

# IMPLEMENTASI METODE *FUZZY K-NEAREST NEIGHBOR* UNTUK KLASIFIKASI PENYAKIT TANAMAN KOPI BERDASARKAN CITRA DAUN

Nurul A'yun Hanifiyah<sup>1)</sup>, Imam Cholissodin, S.Si., M.Kom.<sup>2)</sup>,  
Drs. Marji, M.T.<sup>3)</sup>

Program Studi Teknik Informatika  
Program Teknologi Informasi dan Ilmu Komputer  
Universitas Brawijaya, Malang 65145, Indonesia  
email: ayunhanifiyah[at]gmail.com<sup>1)</sup>, imam.cholissodin[at]ub.ac.id<sup>2)</sup>, marji[at]ub.ac.id<sup>3)</sup>

## Abstrak

Pada perkembangan teknologi saat ini, penerapan teknologi informasi dan komunikasi dilakukan terhadap semua bidang, termasuk bidang pertanian dan perkebunan untuk tanaman kopi. Rendahnya tingkat produksi petani kopi hanya difokuskan pada pengolahan akhir kopi sedangkan penyebab utama rendahnya produktivitas adalah pada proses pembudidayaan dan proses tanam termasuk penanggulangan penyakit. Padahal penyakit kopi akan mudah diatasi jika dideteksi lebih cepat dari gejala-gejala yang terlihat termasuk yang terlihat pada bagian daun. Salah satu penerapan teknologi informasi dan komunikasi pada bidang pertanian dan perkebunan ialah melakukan komputasi untuk mengidentifikasi dan klasifikasi penyakit kopi berdasarkan gejala yang ditampakan pada bagian daun. Identifikasi penyakit menggunakan teknik Data Mining dengan algoritma *Fuzzy K-Nearest Neighbor*. Citra terlebih dahulu dilakukan proses *Thresholding* untuk memperbaiki kualitas citra untuk proses klasifikasi dengan metode *Otsu*. Dataset yang digunakan berupa citra daun kopi berpenyakit yang meliputi penyakit bercak daun kopi dan penyakit karat daun kopi. Metode yang digunakan adalah *Fuzzy K-Nearest Neighbor* yang melakukan prediksi data uji menggunakan nilai keanggotaan pada data uji tiap kelas, kemudian diambil kelas dengan nilai keanggotaan terbesar dari data uji sebagai hasil kelas prediksi. Pengujian dilakukan terhadap empat variabel pengujian yaitu pengujian terhadap perbandingan data latih dan data uji dengan akurasi tertinggi yaitu 100% menggunakan perbandingan 90:10 dengan komposisi 54 data latih dan 6 data uji. Kemudian pengujian kedua menggunakan variabel  $k$  dengan nilai antara rentang 3 sampai dengan 7 dengan akurasi 94,4%. Dan pengujian terhadap nilai  $m$  dengan nilai 2 menghasilkan akurasi 90,8%, serta pengujian terhadap nilai peubah *Threshold* dengan nilai 10 menghasilkan akurasi 90,8%.

**Kata kunci :** kopi, *Thresholding*, *Otsu*, klasifikasi, *Fuzzy K-Nearest Neighbor*

## Abstract

*In the current technological developments, the application of information and communication technology is made to all fields, including agriculture and plantations for coffee plants. The low level of production of coffee farmers are only focused on the final processing of coffee while the main causes of low productivity is in the process of cultivation and planting processes including disease prevention. Though coffee disease will be easily overcome if detected much faster than the visible symptoms include visible on the leaves. One application of information technology and communication technology in the field of agriculture and farming is to perform computation to identify the disease based on symptoms coffee that is displayed on the leaves. Identification of disease using Data Mining techniques with Fuzzy K-Nearest Neighbor algorithm. First, the system doing image Thresholding to increase image quality for the classification process with Otsu method. Dataset use the image of the diseased coffee leaves include karat daun kopi and bercak daun kopi. The method used is Fuzzy K-Nearest Neighbor who perform predictive data test using data test the value of membership in each class, then taken classes with membership biggest value of the data test as a result of class prediction. Tests conducted on four variables, namely the testing of comparative testing data training and data testing with the highest accuracy of 100% uses the ratio of 90:10 with compositions 54 data training and 6 data testing grades were on the Threshold of achieving the highest accuracy 94.4% in value Threshold 10 with the amount of data training and data testing 36 and 24 tests on the value of  $k$  produce accuracy is 94.4% on  $k$  between 1 to 5 with the amount of data training and data testing 36 and 24, and value of  $m$  produce the highest accuracy 94.4% on the value of  $m$  1 and 2 the amount of data training and data testing 36 and 24.*

**Keywords:** coffee, *Thresholding*, *Otsu*, classification, *Fuzzy K-Nearest Neighbor*

# 1. PENDAHULUAN

## 1.1 Latar Belakang

Kopi (*Coffea* spp. L) merupakan salah satu hasil komoditi perkebunan yang memiliki nilai ekonomis yang cukup tinggi di antara tanaman perkebunan lainnya dan berperan penting sebagai sumber devisa negara. Kopi tidak hanya berperan penting sebagai sumber devisa melainkan juga merupakan sumber penghasilan bagi tidak kurang dari satu setengah juta jiwa petani kopi di Indonesia [8].

Kopi merupakan tanaman dengan nilai ekonomis yang cukup tinggi, akan tetapi tanaman kopi merupakan tanaman yang banyak disukai oleh jenis serangga dan hama, sampai saat ini tercatat lebih dari 900 jenis serangga hama pada tanaman kopi yang tersebar diseluruh dunia. Hama dan penyakit dalam budidaya tanaman kopi dapat menurunkan produksi cukup besar yaitu sampai 40%. Penyakit pada tanaman kopi bisa dideteksi melalui gejala yang ditampakan pada daun seperti penyakit karat daun kopi, bercak daun kopi, jamur upas, dan Jelaga.

Untuk identifikasi penyakit pada tanaman kopi, biasanya petani kopi hanya mengamati pada tampilan luar tanaman. Oleh karena pengamatan dilakukan dengan mata telanjang, maka penyakit tanaman kopi sering kali tidak teridentifikasi dengan akurat, sehingga menyebabkan penyakit yang lebih serius pada tanaman. Identifikasi yang benar perlu dilakukan untuk tindakan pencegahan dan pengobatan pada penyakit [3]. Dengan menerapkan teknologi informasi dan komunikasi di bidang pertanian, seharusnya masalah seperti ini dapat ditangani.

