

Identifikasi Jenis Perlakuan Kematangan Pisang Barangan menggunakan Metode *k-Nearest Neighbor*

Identification of Barangan Banana Ripeness Treatment Types using k-Nearest Neighbor

Abdullah*, Rendi Azrian

Program Studi Sistem Informasi, Fakultas Teknik dan Ilmu Komputer, Universitas Islam Indragiri,
Jl. Lintas Provinsi Kabupaten Indragiri Hilir, Riau Indonesia

*Email: abdialam@gmail.com

(received: 16 Agustus 2022, revised: 24 Agustus 2022, accepted: 19 September 2022)

Abstrak

Buah pisang diminati oleh masyarakat dikarenakan buah pisang kaya akan nutrisi yang memang diperlukan oleh tubuh kita. Salah satu buah pisang yang diminati masyarakat adalah pisang jenis barangan. Barangan yang dijual di pasar memiliki berbagai jenis kematangan yang berbeda-beda berdasarkan perlakuannya. Tujuan penelitian ini adalah melakukan identifikasi terhadap jenis perlakuan kematangan barangan. Identifikasi dilakukan berdasarkan analisa terhadap citra pisang barangan dengan menggunakan fitur warna dan tekstur. Dalam melakukan identifikasi digunakan metode *K-Nearest Neighbor* (*k-NN*). *k-NN* membandingkan kemiripan antara data yang belum diketahui dengan data sampel. Nilai *k* yang digunakan pada penelitian ini adalah *k*=1, *k*=3, dan *k*=5. Pengukuran kemiripan menggunakan *Euclidean Distance*, yang merupakan jarak antara 2 vektor fitur. Pengujian klasifikasi menggunakan metode *Holdout*, dimana persentase jumlah data sampel dan data uji adalah 66,67 % data latih dan 33,33 % data uji. Akurasi yang diperoleh pada *k*=1 sebesar 86,67 %, pada *k*=3 sebesar 76,67 %, dan pada *k*=5 sebesar 80%. Akurasi terbaik untuk identifikasi jenis perlakuan kematangan pisang barangan menggunakan metode *k-Nearest Neighbor* didapat pada *k*=1 dengan akurasi mencapai 86,67 %.

Kata kunci: Klasifikasi, *K-Nearest Neighbor*, *Background Subtraction*, *Contrast Stretching*, Matlab

Abstract

Bananas are favored by the public because bananas are rich in nutrients that our bodies need. One of the bananas that people are interested in is the Barangan. Bananas sold in the market have various types of ripeness that vary based on their treatment. This study aimed to identify the type of treatment for Barangan. Identification is carried out based on an analysis of the image of Barangan using color and texture features. The k-Nearest Neighbor (k-NN) method is used in the identification. The k-NN compares the similarity between the unknown data and the sample data. The k values used in this study are k=1, k=3, and k=5. The Euclidean Distance is used to measure the distance between 2 feature vectors. The classification test uses the holdout method, where the percentage of the amount of sample data and test data is 66.67% of training data and 33.33% of test data. The accuracy obtained at k=1 is 86.67%, at k=3 is 76.67%, and at k=5 is 80%. The best accuracy for identifying banana ripeness treatment types using the k-Nearest Neighbor method is obtained at k = 1, with accuracy reaching 86.67%.

Keywords: Classification, *k-Nearest Neighbor*, *Background Subtraction*, *Contrast Stretching*, Matlab

1 Pendahuluan

Pisang barangan (Gambar 1) dijadikan masyarakat sebagai buah konsumsi harian. Pisang barangan yang dijual di pasar memiliki jenis kematangan yang berbeda-beda. Hal yang mempengaruhi kematangan pisang barangan yaitu pisang tersebut ada yang matang dengan kematangan alami (tidak diberi zat pemacu kematangan), ada yang matang dengan kematangan karbit

<http://sistemasi.ftik.unisi.ac.id>

(zat pemacu kematangan) dan ada pula pisang barangan yang matang dengan kematangan ethrel (zat pemacu kematangan).

Klasifikasi jenis kematangan pisang barangan yang berlangsung di pasar selama ini dilakukan secara manual, yakni disortir dengan penglihatan (pengamatan) mata terhadap pisang itu sendiri. Klasifikasi jenis kematangan pisang barangan yang dilakukan secara manual tentu akan mengakibatkan kurang teliti dikarenakan keterbatasan visual, kelelahan, perbedaan persepsi masing-masing pengamat dan menimbulkan kesulitan dalam mengklasifikasikan jenis kematangan pisang barangan tersebut, terlebih pada saat jumlahnya cukup banyak (besar).



Gambar 1. Pisang barangan

Untuk mendapatkan solusi dari permasalahan klasifikasi dan menentukan nilai keakuratan klasifikasi pisang barangan ke dalam kelas kematangan tertentu, maka diperlukan suatu metode yang mampu melakukan klasifikasi pisang barangan. Saat ini, teknik pengolahan citra digital memungkinkan untuk digunakan lebih efektif dan efisien dalam mengidentifikasi kematangan buah-buahan [1]. Adanya kemampuan pengolah citra digital didukung oleh konsep pengenalan dan klasifikasi pola, maka diharapkan klasifikasi pisang barangan berdasarkan jenis kematangan berdasarkan warna *Red*, *Green*, *Blue* (RGB) [2][3] dan tekstur *Gray Level Co-occurrence Matrix* (GLCM) [4] dengan bantuan komputer dapat direalisasikan.

Penelitian ini bertujuan untuk menguji kemampuan *k-NN* [5] dalam mengidentifikasi jenis perlakuan kematangan pisang barangan. Identifikasi dilakukan berdasarkan analisa warna RGB dan tekstur (*energy*, *contrast*, *correlation* dan *homogeneity*). Tingkat keakuratan *k-NN* dihitung dalam melakukan prediksi ke dalam kelas atau kategori dengan persentase kebenarannya. Untuk pengujian tersebut sebuah sistem klasifikasi dibangun dengan menggunakan Bahasa pemrograman Matlab. Pengujian dilakukan dengan beberapa tahapan untuk beberapa nilai *k* yang berbeda. Manfaat yang diharapkan adalah metode ini dapat diaplikasi untuk membantu pendeteksian jenis perlakuan kematangan pisang, khususnya pisang barangan.

