
0.0405	0.0399	0.0656	0.0521

Gambar 2.10 Perhitungan nilai rata-rata

Berikut adalah nilai matriks i dan j yang nantinya akan digunakan sebagai pelengkap perhitungan.

i				j			
1	2	3	4	1	2	3	4
1	2	3	4	1	2	3	4
1	2	3	4	1	2	3	4
1	2	3	4	1	2	3	4

(a) (b)

Gambar 2.11 (a) nilai Matriks variable i (b) nilai matriks variable j

Setelah memperoleh nilai rata-rata, kita dapat menghitung ciri statistik orde dua yang merepresentasikan citra yang diamati. Menurut Haralick et al (1973) mengusulkan berbagai jenis ciri tekstural yang dapat diekstraksi dari matriks kookurensi. Dalam modul ini dicontohkan perhitungan 6 ciri statistik orde dua, yaitu *Angular Second Moment*, *Contrast*, *Correlation*, *Variance*, *Inverse Difference Moment*, dan *Entropy* [PCD4].

a. Angular Second Moment (ASM)

Menunjukkan ukuran sifat homogenitas citra.

$$ASM = \sum_i \sum_j \{p(i,j)\}^2 \dots\dots\dots (2.2)$$

Dimana :

$p(i,j)$: merupakan menyatakan nilai pada baris i dan kolom j pada matriks kookurensi.

Berikut adalah perhitungan nilai ASM

0.0938	0.0690	0.0893	0.0405
0.0690	0.0521	0.0655	0.0399
0.0893	0.0655	0.0625	0.0656
0.0405	0.0399	0.0656	0.0521

(a)

0.0088	0.0048	0.0080	0.0016
0.0048	0.0027	0.0043	0.0016
0.0080	0.0043	0.0039	0.0043
0.0016	0.0016	0.0043	0.0027

(b)

Keterangan:

- (a) Adalah nilai dari penjumlahan masing-masing sudut dibagi dengan banyaknya sudut (nilai rata-rata)
- (b) Adalah nilai dari masing-masing *pixel* yang sudah dipangkat 2
- (c) Jika nilai (b) dijumlahkan, maka hasil yang diperoleh yakni nilai $ASM = 0.0672$

b. Contrast

Menunjukkan ukuran penyebaran (momen inersia) elemen-elemen matriks citra. Jika letaknya jauh dari diagonal utama, nilai kekontrasan besar. Secara visual, nilai kekontrasan adalah ukuran variasi antar derajat keabuan suatu daerah citra. Berikut adalah adalah perhitungan nilai CON

$$CON = \sum_i k^2 [\sum_i \sum_j p(i,j)] \dots\dots\dots (2.3)$$

(i-j) ² =k			
0	1	4	9
1	0	1	4
4	1	0	1
9	4	1	0

(a)

k*rata-rata			
0	0.0690	0.3571	0.3643
0.0690	0	0.0655	0.1595
0.3571	0.0655	0	0.0656
0.3643	0.1595	0.0656	0

(b)

Keterangan:

- (a) Hasil pengurangan nilai dari variable i dengan nilai variable j kemudian dikuadratkan, perhitungan ini digunakan sebagai nilai dari variable k
- (b) Hasil perkalian dari nilai variable k dengan nilai dari variable rata-rata
- (c) Sehingga jika dikerjakan sesuai dengan rumus 2.9 maka nilai yang diperoleh untuk $CON = 2.1622$

c. Correlation

Menunjukkan ukuran ketergantungan linear derajat keabuan citra sehingga dapat memberikan petunjuk adanya struktur linear dalam citra. Berikut adalah perhitungan nilai COR.

$$COR = \frac{\sum_i \sum_j (i \cdot j) \cdot p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \dots\dots\dots (2.4)$$

Dimana:

- μ_x : adalah nilai rata-rata elemen kolom pada matriks $p(i, j)$
- μ_y : adalah nilai rata-rata elemen baris pada matriks $p(i, j)$
- σ_x : adalah nilai standar deviasi elemen pada kolom $p(i, j)$
- σ_y : adalah nilai standar deviasi elemen pada baris $p(i, j)$