Salah satu penerapan teknologi informasi dan komunikasi di bidang pertanian yaitu dapat digunakan sebagai alat bantu klasifikasi penyakit tanaman kopi. Klasifikasi merupakan suatu proses penggabungan atau pengelompokan dua atau lebih data sesuai dengan kesamaan suatu kriteria tertentu. *K-nearest neighbor* (*KNN*) merupakan metode klasifikasi termasuk kelompok *instance-based learning*. Algoritma ini juga merupakan salah satu teknik *lazy learning*. *KNN* dilakukan dengan mencari kelompok *k* objek dalam data *training* yang paling dekat (mirip) dengan objek pada data baru atau data testing. Salah satu penerapan penelitian yang menggunakan metode *K-nearest neighbor* (*KNN*) adalah penggunaan metode *K-nearest neighbor* untuk pengklasifikasian penyakit tanaman jeruk keprok dengan citra daun. Dalam penelitian ini, metode *K-nearest neighbor* (*KNN*) dipakai untuk klasifikasi data uji yaitu daun jeruk sebagai objeknya, dan hasil dari penelitian tersebut adalah identifikasi penyakit tanaman jeruk dengan akurasi mencapai 96,67% [6]. Akan tetapi modifikasi penggunaan metode *KNN* perlu ditingkatkan untuk mengetahui tingkat akurasi metode tersebut apabila digabungkan dengan metode lain. Salah satu cara untuk menemukan

keakurasian tersebut adalah dengan menggabungkannya dengan metode *Fuzzy*. Contoh penerapan *Fuzzy K-NN* adalah pada sebuah penelitian dengan penerapan metode *Fuzzy K-NN* pada klasifikasi jenis mangga berdasarkan tekstur daun yang mengambil objek tekstur tanaman buah mangga dan diekstraksi fitur citranya dan menghasilkan akurasi 81,1574% [1].

Oleh karena itu, pada skripsi ini penulis ingin mencoba mengimplementasikan algoritma *Fuzzy K-nearest neighbor* (*FK-NN*) untuk klasifikasi penyakit tanaman kopi. Dengan penerapan metode ini diharapkan akan menghasilkan akurasi yang lebih tinggi daripada penelitian yang sudah ada. Berdasarkan latar belakang yang telah dijelaskan di atas, maka penulis mengambil judul untuk skripsi ini adalah "Implementasi Metode *FUZZY K-NEAREST NEIGHBOR* Untuk Klasifikasi Penyakit Tanaman Kopi Berdasarkan Citra Daun".

## 1.2 Permasalahan

Dari paparan pendahuluan, penelitian ini merumuskan permasalahan Bagaimana mengimplementasikan metode *Fuzzy K-Nearest Neighbor* (*FK-NN*) untuk klasifikasi penyakit tanaman kopi berdasarkan citra daun dan Bagaimana akurasi yang diperoleh dari klasifikasi menggunakan metode *Fuzzy K-Nearest Neighbor* (*FK-NN*) pada tanaman kopi berdasarkan citra daun.

# 2. TINJAUAN PUSTAKA

## 2.1 Studi terkait

*Application Of support Vector Machine for Rice Diseases using shape and color* yang dilakukan oleh Qing Yao pada tahun 2009 yaitu pembuatan perangkat lunak untuk identifikasi penyakit padi berdasarkan citra daun padi yang telah terinfeksi penyakit. Objek yang digunakan adalah daun dari padi yang memiliki penyakit hawar, hawar pada pelepah padi, dan blas. *Preprocessing* citra yang digunakan *Image acquisition*, *Preprocessing image*, *Segmentasi image*, *Feature extraction*, Klasifikasi dengan *Support Vector Machine*. Hasil dari penelitian ini adalah klasifikasi penyakit bakteri hawar, hawar pada pelepah padi, dan blas dengan tingkat akurasi 97,2% [10].

Metode *Fuzzy K-NN* juga digunakan Çiğdem Beyan, Hasan Oğul dengan penelitian berjudul *A Fuzzy K-NN Approach For Cancer Diagnosis With Microarray Gene Expression Data*. Dataset yang kita digunakan dalam penelitian ini adalah dataset yang diproduksi oleh *oligonucleotidebased teknologi microarray*. Menggunakan 6 dataset yang memiliki 2-9 kategori yang berbeda, 50-102 sampel dan 2308- 10.509 gen. Tujuan dari penelitian ini adalah untuk mendiagnosa tumor dan kanker. *Fuzzy K-NN* digunakan digunakan untuk

memecahkan masalah tersebut. Penelitian ini mencapai tingkat akurasi 94.0% pada objek *SRBCT* [2].

## 2.2 Dasar Teori

### Kopi

Tanaman kopi termasuk dalam Kingdom Plantae, Sub kingdom *Tracheobionta*, Super divisi *Spermatophyta*, Divisi *Magnoliophyta*, Class *Magnoliopsida/Dicotyledons*, Sub class *Asteridae*, Ordo *Rubiales*, Famili *Rubiaceae*, Genus *Coffea*, Spesies *Coffea arabica L*. Tanaman kopi termasuk dalam famili *Rubiaceae* dan terdiri atas banyak jenis, yakni *Coffea arabica*, *Coffea robusta* dan *Coffea liberica*. Beberapa penyakit kopi yang bisa diidentifikasi dari daun adalah sebagai berikut :

#### Karat daun kopi

Penyakit karat daun kopi (*coffee leaf rust*) yang disebabkan oleh jamur *Hemileia vastatrix* B. et Br. adalah penyakit kopi paling penting dan paling berpengaruh di seluruh dunia, hal ini juga berlaku untuk tanaman kopi arabika di Indonesia. Penyakit ini dapat menyebabkan penurunan produksi mencapai 50% (Zambolim *et al.*, 1997 *cit. Haddad et al.*, 2009). Gejala dari penyakit karat daun adalah sisi bawah daun yang terserang karat menunjukkan adanya bercak-bercak yang semula berwarna kuning muda yang akhirnya akan menjadi kuning tua dan coklat. Pada bercak terbentuk tepung berwarna jingga cerah (*bright orange*) yang terdiri atas *urediospora* jamur. Bercak tua berwarna coklat tua sampai hitam dan mengering, daun akhirnya gugur sehingga pohon menjadi gundul



Gambar 1 Citra daun berpenyakit karat daun

#### Bercak daun kopi

Penyakit ini disebabkan oleh jamur *C. coffeicola* yang dapat muncul di pembibitan sampai tanaman dewasa serta menyerang buah kopi. Daun yang sakit timbul bercak berwarna kuning yang tepinya dikelilingi halo (lingkaran) berwarna kuning. Penyakit ini umumnya dijumpai di pertanaman yang kurang mendapat pemeliharaan.

