

Classification of Oil Palm Fruit Ripeness Based on Color Using the K-Nearest Neighbor (KNN) Method

Klasifikasi Kematangan Buah Kelapa Sawit Berdasarkan Warna Menggunakan Metode K-Nearest Neighbor

Nova Elija Barutu, Dafwen Toresa

Universitas Lancang Kuning, Indonesia

E-mail: novaelija079@gmail.com

*Correspondence: novaelija079@gmail.com

abstract

The K-Nearest Neighbor (K-NN) algorithm is a simple machine learning algorithm used for classification and regression. This study aims to implement the K-NN algorithm in classifying the ripeness level of oil palm fruit based on color. The data used consisted of 270 images of dura, tenera, and pisifera oil palm fruits taken using a smartphone camera. The results showed that the K value in the K-NN algorithm plays an important role in determining the classification performance. With $K = 3$, the model achieved the highest accuracy of 93.67%, while the lowest accuracy was 80.05% with a value of $K = 25$. Compared to previous studies that obtained the highest accuracy of 92% at $K = 7$, this study shows an increase in classification performance. Classification data analysis showed that 56 image data were correctly classified and 25 image data were incorrectly classified from a total of 81 test image data. This study proves that K-NN with RGB color images can be effectively used for classification of the ripeness level of oil palm fruit

Keywords : K-Nearest Neighbor (K-NN), Oil Palm Fruit, Image Processing

Abstrak

Algoritma K-Nearest Neighbor (K-NN) adalah algoritma pembelajaran mesin yang sederhana digunakan untuk klasifikasi dan regresi. Penelitian ini bertujuan untuk mengimplementasikan algoritma K-NN dalam klasifikasi tingkat kematangan buah kelapa sawit berdasarkan warna. Data yang digunakan terdiri dari 270 citra buah kelapa sawit jenis dura, tenera, dan pisifera yang diambil menggunakan kamera smartphone. Hasil penelitian menunjukkan bahwa nilai K pada algoritma K-NN berperan penting dalam menentukan performa klasifikasi. Dengan $K=3$, model mencapai akurasi tertinggi sebesar 93,67%, sedangkan akurasi terendah sebesar 80,05% nilai $K=25$. Dibandingkan dengan penelitian sebelumnya yang memperoleh akurasi tertinggi 92% pada $K=7$, penelitian ini menunjukkan peningkatan performa klasifikasi. Analisis data klasifikasi menunjukkan 56 data citra terklasifikasi dengan benar dan 25 data citra salah dari total 81 data citra uji. Penelitian ini membuktikan bahwa K-NN dengan citra warna RGB dapat efektif digunakan untuk klasifikasi tingkat kematangan buah kelapa sawit

Kata Kunci: K-Nearest Neighbor (K-NN), Buah kelapa sawit, Pengolahan Citra

Received: 02-01-2026 | Revised: 01-02-2026 | Accepted: 11-02-2026

I-JAICL is licensed under a Creative Commons Attribution-Share Alike 4.0 International License



1. Pendahuluan

Kelapa sawit (*Elaeis guineensis*) merupakan tanaman yang mempunyai peranan penting dalam perekonomian Indonesia. Kelapa sawit memiliki potensi ekonomi yang besar karena

minyak nabati yang disebut minyak sawit diekstraksi dari buahnya. Menurut Gabungan Pengusaha Kelapa Sawit Indonesia (GAPKI), konsumsi minyak sawit dalam negeri pada tahun 2021 diperkirakan mencapai 18,422 miliar ton, meningkat 6 ton dibandingkan tahun lalu (Gabungan Pengusaha Kelapa Sawit Indonesia, 2022). Riau merupakan salah satu provinsi di Indonesia yang mempunyai peranan penting dalam industri kelapa sawit. Riau terletak di Pulau Sumatera dan kondisi geografisnya sangat cocok untuk budidaya kelapa sawit. Iklim tropis dan tanah yang subur menjadikan Riau salah satu daerah unggulan perkebunan kelapa sawit. Industri kelapa sawit di Indonesia sedang mengalami pertumbuhan yang pesat. Perkebunan kelapa sawit juga berkontribusi terhadap penciptaan lapangan kerja di berbagai daerah.

Kualitas minyak sawit sangat dipengaruhi oleh standar buah sawit yang digunakan dalam proses produksinya. Kematangan merupakan faktor penting yang menentukan kualitas buah. Buah sawit yang diolah harus mencapai tingkat kematangan optimal. Kematangan ini berdampak besar terhadap kualitas hasil panen, termasuk kualitas minyak sawit yang dihasilkan. Petani biasanya menilai secara manual tingkat kematangan berdasarkan warna dan jumlah buah yang jatuh. Buah yang dipanen dikumpulkan dan diangkut dengan truk ke pabrik. Buah yang dapat diangkut meliputi buah yang matang dan terlalu matang, serta buah yang mentah dan buah yang tidak layak untuk diolah. Buah segar tidak menghasilkan minyak apapun pada saat pengolahan di pabrik, sehingga pengangkutan buah mentah ke pabrik akan mengakibatkan kerugian.