2 Tinjauan Literatur

Tugas klasifikasi dengan bantuan komputer melalui teknik pengolahan citra digital telah dilakukan oleh beberapa peneliti terdahulu. Penelitian [6] membahas mengenai klasifikasi terhadap mutu jambu biji dilihat dari luasnya area cacat buah. Klasifikasi dilakukan melalui pengolahan citra digital terhadap citra jambu biji untuk mendapatkan komposisi nilai warna *Red Green Blue* (RGB) dan nilai tekstur *Gray Level Co-occurrence Matrix* (GLCM) [4]. Dipilihnya fitur RGB dan GLCM dikarenakan cacat buah ditandai dengan adanya perubahan warna dan tekstur pada permukaan buah. Penelitian ini menggunakan metode *k-Nearest Neighbor* (*k-NN*). Nilai *k* yang digunakan adalah $k=1$, $k=3$, $k=5$, $k=7$, dan $k=9$. Proses segmentasi citra dilakukan dengan bantuan warna latar belakang (putih, hijau, biru, kuning, merah, dan jingga). Pada penelitian ini segmentasi yang paling baik adalah warna merah. Hal ini disebabkan warna merah memiliki kontras yang paling jelas dari warna jambu biji. Penelitian [7] membahas mengenai klasifikasi terhadap mutu telur ayam ras yang ditinjau dari kebersihan kerabang telur. Penelitian ini menggunakan teknik pengolahan citra digital dengan preprocessing citra menghasilkan citra *grayscale* yang selanjutnya diekstraks cirinya. Metode klasifikasi yang digunakan adalah *k-Nearest Neighbor* (*k-NN*). Penelitian ini menggunakan ekstraksi

ciri orde pertama dan orde kedua. Ekstraksi ciri orde pertama ini untuk mendapatkan nilai fitur *mean*, *variance*, *skewness*, *curtosis*, dan *entropy*. Sedangkan ekstraksi ciri orde kedua untuk mendapatkan nilai fitur *angular second moment*, *contrast*, *correlation*, *variance*, *invers different moment* dan *entropy*. Citra input hanya diubah menjadi citra *grayscale* pada tahap preprocessing dan selanjutnya langsung dilakukan ekstraksi ciri citra. Nilai k yang digunakan adalah $k=1$, $k=3$, $k=5$, dan $k=7$.

Setelah merujuk dua literatur di atas, berikut ini Tabel 1 yang memuat perbedaan dan persamaan penelitian terdahulu dengan penelitian yang dilakukan.

Tabel 1. Persamaan dan Perbedaan Penelitian dengan Penelitian Sebelumnya

No.	Judul Penelitian	Persamaan	Perbedaan
1	Metode Klasifikasi Mutu Jambu Biji Menggunakan KNN Berdasarkan Fitur Warna dan Tekstur [6]	<ul style="list-style-type: none"> - Menggunakan metode klasifikasi <i>k-Nearest Neighbor</i> - Ekstraksi ciri warna RGB dan tekstur GLCM - Aplikasi Matlab 	<ul style="list-style-type: none"> - Pada terdahulu pengujian sistem klasifikasi menggunakan <i>Cross Validation</i>, penelitian ini menggunakan <i>Holdout</i>
2	Klasifikasi Mutu Telor Berdasarkan Kebersihan Kerabang Telor Menggunakan K-Nearest Neighbor [7]	<ul style="list-style-type: none"> - Menggunakan metode klasifikasi <i>k-Nearest Neighbor</i> - Pengujian sistem menggunakan <i>Holdout</i> 	<ul style="list-style-type: none"> - Penelitian terdahulu menggunakan metode ekstraksi ciri statistik orde satu dan orde dua, penelitian ini menggunakan ciri warna RGB dan tekstur GLCM.

Citra digital merupakan image $f(x,y)$ yang telah didigitalisasikan (dibentuk menjadi diskrit) baik koordinat spasialnya atau tingkat intensitasnya [8]. Citra yang ada didigitalisasi agar mudah diolah dan disimpan dalam media penyimpanan seperti memori komputer, hard disk atau media lain. Setiap piksel menyatakan kombinasi komponen warna red, green dan blue (RGB). Model RGB adalah model warna yang terdiri atas tiga kanal warna yaitu merah (R), hijau (Green) dan biru (Blue). Setiap kanal mempunyai kedalaman warna, yang disebut juga sebagai bit depth yang digunakan untuk mengukur jumlah warna yang ditampilkan dalam tiap piksel citra. Kedalaman warna yang cukup sering digunakan adalah 24 bit (true colors) yang didapat dari gabungan 8 bit (256 derajat intensitas yaitu 0-255 pada masing-masing kanal R,G dan B). Sehingga dengan demikian dimiliki warna sebanyak $256 \times 256 \times 256$ atau 2^{24} sama dengan 16.777.216 atau dikenal dengan citra 16 juta warna.

Objek yang ditangkap melalui kamera digital disimpan dalam bentuk digital. Intensitas RGB pada masing-masing piksel yang membentuk citra dapat diilustrasikan sebagaimana matriks di bawah ini:

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ r_{21} & r_{22} & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m1} & r_{m2} & \dots & r_{mn} \end{bmatrix} \quad G = \begin{bmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ g_{21} & g_{22} & \dots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{m1} & g_{m2} & \dots & g_{mn} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix}$$

Setiap kanal warna R (red), G (green) dan B (blue) memiliki histogram. Histogram adalah fungsi yang mendistribusikan jumlah piksel yang ada pada suatu citra untuk setiap level intensitas. Absisnya yang merupakan sumbu x adalah tingkat intensitas, dan ordinatnya yang merupakan sumbu y adalah frekuensi kemunculan atau banyaknya piksel dengan nilai intensitas tertentu.

Vektor fitur objek berupa kombinasi mean warna (1). Jika sebuah citra dengan jumlah piksel P maka mean warna red (R), green (G) dan Blue (B) berturut-turut diberikan oleh (2),(3) dan (4).

$$\bar{x} = (\bar{r}, \bar{g}, \bar{b})^T \quad (1)$$

$$\bar{r} = \frac{\sum r(p)}{P} \quad (2)$$

$$\bar{g} = \frac{\sum g(p)}{P} \quad (3)$$

$$\bar{b} = \frac{\sum b(p)}{P} \quad (4)$$

$r(p)$, $r(g)$, dan $b(p)$ merupakan intensitas pada setiap kanal r (red), g (green) dan b (blue) yang ada pada piksel p .

GLCM melakukan tabulasi berkaitan dengan frekuensi intensitas piksel pada berbagai posisi dalam suatu citra [9], [10]. Ada empat macam fitur tekstur yang sering digunakan dalam GLCM ialah energi, kontras, korelasi, dan homogenitas [11]. Energi merupakan fitur yang diukur melalui konsentrasi pasangan intensitas pada matriks *co-occurrence*. Matriks *co-occurrence* ialah suatu matriks yang menunjukkan frekuensi kemunculan pasangan piksel pada citra dengan intensitas, jarak dan arah tertentu. Persamaan 5 digunakan untuk menghitung energi. Kontras ialah fitur yang menggambarkan sejauh mana kekuatan perbedaan tingkat intensitas pada suatu citra. Nilai kontras berbanding lurus dengan variasi intensitas citra. Formula (6) digunakan untuk mengukur kontras. Korelasi menggambarkan sejauh mana tingkat keteraturan pada citra, apabila nilai pada matriks sama atau merata, maka nilai korelasi yang dihasilkan adalah minimum, sebaliknya apabila nilainya semakin tidak merata, maka korelasi akan maksimum. Korelasi dapat dikalkulasikan melalui (7). Homogenitas ialah fitur yang menggambarkan homogen tidaknya variasi intensitas pada citra. Homogenitas akan membesar jika variasi mengecil dan sebaliknya. Homogenitas dihitung dengan menggunakan (8).