Penyebaran penyakit dibantu oleh keadaan lingkungan yang lembab dan pola tanam yang kurang baik. Penyebaran penyakit melalui spora yang terbawa angin dan aliran air hujan serta alat-alat pertanian.



Gambar 2 Citra daun berpenyakit bercak daun

### Citra Digital

Citra adalah suatu representasi (gambaran), kemiripan, atau imitasi dari suatu objek. Citra terbagi 2 yaitu ada citra yang bersifat analog dan ada citra yang bersifat digital. Citra analog adalah citra yang bersifat kontinu seperti gambar pada monitor televisi, foto sinar X, hasil *CT Scan* dll. Sedangkan pada citra digital adalah citra yang dapat diolah oleh komputer [9]. Jenis-jenis Citra Digital adalah :

Citra Biner atau monokrom adalah citra digital yang hanya memiliki dua kemungkinan nilai pixel yaitu hitam dan putih. Citra biner juga disebut sebagai citra B&W (*black and white*) atau citra monokrom [7].

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 dimiliki adalah warna hitam, keabuan, dan putih [7].

Citra Warna (8 bit) merupakan citra digital yang memiliki setiap pixel dari citra warna (8 bit) hanya diwakili oleh 8 bit dengan jumlah warna maksimum yang dapat digunakan adalah 256 warna. Ada 2 jenis citra warna 8 bit [7].

### Metode Otsu

Pendekatan yang dilakukan oleh metode *Otsu* adalah dengan melakukan analisis diskriminan yaitu menentukan suatu variabel yang dapat membedakan antara dua atau lebih kelompok yang muncul secara alami. Analisis Diskriminan akan memaksimalkan variabel tersebut agar dapat membagi objek latar depan (*foreground*) dan latar belakang (*background*).

Formulasi dari metode *Otsu* adalah dengan melibatkan Nilai Ambang yang akan dicari dari suatu citra *gray level* dinyatakan dengan  $k$ . Nilai  $k$  berkisar antara 0 sampai dengan  $L$ , dengan nilai  $L = 255$  [4]. Probabilitas setiap pixel pada level ke  $i$  dapat dinyatakan pada persamaan 1.

$$P_i = n_i/N \quad (1)$$

Dimana:

$P_i$  adalah probabilitas pixel ke- $i$  menyatakan jumlah pixel pada level ke  $i$ ,  $N$  menyatakan total jumlah pixel pada citra. Nilai *Zeroth cumulative*

moment, First cumulative moment, Dan Total Nilai Rata-Rata berturut-turut dapat dinyatakan dengan rumus berikut.

$$\omega(k) = \sum_{i=1}^k P_i \quad (2)$$

$$\mu(k) = \sum_{i=1}^k i \cdot P_i \quad (3)$$

$$\mu_t = \sum_{i=1}^L i \cdot P_i \quad (4)$$

Dimana:

$\omega(k)$  adalah Momen Kumulatif ke-0

$\mu(k)$  adalah Momen Kumulatif ke-1

$\mu_t$  adalah nilai rata-rata

Nilai ambang  $k$  dapat diperoleh dengan menggunakan persamaan

$$\sigma_B^2(k) = \frac{[\mu_t \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]} \quad (5)$$

Untuk mencari nilai *Red*, *Green*, *Blue* dari sebuah citra dengan metode *Otsu*, hal yang pertama dilakukan adalah yaitu mencari tingkat keabuan ke- $i$  yaitu  $y_1$  dan  $y_2$  yang difungsikan di persamaan

$$y_1 = 2G - R - B \quad (6)$$

$$y_2 = 2R - G - B \quad (7)$$

### Fuzzy K-Nearest Neighbor

Metode *Fuzzy K-Nearest Neighbor (Fk-Nn)* diperkenalkan oleh Keller *et al* (1985) dengan mengembangkan *K-NN* yang digabungkan dengan teori *Fuzzy* dalam menyampaikan definisi pemberian label kelas pada data uji yang diprediksi. Seperti halnya pada teori *Fuzzy*, sebuah data mempunyai nilai keanggotaan pada setiap kelas yang artinya sebuah data bisa dimiliki oleh kelas yang berbeda dengan nilai derajat keanggotaan dalam interval  $[0,1]$  [5].

Rumus yang digunakan adalah :

$$u(x, c_i) = \frac{\sum_{k=1}^K u(x_k, c_i) * d(x, y)^{\frac{-2}{(m-1)}}}{\sum_{k=1}^K d(x, y)^{\frac{-2}{(m-1)}}} \quad (8)$$

Dimana :

$u(x, c_i)$  : nilai keanggotaan data  $x$  ke kelas  $c_i$

$K$  : jumlah tetangga terdekat yang digunakan

$u(x_k, c_i)$  : nilai keanggotaan data tetangga dalam  $K$  tetangga pada kelas  $c_i$ , nilainya 1 jika data latih  $x_k$  milik kelas  $c_i$  atau 0 jika bukan milik kelas  $c_i$

$d(x, y)$  : jarak dari data  $x$  ke data  $y$  dalam  $K$  tetangga terdekat

$m$  : bobot pangkat (*weight exponent*), yang besarnya :  $m > 1$

Untuk mencari nilai *Red*, *Green*, *Blue* dari sebuah citra dengan metode *Otsu*, hal yang pertama dilakukan adalah yaitu mencari tingkat keabuan ke- $i$  yaitu  $y_1$  dan  $y_2$  yang difungsikan di Persamaan 9 dan Persamaan 10.