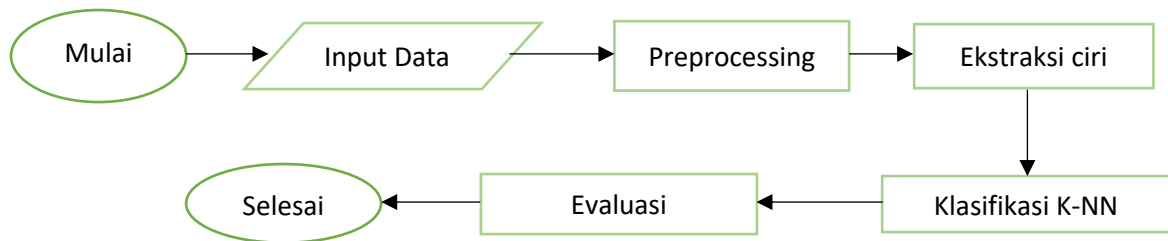
Pada penelitian sebelumnya, Aplikasi Pendektesian kematangan tandan buah segar (TBS) kelapa sawit berdasarkan komposisi warna menggunakan algoritma K-NN. Pada penelitian ini membuat aplikasi berbasis web yang dikelola admin untuk mengelola data training dan aplikasi berbasis android yang dapat memprediksi TBS yang layak dan tidak layak angkut ke pabrik. Ekstraksi ciri yang digunakan adalah warna RGB, nilai numerik yang didapat dari ekstraksi ciri kemudian diolah dengan algoritma. Algoritma yang digunakan pada aplikasi ini adalah k-nearest neighbor. Berdasarkan pengujian yang dilakukan, aplikasi memiliki tingkat akurasi mencapai 92% dengan nilai $K=7$ (Rifqi et al., 2020).

Di era kemajuan teknologi, kini dapat digunakan untuk lebih mudah mendeteksi kematangan buah kelapa sawit. Salah satu teknologi yang dapat digunakan adalah pengolahan citra digital. Data gambar buah dari PT. Bina Pitri Jaya memotret langsung dengan kamera smartphone. Petani dan pekerja pabrik masih melakukan pengecekan kematangan buah kelapa sawit secara manual. Produk buah kelapa sawit yang diolah di pabrik ini terdiri dari tiga jenis tanaman kelapa sawit yaitu Dura, Tenera dan Pisifera. Perbedaan ketiga spesies ini berkaitan dengan struktur buah dan produksi minyak. Tanaman tenera sering digunakan di perkebunan karena kulitnya yang tipis dan hasil minyak yang baik.

Pada penelitian ini, peneliti menggunakan metode K-Nearest Neighbor (Knn) untuk mengklasifikasikan objek gambar buah kelapa sawit dan mengenali buah kelapa sawit berdasarkan warnanya. Artinya gambar buah berisi buah mentah, buah matang, dan buah terlalu matang, serta diproses dengan perangkat lunak Matlab. K-Nearest Neighbors (KNN) adalah algoritma pembelajaran mesin yang sederhana dan mudah diimplementasikan yang dapat digunakan untuk memecahkan masalah klasifikasi dan regresi. Algoritma K-NN menggunakan klasifikasi lingkungan sebagai prediktor sampel uji baru. Biasanya tetangga dekat atau jauh dihitung berdasarkan jarak Euclidean.

Tujuan Penelitian Mengimplementasikan Algoritma K-Nearest Neighbor (K-NN) dalam Klasifikasi Kematangan Buah Kelapa Sawit berdasarkan warna, dan Menghitung hasil tingkat akurasi algoritma K-Nearest Neighbor (K-NN) pada klasifikasi tingkat kematangan buah kelapa sawit berdasarkan warna.

2. Metode Penelitian



Gambar 1. Tahapan Penelitian

1. Input Data

Pada tahap awal adalah menginput data buah kelapa sawit. Jumlah seluruh dataset adalah 270 data citra. Data buah kelapa sawit yang digunakan saat input data adalah data uji (test) yang terdiri dari 73 Matang, 2 Mentah dan 6 Terlalu Matang dengan total data test keseluruhan adalah 81 data uji (test).

2. Preprocessing

Pada tahap Preprocessing dilakukan untuk menghilangkan bagian yang tidak penting dan juga untuk menghilangkan permasalahan saat pemrosesan data seperti membersihkan data dengan membuang gambar yang rusak atau cacat. Proses ini sangat penting karena data yang tidak bersih dapat menyebabkan hasil yang tidak akurat.

3. Ekstraksi Ciri

Pada tahap Ekstraksi ciri merupakan tahap yang dilakukan proses mengekstrak ciri atau informasi yang dimiliki objek citra untuk membedakan objek lainnya. Ekstraksi ciri atau fitur menggunakan model warna RGB. RGB tersusun atas tiga komponen, yaitu Red, Green dan Blue, kemudian dilakukan proses thresholding dengan menetapkan nilai Ambang batas sehingga dapat mengidentifikasi objek citra dengan lebih akurat dan memperoleh informasi warna yang lebih rinci dengan menghitung ciri atau fitur seperti jumlah piksel RGB, mean RGB, dan standar deviasi RGB.

4. Klasifikasi

Pada tahap ini melakukan klasifikasi menggunakan metode K-NN. Cara kerja algoritma K-Nearest Neighbor (K-NN), yaitu:

1. Tentukan nilai jumlah tetangga (K) kedekatan objek ke objek data.
2. Hitung jarak setiap tetangga dari data, lalu urutkan jarak mulai dari minimum hingga maksimum.
3. Mengambil sejumlah K data dengan jarak terdekat, lalu diberikan label.

5. Evaluasi Hasil












Pada tahap ini dilakukan analisis hasil dari data yang sudah di klasifikasi untuk mendapatkan nilai akurasi yang di uji menggunakan GUI di matlab App Design.

