

## PENERAPAN METODE K- NEAREST NEIGHBOR KLASIFIKASI JENIS BUAH SEMANGKA DENGAN IMAGE PROCESSING

<sup>1</sup>Idi Jangcik, <sup>2</sup>Sasmita, <sup>3</sup>Muhammad Baqi Billah, <sup>4</sup>Yuhelmi

<sup>1,2,3</sup>Program Studi Teknik Informatika, Institut Teknologi Kota PagarAlam, Indonesia

<sup>4</sup>Program Studi Sistem Informasi Fakultas Ilmu Komputer Universitas Lancang Kuning, Pekanbaru

Email: [idijangcik@itpa.ac.id](mailto:idijangcik@itpa.ac.id)<sup>2</sup>, [sasmita@gmail.com](mailto:sasmita@gmail.com)<sup>2</sup>, [baqibillah340@gmail.com](mailto:baqibillah340@gmail.com)<sup>3</sup>, [yuhelmi@unilak.ac.id](mailto:yuhelmi@unilak.ac.id)<sup>4</sup>

### Abstrak

Tujuan dari studi ini ialah untuk menciptakan Sistem Klasifikasi Jenis Buah Semangka melalui Pengolahan Citra dengan menggunakan Metode K-Nearest Neighbor di Dinas Pertanian Kota Pagar Alam. Penelitian ini dilakukan karena proses pengklasifikasian jenis-jenis semangka masih dilakukan secara tradisional dan belum diotomatiskan. Saat ini, klasifikasi buah semangka masih bergantung pada pengalaman manusia, penilaian warna, dan bentuk buah semangka. Namun, cara ini memakan waktu yang cukup lama dan seringkali menghasilkan kesalahan. Oleh karena itu, penelitian ini bertujuan untuk membantu dalam pengklasifikasian buah semangka dengan menggunakan metode yang lebih efisien dan akurat melalui teknik pengolahan citra. Dengan menggunakan metode K-Nearest Neighbor dengan cepat. Sistem yang dibangun menggunakan Software MATLAB, dalam metode pengembangan sistem dalam penelitian ini adalah metode SDLC (Software Development Life Cycle), dimana tahapan meliputi analisis, desain, pengkodean dan pengujian, untuk metode pengujian menggunakan Euclidean Distance dilakukan pengujian pada seluruh data testing yang telah disiapkan hasil dari penelitian ini yakni sistem klasifikasi jenis buah semangka dengan metode Euclidean Distance dengan Image Processing dimana pada 80 data training, menghasilkan 69 data berhasil dikenali dan 11 data tidak berhasil dikenali, sehingga mendapat persentase sebesar 86,25%. Kemudian setelah dilakukan holdout validation dengan 20 data testing, menghasilkan 16 data berhasil dikenali dan 4 data tidak berhasil dikenali, sehingga mendapat persentase sebesar 80%. Akhirnya sistem yang menerapkan Euclidean Distance terhadap klasifikasi jenis buah semangka dengan image processing mendapat akurasi yang tinggi

**Kata Kunci:** Image Processing, K-Nearest Neighbor, Euclidean Distance, SDLC

### Abstract

The purpose of this study is to create a Watermelon Type Classification System through Image Processing using the K-Nearest Neighbor Method at the Agriculture Service Office of Pagar Alam City. This research was conducted because the process of classifying the types of watermelon is still done traditionally and has not been automated. Currently, the classification of watermelons still depends on human experience, judging the color and shape of the watermelon. However, this method takes a long time and often results in errors. Therefore, this study aims to assist in classifying watermelons using a more efficient and accurate method through image processing techniques. By using the method K-Nearest Neighbor quickly. The system built using Software MATLAB, in the system development method in this study is the method SDLC (Software Development Life Cycle), where the stages include analysis, design, coding and testing, for the testing method using Euclidean Distance tested on all data testing. The results of this study have been prepared, namely the classification system for watermelon types using the method Euclidean Distance with Image Processing where at 80 data training, resulted in 69 data

*being recognized successfully and 11 data not being recognized, so that a percentage of 86.25% was obtained. Then once done holdout validation with 20 data testing, resulted in 16 data that were successfully recognized and 4 data that were not successfully recognized, so that a percentage of 80% was obtained. Finally a system that implements Euclidean Distance on the classification of types of watermelon with image processing get high accuracy*

**Keywords:** *Image Processing, K-Nearest Neighbor, Euclidean Distance, SDLC*

## 1. PENDAHULUAN

Perkembangan Teknologi yang ditandai dengan Revolusi Industri 1.0 sampai dengan Revolusi Industri 4.0 saat ini merupakan salah satu faktor yang mempengaruhi perilaku kehidupan masyarakat. Pada penelitian yang dilakukan oleh Prasetyo dan Trisyanti bahwa dengan melakukan kajian sosial yang terjadi dalam Revolusi Industri 4.0, menghasilkan suatu anggapan bahwa Revolusi Industri tidak hanya mendistrupsi bidang teknologi saja, namun merambah pada bidang sosial, ekonomi dan hukum dan untuk menanggulangi distrupsi dampak Revolusi Industri perlu adanya kajian humaniora untuk mengawal perkembangan teknologi agar tidak keluar dari norma-norma kemanusiaan yang berlaku. [1]. Salah satu perkembangan teknologi yakni pada bidang Artificial Intelligence.