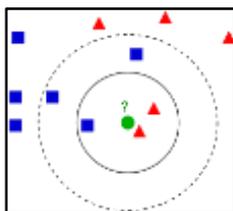
$$E = \sum_{i_1} \sum_{i_2} p^2(i_1, i_2) \quad (5)$$

$$I = \sum \sum (i_1 - i_2)^2 p(i_1, i_2) \quad (6)$$

$$C = - \sum_{i_1} \sum_{i_2} p(i_1, i_2) \log p(i_1, i_2) \quad (7)$$

$$H = \sum_{i_1} \sum_{i_2} \frac{p(i_1, i_2)}{1 + |i_1 - i_2|} \quad (8)$$

K-Nearest Neighbor (*k*-NN) merupakan salah satu algoritma klasifikasi yang populer, efektif dan efisien [12]. Metode ini menggunakan pendekatan kemiripan yaitu objek yang mirip berada pada kategori yang sama. Kemiripan diukur berdasarkan jarak terdekat antara data sampel dengan objek yang diamati. Objek diklasifikasikan berdasarkan mayoritas dari kategori pada *k*-NN, dimana parameter *k* menunjukkan jumlah tetangga terdekat. Gambar 2 menunjukkan ilustrasi *k*-NN bekerja. Termasuk kategori apakah lingkaran hijau? Apakah kategori kotak biru atau segitiga merah? Jika $k=3$ (perhatikan lingkaran dalam), maka lingkaran hijau adalah kategori segi tiga merah, karena ada 2 segitiga merah dan hanya 1 kotak biru di dalam lingkaran. Jika $k=5$ (lingkaran luar), maka lingkaran hijau termasuk kategori segi empat biru (karena ada 3 segi empat biru dan hanya ada 2 segitiga merah).



Gambar 2. Ilustrasi k-NN (Sumber: Wikipedia)

Algoritma k-NN terdiri atas dua langkah utama yaitu: 1) temukan sejumlah k objek dalam sampel yang terdekat dengan objek yang akan diprediksi dengan menggunakan ukuran jarak vektor 2) lakukan voting sejumlah k objek terdekat tersebut untuk menentukan kelas objek target. Akurasi k-NN bergantung kepada ukuran jarak yang digunakan dan nilai k yang diambil. Pada umumnya ukuran jarak (*distance metric*) yang digunakan adalah *euclidean distance* [13] seperti yang ditunjukkan oleh persamaan 9. Jika diketahui dua buah vektor sebagai berikut: $x = [x_1, x_2, x_3, \dots, x_n]$ dan $y = [y_1, y_2, y_3, \dots, y_n]$ maka jarak kedua vektor tersebut adalah:

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (9)$$

3 Metode Penelitian

Penelitian ini bertujuan untuk membangun sistem klasifikasi pisang barangan. Untuk mencapai tujuan tersebut dilakukan langkah-langkah sebagaimana Gambar 3 berikut:



Gambar 3. Langkah-langkah Penelitian

3.1. Akuisisi Data

Sejumlah bahan dan peralatan digunakan dalam akuisisi data yang selanjutnya data tersebut dijadikan sebagai dataset. Bahan dan peralatannya sebagai berikut: (1) Sembilan puluh (90) sampel pisang barangan diakuisi citranya yang selanjutnya dijadikan sebagai dataset. Ada 3 (tiga) kategori pisang barangan yang diambil yaitu ke dalam kategori A (matang alami), kelas B (matang karbit) dan kelas C (matang ethrel). Masing-masing kategori terdiri atas 30 sampel. (2) Sebuah tiang penyangga (tripod) digunakan agar pengambilan citra berjarak sama. (3) Selembar kain hitam sebagai latar pengambilan citra untuk menghindari bayangan dari objek. (4) Intensitas pencahayaan secukupnya saat pengambilan citra yang dibuat sama kualitasnya. (5) Sebuah kamera digital Sony DSC-W610 14.1 Mega Pixels digunakan untuk melakukan akuisisi citra dengan ukuran 4320 x 3240 piksel. (6) Sistem komputer yang digunakan terdiri dari perangkat lunak bahasa pemrograman MATLAB versi 2015b dan sistem operasi windows 10 home. Sedangkan untuk perangkat kerasnya adalah komputer dengan processor intel core i7 generasi terbaru 4.1 GHz dan RAM 32 GB.

3.2. Praproses

Semua citra dikonversi ke format BMP dan diubah ukuran resolusinya menjadi 640 x 480 piksel. Citra tersebut kemudian disimpan dengan intensitas 24 bit ($2^{24}=16$ juta warna) atau biasa disebut *true color*, dimana masing-masing kanal RGB memiliki panjang 8 bit ($2^8 = 256$ varian warna). Hal ini dilakukan untuk meringankan kinerja *software* MATLAB [14]. Praproses (preprocessing) bertujuan untuk memperbaiki mutu citra tersebut dengan menggunakan teknik pengolahan citra sehingga mudah untuk dilakukan proses selanjutnya. Pengolahan citra mentransformasi citra yang asalnya kurang baik menjadi citra keluaran yang lebih baik. Beberapa pengolahan citra yang dilakukan adalah: (1) peregangan kontras (*contrast stretching*) [15], yang bertujuan agar intensitas cahaya pada citra tersebar merata dari 0-255, sehingga citra tampak lebih baik dan objek tampak lebih jelas. (2) Segmentasi citra yang bertujuan untuk mendeteksi objek atau foreground dari background, menggunakan pendekatan *background subtraction*, Teknik yang digunakan adalah pengurangan piksel (*pixel subtraction*) [16][17]. *Pixel subtraction* bekerja dengan cara dilakukan operasi pengurangan matriks intensitas piksel citra objek dikurangi dengan matriks intensitas citra latar. (3) ekstraksi ciri mengawali proses pembentukan basisdata pengetahuan yaitu dengan melakukan proses ekstraksi fitur warna dan tekstur. Setelah dilakukan proses ekstraksi fitur citra sampel, vektor fitur masing-masing citra sampel disimpan di dalam basis data pengetahuan

3.3. Klasifikasi

Proses yang dilakukan dalam klasifikasi kematangan pisang barangan meliputi dua proses utama yaitu proses pembentukan kelas dan proses klasifikasi. Untuk melakukan masing-masing proses, maka akan dikembangkan dua buah sistem, dimana Matlab R2015b, digunakan dalam membangun Sistem. (1) Sistem pembentuk kelas yaitu berfungsi untuk membentuk basis data pengetahuan. Proses pembentukan basis data pengetahuan diawali dengan melakukan proses ekstraksi fitur. Setelah itu vektor fitur masing-masing citra sampel disimpan di dalam basis data. Atribut atau fitur yang digunakan adalah warna RGB (red, green, blue) dan tekstur GLCM (energy, contrast, correlation, homogeneity). (2) sistem klasifikasi yaitu berfungsi untuk melakukan identifikasi citra pisang barangan yang belum diketahui kategorinya. Proses identifikasi dilakukan dengan menggunakan metode k-Nearest Neighbor [18]. Vektor fitur citra pisang barangan yang belum diketahui kategorinya dibandingkan dengan vektor fitur citra sampel yang tersimpan dalam basis data pengetahuan. Keputusan diambil berdasarkan kemiripan. Kemiripan dihitung berdasarkan jarak dua buah vektor menggunakan *euclidean distance* [19]. Semakin kecil jarak vektor yang diperoleh menunjukkan bahwa objek tersebut semakin mirip dengan sampel.