3. Hasil dan Pembahasan

3.1 Pengumpulan Data









Pengumpulan data dilakukan dengan menggunakan objek buah kelapa sawit yang diambil dari petani sebelum di sortir ke pabrik menggunakan kamera smartphone yang diberikan alas dengan karton berwarna putih. Dataset tersebut terdiri dari tiga jenis, yaitu Dura, Tenera, dan Pisifera yang kemudian dikategorikan berdasarkan kematangan menjadi Mentah, Matang, dan Terlalu Matang. Data awal yang sudah diberikan label kematangan tersebut terdiri dari Mentah, Matang, dan Terlalu Matang dengan total jumlah 270 citra buah sawit. Berikut ini data awal pada tabel 1

Tabel 1. Data Awal

No	Gambar	Jenis	Tingkat Kematangan(klasifikasi)
1		Dura	Mentah
2		Dura	Mentah
3		Dura	Mentah
4		Dura	Mentah
5		Tenera	Mentah
6		Tenera	Mentah
7		Dura	Matang
...
267		Tenera	Terlalu Matang
268		Tenera	Terlalu Matang
269		Tenera	Terlalu Matang
270		Tenera	Terlalu Matang

Pada data awal tersebut terdapat 6 Mentah, 244 Matang dan 20 Terlalu Matang lalu di bagi menjadi 70 % data latih(train) dan 30% data uji (test). Data latih(train) digunakan untuk melatih model. Proses pelatihan ini melibatkan pencarian pola atau hubungan antara ciri-ciri atau fitur-fitur dalam data latih. Data uji(test) digunakan untuk menguji performa model setelah melalui proses pelatihan. Pada pembagian data tersebut juga dilakukan resize karena memiliki dengan ukuran besar. data citra tersebut di ubah ukurannya sebesar 500 × 500 piksel dengan mengubah ukuran gambar dapat membuat ukuran input yang konsisten. Pada tabel 2 terdapat citra data train .


Tabel 2. Data Train

No	Gambar	Jenis	Tingkat Kematangan (Klasifikasi)
1		Dura	Mentah
2		Dura	Mentah
...
184		Pisifera	Matang
185		Pisifera	Matang
186		Pisifera	Matang
187		Tenera	Terlalu Matang
188		Tenera	Terlalu Matang
189		Tenera	Terlalu Matang

Pada tabel 2. Data train terdiri dari 171 Matang, 4 Mentah dan 14 Terlalu Matang dengan total data train keseluruhan tersebut adalah 189 data latih(train). pada tabel 5.3 data test dibawah ini terdiri dari 73 Matang, 2 Mentah dan 6 Terlalu Matang dengan total data test keseluruhan adalah 81 data uji(test).

Tabel 3. Data Test

No	Gambar	Jenis	Tingkat Kematangan (Klasifikasi)
1		Dura	Mentah
2		Dura	Mentah
3		Dura	Matang
4		Dura	Matang
5		Dura	Matang
...
80		Tenera	Terlalu Matang

No	Gambar	Jenis	Tingkat Kematangan (Klasifikasi)
81		Tenera	Terlalu Matang

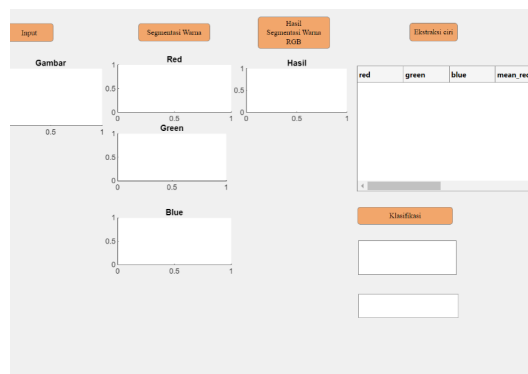
A. Preprocessing

Preprocessing adalah langkah dalam pengolahan data sebelum diterapkan pada model K-NN. Pengolahan data klasifikasi kematangan buah kelapa sawit berdasarkan warna menggunakan metode K-Nearest Neighbor (KNN) ini melibatkan serangkaian langkah penting untuk mengidentifikasi dan mengklasifikasikan buah kelapa sawit ke dalam tingkat kematangan berdasarkan ciri-ciri atau fitur-fitur warna yang diekstraksi dari citra. Preprocessing yang efektif dan tepat akan meningkatkan akurasi. Preprocessing pada saat melatih model dari data latih (train) adalah serangkaian langkah atau teknik yang diterapkan pada data sebelum data tersebut digunakan untuk melatih model.

Setelah preprocessing dibuat model pelatihan dari data train menggunakan teknik validasi silang dengan 5 lipatan (*5-fold cross-validation*). validasi silang adalah Teknik membantu dalam mendeteksi masalah seperti overfitting dengan cara membagi data latih(train) menjadi lima subset atau lipatan (*folds*). Data train dibagi menjadi 5 lipatan yang kurang lebih sama besar. Proses ini dilakukan sebanyak 5 kali (iterasi). Pada setiap iterasi, satu lipatan digunakan sebagai data validasi sementara empat lipatan lainnya digunakan sebagai data pelatihan. Validasi silang dengan 5 lipatan memberikan gambaran yang lebih akurat mengenai kinerja model pada data yang belum pernah dilihat sebelumnya.

B. Proses Klasifikasi

Proses pada hasil program klasifikasi kematangan buah kelapa sawit berdasarkan warna menggunakan metode K-Nearest Neighbor ini dilakukan dengan cara input citra yang di ambil dari data test dengan yang sudah lalu akan di segmentasi warna di tampilkan Dalam model warna RGB, setiap piksel direpresentasikan oleh tiga komponen warna: merah (R), hijau (G), dan biru (B). pada gambar 5.1 adalah tampilan awal pengujian program :

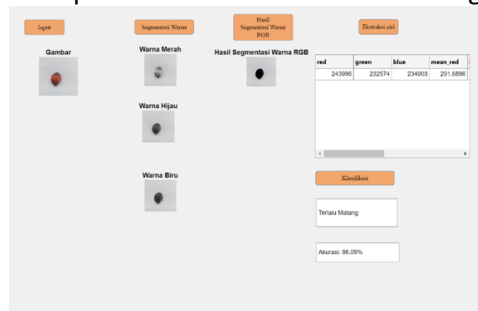


Gambar 1. Tampilan Awal

Setelah melakukan segmentasi warna di tiga sumbu berbeda (R, G, B) dalam citra buah kelapa sawit, dapat menggabungkan informasi karnal warna RGB menjadi satu dapat memvisualisasikan hasilnya dalam satu sumbu atau dimensi untuk melihat bagaimana piksel-piksel dalam citra telah dikelompokkan berdasarkan kriteria warna yang telah ditentukan. satu dimensi berdasarkan skala warna atau nilai tertentu yang menunjukkan tingkat kematangan.