*Artificial Intelligence* adalah kecerdasan yang ditambahkan kepada suatu system yang bisa diatur dalam konteks ilmiah atau bisa disebut juga Artificial Intelligence Atau biasa disebut kecerdasan buatan. [2] Seiring berkembangnya sistem kecerdasan buatan dalam berbagai bidang yakni salah satunya dalam bidang pertanian yang dapat mempermudah pihak instansi dalam melakukan pendataan, sehingga data yang diolah dapat menghasilkan sebuah informasi yang akurat. Salah satunya dalam pengolahan hasil pertanian dan perkebunan khususnya pada produksi buah semangka. Buah semangka merupakan tanaman semusim yang tumbuh merambat yang menjadi salah satu buah yang di gemari masyarakat. Dengan adanya perkembangan sistem dengan kecerdasan buatan, seperti sistem pengklasifikasian jenis buah semangka dengan memanfaatkan pengolahan citra digital (*image processing*)

Berdasarkan studi pendahuluan melalui observasi, wawancara dan dokumentasi pada Dinas Pertanian Kota Pagar Alam, saat ini pengklasifikasian buah semangka dilakukan secara manual yaitu dilakukan berdasarkan pengalaman di pilih secara langsung tanpa menggunakan sistem/alat untuk menentukan jenis buah semangka tersebut, yakni dengan cara ditentukan berdasarkan warna, bentuk, dan besar atau kecilnya buah semangka untuk di klasifikasikan

Dari penelitian yang dilakukan oleh [3] dengan judul “Klasifikasi Jeruk Nipis Terhadap Tingkat Kematangan Buah Berdasarkan Fitur Warna Menggunakan *K-Nearest Neighbor*”, dari penelitian yang dilakukan oleh [4] dengan judul “Klasifikasi Kematangan Buah Pisang Ambon Menggunakan Metode KNN dan PCA Berdasarkan Citra RGB dan HSV “ pada penelitian ini mendapatkan hasil berupa sistem klasifikasi tingkat kematangan buah pisang ambon atau *Cavendish* yang dikembangkan menggunakan *software* Matlab R2019b, Dan penelitian yang dilakukan oleh [5] dengan judul “Usaha tani Buah Semangka Di Desa Manggis, Mojosongo, Boyolali

Dari Penelitian terdahulu diatas maka peneliti membuat sistem klasifikasi jenis buah semangka menggunakan *K-Nearest Neighbor* dengan *Image Processing*. Oleh sebab itu peneliti mengangkat judul “Penerapan Metode *K- Nearest Neighbor* Terhadap Klasifikasi Jenis Buah Semangka Dengan *Image Processing*”

## 2. METODOLOGI PENELITIAN

### 2.1 Metode Pengumpulan Data

Dalam mendapatkan data yang diperlukan untuk penulisan laporan ini ada beberapa metode yang diperlukan antara lain :

#### 1. Observasi

Pengumpulan data dengan mengadakan pengamatan-pengamatan secara langsung atau seksama pada pelaksanaan objek yang diteliti.

**Tabel 1.** Rangkuman: Kerangka Lembar Observasi

No	Aspek	Indikator
1	Teknik Pelaksanaan	1. Metode untuk mengidentifikasi cara pengelompokkan atau pengklasifikasi jenis-jenis Buah Semangka di Dinas Pertanian Kota Pagar Alam.

#### 1. Wawancara

Wawancara merupakan strategi pengumpulan data yang melibatkan penyampaian sejumlah pertanyaan secara langsung kepada sumber informasi yang relevan dan objek penelitian, dalam hal ini diarahkan kepada Dinas Pertanian Kota Pagar Alam.

**Tabel 2.** Kisi-kisi Pedoman Wawancara

No	Aspek	Pertanyaan
1	Proses Berjalan	1. Bagaimana proses mengelompokkan / klasifikasi jenis Buah Semangka?
2	Kendala	1. Sudahkah pelaksanaan proses pengumpulan data dan klasifikasi varietas Buah Semangka berjalan dengan efisiensi dan efektivitas yang diharapkan? 2. Pernahkah terdapat situasi di mana penentuan kategori jenis Buah Semangka mengalami kesalahan?

#### 1. Studi Pustaka

Studi Pustaka merupakan upaya untuk memahami teori yang terdapat dalam literatur, baik dalam bentuk buku maupun jurnal cetak dan digital, yang memiliki kaitan dengan subjek penelitian.