3.4. Evaluasi

Evaluasi dilakukan untuk mengestimasi performansi tingkat keberhasilan klasifikasi dengan menggunakan metode *holdout* [20][21]. Data set pisang berangan dibagi menjadi dua bagian yaitu data sampel dan data uji. Setiap bagian diambil secara random bertingkat masing-masing, yaitu 2/3 sebagai data sampel dan 1/3 data uji. Setiap kategori pada data sampel memiliki jumlah citra yang sama yaitu 20, sehingga total data sampel adalah 60 (3 kategori x 20 citra) seperti terlihat pada Tabel 2. Setiap kategori pada data uji berjumlah yang sama yaitu 10 citra, sehingga jumlah data uji berjumlah 30 (3 kategori x 10 citra) seperti terlihat pada Tabel 3.

Tabel 2. Citra Sampel

No.	Kelas	Jumlah
1	Kelas A	20
2	Kelas B	20
3	Kelas C	20
Total		60

Tabel 3 Citra Uji

No.	Kelas	Jumlah
1	Kelas A	10
2	Kelas B	10
3	Kelas C	10
Total		30

Pengujian dilakukan dengan variasi jumlah k pada fungsi *k-Nearest Neighbor*. Nilai k merupakan jumlah tetangga terdekat. Nilai k yang diujikan adalah $k=1$, $k=3$, dan $k=5$. Dipilih nilai ganjil untuk menghindari kesamaan kedekatan pada dua titik yang berbeda kelas. Karena *k-Nearest Neighbor* akan mengklasifikasikan berdasarkan *voting* kelas terbanyak.

Akurasi adalah tingkat keberhasilan sistem dalam melakukan identifikasi dan klasifikasi pisang barangan. Akurasi yang dimaksud adalah ketepatan hasil klasifikasi yang dilakukan program klasifikasi terhadap seluruh data uji. Untuk identifikasi ketepatan hasil klasifikasi digunakan Matriks konfusi (*confusion matrix*) [22]. Ukuran akurasi (accuracy) sistem yang digunakan adalah perbandingan antara jumlah objek yang diklasifikasikan dengan benar dengan jumlah seluruh prediksi yang dilakukan. Penghitungan akurasi menggunakan persamaan sebagai berikut:

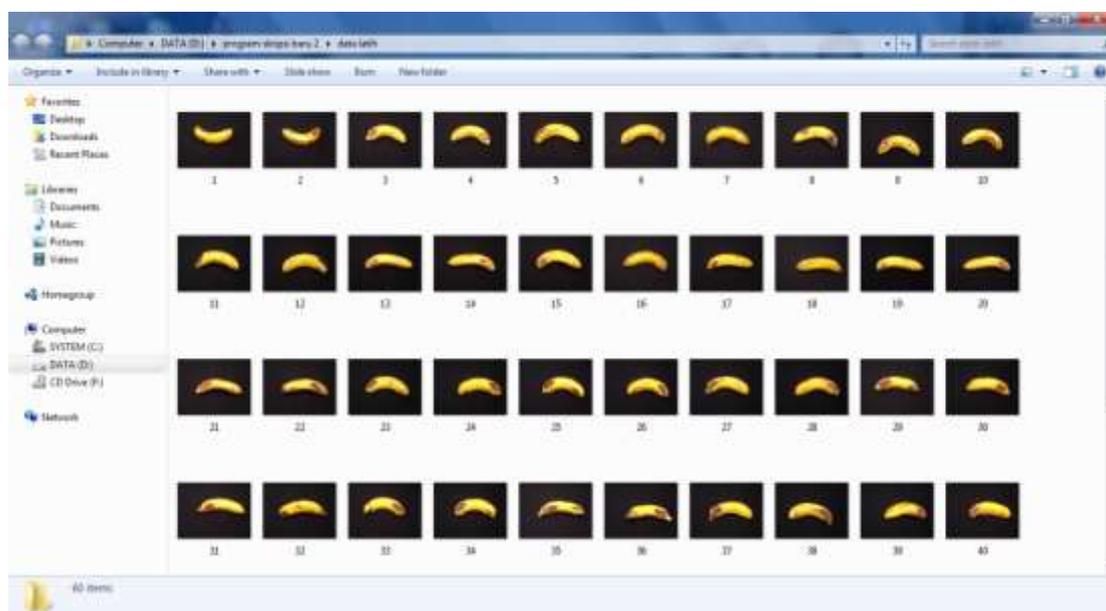
$$\text{akurasi} = \frac{\text{jumlah klasifikasi benar}}{\text{total jumlah eksperimen}} \times 100\%$$

4 Hasil dan Pembahasan

Hasil dan pembahasan diberikan sesuai dengan urutan langkah-langkah penelitian yang telah dilakukan. Hasil yang diperoleh langsung dilakukan pembahasan sebagai berikut:

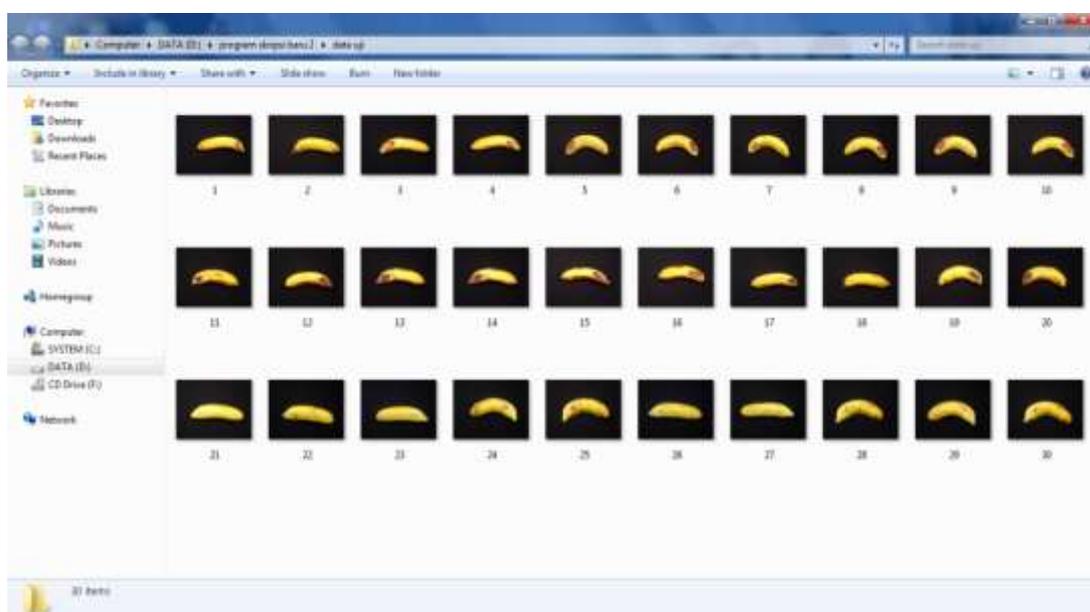
4.1 Hasil Akuisisi Data

Sembilan puluh 90 citra pisang barangan berhasil diakuisi citranya, selanjutnya 60 dijadikan sebagai data latih. Distribusi data latih terdiri atas 20 kategori A (matang alami), 20 kategori B (matang karbit) dan 20 kategori C (matang ethrel). Gambar 4 menunjukkan citra data latih sebanyak 60 citra namun, dalam gambar hanyalah sebanyak 40 item data yang ditampilkan.