$$y_1 = 2G - R - B \quad (9)$$

$$y_2 = 2R - G - B \quad (10)$$

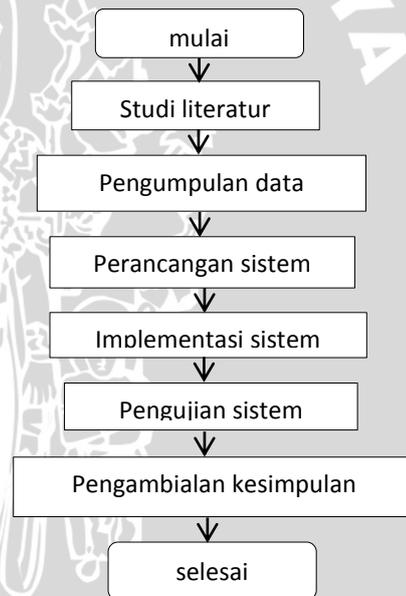
Dari nilai  $y_1$  dan  $y_2$  yang diterapkan pada metode *Otsu*, maka didapatkan dua nilai *Threshold* ( $T_1$  dan  $T_2$ ). Dengan tujuan untuk meminimalkan *noise* yang terdapat pada citra, maka proses segmentasi juga perlu dimodifikasi dengan fungsi sebagai berikut [10].

$$P_s = \begin{cases} (0,0,0) & y_1 > T_1 + 10 \text{ dan } y_2 < T_2 - 10 \\ P_r & y_1 < T_1 + 10 \text{ dan } y_2 > T_2 - 10 \end{cases} \quad (11)$$

Dimana  $P_r$  adalah nilai *Red*, *Green*, dan *Blue* (*RGB*) disetiap piksel asli pada citra, dan  $10$  adalah peubah nilai *Threshold* ( $T_1$  dan  $T_2$ ).

### 3. METODE PENELITIAN

Pada penelitian dalam bidang perangkat lunak, umumnya menggunakan tahapan-tahapan penelitian. Gambar 3 adalah tahapan-tahapan penelitian untuk membangun jenis penyakit kopi berdasarkan citra daun.



Gambar Error! No text of specified style in document.. Tahap penelitian

Gambar 3 merupakan tahapan-tahapan yang dilakukan dalam penelitian. Berikut penjelasan dari masing-masing tahapan:

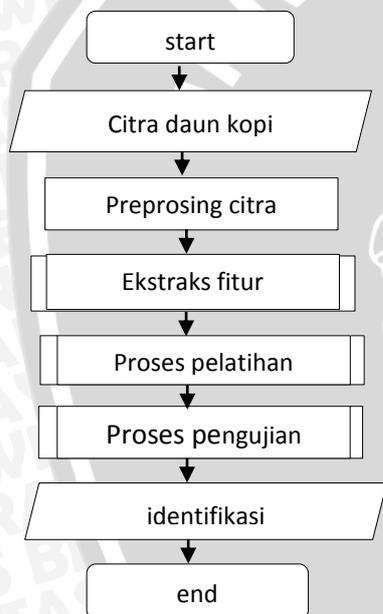
1. Studi literatur merupakan tahap metode penelitian yang bertujuan untuk mengumpulkan dan mempelajari teori-teori dasar yang selanjutnya akan diterapkan dalam penelitian.
2. Pengumpulan data yaitu proses Data yang dikumpulkan pada penelitian ini yaitu citra daun tanaman kopi yang berpenyakit.
3. Perancangan sistem Melakukan perancangan sistem yaitu menerapkan metode *Fuzzy K-*



- Nearest Neighbor* untuk klasifikasi penyakit pada citra daun kopi .
4. Implementasi sistem Implementasi sistem adalah proses pembuatan aplikasi atau perangkat lunak untuk mengklasifikasikan penyakit kopi berdasarkan citra daun.
  5. Pengujian sistem Pengujian sistem digunakan untuk melihat presentase akurasi sistem terhadap citra yan diinputkan.
  6. Pengambilan kesimpulan adalah tahap untuk menyimpulkan hasil keseluruhan penelitian dari awal sampai akhir dan berdasarkan tingkat akurasi sistem yang telah dilakukan.

Dari tahapan-tahapan yang telah dipaparkan, telah dibangun aplikasi klasifikasi untuk penyakit tanaman kopi berdasrakan citra daun. Sedangkan untuk diagram alir sistem, ditunjukkan pada gambar 4.

#### 4. PERANCANGAN



**Gambar** Error! No text of specified style in document. Diagram Alir Sistem

Berdasarkan gambar 2, proses pada sistem meliputi 4 proses yaitu:

##### 1. *Preprocessing* citra

Dimisalkan sebuah citra daun kopi berukuran 8x 8 piksel, setiap piksel citra dibuat sebuah matrik R, G, B.



**Gambar 5** Contoh citra asli

Dari sebuah gambar, akan diketahui nilai *red*, *green*, *blue* dari citra tersebut, dari nilai tersebut maka dilakukan proses untuk memperbaiki citra yaitu dengan *thresholding* dengan menggunakan metode *Otsu*. Pada gambar diketahui piksel pertama memiliki nilai *red* 132, *blue* 118, dan *green* 159. Perhitungan ini berlaku untuk keseluruhan piksel citra. Apabila sebuah citra terdiri dari 36 piksel, maka terdapat 36 nilai *red*, *green*, dan *blue* yang diketahui. Langkah pertama yaitu intuk mencari nilai *y1* dan *y2* sesuai persamaan 6.

$$\begin{aligned}
 y1 &= 2G - R - B \\
 &= 2.159 - 132 - 118 \\
 &= 68
 \end{aligned}$$

**Tabel 1** Nilai *y1*

68	68	60	54	50	33	52	42
60	81	78	47	59	38	44	68
49	65	79	0	0	32	47	37
47	63	0	0	0	0	62	34
53	54	0	0	0	0	14	53
55	60	0	0	0	0	0	55
24	13	26	0	0	0	47	48
4	9	26	51	56	52	36	33

Tabel 1 merupakan hasil dari normalisasi perhitungan *y1*, untuk *y1* yang bernilai lebih kecil daripada 0, maka dibulatkan menjadi 0, dan untuk *y1* yang bernilai lebih besar dari pada 255 maka dibulatkan menjadi 255. Normalisasi ini juga berlaku untuk perhitungan *y2*.

Untuk mencari nilai *y2* menggunakan persamaan 7.

$$\begin{aligned}
 y2 &= 2R - G - B \\
 &= 2.132 - 159 - 118 \\
 &= -13
 \end{aligned}$$

**Tabel 2** Nilai *y2*

0	0	0	9	14	0	0	0
0	0	0	23	35	17	0	0
0	0	46	100	115	59	0	0
0	21	77	137	162	135	35	4
0	21	118	129	141	128	59	5
0	0	115	151	138	136	61	0
0	0	26	111	115	78	17	0
0	0	0	0	0	0	0	0

Untuk mencari nilai *ni1* (*histogram1*) yaitu dengan menghitung frekuensi piksel yang digunakan oleh perhitungan *y1* Untuk hasil perhitungan *ni1* (*histogram1*) ditunjukkan pada tabel 3.