#### 2. Dokumentasi

Teknik dokumentasi adalah pendekatan yang digunakan untuk mengumpulkan informasi terkait dengan studi. Proses dokumentasi yang dilakukan oleh peneliti melibatkan pengambilan foto-foto yang relevan untuk keperluan penyusunan penelitian ini. [6]

## 2.2 Pengumpulan Data

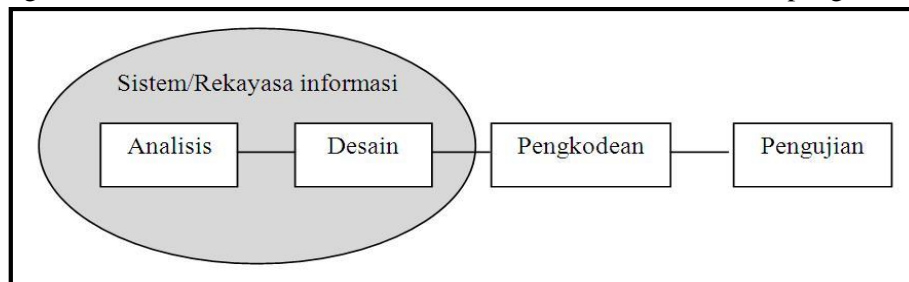
Sistem yang diajukan dalam penelitian ini merupakan bentuk dari sistem informasi yang bertujuan untuk mengklasifikasikan jenis buah semangka di Kota Pagar Alam. Sistem pengklasifikasian ini diharapkan mampu memberikan dukungan kepada Dinas Pertanian atau pengguna lainnya dalam mengidentifikasi informasi terkait jenis-jenis buah semangka..



**Gambar 1 :** Use Case Sistem yang Diusulkan

## 1.3 Metode Pengembangan Sistem

Model SDLC (Software Development Life Cycle) atau (System Development Life Cycle) air terjun (Waterfall) sering juga di sebut model sekuensial linier (sequential linear) atau alur hidup klasik (classic life cycle). Model air terjun menyediakan pendekatan alur hidup perangkat lunak secara sekuensial atau terurut dimulai analisis, desain, pengkodean, pengujian.



**Gambar 2.** Metode Waterfal

## 1.4 Metode K-Nearest Neighbor

K-Nearest Neighbor merupakan salah satu algoritma yang digunakan dalam pengklasifikasian. Prinsip kerja K-Nearest Neighbor (KNN) adalah mencari jarak terdekat antara data yang akan dievaluasi dengan K-Nearest Neighbor terdekatnya dalam data pelatihan. Adapun langkah-langkah penggunaan metode K-NN ini dijelaskan sebagai berikut:

1. Tentukan parameter K
2. Hitung jarak antara data yang akan dievaluasi dengan semua pelatihan.
3. Urutkan jarak yang terbentuk (urutan dari nilai yang terkecil hingga nilai yang terbesar).
4. Tentukan jarak terdekat sampai urutan K
5. Pasangkan kelas yang bersesuaian
6. Cari jumlah kelas dari tetangga yang terdekat dan tetapkan kelas tersebut sebagai kelas data yang akan dievaluasi.

Ada beberapa Metode Pendekatan K-NN. Dalam penelitian ini Metode Pendekatan yang digunakan adalah Metode Euclidean Distance dan Manhattan Distance. Berikut adalah rumus untuk mencari nilai akar dari kuadrat dua vector. [7]

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2}$$

Keterangan:

x1 = Sampel data

x2 = Data uji atau data testing

i = Variabel data

d = Jarak

p = Jumlah data training

$$d_{ij} = \sum W_k |X_{ik} - C_{jk}|$$

Dimana dij dengan rumus adalah jarak antara data testing dan data training dengan semua parameternya. W merepresentasikan dari jumlah bobot. X adalah data testing. C adalah training dalam case base.

### 1.5 Pengujian *Confusion Matrix*s

Confusion matrix adalah pengukuran performa untuk masalah klasifikasi machine learning dimana keluaran dapat berupa dua kelas atau lebih. *Confusion matrix* berupa tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah.

**Tabel 3.:** *Confusion Matrix*

<i>Confusion Matrix</i>		Kelas Hasil Prediksi	
		<i>Positive</i>	<i>Negative</i>
Kelas	True	TP = 69	FP = 11
Sebenarnya	<i>False</i>	FN = 0	TN = 0

Keterangan :

TP (*True Positive*) = Jumlah dokumen dari kelas 1 yang benar diklasifikasikan sebagai kelas 1

TN (*True Negative*) = Jumlah dokumen dari kelas 0 yang benar diklasifikasikan sebagai kelas 0

FP (*False Positive*) = Jumlah dokumen dari kelas 0 yang salah diklasifikasikan sebagai kelas 1

FN (*False Negative*) = Jumlah dokumen dari kelas 1 yang salah diklasifikasikan sebagai kelas 0

Rumus confusion matrix untuk menghitung accuracy, precision, recall dan F-1 Score. Nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Accuracy menggambarkan seberapa akurat model dalam mengklasifikasikan dengan benar.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \%$$