Gambar 4. Citra data latih

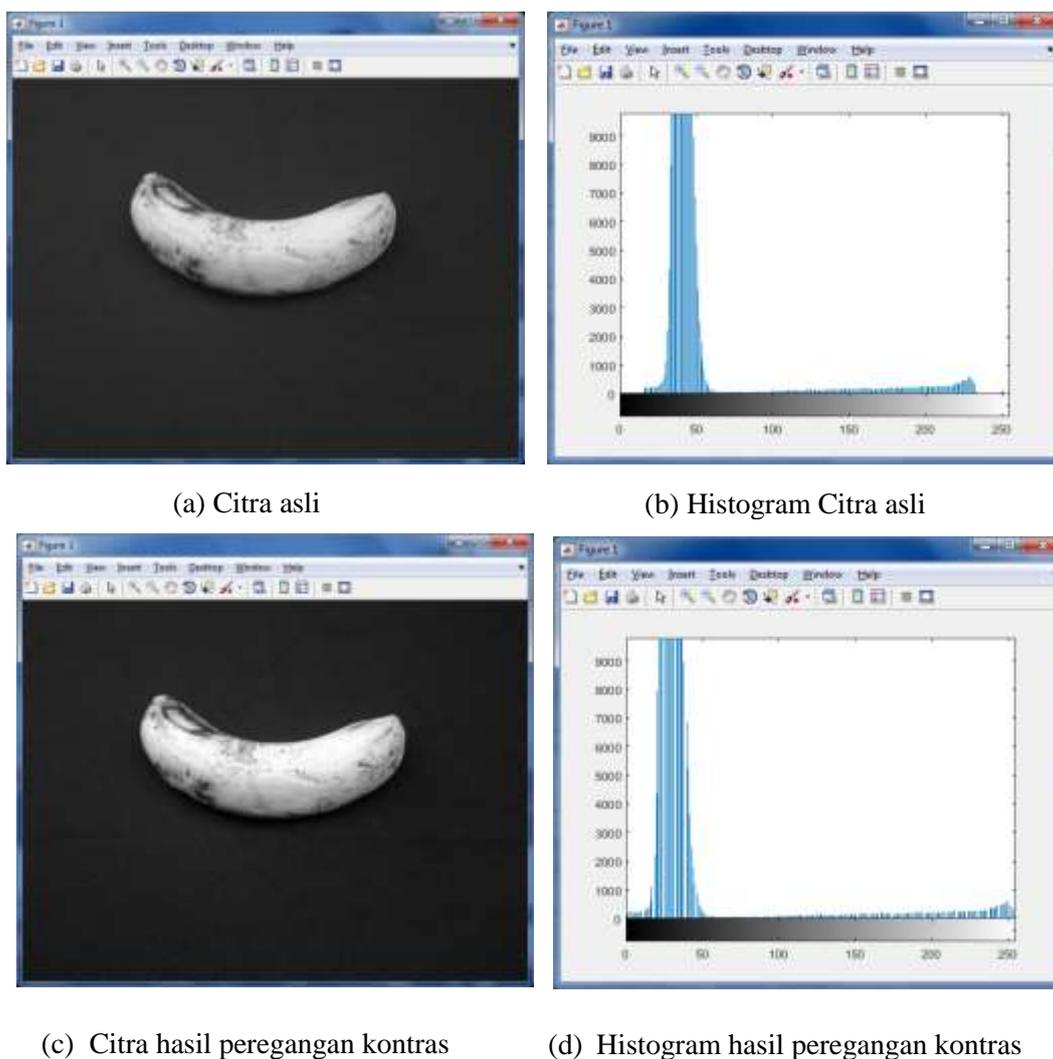
Selanjutnya 30 citra pisang dijadikan sebagai data uji. Distribusi data uji terdiri atas 10 kategori A (matang alami), 10 kategori B (matang karbit) dan 10 kategori C (matang ethrel). Gambar 5 menunjukkan citra data uji sebanyak 30 item.



Gambar 5. Citra data Uji

4.2 Hasil Praproses

Sistem klasifikasi pisang barangan yang dibangun pada penelitian ini ditambah beberapa fitur pada tahapan praproses yaitu: (1) teknik peregangan kontras atau *contrast stretching* pada tahap praproses citra, (2) teknik segmentasi citra menggunakan metode *Background Subtraction* dengan teknik *pixel subtraction* yang bertujuan memisahkan objek pisang barangan dengan latar belakangnya agar hanya area objek pisang saja yang akan diekstraksi cirinya. (3) ekstraksi ciri warna dan tekstur untuk pembuatan vektor fitur yang berguna pada saat pembuatan basis data maupun pada saat proses klasifikasi. (4) fitur penolakan pada sistem yang dibangun agar sistem menolak dengan menampilkan pesan atau *Messagebox* pada saat data yang diujikan bukanlah pisang barangan. Pada tahapan klasifikasi terdiri atas sistem pembentukan kelas dan sistem klasifikasi. Pada tahapan terakhir adalah evaluasi, untuk mengetahui tingkat akurasi. Gambar 6 menunjukkan proses peregangan kontras. Citra asli dan citra hasil peregangan tampak berbeda kualitasnya. Untuk lebih jelas dapat dilihat perbedaan pada histogram citra asli dan citra hasil.



Gambar 6. Peregangan Kontras Pada Citra Grayscale

Langkah awal dalam melakukan peregangan kontras terhadap seluruh data citra adalah dengan melakukan eksperimen terhadap suatu sampel citra untuk mengetahui histogramnya. Hal ini bisa dilihat pada script matlab di bawah ini.

```
i = imread('pisang.bmp');  
iGray = rgb2gray(i);  
j = imadjust(iGray, [15/255 234/255], [0/255 255/255]);  
figure, imshow(i);
```

<http://sistemasi.ftik.unisi.ac.id>

```
figure, imhist(i);  
figure, imshow(j);  
figure, imshit(j);
```

Citra pisang yang disimpan dalam variabel *i*, berdasarkan histogramnya diketahui bahwa citra tersebut tidak memiliki piksel dengan intensitas di bawah 15 dan di atas 234. Untuk memperbaikinya, dilakukan dengan fungsi *intensity adjustment*. *Intensity adjustment* bekerja dengan melakukan pemetaan *linear* terhadap nilai intensitas pada histogram awal menjadi nilai intensitas pada histogram yang baru, sehingga diperoleh sebuah citra baru yang memiliki rentang histogram 0 hingga 255.

Setelah peregangannya diterapkan pada citra grayscale, selanjutnya adalah menerapkan teknik peregangannya kontras di atas pada citra berwarna RGB.

```
i = imread('pisang.bmp');  
j = imadjust(i, [15/255 234/255], [0/255 255/255]);  
figure, imshow(i);  
figure, imshow(j);
```



(a) Citra asli

(b) Citra hasil peregangannya kontras

Gambar 7. Peregangannya Kontras Pada Citra Warna (RGB)

Pada Gambar 7, terlihat bahwa setelah dilakukan peregangannya intensitas kekontrasannya (intensitas piksel 0-255), secara visual citra nampak lebih baik dan objek tampak lebih jelas.