**Tabel 3** Nilai *ni1*

0	1	2	3	4	.....	.....	.....	.....	255
18	0	0	0	1	.....	.....	.....	.....	0

Untuk menghitung nilai *ni2* (*histogram2*) yaitu dengan menghitung jumlah frekuensi piksel yang digunakan di perhitungan *y2* untuk perhitungan *ni2* (*histogram2*) ditunjukkan pada tabel 4.

**Tabel 4 Nilai *ni2***

0	1	2	3	4	5	.....	.....	.....	255
31	0	0	0	1	1	.....	.....	.....	0

Langkah berikutnya adalah mencari nilai Total Rata-rata dengan menggunakan persamaan 4, nilai Momen Kumulatif ke-0 dengan persamaan 2, nilai Momen Kumulatif ke-1 dengan persamaan 3. Setelah semua diketahui, maka barulah nilai *threshold1* (*t1*) dan *threshold2* (*t2*) dapat dihitung.

Perhitungan pertama adalah menghitung nilai total rata-rata dan , cara perhitungan nilai rata-rata berdasarkan contoh dibawah ini.

$$\begin{aligned} \mu_t &= \sum_{i=1}^L i \cdot P_i \\ &= \sum_{i=1}^{256} 0 \cdot \left(\frac{0}{64}\right) \\ &= 0 \end{aligned}$$

Kemudian menghitung Momen Kumulatif ke-1 cara perhitungan Momen Kumulatif ke-1 berdasarkan contoh dibawah ini.

$$\begin{aligned} \mu(k) &= \sum_{i=1}^k i \cdot P_i \\ \mu(1) &= \sum_{i=1}^1 0 \cdot \left(\frac{0}{64}\right) \\ &= 0 \end{aligned}$$

Maka, saat warna ke-0 dari perhitungan *ni1* menghasilkan nilai total dan nilai kumulatif ke-1 yang sama yaitu 0. Untuk keseluruhan perhitungan nilai Total Rata-Rata dan Nilai Momen Kumulatif berdasarkan hasil perhitungan *ni1* ditunjukkan pada tabel 5.

**Tabel 5. Nilai momen kumulatif ke-1**

0	1	2	3	4	5	.....	.....	.....	255
0	0	0	0	0,1	0,1	.....	.....	.....	34,2

Dari Tabel 5 adalah diketahui bahwa Total Rata-rata adalah 34,2. Dengan perhitungan yang sama didapat Total Rata-rata berdasarkan hasil perhitungan *ni2* adalah 39,7. Langkah berikutnya adalah menghitung nilai Momen Kumulatif berdasarkan hasil perhitungan *ni1*. Sebagai contoh perhitungan ditunjukkan dibawah ini.

$$\begin{aligned} \omega(k) &= \sum_{i=1}^k P_i \\ \omega(1) &= \sum_{i=1}^1 \left(\frac{18}{64}\right) \\ &= 0,3 \end{aligned}$$

**Tabel 6 Nilai Momen Kumulatif ke-0 berdasarkan *ni1***

0	1	2	3	..	.....	.....	255
0,3	0,3	0,3	0,3	..	.....	.....	1

Tabel 6 adalah hasil perhitungan Nilai Momen Kumulatif ke-0 berdasarkan *ni1*. Dengan persamaan dan perhitungan juga dapat diketahui Nilai Momen Kumulatif ke-0 berdasarkan *ni2*. Setelah mendapatkan nilai total rata-rata, nilai Momen Kumulatif ke-0, dan nilai Momen Kumulatif ke-1, maka perhitungan varian, maksimum varian, dan nilai *threshold* dapat dilakukan. Berikut contoh perhitungan varian saat warna ke-0 dari hasil nilai total rata-rata, nilai Momen Kumulatif ke-1.

$$\begin{aligned} \sigma_B^2(k) &= \frac{[34,2 \cdot 0,3 - 0]^2}{0,3 \cdot [1 - 0,3]} \\ &= 235,84 \end{aligned}$$

Dan berikut contoh perhitungan varian saat warna ke-0 dari hasil nilai total rata-rata, nilai Momen Kumulatif ke-0.

$$\begin{aligned} \sigma_B^2(k) &= \frac{[39,7 \cdot 0,5 - 0]^2}{0,5 \cdot [1 - 0,5]} \\ &= 1477 \end{aligned}$$

Hasil keseluruhan dari perhitungan dari keduanya akan dibuat tabel yang nantinya dicari nilai varian terbesar yang akan menjadi *t1* dan *t2*. Tabel hasil varian berdasarkan nilai Momen Kumulatif ke-1 adalah pada piksel ke-26 dan nilai Momen Kumulatif ke-0 adalah pada piksel ke-61. Setelah nilai *t1* dan *t2* didapatkan, maka nilai *t1* ditambahkan dan dikurangkan dengan nilai peubah *Threshold* yang pada perhitungan ini menggunakan nilai 10 dan didapat dilai *t1* baru 36 dan *t2* baru 51. Kemudian akan dibandingkan dengan nilai *y1* dan nilai *y2*. Nilai *Red, Green, Blue* piksel pertama akan akan bernilai tetap jika *t1* lebih besar dari *y1*, dan nilai *t2* lebih kecil daripada *y2* pada piksel pertama. jika kondisi tersebut tidak terpenuhi maka nilai *Red, Green, Blue* akan diubah menjadi 0.



**Gambar 6** Contoh citra otsu

Dari citra yang telah dibuat menjadi bentuk otsu, maka citra tersebut diekstraksi nilai *red, green, dan blue* dengan membagi nilai warna dengan jumlah piksel.

$$\text{Rata-rata } R = \frac{(166+144+150+148+157+168+176+166+158+175+175+188+188+177+181+178+152)}{17}$$

$$=166$$

$$\text{Rata-rata } G = \frac{(1138+78+111+72+83+82+116+97+74+106+100+106+104+99+90+118)}{17}$$

$$=105$$

$$\text{Rata-rata } B = \frac{(117+92+87+85+105+124+119+105+104+114+135+138+113+115+130+105)}{17}$$

$$=114$$

Dengan cara yang sama, gunakan beberapa data dan ambil nilai fitur, *red*, *gree*, dan *blue* seperti pada Tabel 7.