0.0938	0.0690	0.0893	0.0405
0.0690	0.0521	0.0655	0.0399
0.0893	0.0655	0.0625	0.0656
0.0405	0.0399	0.0656	0.0521

(a)

μ_x	2.3865	σ_x	1.1038
μ_y	2.3865	σ_y	1.1038
$\mu_x \cdot \mu_y$	5.6952	$\sigma_x \cdot \sigma_y$	1.2184

(b)

$i \cdot j$			
1	2	3	4
2	4	6	8
3	6	9	12
4	8	12	16

(c)

$(i \cdot j) \cdot a$			
0.0938	0.1381	0.2679	0.1619
0.1381	0.2083	0.3929	0.3190
0.2679	0.3929	0.5625	0.7875
0.1619	0.3190	0.7875	0.8333

(d)

Keterangan:

- (a) Adalah nilai dari penjumlahan masing-masing sudut dibagi dengan banyaknya sudut (nilai rata-rata)
- (b) Adalah nilai rata-rata dan standar deviasi
- (c) Adalah nilai perkalian matriks i dengan j
- (d) Adalah nilai dari matriks (c) dikalikan dengan nilai dari matriks (a)
- (e) Sehingga jika dikerjakan sesuai dengan rumus 2.10 maka nilai yang diperoleh untuk COR = 0.1127

d. Variance

Menunjukkan variasi elemen-elemen matriks kookurensi. Citra dengan transisi derajat keabuan kecil akan memiliki variansi yang kecil pula. Berikut adalah perhitungan nilai VAR

$$VAR = \sum_i \sum_j (i - \mu_x)(j - \mu_y)p(i, j) \dots \dots \dots (2.5)$$

<i>i - μ_x</i>			
-1.3865	-1.3865	-1.3865	-1.3865
-0.3865	-0.3865	-0.3865	-0.3865
0.6135	0.6135	0.6135	0.6135
1.6135	1.6135	1.6135	1.6135

(a)

<i>j - μ_y</i>			
-1.3865	-1.3865	-1.3865	-1.3865
-0.3865	-0.3865	-0.3865	-0.3865
0.6135	0.6135	0.6135	0.6135
1.6135	1.6135	1.6135	1.6135

(b)

<i>(i - μ_x) * (j - μ_y)</i>			
1.9223	0.5358	-0.8506	-2.2371
0.5358	0.1494	-0.2371	-0.6236
-0.8506	-0.2371	0.3764	0.9900
-2.2371	-0.6236	0.9900	2.6035

(c)

0.1802	0.0370	-0.0760	-0.0905
0.0370	0.0078	-0.0155	-0.0249
-0.0760	-0.0155	0.0235	0.0650
-0.0905	-0.0249	0.0650	0.1356

(d)

Keterangan:

1. Adalah hasil pengurangan antara matriks pada variable *i* dengan nilai pada variable μ_x
2. Adalah hasil pengurangan antara matriks pada variable *j* dengan nilai pada variable μ_y
3. Adalah hasil perkalian dari (a) dengan (b)

4. Adalah hasil perkalian antara nilai pada hasil (c) dengan nilai rata-rata awal
5. Sehingga jika dikerjakan sesuai dengan rumus 2.11 maka nilai yang diperoleh untuk COR = 0.1373

e. Inverse Difference Moment

Menunjukkan kehomogenan citra yang berderajat keabuan sejenis. Citra homogen akan memiliki harga IDM yang besar. Berikut adalah perhitungan nilai IDM

$$IDM = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j) \dots\dots\dots (2.6)$$

$1+(i-j)^2$			
1	2	5	10
2	1	2	5
5	2	1	2
10	5	2	1

(a)

$\frac{1}{1+(i-j)^2} p(i,j)$			
0.0938	0.0345	0.0179	0.0040
0.0345	0.0521	0.0327	0.0080
0.0179	0.0327	0.0625	0.0328
0.0040	0.0080	0.0328	0.0521

(b)

Keterangan:

- (a) Adalah hasil 1 ditambah dengan matriks variable i dikurang dengan matriks variable j kemudian dikuadratkan
- (b) Adalah hasil dari 1 dibagi dengan hasil (a) kemudian dikalikan dengan nilai rata-rata awal
- (c) Sehingga jika dikerjakan sesuai dengan rumus 2.12 maka nilai yang diperoleh untuk IDM = 0.5203

f. Entropy

Menunjukkan ukuran ketidakteraturan bentuk. Harga ENT besar untuk citra dengan transisi derajat keabuan merata dan bernilai kecil jika struktur citra tidak teratur (bervariasi). Berikut adalah perhitungan nilai ENT

$$ENT_2 = - \sum_i \sum_j p(i,j) \cdot {}^2\log p(i,j) \dots\dots\dots (2.7)$$

$$-\sum_i p(i,j)$$

-0.0938	-0.0690	-0.0893	-0.0405
-0.0690	-0.0521	-0.0655	-0.0399
-0.0893	-0.0655	-0.0625	-0.0656
-0.0405	-0.0399	-0.0656	-0.0521

$${}^2\log p(i,j)$$

-3.4150	-3.8563	-3.4854	-4.6268
-3.8563	-4.2630	-3.9329	-4.6482
-3.4854	-3.9329	-4.0000	-3.9296
-4.6268	-4.6482	-3.9296	-4.2630

(a) (b)

$$-\sum_i p(i,j) * {}^2\log p(i,j)$$

0.3202	0.2663	0.3112	0.1873
0.2663	0.2220	0.2575	0.1854
0.3112	0.2575	0.2500	0.2579
0.1873	0.1854	0.2579	0.2220

(c)

Keterangan:

- (a) Adalah hasil dari nilai rata-rata awal dengan ditambah min
- (b) Adalah hasil dari ${}^2\log p(i,j)$
- (c) Adalah hasil perkalian antara nilai hasil pada (a) dengan nilai hasil pada (b)

Sehingga jika dikerjakan sesuai dengan rumus 2.12 maka nilai yang diperoleh untuk IDM = 3.9452

2.3.3. Square Euclidean

Square euclidean digunakan untuk menentukan perhitungan jarak terdekat nilai vektor ciri citra uji dengan citra acuan. Nilai *square euclidean* yang mendekati nilai nol, akan menunjuk pada citra tertentu. Nilai vektor ciri citra masukan yang memiliki nilai vektor ciri yang sama dengan vektor ciri citra tertentu akan memiliki nilai *square euclidean* yang mendekati nol.

Rumus menghitung *square euclidean* sebagai berikut :

Rumus menghitung jarak euclidian

$$d_i = \sqrt{\sum_{l=1}^n (X_{ij} - P_j)^2} \dots\dots\dots (2.8)$$

Keterangan:

- d_i = Jarak sampel
 X_{ij} = Data sampel pengetahuan
 P_j = Data input variabel ke-j
 n = Jumlah sample

Berikut adalah contoh perhitungan menggunakan *Square Euclidean*:

$$\text{Hitung} = \text{SQRT} = (\text{ASM_Latih} - \text{ASM_Uji})^2 + (\text{CON_Latih} - \text{CON_Uji})^2 + (\text{COR_Latih} - \text{COR_Uji})^2 + (\text{VAR_Latih} - \text{VAR_Uji})^2 + (\text{IDM_Latih} - \text{IDM_Uji})^2 + (\text{ENT_Latih} - \text{ENT_Uji})^2$$

2.4. Metode K-Nearest Neighbor

Metode komputasi statistik telah diperkenalkan untuk mengekstrak informasi tutupan lahan secara digital di citra penginderaan jauh. Beberapa diantaranya mampu bekerja tanpa didasarkan pada *training area*, akan tetapi beberapa metode menuntut adanya daerah sampel atau *training area* untuk mengekstrak informasi tutupan lahan tersebut, metode seperti ini disebut *supervised classification*. Algoritma KNN terlebih dahulu mengidentifikasi daerah sampel sebelum melakukan penunjukkan kelas. Dalam proses identifikasi daerah-daerah sampel ini sebagian algoritma akan mengambil beberapa parameter statistik tertentu dan masing-masing kelas daerah sampel, seperti ini rata-rata, variasi, dan sebagainya. Algoritma-algoritma seperti ini dikenal sebagai algoritma klasifikasi parametric.