Precision menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang telah diberikan oleh model.

$$\text{Precision} = \frac{TP}{TP+FP} \times 100 \%$$

Recall atau sensitivity menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi.

$$Recall = \frac{TP}{TP + FN} \times 100 \%$$

F-1 Score menggambarkan perbandingan rata-rata precision dan recall yang dibobotkan. Accuracy tepat kita gunakan sebagai acuan performansi algoritma jika dataset memiliki jumlah data

False Negative dan False Positive yang sangat mendekati (symmetric). Namun, jika jumlahnya tidak mendekati, maka sebaiknya menggunakan F1-Score Sebagai acuan

$$F-1 \text{ Score} = \frac{2 \times Recall \times Precision}{Recall + Precision}$$

### 3. HASIL DAN PEMBAHASAN

#### 3.1 Hasil

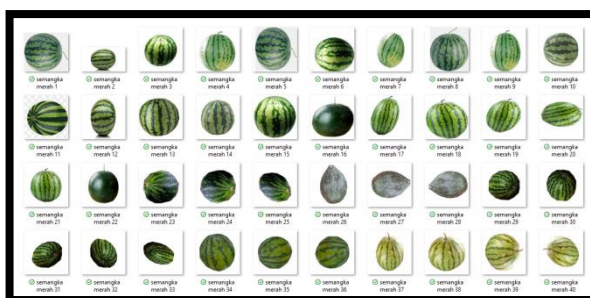
Dari hasil penelitian, diperoleh hasil yang menjelaskan bagaimana implementasi sistem menggunakan metode K-Nearest Neighbor untuk mengklasifikasikan jenis-jenis Buah Semangka melalui pengolahan citra. Pengujian sistem dilakukan dengan menerapkan metode Euclidean Distance. Proses pengembangan sistem ini mengadopsi metode pengembangan Waterfall. Sistem tersebut memberikan panduan tentang bagaimana mengantisipasi jenis Buah Semangka melalui solusi yang diterapkan. Untuk merumuskan sistem prediksi yang tepat, metode K-Nearest Neighbor digunakan dengan cara menghimpun informasi awal atau data awal, kemudian diproses untuk menghasilkan informasi prediksi berdasarkan tanggapan yang diberikan oleh pengguna.

##### 3.1.1. Implementasi Sistem

Pada sub bab ini akan memperlihatkan tampilan yang digunakan pada program dan cara menjalankannya, kemudian pembahasan mengenai fungsi-fungsi yang digunakan dan metode yang diterapkan.

##### 1. Data Buah Semangka

Data buah semangka yang digunakan penulis pada penelitian ini yakni terdiri dari 2 jenis buah semangka, jenis-jenis buah semangka tersebut ialah buah semangka kuning dan semangka merah.



Gambar 4. Data Semangka Merah



Gambar 5. Data Semangka Kuning



## 2. Resizing Image Buah Semangka

Tahap selanjutnya yakni resizing semua ukuran gambar buah semangka. Resizing dilakukan untuk menyamakan semua ukuran gambar buah semangka, sehingga nantinya data buah semangka mempunyai ukuran yang sama secara merata. Ukuran image buah semangka yang digunakan ialah 130x130 pixel, 130 sebagai panjang dan lebar dari gambar yang di resizing. Tahap resizing dilakukan pada software Adobe Photoshop Cs6 dengan cara menginput satu per satu hasil foto buah semangka dari kamera smartphone kedalam software Adobe Photoshop Cs6 dan diubah ukuran pixel gambar sesuai dengan yang telah ditentukan

## 3. Grayscale Image Buah Semangka

Tahap selanjutnya yakni grayscale image buah semangka untuk melihat komposisi bentuk dari buah semangka yang menyerupai keasliannya. Grayscale adalah warna-warna pixel yang berada dalam rentang gradasi warna hitam dan putih. Adapun coding untuk melakukan grayscale pada Matlab sebagai berikut.

```
% mengkonversi citra rgb menjadi grayscale
Img_gray = rgb2gray(Img);
```

Gambar 6. Coding

## 4. Ekstraksi Ciri

Tahap ekstraksi ciri ini, penulisan menggunakan 2 neuron masukan yaitu *Corelation & Energy* untuk ekstraksi ciri nilai 2 neuron tersebut menggunakan image hasil yang telah di konversi dari Asli ke *grayscale*. Sehingga mendapat hasil berikut.