Selanjutnya di ekstraksi ciri untuk proses mengidentifikasi ciri-ciri atau fitur-fitur penting dari setiap piksel yang akan digunakan untuk membedakan tingkat kematangan. Dalam konteks ini, ciri-ciri atau fitur-fitur ini berdasarkan nilai-nilai RGB dari piksel. Ekstraksi ciri berdasarkan warna ini untuk proses mengambil informasi penting dari citra yang merepresentasikan karakteristik warna dari suatu objek. pada penelitian ini terdapat 9 ekstraksi ciri atau fitur jumlah piksel Red, jumlah piksel Green, jumlah piksel Blue, mean Red, mean Green, mean Blue dan standar Red, standar Green, standar Blue yang ditampilkan pada tabel ekstraksi ciri seperti gambar 5.2.

Metode seperti KNN digunakan untuk mengklasifikasikan setiap piksel dalam citra ke salah satu dari kategori kematangan di atas berdasarkan ciri-ciri atau fitur warna RGB yang diekstraksi sebelumnya lalu mengambil keputusan berdasarkan jarak terdekat dari tetangga terdekat dalam ruang fitur RGB. berikut ini pada gambar 5.2 tampilan pengujian program klasifikasi kematangan buah kelapa sawit berdasarkan warna menggunakan metode K-NN :



Gambar 2. Tampilan akhir Pengujian

C. Evaluasi Hasil Akurasi

Akurasi adalah ukuran utama untuk mengevaluasi seberapa baik model dapat mengklasifikasikan kematangan buah kelapa sawit dari citra berdasarkan warna.

```

1 % Baca gambar
2 img = imread('databuah\train\Matang\psiferai.jpg');
3
4 % Tampilkan gambar
5 imshow(img);
6
7 % Ambil nilai RGB
8 [height, width, ~] = size(img);
9 centerX = round(width / 2);
10 centerY = round(height / 2);
11
12 r = img(centerY, centerX, 1); % Nilai merah
13 g = img(centerY, centerX, 2); % Nilai hijau
14 b = img(centerY, centerX, 3); % Nilai biru
15
16 % Menampilkan nilai RGB
17 fprintf('Nilai RGB : (%d, %d, %d)\n', r, g, b);

```

Command Window

```

Nilai RGB : (221, 133, 109)
>> rgb6
Nilai RGB : (166, 87, 74)
>> rgb6
Nilai RGB : (153, 121, 108)
>> rgb6
Nilai RGB : (111, 27, 25)

```

Gambar 1. Kode nilai RGB data citra latih

Pada gambar 3 terdapat kode yang untuk mendapatkan nilai RGB pada citra data latih(train). Kode ini berfungsi untuk membaca setiap gambar, mengidentifikasi nilai-nilai intensitas warna merah (R), hijau (G), dan biru (B).

```

1 % Baca gambar
2 img = imread('databuah\test\Matang\tenera26.jpg');
3
4 % Tampilkan gambar
5 imshow(img);
6
7 % Ambil nilai RGB
8 [height, width, ~] = size(img);
9 centerX = round(width / 2);
10 centerY = round(height / 2);
11
12 r = img(centerY, centerX, 1); % Nilai merah
13 g = img(centerY, centerX, 2); % Nilai hijau
14 b = img(centerY, centerX, 3); % Nilai biru
15
16 % Menampilkan nilai RGB
17 fprintf('Nilai RGB : (%d, %d, %d)\n', r, g, b);

```

Command Window

```

Nilai RGB : (186, 73, 41)
>> rgb6
Nilai RGB : (130, 74, 73)
>> rgb6
Nilai RGB : (196, 87, 58)
>> rgb6
Nilai RGB : (129, 49, 40)

```

Gambar 2. Kode Nilai RGB data citra test

Pada gambar 4 tersebut kode untuk mendapatkan nilai RGB dari data citra uji(test). Nilai-nilai RGB yang diperoleh ini kemudian digunakan dalam perhitungan manual untuk mengukur jarak Euclidean antara citra data uji dan citra data latih.

Berikut ini langkah-langkah pada perhitungan manual K-Nearest Neighbors (K-NN) dan hasil akurasi berdasarkan sampel 7 data citra latih dan 5 data uji ,yaitu :

1. Langkah pertama :Tentukan Nilai RGB Data Latih dan Data uji
Terdiri dari 7 sampel citra data latih dengan label kematangan buah:

Citra 1 (Matang): RGB = (221, 133, 109)
 Citra 2 (Mentah): RGB = (166, 87, 74)
 Citra 3 (Terlalu Matang): RGB = (153, 121, 108)
 Citra 4 (Matang): RGB = (111, 27, 25)
 Citra 5 (Mentah): RGB = (127, 83, 722)
 Citra 6 (Terlalu Matang): RGB = (205, 109, 58)
 Citra 7 (Mentah): RGB = (104, 65, 58)

Terdiri dari 5 sampel citra data uji dengan label kematangan buah:

Citra Uji 1(Matang): RGB = (186, 73, 41)
 Citra Uji 2 (Mentah): RGB = (130, 74, 73)
 Citra Uji 3(Terlalu Matang): RGB = (196,87, 58)
 Citra Uji 4(Matang): RGB = (129, 49, 40)
 Citra Uji 5(Mentah): RGB = (191, 101, 77)

2. Langkah Kedua : Hitung Jarak Euclidean

Perhitungan jarak Euclidean dari setiap citra uji ke setiap citra latih. Rumus jarak Euclidean adalah:

$$d(i, j) = \sqrt{|x_{i1} - x_{j1}|^2 + (|x_{i2} - x_{j2}|)^2 + \dots + |x_{ip} - x_{jp}|^2}$$

1. Citra Uji 1 (RGB = 186, 73, 41):

1. Jarak ke Data Latih 1 (Matang):

$$\sqrt{(186 - 221)^2 + (73 - 133)^2 + (41 - 109)^2}$$

$$= \sqrt{-35^2 + (-60)^2 + (-68)^2} = 97,2$$

2. Jarak ke Data Latih 2 (Mentah):

$$\sqrt{(186 - 166)^2 + (73 - 87)^2 + (41 - 74)^2}$$

$$= \sqrt{20^2 + (-14)^2 + (-33)^2} = 41,04$$

3. Jarak ke Data Latih 3 (Terlalu Matang):

$$\sqrt{(186 - 153)^2 + (73 - 121)^2 + (41 - 108)^2}$$

$$= \sqrt{33^2 + (-48)^2 + (-67)^2} = 88,7$$

4. Jarak ke Data Latih 4 (Matang):

$$\sqrt{(186 - 111)^2 + (73 - 27)^2 + (41 - 25)^2}$$

$$= \sqrt{75^2 + 46^2 + 16^2} = 89,4$$

5. Jarak ke Data Latih 5 (Mentah):

$$\sqrt{(186 - 127)^2 + (73 - 83)^2 + (41 - 72)^2}$$

$$= \sqrt{59^2 + (-10)^2 + (-31)^2} = 67,3$$

6. Jarak ke Data Latih 6 (Terlalu Matang):

$$\sqrt{(186 - 205)^2 + (73 - 109)^2 + (41 - 58)^2}$$

$$= \sqrt{(-19)^2 + (-36)^2 + (-17)^2} = 44,1$$

7. Jarak ke Data Latih 7 (Mentah):

$$\sqrt{(186 - 104)^2 + (73 - 65)^2 + (41 - 58)^2}$$

$$= \sqrt{82^2 + 8^2 + (-17)^2} = 84,1$$

2. Citra Uji 2 (RGB = 130, 74, 73):

1. Jarak ke Data Latih 1 (Matang):

$$\sqrt{(130 - 221)^2 + (74 - 133)^2 + (73 - 109)^2}$$

$$= \sqrt{(-91)^2 + (-59)^2 + (-36)^2} = 114,2$$

2. Jarak ke Data Latih 2 (Mentah):

$$\sqrt{(130 - 166)^2 + (74 - 87)^2 + (73 - 74)^2}$$

$$= \sqrt{(-36)^2 + (-13)^2 + (-1)^2} = 38,2$$

3. Jarak ke Data Latih 3 (Terlalu Matang):

$$\sqrt{(130 - 153)^2 + (74 - 121)^2 + (73 - 108)^2}$$

$$= \sqrt{(-23)^2 + (-47)^2 + (-35)^2} = 62,9$$

4. Jarak ke Data Latih 4 (Matang):

$$\sqrt{(130 - 111)^2 + (74 - 27)^2 + (73 - 25)^2}$$

$$= \sqrt{19^2 + 47^2 + 48^2} = 69,8$$

5. Jarak ke Data Latih 5 (Mentah):

$$\sqrt{(130 - 127)^2 + (74 - 83)^2 + (73 - 72)^2}$$

$$= \sqrt{3^2 + (-9)^2 + 1^2} = 9,5$$

6. Jarak ke Data Latih 6 (Terlalu Matang):

$$\sqrt{(130 - 205)^2 + (74 - 109)^2 + (73 - 58)^2}$$

$$= \sqrt{(-75)^2 + (-35)^2 + 15^2} = 84,1$$

7. Jarak ke Data Latih 7 (Mentah):

$$\sqrt{(130 - 104)^2 + (74 - 65)^2 + (73 - 58)^2}$$

$$= \sqrt{26^2 + 9^2 + 15^2} = 31,3$$

3. Citra Uji 3 (RGB = 196,87, 58):

1. Jarak ke Data Latih 1 (Matang):

$$\sqrt{(196 - 221)^2 + (87 - 133)^2 + (58 - 109)^2}$$

$$= \sqrt{-25^2 + (-46)^2 + (-51)^2} = 73,08$$

2. Jarak ke Data Latih 2 (Mentah):

$$\sqrt{(196 - 166)^2 + (87 - 87)^2 + (58 - 74)^2}$$

$$= \sqrt{30^2 + 0^2 + (-16)^2} = 34$$

3. Jarak ke Data Latih 3 (Terlalu Matang):

$$\sqrt{(196 - 153)^2 + (87 - 121)^2 + (58 - 108)^2}$$

$$= \sqrt{43^2 + (-34)^2 + (-50)^2} = 74,19$$

4. Jarak ke Data Latih 4 (Matang):

$$\sqrt{(196 - 111)^2 + (87 - 27)^2 + (58 - 25)^2}$$

$$= \sqrt{85^2 + 60^2 + 33^2} = 109,1$$

5. Jarak ke Data Latih 5 (Mentah):

$$\sqrt{(196 - 127)^2 + (87 - 83)^2 + (58 - 72)^2}$$

$$= \sqrt{69^2 + 4^2 + (-14)^2} = 70,5$$

6. Jarak ke Data Latih 6 (Terlalu Matang):

$$\sqrt{(196 - 205)^2 + (87 - 109)^2 + (58 - 58)^2}$$

$$= \sqrt{(-9)^2 + (-22)^2 + 0^2} = 23,7$$

7. Jarak ke Data Latih 7 (Mentah):

$$\sqrt{(196 - 104)^2 + (87 - 65)^2 + (58 - 58)^2}$$

$$= \sqrt{92^2 + 22^2 + 0^2} = 94,5$$

4 Citra Uji 4 (RGB = 129, 49, 40):

1. Jarak ke Data Latih 1 (Matang):