Algoritma klasifikasi parametrik biasanya disertai dengan sejumlah asumsi tertentu, yang sepenuhnya telah ditentukan sebelum proses k berlangsung. Misalnya sebagian classifier berasumsi bahwa data (nilai spektral) homogen atau terdistribusi menurut pola tertentu. Dikarenakan algoritma klasifikasi parametrik berdasarkan atas asumsi tertentu yang dibuat sebelum proses klasifikasi, algoritma-algoritma ini dinilai memiliki keterbatasan jika kemudian asumsi yang digunakan tidak sesuai dengan keadaan atau kondisi data yang sebenarnya. Sejumlah ahli penginderaan jauh menyatakan bahwa seringkali asumsi dasar yang digunakan oleh algoritma klasifikasi parametrik tertentu tidak sesuai dengan kondisi data yang sebenarnya.

Algoritma klasifikasi non-parametrik dinilai lebih siap untuk dihadapkan pada berbagai kondisi data. Algoritma klasifikasi non-parametrik memiliki asumsi dasar yang relative lebih sedikit dibandingkan algoritma klasifikasi parametrik, lebih jauh lagi algoritma ini tidak memerlukan parameter statistik tertentu dan daerah sampel, seperti rata-rata atau variasi.

Algoritma klasifikasi non-parametrik yang memiliki formulasi statistik paling sederhana dan paling mudah diimplementasikan adalah (*K-Nearest Neighbour*). K-NN masuk kategori supervised algorithm karena label data tetap dipakai dalam proses training, yaitu optimisasi nilai k (dan juga saat klasifikasi). Yang membedakannya adalah ke neural network, misalnya, pasangan data-label dipakai untuk menentukan parameter vector (*weights*) dalam proses training. Setelah proses training selesai, training data itu dibuang, tidak diperlukan lagi dalam proses klasifikasi. Adapun K-NN termasuk kategori “memory based method”, yaitu seluruhnya atau sebagian data training disimpan dan dipakai dalam proses klasifikasi. Ciri dan memory based method proses trainingnya berlangsung cepat, sedangkan proses klasifikasinya lambat.

Algoritma K-NN (*K-Nearest Neighbor*) sensitif terhadap keberadaan noise & irrelevant input, dibandingkan dengan neural network. Dalam proses training neural network (*perceptron*), weight untuk atribut yang tidak memiliki kontribusi terhadap klasifikasi akan di-adjust agar nilainya mendekati 0, tetapi dalam K-NN, tidak ada perbedaan bobot untuk tiap atribut.

Data pembelajaran diproyeksikan ke ruang berdimensi banyak, dimana masing-masing dimensi merepresentasikan fitur dan data. Ruang ini dibagi menjadi bagian-bagian berdasarkan klasifikasi data pembelajaran. Sebuah titik pada ruang ini ditandai kelas c jika kelas c merupakan klasifikasi yang paling banyak ditemui pada k buah tetangga terdekat titik tersebut. Dekat atau jauhnya tetangga biasanya dihitung berdasarkan jarak Euclidean.

Algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi dari data pembelajaran. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk data test (yang klasifikasinya tidak diketahui). Jarak dan vektor yang baru ini terhadap seluruh vektor data pembelajaran dihitung, dan sejumlah k

buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak da titik-titik tersebut.

Nilai k yang terbaik untuk algoritma ini tergantung pada data; secara umumnya, nilai k yang tinggi akan mengurangi efek noise pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi lebih kabur, Nilai yang bagus dapat dipilih dengan optimasi parameter, misalnya dengan menggunakan cross-validation. Kasus khusus di mana klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain, $k = 1$) disebut algoritma nearest neighbor

Ketepatan algoritma K-NN ini sangat dipengaruhi oleh ada atau tidaknya fitur-fitur yang tidak relevan. atau jika bobot fitur tersebut tidak setara dengan relevansinya terhadap idasifikasi. Riset terhadap algoritma ini sebagian besar membahas bagaimana memilih dan memberi bobot terhadap fitur, agar performa - klasifikasi menjadi lebih baik. Algoritma k-NN ini memiliki konsistensi yang kuat. Ketika jumlah data mendekati tak hingga, algoritma ini menjamin error rate yang tidak lebih dari dua kali Bayes *error rate* (*error rate* minimum untuk distribusi data tertentu).