**Tabel 7 Data Latih**

Data Lat	R	G	B	KELAS
1	179	130	60	1
2	210	175	140	1
3	201	141	105	1
4	165	119	38	1
...	.....	.....	.....	.....
...	.....	.....	.....	.....
27	161	107	90	2
28	110	69	55	2
29	129	84	74	2
30	210	166	152	2

Masuk ke kelas apakah data dengan fitur R= 166, G= 105, B=144.

Untuk langkah pertama adalah mencari jarak euclidean dengan persamaan

$$D(x,y) = \sqrt{\sum_{k=1}^n (x_k - y_k)^2}$$

$$= \sqrt{(166 - 179)^2 + (105 - 130)^2 + (114 - 60)^2}$$

$$= 60,90$$

Sehingga didapat nilai *euclidean* seperti di tabel

**Tabel 8 Jarak terhadap data latih**

Data Latih	jarak	rank min distance	kelas
1	60,90977	19	1
2	86,67179	24	1
3	51,0098	18	1
4	77,28519	20	1
...	.....	.....	.....
27	24,59675	7	2
28	88,95504	28	2
29	58,39521	26	2
30	84,26743	25	2

Kemudian hasil tersebut diurutkan berdasarkan *rank min distance* terendah. Kemudian dihitung nilai derajat keanggotaan berdasarkan *K* yang telah ditentukan, semisal nilai *k* adalah 15.

**Tabel 9 Nilai keanggotaan kelas 1**

Jarak	Kelas	<i>u</i> (kelas1)	Nilai Keanggotaan
27,38612788	2	0	0,16
34,42382896	2	0	
41,82104733	2	0	
45,7165178	2	1	
.....	.....	.....	
.....	.....	.....	
71,79136438	1	1	
74,47147105	1	1	
75,63729239	2	0	

Pada tabel 8 diatas nilai *u* adalah nilai keanggotaan data kedalam kelas, nilainya 1 jika jarak milik kelas 1 dan 0 jika jarak milik kelas 2. Selanjutnya menghitung nilai keanggotaan data terhadap kelas 1 menggunakan persamaan 8.

$$u(x, c_1) = \frac{0 \cdot 27,38612^{\frac{-2}{(2-1)}} + 0 \cdot 34,423827^{\frac{-2}{(2-1)}} + 0 \cdot 41,8210^{\frac{-2}{(2-1)}} + \dots + 0 \cdot 45,7165178^{\frac{-2}{(2-1)}}}{27,38612^{\frac{-2}{(2-1)}} + 34,423827^{\frac{-2}{(2-1)}} + 41,8210^{\frac{-2}{(2-1)}} + \dots + 75,63729^{\frac{-2}{(2-1)}}}$$

$$= 0,16$$

**Tabel 9 Nilai keanggotaan kelas 1**

Jarak	Kelas	<i>u</i> (kelas1)	Nilai Keanggotaan
27,38612788	2	0	0,83
34,42382896	2	0	
41,82104733	2	0	
45,7165178	2	1	
49,49747468	2	0	
.....	.....	.....	
.....	.....	.....	
74,47147105	1	1	
75,63729239	2	0	



Pada tabel 9 diatas nilai  $u$  adalah nilai keanggotaan data kedalam kelas, nilainya 1 jika jarak milik kelas 2 dan 0 jika jarak milik kelas 1. Selanjutnya menghitung nilai keanggotaan data terhadap kelas 2

$$u(x, c_2) = \frac{11 \cdot 27,38612 \frac{-2}{(2-1)} + 1 \cdot 34,423827 \frac{-2}{(2-1)} + 0 \cdot 41,8210 \frac{-2}{(2-1)} + \dots + 1 \cdot 75,63729 \frac{-2}{(2-1)}}{27,38612 \frac{-2}{(2-1)} + 34,423827 \frac{-2}{(2-1)} + 41,8210 \frac{-2}{(2-1)} + \dots + 75,63729 \frac{-2}{(2-1)}} = 0,83$$

Untuk data uji diperoleh nilai keanggotaan pada kelas 1 sebesar 0.18 dan nilai keanggotaan pada kelas 2 sebesar 0.83. Nilai keanggotaan terbesar dipilih sebagai kelas target yaitu kelas 2 atau penyakit karat daun, sehingga data uji masuk kedalam kelas target.

## 5. IMPLEMENTASI

### 5.1 lingkungan perangkat keras

Untuk spesifikasi perangkat keras yang digunakan adalah :

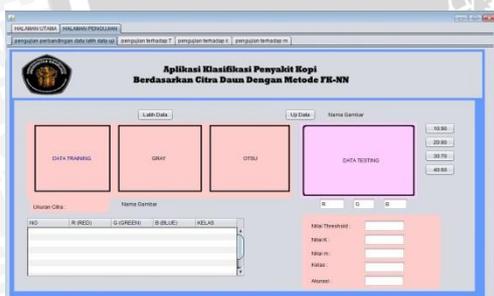
1. Processor Intel (R) Pentium (R) CPU B960 @ 2.20 Ghz
2. RAM 2.00 GB
3. Hardisk 320 GB

### 5.2 lingkungan perangkat lunak

Untuk spesifikasi perangkat lunak yang digunakan adalah :

1. Sistem Operasi Windows 7 Enterprise (32-bit).
2. Bahasa pemrograman Java JDK 6
3. Tools NetBeans IDE 7.4

### 5.3 Implementasi Interface



Gambar 7 Implementasi Interface

## 6. PENGUJIAN

### 6.1 Skenario pengujian

Pengujian sistem meliputi beberapa skenario yaitu :

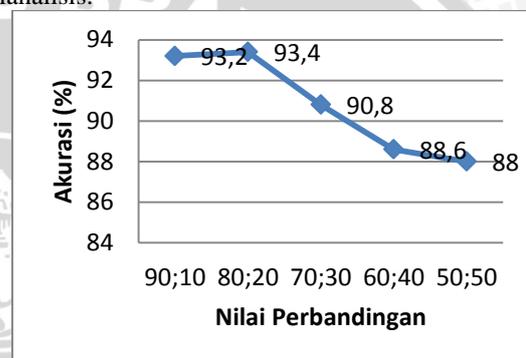
1. pengujian terhadap perbandingan data latih dan data uji  
Pengujian ini bertujuan untuk mengetahui perbandingan data latih dan data uji yang optimal dari data.
2. pengujian terhadap nilai peubah threshold  
Peubah *threshold* ( $T$ ) adalah variabel yang digunakan pada *Preprocessing* citra dengan metode *Otsu*. Pengujian ini bertujuan untuk

mencari nilai peubah *threshold* ( $T$ ) yang paling optimal untuk proses *Preprocessing* sehingga menghasilkan akurasi terbaik.