Variables - data_latih		
data_latih 80x2 double		
	1	2
1	0.9810	0.2372
2	0.9572	0.1896
3	0.9549	0.2642
4	0.9713	0.2842
5	0.9486	0.2811
6	0.9688	0.3147
7	0.9671	0.1946
8	0.8917	0.2623
9	0.9621	0.2345
10	0.9643	0.2884
11	0.9660	0.2854
12	0.9615	0.3149
13	0.9709	0.2404
14	0.8683	0.3516
15	0.8801	0.2906
16	0.9535	0.3913
17	0.9534	0.3914

(a)

data_uji 20x2 double		
	1	2
1	0.9810	0.2372
2	0.9572	0.1896
3	0.9615	0.3148
4	0.9853	0.2255
5	0.9864	0.2811
6	0.9688	0.3147
7	0.9461	0.2874
8	0.9878	0.3471
9	0.9466	0.1626
10	0.9864	0.2811
11	0.9531	0.1369
12	0.9590	0.2757
13	0.9412	0.2518
14	0.9504	0.1266
15	0.9446	0.1641
16	0.9554	0.2766
17	0.9523	0.2397

(b)

Gambar7. (a) Ekstraksi Data Latih

(b) Ekstraksi Data Uji

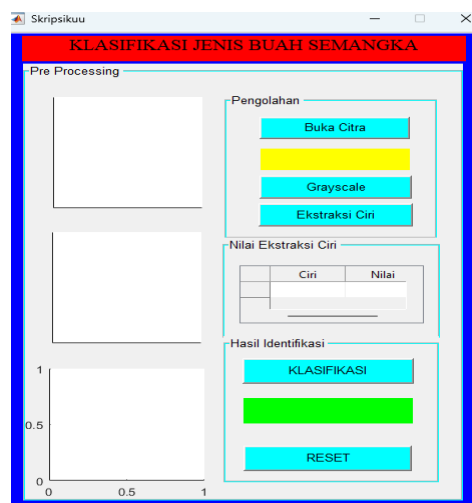
Keterangan :

1. *Energy* merupakan besaran nilai *pixel* pada gambar.
2. *Correlation* merupakan nilai yang menunjukkan kekuatan dan arah hubungan linier.

## 3.2 Pembahasan

### 3.2.1 Implementasi User Interface

Pada bagian ini akan membahas mengenai tampilan antarmuka dari program klasifikasi jenis buah semangka menggunakan jaringan syaraf tiruan. Berikut ini merupakan tampilan keseluruhannya.



**Gambar 8.** Tampilan *User Interface* Keseluruhan

Adapun keterangan component library mengenai tampilan antarmuka dari program klasifikasi jenis buah semangka menggunakan jaringan syaraf tiruan. Berikut ini keterangannya

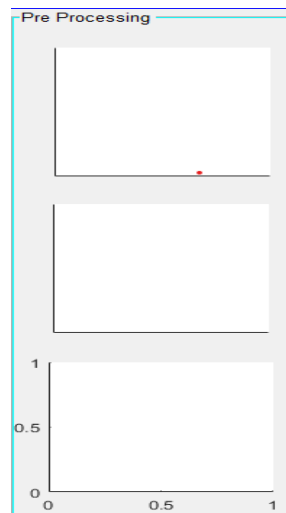
**Tabel 4.** Keterangan Component Library

No	Componet Library	Name	Fungsi
1	Label	Klasifikasi Jenis Buah Semangka Menggunakan Metode <i>K-Nearest Nearest Neighbor</i>	Judul
2	Axes (2D)		Menampilkan gambar asli dari input
3	Axes (2D)		Menampilkan gambar pengolahan
4	Axes (2D)		Menampilkan grafik batang
5	Button	Buka Citra	Tombol untuk memasukan citra gambar
6	Edit Field (Text)		Nama citra gambar inputan
7	Button	Grayscale	Proses grayscale
8	Button	Ekstraksi Ciri	Proses ekstraksi ciri
9	Table		Menampilkan hasil dari ekstraksi ciri
10	Button	Klasifikasi	Tombol untuk mengklasifikasikan
11	Label	Hasil Identifikasi	Title
12	Edit Field (Text)		Menampilakn hasil identifikasi pengenalan citra gambar
13	Button	Reset	Tombol untuk mereset data yang telah di uji



1. Bagian *Pre-Processing*

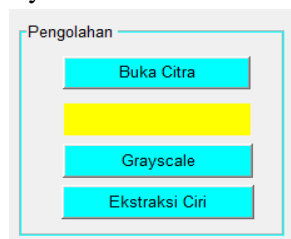
Bagian *pre-processing* ini, akan ada tiga Axes(2D) yang akan menampilkan citra asli dan citra yang telah di olah untuk hasil ekstraksi ciri. Berikut ini merupakan tampilannya.



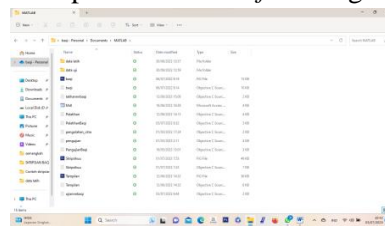
**Gambar 9.** Tampilan *User Interface* Bagian *Pre-Processing*

2. Bagian Pengolahan

Pada bagian pengolahan ini, user diminta memilih file citra asli yang akan di proses. Berikut ini merupakan tampilannya.