$$\sqrt{(129 - 221)^2 + (49 - 133)^2 + (40 - 109)^2}$$

$$= \sqrt{(-92)^2 + (-84)^2 + (-69)^2} = 142,4$$

2. Jarak ke Data Latih 2 (Mentah):

$$\sqrt{(129 - 166)^2 + (49 - 87)^2 + (40 - 74)^2}$$

$$= \sqrt{(-37)^2 + (-38)^2 + (-34)^2} = 63$$

3. Jarak ke Data Latih 3 (Terlalu Matang):

$$\sqrt{(129 - 153)^2 + (49 - 121)^2 + (40 - 108)^2}$$

$$= \sqrt{(-23)^2 + (-91)^2 + (-104)^2} = 140,09$$

4. Jarak ke Data Latih 4 (Matang):

$$\sqrt{(129 - 111)^2 + (49 - 27)^2 + (40 - 25)^2}$$

$$= \sqrt{18^2 + 22^2 + 15^2} = 32,14$$

5. Jarak ke Data Latih 5 (Mentah):

$$\sqrt{(129 - 127)^2 + (49 - 83)^2 + (40 - 72)^2}$$

$$= \sqrt{2^2 + (-34)^2 + (-32)^2} = 46,73$$

6. Jarak ke Data Latih 6 (Terlalu Matang):

$$\sqrt{(129 - 205)^2 + (49 - 109)^2 + (40 - 58)^2}$$

$$= \sqrt{(-76)^2 + (-60)^2 + (-18)^2} = 98,4$$

7. Jarak ke Data Latih 7 (Mentah):

$$\sqrt{(129 - 104)^2 + (49 - 65)^2 + (40 - 58)^2}$$

$$= \sqrt{25^2 + (-16)^2 + (-18)^2} = 34,7$$

5 Citra Uji 5 (RGB = 191, 101, 77):

1. Jarak ke Data Latih 1 (Matang):

$$\sqrt{(191 - 221)^2 + (101 - 133)^2 + (77 - 109)^2}$$

$$= \sqrt{(-30)^2 + (-32)^2 + (-32)^2} = 54,29$$

2. Jarak ke Data Latih 2 (Mentah):

$$\sqrt{(191 - 166)^2 + (101 - 87)^2 + (77 - 74)^2}$$

$$= \sqrt{25^2 + 14^2 + 3^2} = 28,8$$

3. Jarak ke Data Latih 3 (Terlalu Matang):

$$\sqrt{(191 - 153)^2 + (101 - 121)^2 + (77 - 108)^2}$$

$$= \sqrt{38^2 + (-20)^2 + (-31)^2} = 52,96$$

4. Jarak ke Data Latih 4 (Matang):

$$\sqrt{(191 - 111)^2 + (101 - 27)^2 + (77 - 25)^2}$$

$$= \sqrt{80^2 + 74^2 + 52^2} = 120,7$$

5. Jarak ke Data Latih 5 (Mentah):

$$\sqrt{(191 - 127)^2 + (101 - 83)^2 + (77 - 72)^2}$$

$$= \sqrt{64^2 + 18^2 + 5^2} = 66,6$$

6. Jarak ke Data Latih 6 (Terlalu Matang):

$$\sqrt{(191 - 205)^2 + (101 - 109)^2 + (77 - 58)^2}$$

$$= \sqrt{(-14)^2 + (-8)^2 + 19^2} = 24,91$$

7. Jarak ke Data Latih 7 (Mentah):

$$\sqrt{(191 - 104)^2 + (101 - 65)^2 + (77 - 58)^2}$$

$$= \sqrt{87^2 + 36^2 + 19^2} = 96,05$$

1. Langkah Ke tiga : Tentukan Kelas Berdasarkan Jarak Terdekat

1. Jarak Citra uji 1: 41,04

Prediksi : Benar

2. Jarak Citra uji 2 : 31,3

Prediksi : Benar

3. Jarak Citra uji 3 : 23,7

Prediksi : Benar

4. Jarak Citra uji 4: 32,14

Prediksi : Benar

5. Jarak Citra uji 5: 24,91

Prediksi : Salah

2. Langkah 4: Menghitung Akurasi

$$Akurasi = \left(\frac{\text{Jumlah klasifikasi Benar}}{\text{jumlah data Uji}} \right) \times 100\%$$

$$= \left(\frac{4}{5} \right) \times 100\% = 80\%$$

Pada perhitungan manual tersebut yang menggunakan sampel 7 data train dan 5 data test hasil akurasi yang didapatkan adalah 80%. Hal ini menunjukkan bahwa dari 5 citra yang diuji, 4 citra berhasil diklasifikasikan dengan benar sesuai label yang sebenarnya. Langkah-langkah perhitungan ini melibatkan pengukuran jarak Euclidean antara nilai RGB dari citra-citra uji dengan citra-citra latih, serta penentuan kelas berdasarkan mayoritas dari K tetangga terdekat.

```

70         actualLabel = '3';
71     end
72     % Hitung akurasi
73     if predictedLabel == actualLabel
74         numCorrect = numCorrect + 1;
75     end
76 end
77 % Hitung akurasi akhir
78 accuracy = (numCorrect / numTotal) * 100;
79 % Tampilkan hasil perhitungan manual
80 fprintf('Jumlah prediksi benar: %d\n', numCorrect);
81 fprintf('Jumlah total data uji: %d\n', numTotal);
82 fprintf('Akurasi Total: %.2f%%\n', accuracy);
83
Command Window
Trial License -- for use to evaluate programs for possible purchase as an end-user only.
>> knn_script2
Jumlah prediksi benar: 72
Jumlah total data uji: 81
Akurasi Total: 88.89%
    
```