Setelah dilakukan proses peregangannya kontras pada citra sehingga objek pisang barangan tampak lebih jelas, tahap selanjutnya yaitu melakukan segmentasi citra yang bertujuan memisahkan objek dan latar belakangnya. Gambar 8 menunjukkan citra hasil peregangannya dan citra latar.



(a) Citra Pisang Barangan

(b) Citra Latar

Gambar 8. Citra Hasil Peregangannya Kontras dan Citra Latar Belakang (Background)



Gambar 9. Citra pisang barangan tersegmentasi dari latar belakangnya

Setelah operasi pengurangan nilai intensitas piksel RGB matriks intensitas citra objek pisang terhadap matriks intensitas citra latar belakang, selanjutnya dilakukan proses *threshold* yang bertujuan untuk memperbaiki hasil *pixel subtraction* yang telah dilakukan sebelumnya agar pemisahan tersebut benar-benar optimal, sehingga objek buah pisang barangan benar-benar tersegmentasi dari latarnya. Hasil segmentasi citra tersebut bisa dilihat pada Gambar 9. Terlihat hanya objek pisang saja yang tampak pada citra, sehingga pada proses ekstraksi ciri hanya area objek pisang saja yang akan dihitung cirinya.

Pada tahap ekstraksi ciri, fitur warna RGB dan fitur tekstur GLCM citra pisang barangan diekstraksi untuk mendapatkan vektor fitur. Vektor fitur ini berperan mewakili citra pisang barangan tersebut seperti yang terlihat pada Gambar 10.

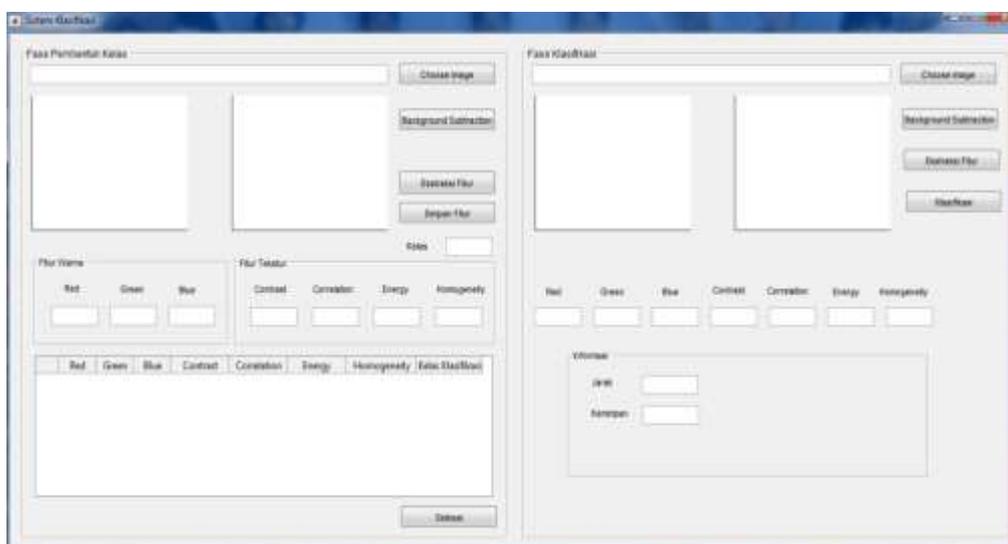
Fitur Warna			Fitur Tekstur			
Red	Green	Blue	Contrast	Correlation	Energy	Homogeneity
171	138	20	0.14819	0.94123	0.8057	0.98198

Gambar 10. Tampilan hasil ekstraksi ciri citra berupa vektor fitur

Vektor fitur ini disimpan pada fase pembentukan kelas sebagai basis data pengetahuan sistem. Setelah sistem memiliki pengetahuan, barulah proses pengujian atau klasifikasi dapat dilakukan.

4.3 Hasil Klasifikasi

Sebuah sistem klasifikasi berhasil dibangun yang terdiri dari dua buah sub sistem yaitu sistem pembentuk kelas yang digunakan untuk membangun basis data pengetahuan dan sistem prediksi yang digunakan untuk memprediksi kategori pisang barangan yang belum diketahui kategorinya. Fase pembentukan kelas dan fase pengujian atau klasifikasi dibuat dalam satu antarmuka yang sama dan dipisahkan oleh panel program.

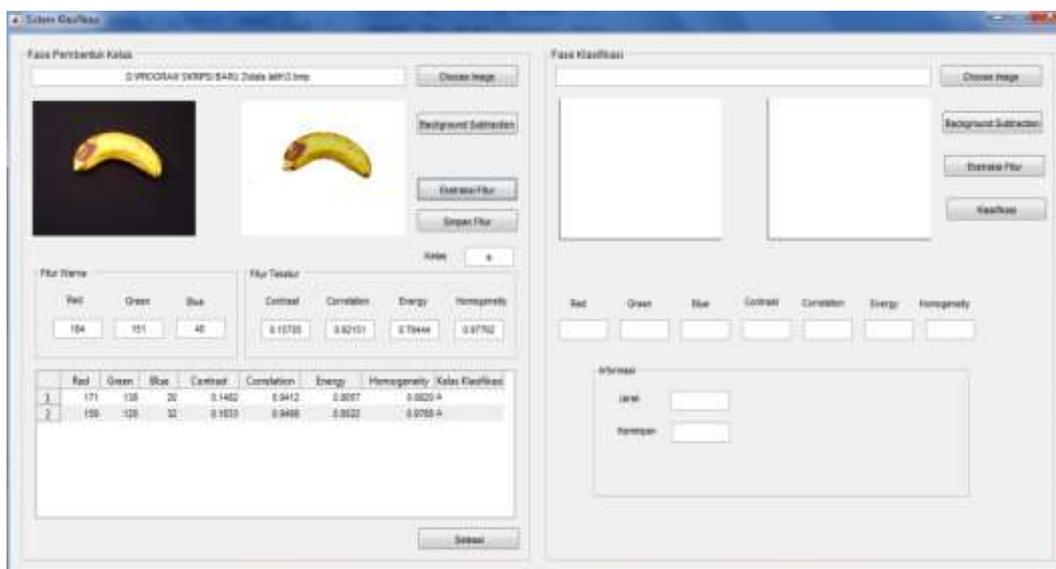


Gambar 11. Antar muka (Pembentuk kelas dan klasifikasi)

Pada antar muka klasifikasi Gambar 11 terdapat 2 panel utama, yakni panel fase pembentuk kelas dan panel fase klasifikasi. Panel fase pembentuk kelas ditujukan untuk membentuk kelas atau basis data pengetahuan berdasarkan data sampel, sedangkan panel klasifikasi ditujukan untuk melakukan identifikasi terhadap suatu data yang belum diketahui. Sebelum bisa dilakukan identifikasi terhadap suatu objek tertentu, terlebih dahulu harus dilakukan pembentukan kelas pada fase pembentuk kelas.

4.3.1 Sistem Pembentukan Kelas

Pada tahap ini, setiap citra sampel akan diberi label kelas untuk membentuk basis data pengetahuan.