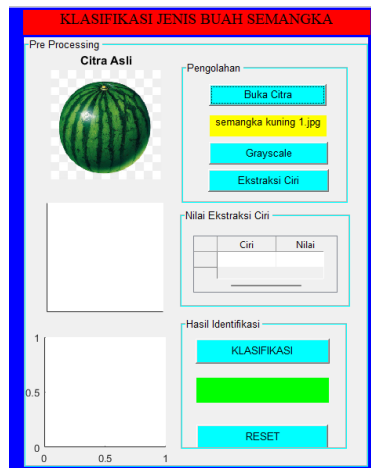


**Gambar 10.** Tampilan *User Interface* Bagian Pengolahan



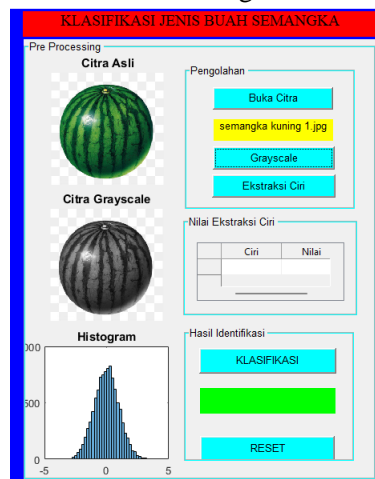
**Gambar 3.8** Data Gambar Pada Browe File

Setelah selesai memilih file citra asli maka program akan menampilkan user interface Sebagai berikut:



**Gambar 11.** Tampilan *User Interface* Citra Asli

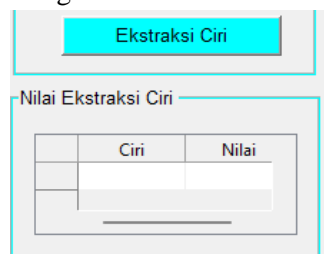
Setelah file citra asli telah di input kedalam program, dan pada bagian pengoahan terdapat pushbutton “Grayscale” untuk mandapatkan nilai untuk ekstraksi ciri maka program akan menampilkan user interface sebagai berikut.



**Gambar 12.** Tampilan *User Interface* Grayscale

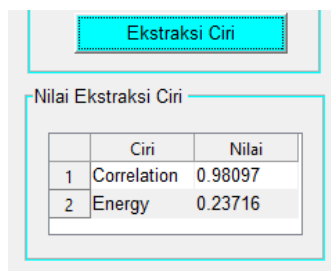
### 3. Bagian Ekstraksi Ciri

Pada bagian ekstraksi ciri ini, user diminta untuk mengklik pushbutton “Ekstraksi Ciri”, dimana akan menampilkan ciri dan nilai dari citra asli yang telah di lakukan pengolahan sebelumnya. Ciri dari ekstraksi ciri ada 2 ciri yaitu Energy, dan Correlation. Berikut tampilan user interface bagian ekstraksi ciri.



**Gambar 13 :** Tampilan *User Interface* bagian Ekstraksi Ciri

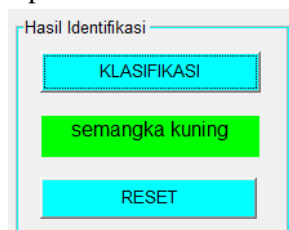
Setelah Pushbutton “Ekstraksi Ciri” di klik maka program akan menampilkan tampilan user interface sebagai berikut.



**Gambar 14.** Tampilan *User Interface*

#### 4. Bagian Hasil

Pada bagian hasil klasifikasi, masih kosong, kemudian untuk mengenali gambar tersebut, user diminta untuk menekan tombol pushbutton “Klasifikasi” pada bagian hasil klasifikasi, untuk mengetahui jenis tersebut. Dan juga terdapat tombol pushbutton “Reset” untuk menghapus dan mengulangi proses pada citra gambar asli yang lainnya. Berikut tampilannya setelah tombol pushbutton “Klasifikasi” tersebut di klik.



**Gambar 15.** Tampilan *User Interface* Ekstraksi Ciri

#### 4.2.2 Pengujian

##### 1. Pengujian Terhadap Data Training

Pengujian Terhadap Data Training dilakukan untuk mendapatkan nilai akurasi pada program yang telah dibuat dengan data yang sudah ada. Proses pengujian data training dilakukan menggunakan arsitektur jaringan syaraf tiruan dengan data sebanyak 100 gambar dari 2 jenis semangka kuning dan semangka merah dari 80 gambar, 69 gambar benar diklasifikasikan, dan 11 gambar salah. Program klasifikasi ini bisa dikatakan baik jika dilihat dari pengujian diatas.

##### 1. Evaluasi Model Performa Menggunakan Confusion Matrix

Confusion matrix adalah pengukuran performa untuk masalah klasifikasi machine learning dimana keluaran dapat berupa dua kelas atau lebih. Confusion matrix berupa tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah.