3. pengujian terhadap nilai  $k$   
Pengujian ini bertujuan untuk mencari nilai  $k$  yang paling optimal untuk proses klasifikasi citra sehingga klasifikasi yang dihasilkan memiliki akurasi yang tinggi.
4. pengujian terhadap nilai  $m$   
Pengujian ini bertujuan untuk mencari nilai  $k$  yang paling optimal untuk proses klasifikasi citra sehingga klasifikasi yang dihasilkan memiliki akurasi yang tinggi.

### 6.2 Hasil pengujian

Hasil pengujian dilakukan berdasarkan pada skenario pengujian yang telah dibuat kemudian dianalisis.

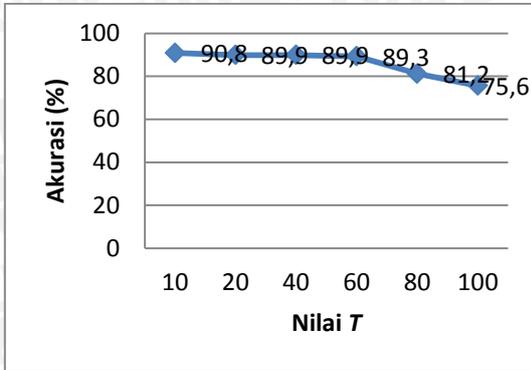


Gambar 8 pengujian terhadap perbandingan Data latih dan data uji

Pada gambar 8 terlihat Grafik garis semakin menurun yang menghubungkan semua skenario perbandingan data latih dan data uji. Perbandingan pertama yaitu 90:10 yaitu dengan komposisi 54 data latih dan 6 data uji menghasilkan rata-rata akurasi sebesar 93,2%, perbandingan kedua dengan komposisi 48 data latih dan 12 data uji menghasilkan akurasi rata-rata sebesar 93,4%, perbandingan ketiga dengan komposisi 42 data latih dan 18 data uji menghasilkan akurasi rata-rata sebesar 90,8%, perbandingan keempat dengan komposisi 36 data latih dan 24 data uji menghasilkan akurasi rata-rata sebesar 88,6% dan perbandingan skenario pengujian terakhir dengan jumlah komposisi data latih dan data uji masing-masing 50 menghasilkan rata-rata akurasi sebesar 88%. Berdasarkan gambar didapatkan kesimpulan bahwa akurasi tertinggi didapatkan saat pengujian pertama dan kedua yaitu dengan berbandingan 90:10 dan 80:10.

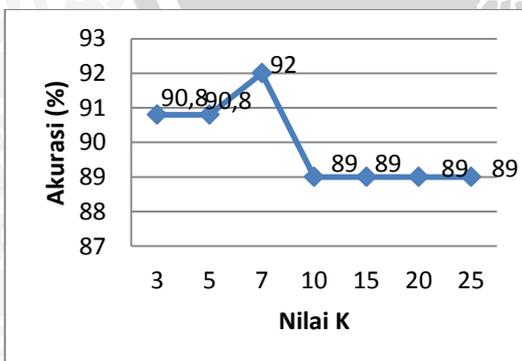
Berdasarkan Gambar 8, perbedaan akurasi yang dihasilkan dipengaruhi oleh perbandingan data latih dan data uji, semakin banyak data latih dan semakin sedikit data uji, maka akurasi yang dihasilkan semakin tinggi. Sebaliknya, semakin sedikit data latih dan semakin banyak data uji maka akurasi yang dihasilkan semakin rendah. Hal ini

disebabkan karena proses pelatihan menggunakan banyak data sehingga variasi hasil latih yang didapatkan lebih beragam sehingga memudahkan pada proses pengujian yang menggunakan data lebih sedikit dalam memutuskan klasifikasi data citra sehingga sehingga akurasi yang dihasilkan tinggi.



Gambar 9 pengujian terhadap nilai threshold (T)

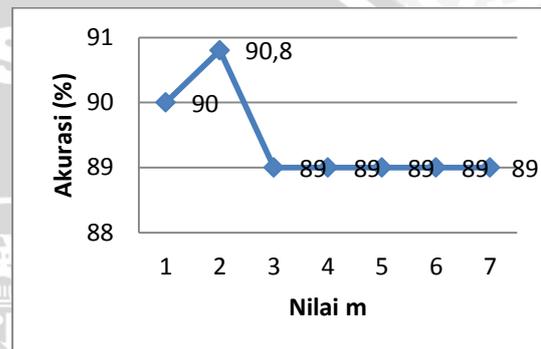
Gambar 9 menunjukkan garis semakin menurun pada tiap pengujian. Skenario pengujian pertama mendapatkan 90,8% sebagai rata-rata akurasi, skenario pengujian kedua mendapatkan rata-rata akurasi sebesar 89,9%, skenario pengujian ketiga memiliki rata-rata akurasi yang sama dengan skenario kedua yaitu 89,9%, kemudian skenario berikutnya dengan nilai T 60 mendapatkan rata-rata akurasi sebesar 89,3%, untuk skenario pengujian dengan nilai T 80 mempunyai rata-rata akurasi sebesar 81,2% dan skenario pengujian terakhir dengan nilai T 100 mendapatkan rata-rata akurasi terendah yaitu 75,6%. Dari Gambar 6.2 dapat disimpulkan bahwa semakin besar T, maka akurasi yang dihasilkan semakin rendah. Hal ini disebabkan semakin besar T, maka nilai perbandingan *histogram* akan semakin besar yang menyebabkan terbentuknya banyak *noise*. Sehingga banyak bagian citra tidak terdeteksi atau terdeteksi salah.