Gambar 12. Pemberian label kelas klasifikasi

Pada Gambar 12, pemberian label kelas klasifikasi telah dibatasi dengan nilai inputan (A/ B/ C) saja. Hal ini sesuai dengan kelas tujuan klasifikasi yang terdiri dari kelas A (matang alami), kelas B (matang karbit) dan kelas C (matang ethrel).

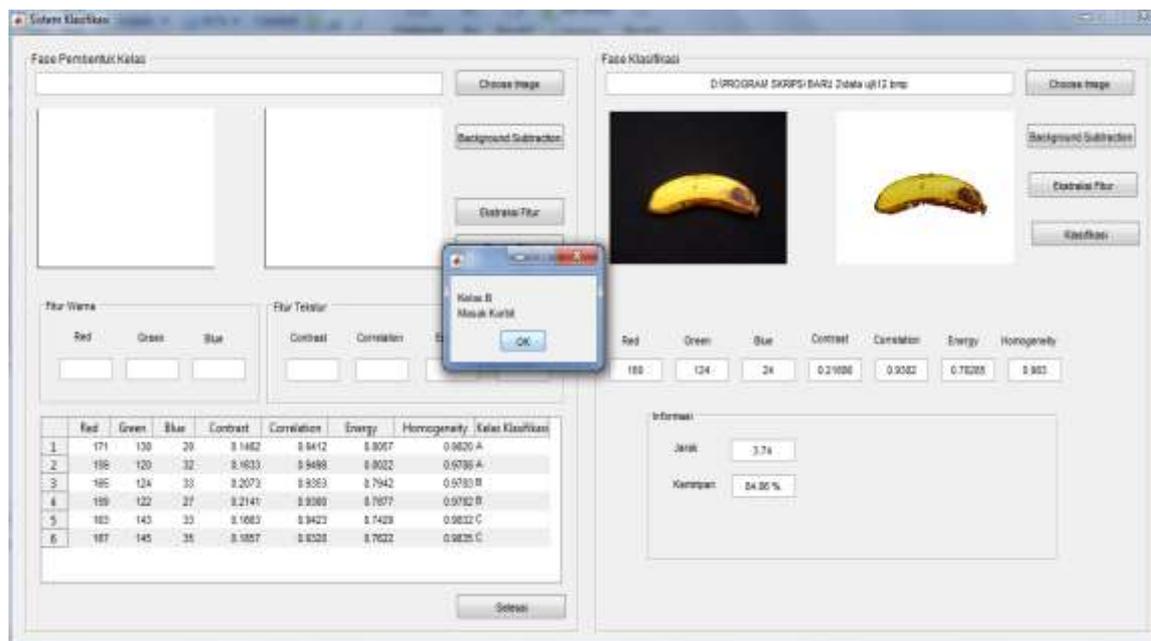


Gambar 13. Pesan kesalahan dalam pemberian label kelas

Jika terjadi input diluar dari nilai A, B atau C maka sistem akan menampilkan pesan kesalahan seperti Gambar 13.

4.3.2 Sistem Klasifikasi

Klasifikasi data citra uji menggunakan metode *k-Nearest Neighbor* dengan perhitungan jarak *euclidean distance* menampilkan informasi persentase kemiripan data uji terhadap data sampel seperti terlihat pada Gambar 14. Hal ini dilihat dari besar kecilnya nilai jarak hasil perhitungan dengan *euclidean distance* antara vektor fitur citra uji dengan vektor fitur data sampel yang ada di dalam *database*. Semakin kecil jarak antara keduanya, maka kemiripan semakin besar.



Gambar 14. Antarmuka klasifikasi pisang barangan

Selanjutnya dilakukan pengujian terhadap sistem penolakan (*rejection*) bila mana kemiripan yang diperoleh dari citra query kurang dari nilai ambang yakni sebesar 60 %. Untuk itu dilakukan pengujian terhadap sistem penolakan (*rejection*) bila mana kemiripan yang diperoleh dari citra uji kurang dari nilai ambang batas atau *threshold* yakni sebesar 60 %.



Nama Kelas : DITOLAK
Jarak : 26.48
Kemiripan : 56 %

Gambar 15. Pengujian sistem penolakan (*rejection*)

Berdasarkan Gambar 15, diketahui bahwa apabila kemiripan yang diperoleh dari citra uji kurang dari 60%, maka sistem klasifikasi pisang barangan akan melakukan penolakan (*rejection*). Dapat ditarik kesimpulan bahwa sistem yang dibangun mampu untuk mengenali pisang barangan, sedangkan buah durian ditolak oleh sistem, karena tidak mirip dengan sampel yang ada dalam basis data pengetahuan sistem.

4.4 Hasil Evaluasi

Evaluasi dilakukan untuk mengetahui sejauh mana kemampuan *k-NN* dalam melakukan identifikasi jenis perlakuan kematangan pisang barangan. Pengujian dilakukan dengan beberapa tahapan untuk beberapa nilai *k* yang berbeda.

Tabel 4. Matriks konfusi pengujian pada $k = 1$

f_{ij}		Kelas Hasil Klasifikasi		
		Kelas A	Kelas B	Kelas C
Kelas Asli	Kelas A	9	1	-
	Kelas B	3	7	-
	Kelas C	-	-	10

Dari Tabel 4 diketahui bahwa klasifikasi yang dilakukan menggunakan $k=1$ dengan data uji sebanyak 30 citra, citra uji yang diklasifikasikan secara benar oleh sistem sebanyak 26 citra uji, sedangkan sisanya sebanyak 4 citra uji diklasifikasikan salah oleh sistem. Dengan demikian dapat diketahui bahwa akurasi yang diperoleh sebesar 86,67 %.

Tabel 5. Matriks Konfusi Pengujian Pada $k = 3$

f_{ij}		Kelas Hasil Klasifikasi		
		Kelas A	Kelas B	Kelas C
Kelas Asli	Kelas A	7	3	-
	Kelas B	3	7	-
	Kelas C	0	1	9

Dari Tabel 5 diketahui bahwa klasifikasi yang dilakukan menggunakan $k=3$ dengan data uji sebanyak 30 citra, citra uji yang diklasifikasikan secara benar oleh sistem sebanyak 23 citra uji, sedangkan sisanya sebanyak 7 citra uji diklasifikasikan salah oleh sistem. Dengan demikian dapat diketahui bahwa akurasi yang diperoleh sebesar 76,67 %.

Tabel 6. Matriks Konfusi Pengujian Pada $k = 5$

f_{ij}		Kelas Hasil Klasifikasi		
		Kelas A	Kelas B	Kelas C
Kelas Asli	Kelas A	8	2	-
	Kelas B	2	8	-
	Kelas C	1	1	8

Dari Tabel 6 diketahui bahwa klasifikasi yang dilakukan menggunakan $k=5$ dengan data uji sebanyak 30 citra, citra uji yang diklasifikasikan secara benar oleh sistem sebanyak 24 citra uji, sedangkan sisanya sebanyak 4 citra uji diklasifikasikan salah oleh sistem. Dengan demikian dapat diketahui bahwa akurasi yang diperoleh sebesar 80 %.

Selanjutnya akan diperlihatkan beberapa sampel hasil klasifikasi pisang barangan yang dilakukan secara benar atau berhasil diklasifikasikan oleh sistem sebagaimana terlihat pada Tabel 7.