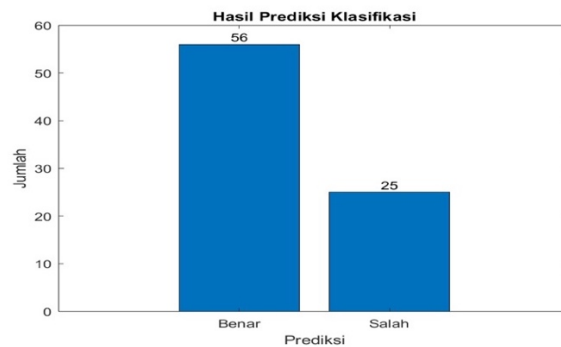
Gambar 3. Hasil akurasi perhitungan manual di program

Pada gambar 5 terdapat kode program untuk perhitungan manual K-NN tersebut mendapatkan akurasi 88,89% dengan jumlah prediksi benar 72 dan total data uji 81. Kode tersebut menghitung jarak Euclidean antara nilai RGB dari citra-citra uji dan data latih, serta menentukan prediksi. Berikut ini tabel 5.4 perbandingan hasil akurasi penelitian ini dengan penelitian terdahulu.

Tabel 4. Perbandingan hasil akurasi

Nilai K (K-NN)	K=3	K=5	K=7	K=9	K=11	K=13	K=15	K=17	K=19	K=21	K=23	K=25
Penelitian terdahulu(Rifqi et al., 2020)	86%	88%	92%	90%	88%	88%	86%	82%	76%	76%	76%	76%
Penelitian sekarang	93,67%	91,00%	90,27%	90,50%	86,37%	87,10%	82,20%	85,64%	86,30%	82,00%	80,54%	80,05%

Pada pelitian pada tabel 4 perbandingan hasil akurasi penelitian terdahulu dan sekarang terdapat hasil akurasi tertinggi pada tabel tersebut adalah 93,67% di K=3 pada buah jenis dura yang mentah sedangkan pada penelitian terdahulu hasil akurasi tertinggi adalah 92% di k=7 pada TBS. Hasil akurasi terendah pada tabel tersebut terletak pada nilai K=17 sampai K= 25 pada penelitian terdahulu adalah 76% untuk TBS sedangkan penelitian ini terletak di nilai K=25 adalah 80,05% pada buah jenis dura yang mentah.pada penelitian terdahulu meneliti tandan buah segar(TBS) dengan kondisi pengambilan gambar TBS yang diletak di dalam kotak yang memiliki akurasi tertinggi sedangkan penelitian ini meneliti buah kelapa sawit yang pengambilan gambar buahnya diluar yang dialas dengan karton berwarna putih dan hasil akurasi tertinggi dengan jenis buah dura yang mentah.



Gambar 6. Hasil Prediksi

Pada gambar 6 merupakan jumlah hasil prediksi klasifikasi kematangan dari data uji. Jumlah yang benar adalah 56 data citra dan yang salah 25 data citra dengan total keseluruhan data uji(test) adalah 81 data citra. Perbedaan antara jumlah data test(uji) klasifikasi yang sudah diberi label dengan hasil prediksi klasifikasi dapat disebabkan oleh beberapa faktor yang terjadi dalam konteks pengolahan citra dan pengklasifikasian menggunakan metode K-Nearest Neighbors (K-NN) seperti model mungkin mengalami overfitting atau underfitting, dimana model terlalu mengikuti data pelatihan atau tidak cukup menangkap pola dari data pelatihan.

D. Pembahasan

Pada penelitian ini, nilai K pada algoritma K-Nearest Neighbors (K-NN) berperan penting dalam menentukan performa model klasifikasi. Dengan K=3, model mempertimbangkan tiga tetangga terdekat untuk melakukan prediksi. Ini berarti keputusan prediksi sangat dipengaruhi oleh data poin yang sangat dekat dengan titik yang diuji, sehingga representasi lokal dari data menjadi sangat kuat. Nilai K yang lebih kecil, seperti K=3, cenderung menghasilkan akurasi yang lebih tinggi karena model lebih fokus pada informasi lokal yang relevan dan mengurangi pengaruh dari data poin yang jauh dan mungkin tidak relevan.

Sebaliknya, nilai K yang lebih besar, seperti K=25, bisa menghasilkan akurasi yang lebih rendah karena model terlalu menggeneralisasi dan memasukkan terlalu banyak tetangga yang mungkin berasal dari berbagai distribusi, mengaburkan batas antara kelas-kelas yang berbeda. Perbedaan tidak hanya karena nilai K (K-NN) bisa juga karena Dataset yang digunakan dalam kedua penelitian berbeda. Perbedaan dalam jumlah sampel, kualitas gambar, variasi kondisi pencahayaan, atau cara pengambilan data dapat mempengaruhi hasil akurasi.

Pada hasil prediksi klasifikasi Jumlah yang benar adalah 56 data citra dan yang salah 25 data citra dengan total keseluruhan data uji(test) adalah 81 data citra. terdapat Perbedaan antara jumlah data test(uji) klasifikasi yang sudah diberi label dengan hasil prediksi klasifikasi dapat disebabkan oleh cara ekstraksi dan representasi ciri atau fitur dari data gambar yang mempengaruhi kemampuan model untuk membedakan antara kematangan yang berbeda (Mentah, Matang, dan Terlalu Matang).