Gambar 10 pengujian terhadap nilai k

Gambar 10 menunjukkan grafik lurus pada nilai k 3 dan 5 dengan akurasi 90,8%, kemudian akurasi tertinggi pada nilai k 7 dengan akurasi 92% dan grafik menurun pada k selanjutnya yaitu 10, 15,

20, dan 25 dengan nilai akurasi yang sama yaitu 89%. Hasil pengujian yang didapatkan pada pengujian terhadap nilai k terhadap tingkat akurasi yaitu semakin bertambahnya nilai k, maka akurasi yang didapatkan semakin menurun dan stabil pada nilai k tertentu. Hal ini disebabkan rentang kelas pada k yang semakin banyak memberikan pengaruh besar pada penentuan prediksi klasifikasi dan nilai keanggotaan yang digunakan untuk penentuan kelas target pada penelitian kemudian akurasi yang dihasilkan oleh nilai k yang berdekatan cenderung stagnan, hal ini disebabkan persebaran nilai fitur pada data, untuk data yang berdekatan lebih sering memiliki kelas yang sama sehingga mempengaruhi nilai derajat keanggotaan.



Gambar 11 pengujian terhadap nilai m

Pada gambar 11 diketahui bahwa skenario pengujian kesatu dan kedua memberikan nilai rata-rata akurasi yang tinggi yaitu berturut-turut 90% dan 90,8%. Dengan nilai m yaitu 1 dan 2. Kemudian pada skenario pengujian ketiga sampai skenario ketujuh, grafik bergerak lurus dengan akurasi yang sama yaitu 89%. Pada pengujian terhadap nilai m, nilai m pada proses pengklasifikasian menggunakan Metode *Fuzzy K-Nearest Neighbor* tidak berpengaruh langsung terhadap nilai akurasi, namun berpengaruh pada nilai derajat keanggotaan tiap-tiap data uji terhadap masing-masing kelas. Variabel m merupakan bobot pangkat yang digunakan untuk mengetahui seberapa besar jarak antar tetangga ketika menghitung pengaruh tetangga pada nilai keanggotaan. Nilai m yang semakin besar akan membuat nilai keanggotaan semakin rendah, sehingga berpengaruh pada penentuan hasil kelas klasifikasi penyakit kopi berdasarkan citra daun dimana hasil klasifikasi tersebut yang mempengaruhi hasil klasifikasi.

## 7. KESIMPULAN DAN SARAN

### 7.1 Kesimpulan

Dari hasil uji dan analisa yang telah dilakukan terhadap klasifikasi penyakit tanaman kopi menggunakan metode *Fuzzy K-Nearest*

*Neighbor*, dapat diambil kesimpulan diantaranya ialah:

1. Algoritma *Fuzzy K-Nearest Neighbor (Fk-Nn)* dapat diterapkan pada data yang berupa citra daun kopi. Terdapat dua proses utama yang dilakukan dalam penentuan klasifikasi tanaman kopi berdasarkan daun, yaitu *Preprocessing* citra dan proses klasifikasi. Proses *Preprocessing* bertujuan untuk memperbaiki kualitas citra sehingga akurasi yang dihasilkan pada proses klasifikasi bisa optimal. Selanjutnya proses klasifikasi melibatkan beberapa tahapan yaitu menentukan  $k$  sebagai tetangga terdekat yang kemudian dicari jarak antar data uji dan data latih, kemudian dicari nilai keanggotaan tiap-tiap kelas sehingga didapat kelas terdekat yang merupakan kelas target dari data uji yang baru.
2. Tingkat akurasi pada *FUZZY K-NEAREST NEIGHBOR (FK-NN)* dipengaruhi oleh beberapa faktor antar lain perbandingan antara jumlah data latih dan data uji, nilai *Threshold*, nilai  $k$ , dan nilai  $m$ . Pada perbandingan data latih dan data uji, akurasi yang dihasilkan mencapai 100% dengan jumlah komposisi data latih lebih banyak dari pada data uji. Untuk pengujian terhadap nilai *Threshold*, dari beberapa percobaan yang dilakukan didapat akurasi tertinggi pada nilai *Threshold* 10 dengan akurasi mencapai 90,8%. Pada pengujian lain menggunakan nilai  $k$ ,  $k$  terbaik adalah pada nilai 7 dengan akurasi yang dihasilkan adalah 82%. Dan pengujian terakhir adalah nilai pangkat  $m$  yang menghasilkan akurasi tertinggi pada  $m = 2$  yaitu mencapai 88,89% .

## 7.2 Saran

1. Salah satu penerapan teknologi informasi dan komunikasi di bidang pertanian yaitu dapat digunakan sebagai alat bantu klasifikasi penyakit tanaman kopi dan tidak hanya di lingkup teknologi tetapi juga terhadap lingkungan.
2. Untuk penelitian selanjutnya perlu penambahan data dengan kondisi daun normal dan variasi data yang lebih banyak dengan menggunakan nilai fitur yang kompleks untuk objek terkait sehingga menghasilkan akurasi yang lebih tinggi.

3. Menggunakan metode segmentasi lain yang lebih baik untuk memperbaiki kualitas citra .

## DAFTAR PUSTAKA

- [1] Achmad Z.M., Ridok Achmad. 2013. *Penerapan Metode Fuzzy K-NN Pada Klasifikasi Jenis Mangga Berdasarkan Tekstur Daun*. Program Teknologi Informasi Dan Komunikasi. Universitas Brawijaya. Malang.
- [2] Beyan Cigdem, Ogul Hasan .A *Fuzzy K-NN Approach For Cancer Diagnosis With Microarray Gene Expression Data*. Dataset
- [3] Katiman. 2015. Wawancara " *Penyakit Tanaman Kopi*" di perkebunan Bumisari. Banyuwangi
- [4] Otsu Nubuyoki. 1979. *A Threshold Selection Method From Gray-Level Histograms*. IEEE transactions ON SYSTEMS, MAN, AND cybernencs, VOL SMC-9, NO. 1.
- [5] Prasetyo Eko, 2012. *FUZZY K-NEAREST NEIGHBOR In Every Class Untuk Klasifikasi*.
- [6] Priambodo Apiladosi, Dewi Candra, Triwiratno Anang. 2014. *Implementasi metode k-nearest neighbour untuk identifikasi penyakit tanaman jeruk keprok berdasarkan citra daun*. Universitas Brawijaya. Malang.
- [7] Putra darama .2010. *pengolahan citra digital*.
- [8] Rahardjo, Pudji. 2012. *Panduan Budidaya dan Pengolahan Kopi Arabika dan Robusta*. Penebar Swadaya. Jakarta
- [9] sutoyo,T,dkk. 2009. *Teori Pengolahan Citra Digital*. Yogyakarta
- [10] Yao Qing. Guan Zexin. Zhou Yingfeng. 2009. *Application Of support Vector Machine for Rice Diseases using shape and color*. China.