**Tabel 5.** Tabel Evaluasi Model Confusion Matrix

Confusion Matrix		Kelas Hasil Prediksi	
		Positive	Negative
Kelas Sebenarnya	True	TP = 69	FP = 11
	False	FN = 0	TN = 0

Keterangan :

TP (True Positive) = Jumlah dokumen dari kelas 1 yang benar diklasifikasikan sebagai kelas 1

TN (True Negative) = Jumlah dokumen dari kelas 0 yang benar diklasifikasikan sebagai kelas 0

FP (False Positive) = Jumlah dokumen dari kelas 0 yang salah diklasifikasikan sebagai kelas 1

FN (False Negative) = Jumlah dokumen dari kelas 1 yang salah diklasifikasikan sebagai kelas 0

Rumus confusion matrix untuk menghitung accuracy, precision, recall dan F-1 Score. Nilai akurasi merupakan perbandingan antara data yang terklasifikasi benar dengan keseluruhan data. Accuracy menggambarkan seberapa akurat model dalam mengklasifikasikan dengan benar.

$$\begin{aligned}\text{Accuracy} &= \frac{TP+TN}{TP+TN+FP+FN} \times 100 \% \\ &= \frac{69+0}{69+0+0+11} \times 100 \% \\ &= \frac{69}{80} \times 100 \% \\ \text{Accuracy} &= 86,25 \%\end{aligned}$$

Precision menggambarkan akurasi antara data yang diminta dengan hasil prediksi yang telah diberikan oleh model.

$$\begin{aligned}\text{Precision} &= \frac{TP}{TP+FP} \times 100 \% \\ &= \frac{69}{69+11} \times 100 \% \\ \text{Precision} &= 86,25 \%\end{aligned}$$

Recall atau sensitivity menggambarkan keberhasilan model dalam menemukan kembali sebuah informasi.

$$\begin{aligned}\text{Recall} &= \frac{TP}{TP+FN} \times 100 \% \\ &= \frac{69}{69+0} \times 100 \% \\ \text{Recall} &= 100 \%\end{aligned}$$

F-1 Score menggambarkan perbandingan rata-rata precision dan recall yang dibobotkan. Accuracy tepat kita gunakan sebagai acuan performansi algoritma jika dataset memiliki jumlah data False Negative dan False Possitive yang sangat mendekati (symmetric). Namun, jika jumlahnya tidak mendekati, maka sebaiknya menggunakan F1-Score Sebagai acuan.

$$\begin{aligned}\text{F-1 Score} &= \frac{2 \times \text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \\ &= \frac{2 \times 1 \times 0,86}{1 + 0,86} \\ &= \frac{1,72}{1,86} \\ &= 0,92 * 100 \%\end{aligned}$$

$$\text{F-1 Score} = 92 \%$$

Dari hasil evaluasi model confusion matrix di atas, diketahui akurasi yakni sebesar 86,25 %, pada Precision 86,25 %, Pada Recall 100% dan pada F-1 Score mendapat hasil sebesar 92 % dengan pengujian dari 80 data training jenis buah semangka.

## 2. Pengujian Terhadap Data Testing

Setelah dilakukan pelatihan pada program dengan data yang sudah ada, maka langkah selanjutnya yakni melakukan pengujian terhadap data uji. Proses pengujian dilakukan menggunakan arsitektur jaringan syaraf tiruan dengan akurasi tertinggi sebelumnya. Data yang akan di ujikan sebanyak 40 data testing gambar jenis semangka.

## 4. KESIMPULAN

### 4.1 Kesimpulan

Dari analisa dan pembahasan yang dilakukan peneliti pada Penerapan Metode *K-Nearest Neighbor* Terhadap Klasifikasi Jenis Buah Semangka Dengan *Image Processing*, dapat disimpulkan bahwa:

1. Sistem klasifikasi ini dapat membantu pengguna (*user*) untuk mempermudah dalam pengklasifikasian informasi tentang jenis semangka dengan algoritma *K-Nearest Neighbor*, Sistem klasifikasi ini juga dapat dijadikan sebagai media penerapan intelegensi seorang ahli atau pakar dalam menentukan klasifikasi jenis semangka dengan algoritma *K-Nearest Neighbor*.

2. Dalam hasil program klasifikasi buah semangka dengan algoritma *K-Nearest Neighbor* yang dicocokkan pada perhitungan manual didapat hasil untuk akurasi data *training* sebesar 86.25 % terhadap 80 data *training*. Kemudian setelah dilakukan proses pengolahan pada interface kemudian didapat akurasi data *testing* sebesar 80 % terhadap 20 data *testing*.
3. Model jaringan syaraf tiruan yang dihasilkan sudah mampu mengenali jenis buah semangka dengan sangat baik. Hal ini dapat dilihat dari hasil uji data yang datanya tidak tergabung dalam proses *training* mampu mengenali 16 yang bernilai benar dan 4 yang bernilai salah.

#### 4.2. Saran

1. Semoga bisa menampilkan fitur-fitur dan tampilan lebih menarik lagi untuk pengguna seperti dibagian hasil prediksi user bisa ditambahkan form cetak.
2. Semoga pembangunan sistem prediksi selanjutnya dibuat dengan tingkatan yang lebih baik lagi, agar sistem prediksi yang dibangun lebih baik.