3. Kesimpulan

Berdasarkan analisis dan hasil pengujian yang telah dilakukan, dapat disimpulkan bahwa metode K-Nearest Neighbor (K-NN) dengan citra warna RGB dapat diterapkan untuk klasifikasi tingkat kematangan buah kelapa sawit. Penelitian ini juga berhasil membangun aplikasi klasifikasi menggunakan MATLAB App Designer (GUI) sehingga proses identifikasi kematangan dapat dilakukan secara lebih praktis melalui antarmuka yang mudah digunakan.

Hasil pengujian menunjukkan akurasi tertinggi sebesar 93,67% pada nilai K=3 untuk buah kelapa sawit mentah jenis dura, sedangkan akurasi terendah sebesar 80,05% pada nilai K=25 untuk buah kelapa sawit mentah jenis dura. Jika dibandingkan dengan penelitian terdahulu pada TBS, akurasi tertinggi yang diperoleh adalah 92% pada K=7 dan akurasi terendah pada rentang K=17 hingga K=25 adalah 76%. Selain itu, pengujian perhitungan manual K-NN pada program menghasilkan akurasi 88,89% dengan total 81 data uji, sedangkan perhitungan manual pada 7 data citra latih dan 5 data uji menghasilkan akurasi 80%. Untuk pengembangan selanjutnya, aplikasi pengolahan citra dapat diuji pada objek lain, serta peneliti berikutnya disarankan mencoba metode berbeda agar dapat dibandingkan dan berpotensi meningkatkan akurasi.

Daftar Pustaka

- [1] S. R. Cholil, T. Handayani, R. Prathivi, and T. Ardianita, "Implementasi Algoritma Klasifikasi K-Nearest Neighbor (KNN) untuk Klasifikasi Seleksi Penerima Beasiswa," *IJCIT (Indonesian Journal on Computer and Information Technology)*, vol. 6, no. 2, 2021.
-

-
- [2] L. Farokhah, "Implementasi K-Nearest Neighbor untuk Klasifikasi Bunga dengan Ekstraksi Fitur Warna RGB," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 7, no. 6, pp. 1129–1136, 2020, doi: 10.25126/jtiik.2020722608.
- [3] A. S. Febrianti, T. A. Sardjono, and A. F. Babgei, "Klasifikasi Tumor Otak pada Citra Magnetic Resonance Image dengan Menggunakan Metode Support Vector Machine," *Jurnal Teknik ITS*, vol. 9, no. 1, 2020, doi: 10.12962/j23373539.v9i1.51587.
- [4] M. Meiriyama, S. Devella, and S. M. Adelfi, "Klasifikasi Daun Herbal Berdasarkan Fitur Bentuk dan Tekstur Menggunakan KNN," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 9, no. 3, pp. 2573–2584, 2022, doi: 10.35957/jatisi.v9i3.2974.
- [5] M. A. Nasution, H. N. Winata, F. Q. Nasution, H. Lydiasari, R. Y. Pasaribu, A. Nasution, and A. Wulandari, "Determinasi Nilai RGB dan Grayscale pada Citra Tandan Buah Segar (TBS) Kelapa Sawit (*Elaeis guineensis* Jacq.) Menggunakan Matlab," *Jurnal Penelitian Kelapa Sawit*, no. 1, pp. 37–48, 2022.
- [6] B. Rahman, F. Fauzi, and S. Amri, "Perbandingan Hasil Klasifikasi Data Iris menggunakan Algoritma K-Nearest Neighbor dan Random Forest," *Journal of Data Insights*, vol. 1, no. 1, pp. 19–26, 2023, doi: 10.26714/jodi.v1i1.135.
- [7] M. Rifqi, M. Akbar, and D. Y. Fitriasia, "Aplikasi Pendeteksian Kematangan Tandan Buah Segar (TBS) Kelapa Sawit Berdasarkan Komposisi Warna Menggunakan Algoritma K-NN," *Jurnal Komputer Terapan*, vol. 6, no. 1, 2020. [Online]. Available: <https://jurnal.pcr.ac.id/index.php/jkt/>
- [8] A. Ritonga, A. A. Munawar, and I. S. Nasution, "(Judul tidak dicantumkan)," vol. 8, pp. 437–445, 2023. ← (data sumber belum lengkap: nama jurnal/prosiding & judul)
- [9] Rosidin, "Citra Digital Menggunakan Matlab: Analisis Pendeteksi Kecocokan Objek pada Citra Digital Menggunakan Matlab dengan Metode Algoritma SIFT," 2018. ← (lengkapi penerbit/jenis dokumen bila ada)
- [10] M. R. Setiawan, Y. A. Sari, and P. P. Adikara, "Klasifikasi Citra Makanan Menggunakan K-Nearest Neighbor dengan Fitur Bentuk Simple Morphological Shape Descriptors dan Fitur Warna Grayscale Histogram," vol. 3, no. 3, pp. 2726–2731, 2019. ← (lengkapi nama jurnal/prosiding)
- [11] S. Raysyah, V. Arinal, and D. I. M., "Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode KNN dan PCA," vol. 8, no. 2, pp. 88–95, 2021. ← (lengkapi nama jurnal/prosiding)
- [12] I. M. A. Swantika, B. Kanata, and I. M. B. Suksmadana, "Perancangan Sistem untuk Mengetahui Kualitas Biji Kopi Berdasarkan Warna dengan K-Nearest Neighbor," vol. 1, no. 2, pp. 25–36, 2020. ← (lengkapi nama jurnal/prosiding)