Tabel 7. Sampel Uji Yang Diklasifikasikan secara Benar

No.	Citra Pisang Barangan	Keterangan	
1		Nama kelas Jarak Kemiripan	: Kelas A : 3,74 : 87,56 %
2		Nama kelas Jarak Kemiripan	: Kelas B : 4,12 : 91,37 %
3		Nama kelas Jarak Kemiripan	: Kelas A : 4,58 : 87,33 %

5 Kesimpulan

Berdasarkan pengujian k -NN dalam melakukan identifikasi jenis perlakuan kematangan pisang barangan menggunakan metode validasi *Holdout*, maka dapat ditarik kesimpulan sebagai berikut: (1) Metode k -Nearest Neighbor dapat melakukan klasifikasi terhadap jenis kematangan buah pisang barangan berdasarkan analisa terhadap citranya. (2) Akurasi yang diperoleh menggunakan $k=1$ sebesar 86,67 %, pada $k=3$ sebesar 76,67 %, dan pada $k = 5$ sebesar 80 %. Akurasi tertinggi didapat pada $k=1$, yakni sebesar 86,67 %. (3) Sistem klasifikasi pisang barangan memiliki kemampuan untuk melakukan penolakan klasifikasi jika citra uji yang tidak dikenal atau diluar pengetahuan sistem. Evaluasi klasifikasi pisang barangan berdasarkan jenis perlakuan kematangan berdasarkan warna dan tekstur akan lebih baik jika dilakukan percobaan maupun perbandingan dengan metode lain. Aplikasi sistem klasifikasi ini dapat dikembangkan dengan berbasis android serta dapat dilakukan perbaikan user interface sehingga memudahkan pengguna.

Referensi

- [1] K. Hameed, D. Chai, and A. Rassau, "A Progressive Weighted Average Weight Optimisation Ensemble Technique for Fruit and Vegetable Classification," in *16th IEEE International Conference on Control, Automation, Robotics and Vision, ICARCV 2020*, 2020, pp. 303–308, doi: 10.1109/ICARCV50220.2020.9305474.
- [2] Z. Alqadi, M. Khrisat, A. Hindi, and M. Dwairi, "Features Analysis of RGB Color Image based on Wavelet Packet Information," 2020.
- [3] Claudio Cusano, P. Napoletano, and R. Schettini, "Combining Multiple Features for Color Texture Classification," *J. Electron. Imaging*, vol. 25, no. 6, pp. 1–9, 2016.
- [4] E. K. Sharma, E. Priyanka, E. A. Kalsh, and E. K. Saini, "GLCM and its Features," *Int. J. Adv. Researc Electron. Commun. Eng.*, vol. 4, no. 8, pp. 2180–2182, 2015.
- [5] Z. Zhang, "Introduction to Machine Learning: k-Nearest Neighbors," *Ann. Transl. Med.*, vol. 4, no. 11, pp. 1–7, 2016, doi: 10.21037/atm.2016.03.37.
- [6] T. Y. Prahudaya and A. Harjoko, "Metode Klasifikasi Mutu Jambu Biji menggunakan KNN berdasarkan Fitur Warna dan Tekstur," *J. Teknosains*, vol. 6, no. 2, pp. 113–123, 2017, doi: 10.22146/teknosains.26972.
- [7] P. R. Trisnaningtyas and Maimunah, "Klasifikasi Mutu Telur berdasarkan Kebersihan Kerabang Telur menggunakan K-Nearest Neighbor," *Konf. Nas. Inform. 2015 Klasifikasi*, no. October 2015, pp. 241–245, 2015.
- [8] M. Bhat, "Digital Image Processing," *Int. J. Sci. Technol. Res.*, vol. 3, no. 1, pp. 272–276, 2014, [Online]. Available: www.ijstr.org.
- [9] C. V. Angkoso, "Pengenalan Jender Berbasis Tekstur Pada Citra Wajah Foto Digital," *Prosiding Konf. Nas. "Inovasi dalam Desain dan Teknol.*, pp. 119–125, 2011.
- [10] I. R. G. A. Sugiarta, M. Sudarma, and I. M. O. Widyantara, "Ekstraksi Fitur Warna , Tekstur dan Bentuk untuk Clustered-Based Retrieval of Images (CLUE)," *Teknol. Elektro*, vol. 16, no. 1, pp. 85–90, 2017.
- [11] A. Halim, H. Hardy, and M. Mytosin, "Aplikasi Image Retrieval dengan Histogram Warna dan Multi- scale GLCM," vol. 16, no. 1, pp. 41–50, 2015.
- [12] M. Kibanov, M. Becker, J. Mueller, M. Atzmueller, A. Hotho, and G. Stumme, "Adaptive kNN using Expected Accuracy for Classification of Geo-Spatial Data," in *Proceedings of Symposium on Applied Computing (SAC)*, 2017, pp. 1–9, doi: 10.1145/3167132.3167226.
- [13] E. López-Iñesta, F. Grimaldo, and M. Arevalillo-Herráez, "Classification Similarity Learning using Feature-Based and Distance-Based Representations: A Comparative Study," *Appl. Artif. Intell.*, vol. 29, no. 5, pp. 445–458, 2015, doi: 10.1080/08839514.2015.1026658.
- [14] "Matrices and Arrays - MATLAB & Simulink," www.mathworks.com, 2022. .
- [15] S. Supiyanto and T. Suparwati, "Perbaikan Citra Menggunakan Metode Contrast Stretching," *J. Siger Mat.*, vol. 2, no. 1, pp. 13–18, 2021, doi: 10.23960/jsm.v2i1.2743.
- [16] S. Kamate and N. Yilmazer, "Application of Object Detection and Tracking Techniques for Unmanned Aerial Vehicles," *Procedia Comput. Sci.*, vol. 61, pp. 436–441, 2015, doi: 10.1016/j.procs.2015.09.183.
- [17] S. Kumar and J. S. Yadav, "Advances in Intelligent Systems and Computing," in *Proceeding*

- of International Conference on Intelligent Communication, Control and Devices*, 2017.
- [18] S. Zhang, X. Li, M. Zong, X. Zhu, and R. Wang, "Efficient k-NN Classification with Different Numbers of Nearest Neighbors," *IEEE Trans. Neural Networks Learn. Syst.*, pp. 1–12, 2017, doi: 10.1109/TNNLS.2017.2673241.
 - [19] L. Liberti, "Distance Geometry and Data Science," *TOP*, vol. 28, no. 2, pp. 271–339, 2020, doi: 10.1007/s11750-020-00563-0.
 - [20] S. Raschka, *Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning*. 2018.
 - [21] P. Galdi and R. Tagliaferri, "Data Mining: Accuracy and Error Measures for Classification and Prediction," in *Reference Module in Life Sciences*, January, Elsevier, 2019, pp. 1–14.
 - [22] P. Thomas, H. Bril El Haouzi, M. C. Suhner, A. Thomas, E. Zimmermann, and M. Noyel, "Using A Classifier Ensemble for Proactive Quality Monitoring and Control: The Impact of the Choice of Classifiers Types, Selection Criterion, and Fusion Process," *Computers in Industry*. 2018.