#### Daftar Pustaka

- [1] H. A. & K. H. A. Mumtaha, "Analisis Dampak Perkembangan Revolusi Industri 4.0 Dan Society 5.0 Pada Perilaku Masyarakat Ekonomi (E-Commerce)," *Jurnal Pilar Teknologi : Jurnal Ilmiah Ilmu Ilmu Teknik*, vol. 4, no. 2, p. 55–60, 2019.
- [2] M. J. C. H. A. K. & V. M. Siahaan, "Penerapan Artificial Intelligence ( AI ) Terhadap Seorang Penyandang Disabilitas Tunanetra.," *Information System and Technology*, vol. 01, no. 02, p. 186–193, 2020.
- [3] Paramita, C., Rachmawanto, E. H., Sari, C. A., & Setiadi, D., "(2019). Klasifikasi Jeruk Nipis Terhadap Tingkat Kematangan Buah Berdasarkan Fitur Warna Menggunakan K-Nearest Neighbor.," *In Jurnal Informatika*, 2019.
- [4] Adenugraha, S. P., Arinal, V., & Mulyana, D. I. , "Klasifikasi Kematangan Buah Pisang Ambon Menggunakan Metode KNN dan PCA Berdasarkan Citra RGB dan HSV," *Jurnal Media Informatika Budidarma*, vol. 6, no. 1, 2022.
- [5] M. (. Village, "Usaha tani Buah Semangka Di Desa," vol. 3, no. 405, pp. 1-6, 2022.
- [6] Firmansyah, I. S., & Marhaeni, ".Pembuatan Website Reservasi Hotel Berbasis Php Dan Mysql," *Jurnal Rekayasa Informasi*, vol. 8, no. 2, pp. 130-137, 2019.
- [7] Dinata, R. K., Akbar, H., & Hasdyna, N., "Algoritma K-Nearest Neighbor dengan Euclidean Distance dan Manhattan Distance untuk Klasifikasi Transportasi Bus.," *ILKOM Jurnal Ilmiah*, vol. 12, no. 2, pp. 104-111, 2020.
- [8] A. Desiani dan M. Arhami, Konsep Kecerdasan Buatan, 1 penyunt., D. Hardjono, Penyunt., Yogyakarta: Penerbit ANDI, 2006.
- [9] Kusrini, Sistem Pakar Teori dan Aplikasi, 1 penyunt., f. Suyantoro, Penyunt., Yogyakarta: Penerbit ANDI, 2006.
- [10] A. Desiani dan M. Arhami, Konsep Kecerdasan Buatan, 1 penyunt., D. Hardjono, Penyunt., Yogyakarta: Penerbit ANDI, 2006.
- [11] I. A. Adriana, Penalaran Komputer Berbasis Kasus (Case Based Reasoning), Yogyakarta: Ardana Media, 2007.
- [12] R. D. R. e. all, Telinga Hidung Tenggorok Kepala dan Leher edisi ketujuh, Jakarta: FK UI, 2012.
- [13] E. M. V. S. T.Sutojo, Kecerdasan Buatan, Yogyakarta: Andi, 2011.
- [14] S. W. Faza Akmal, "SISTEM PPAKAR UNTUK MENDIAGNOSA PENYAKIT LAMBUNG DENGAN IMPLEMENTASI METODE CBR (CASE BASED REASONING) BERBASIS WEB," *Jurnal Sarjana Teknik Informatika*, vol. 2 , no. 1, Februari 2014.

- [15] A. M. M. M. N. W. a. N. F. Adiwijaya, "A comparative study of MFCC-KNN and LPC-KNN for hijaiyyah letters pronunciation classification system," *Information and Communication Technology (ICoIC7)*, pp. (pp. 1-5), 2017.
- [16] M. N. Al-Kabi, G. Kanaan, R. Al-Shalabi, S. Al-Sinjlawi dan R. S. Al-Mustafa, "Al-Hadith Text Classifier," *Journal of Applied Sciences* 5, pp. 584-587, 2005.
- [17] F. Harrag dan E. El-Qawasmah, "Neural Network for Arabic Text Classification," *2009 Second International Conference on the Applications of Digital Information and Web Technologies*, pp. 778-783, 2009.
- [18] E. R. R. J. S. A.-F. dan A. , "Klasifikasi Anjuran, Larangan dan Informasi pada Hadis Sahih Al-Bukhari," *e-Proceeding of Engineering*, p. 4683, 2017.
- [19] A. K. S. A.-F. dan A. , "Klasifikasi Informasi, Anjuran dan Larangan pada Hadits Shahih Bukhari menggunakan Metode Support Vector Machine," *e-Proceeding of Engineering*, p. 5014, 2017.



**ZONasi: Jurnal Sistem Informasi**

is licensed under a [Creative Commons Attribution International \(CC BY-SA 4.0\)](https://creativecommons.org/licenses/by-sa/4.0/)