Konsep dasar da KNN adalah mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga terdekatnya dalam data pelatihan. Jumlah kelas yang paling banyak dengan jarak terdekat tersebut akan menjadi kelas dimana data data evaluasi tersebut berada.

Algoritma KNN :

1. Mulai
 2. Tentukan parameter K
 3. Hitung jarak antara data yang akan dievaluasi dengan semua pelatihan
- Jarak ini didapatkan dari data yang telah di ekstraksi metode *Co-occurrence Matrix* yaitu data baru yang akan di uji di kurangi dengan semua data pelatihan yang di sebut dengan jarak Euclidian.

4. Urutkan jarak yang terbentuk (urut naik)
Setelah diketahui jarak setiap data pelatihan, maka data di urutkan atau di sorting mulai dari data terkecil sampai data yang paling besar lalu cari jarak yang paling kecil (terdekat) dengan data pelatihan.
5. Tentukan jarak terdekat sampai urutan K
6. Pasangkan kelas yang bersesuaian
Setelah di tentukan K terdekatnya berapa, maka bisa di ketahui beberapa kelas yang akan di tujukan pada data uji tersebut, apakah masuk ke kelas pohon kepala sawit muda atau pohon kepala sawit tua
7. Cari jumlah kelas dan tetangga yang terdekat dan tetapkan kelas tersebut sebagai kelas data yang akan dievaluasi
8. Selesai.

K-Nearest Neighbour berasumsi bahwa pixel-pixel yang saling berdekatan satu sama lain (dikenal sebagai pixel-pixel bertetangga) di dalam *feature space* akan tergolong ke dalam kelas yang sama. Dalam proses klasifikasinya, *K-Nearest Neighbour* akan mengukur jarak spektral (*spectral distance*) setiap *pixel* yang ada pada citra terhadap piksel-piksel yang berada di bawah daerah sample

K-Nearest Neighbour dihadapkan pada satu masalah utama, yaitu lamanya waktu proses klasikasi. Hal ini dikarenakan *K-Nearest Neighbour* mengukur jarak spektral setiap pixel ke semua pixel yang ada dalam daerah sample. Waktu proses klasifikasi akan semakin lama seiring semakin banyaknya jumlah saluran (band), bertambah besarnya ukuran daerah sampel dan bertambahnya jumlah kelas yang akan diklasikasikan.

Kemungkinan metode yang dapat dilakukan untuk mengefisiensikan waktu eksekusi *K-Nearest Neighbour* adalah dengan memodifikasi algoritma pengukur jarak spektral pada algoritma ini. Umumnya *K-Nearest Neighbour* menggunakan Euclidean Distance sebagai metode pengukur jarak spektral.

Akan tetapi, metode pengukur jarak spektral lainnya seperti Manhattan Distance juga dapat digunakan dalam *K-Nearest Neighbour*. *Manhattan Distance* memiliki bentuk formulasi yang lebih sederhana dibandingkan dengan Euclidean Distance. Dilihat dari bentuk persamaan matematisnya, sudah dapat diprediksi

bahwa algoritma pengukur jarak spektral ini akan memiliki waktu kalkulasi yang lebih cepat dibanding Euclidean Distance.

2.5. Matlab (*Matrix Laboratory*)

Matlab adalah salah satu *software* aplikasi untuk menyelesaikan berbagai masalah teknis. Matlab mengintegrasikan komputasi, visualisasi, dan pemrograman dalam suatu model yang sangat mudah untuk dipakai, dimana masalah-masalah dan penyelesaiannya diekspresikan dalam notasi matematika yang familiar. Penggunaan Matlab meliputi bidang-bidang:

- Matematika dan Komputasi
- Pembentukan Algoritma
- Akusisi Data
- Pemodelan, Simulasi, dan Pembuatan *Prototipe*
- Analisa Data, Eksplorasi, dan Visualisasi
- Grafik Keilmuan dan Bidang Rekayasa

Matlab merupakan suatu sistem interaktif yang memiliki elemen data dalam suatu *array* sehingga memungkinkan untuk memecahkan banyak masalah teknis yang terkait dengan komputasi, khususnya yang berhubungan dengan matriks dan formulasi vektor.

Fitur-fitur Matlab sudah banyak dikembangkan yang lebih dikenal dengan nama *toolbox*. *Toolbox* merupakan kumpulan dari fungsi-fungsi Matlab (*M-files*) yang telah dikembangkan ke suatu lingkungan kerja Matlab untuk memecahkan masalah. Area-area yang sudah bisa dipecahkan dengan *toolbox* saat ini meliputi pengolahan sinyal, sistem kontrol, *neural networks*, *fuzzy logic*, *wavelets*, dan lain-lain.

2.6. Penelitian Sebelumnya

1. KLASIFIKASI UMUR LAHAN KELAPA SAWIT PADA CITRA FOTO UDARA BERDASARKAN TEKSTUR MENGGUNAKAN FUZZY-KNN. Pada tahun 2015 Siti Jumaidah. dari fakultas Teknik jurusan Informatika Universitas Muhammadiyah Gresik telah melakukan penelitian tersebut sebagai Tugas Akhir (Skripsi). Penelitian ini difungsikan untuk mengklasifikasi umur lahan kelapa sawit berdasarkan tekstur dengan penyelesaian menggunakan metode Co-occurrence Matrix, tingkat keakuratan mencapai 85%.
2. KLASIFIKASI UMUR LAHAN PERKEBUNAN KELAPA SAWIT PADA CITRA FOTO UDARA BERDASARKAN TEKSTUR MENGGUNAKAN NAÏVE BAYES. Pada tahun 2015 Elin Rosalina. dari fakultas Teknik jurusan Informatika Universitas Muhammadiyah Gresik telah melakukan penelitian tersebut sebagai Tugas Akhir (Skripsi). Penelitian ini difungsikan untuk mengklasifikasi umur lahan perkebunan kelapa sawit berdasarkan tekstur dengan penyelesaian menggunakan metode Co-occurrence Matrix, tingkat keakuratan mencapai 71,5%.
3. SYSTEM PENDUKUNG KEPUTUSAN UNTUK MENENTUKAN JURUSAN PADA SISWA SMA DENGAN MENGGUNAKAN METODE K-NEAREST NEIGHBOR. Pada tahun 2010 M. Aminul Fuad dari fakultas Teknik jurusan Informatika Universitas Muhammadiyah Gresik telah melakukan penelitian tersebut sebagai Tugas Akhir (Skripsi). Penelitian ini difungsikan untuk mengklasifikasikan penjurusan siswa SMA berdasarkan nilai yang diperoleh siswa dengan menggunakan metode klasifier K-Nearest Neighbor, tingkat keberhasilan program mencapai 80,55%
4. PENGOLAHAN CITRA IDENTIFIKASI KUALITAS BUAH MENGKUDU (*MORINDA CITRIFOLIA*) BERDASARKAN WARNA DAN TEKSTUR MENGGUNAKAN ANALISIS *CO-OCCURRENCE*

MATRIX. Pada tahun 2012 Hamdan Sobri Andhika. dari fakultas Teknik jurusan Informatika Universitas Muhammadiyah Gresik telah melakukan penelitian tersebut sebagai Tugas Akhir (Skripsi). Penelitian ini difungsikan untuk mengenali kualitas buah mengkudu berdasarkan tekstur dengan penyelesaian menggunakan metode Co-occurrence Matrix. Dalam penyelesaian menggunakan metode tersebut, tingkat keberhasilan program mencapai 71,4%.

5. **KLASIFIKASI UMUR POHON KELAPA SAWIT BERDASARKAN TEKSTUR MENGGUNAKAN METODE K-NEAREST NEIGHBOR**
Pada tahun 2013 Mohamad Hilmy dari fakultas Teknik jurusan Informatika Universitas Muhammadiyah Gresik telah melakukan penelitian tersebut sebagai Tugas Akhir (Skripsi). Penelitian ini difungsikan untuk Pengidentifisian jenis kerang dengan metode K-NN (K-Nearest Neighbor).dengan menggunakan metode KNN yang mengacu pada pennggunaan Square Euclidean memiliki tingkat akurasi 85.21%.