

# KLASIFIKASI TINGKAT KEMATANGAN BUAH PEPAYA CALIFORNIA DALAM RUANG WARNA HSV (HUE SATURATION VALUE) DENGAN ALGORITMA K-NEAREST NEIGHBORS

A Isatul Masruroh<sup>1</sup>, Sorikhi<sup>2</sup>, Achmad Syauqi<sup>3</sup>

<sup>1,2</sup>Program Studi Informatika, Universitas Peradaban, Indonesia

<sup>3</sup>Program Studi Sistem Informasi, Universitas Peradaban, Indonesia

## Article Info

### Article history:

Received Februari 2, 2023

Accepted Februari 10, 2023

Published Februari 24, 2023

### Keywords:

Pepaya, HSV, K- Nearest Neighbor (K-NN), MATLAB.

*Papaya, HSV, K- Nearest Neighbor (K-NN), MATLAB*

## ABSTRACT

Buah pepaya banyak diminati oleh masyarakat di dalam maupun di luar negeri, sehingga membuktikan bahwa produk pertanian yang satu ini sudah menjadi kebutuhan global yang banyak diminati dan dicari. Untuk penentuan panen buah pepaya berdasarkan warna kulit buah, kematangan buah pepaya dimulai dari mentah, mengkal (setengah matang) dan matang sehingga peneliti mengajukan sebuah ide untuk menjawab permasalahan pada penentuan kematangan buah pepaya yang sebagian besar masih dilakukan secara manual masih memiliki beberapa kelemahan dan membutuhkan proses yang cukup lama, memiliki akurasi yang rendah dan tidak konsisten. Berdasarkan permasalahan tersebut dibuatlah sebuah sistem untuk mengklasifikasikan tingkat kematangan buah pepaya dengan memanfaatkan fitur warna HSV dengan menggunakan algoritma K- Nearest Neighbor (K-NN). Klasifikasi ini menggunakan pengolahan citra dengan memanfaatkan software MATLAB untuk pembuatan sistem klasifikasi dengan tiga kelas klasifikasi yaitu mentah, mengkal (setengah matang) dan matang. Klasifikasi yang dihasilkan dengan menggunakan algoritma K-Nearest Neighbor (K-NN) menunjukkan hasil akurasi sebesar 86,6667 %.

*Papaya fruit is in great demand by people at home and abroad, thus proving that this one agricultural product has become a global need that is in great demand and sought after. To determine the papaya harvest based on the color of the fruit skin, the ripeness of the papaya starts from unripe, unripe (half-ripe) and overripe so that the researchers put forward an idea to answer the problem in determining the ripeness of papaya fruit, which is mostly done manually, still has some weaknesses and requires the process is quite long, has low accuracy and is inconsistent. Based on these problems, a system was created to classify the ripeness level of papaya by utilizing the HSV color features using the K-Nearest Neighbor (K-NN) algorithm. This classification uses image processing by utilizing MATLAB software to create a classification system with three classification classes, namely raw, half cooked and cooked. The classification generated using the K-Nearest Neighbor (K-NN) algorithm shows an accuracy of 86.6667%.*



## Corresponding Author:

A Isatul Masruroh,

Program Studi Informatika, Universitas Peradaban,

Jl. Raya Pagojengan Km 3 Paguyangan, Brebes, Jawa Tengah 52276.

Email: \*nengaiz2909@gmail.com

## 1. PENDAHULUAN

Pepaya merupakan tanaman yang banyak dibudidayakan di Indonesia. Masyarakat Indonesia biasa menanam tanaman ini di pekarangan atau di sawah. Namun, pada umumnya masyarakat menanam tanaman ini hanya sebatas digunakan untuk memenuhi kebutuhan sayur atau buah dalam rumah tangga. Kompleksnya mengandung senyawa yang terkandung dalam biji menjadikan biji pepaya sebagai bahan yang dapat dimanfaatkan menjadi bahan olahan yang dapat dikonsumsi dan memberikan manfaat bagi kesehatan manusia. Biji pepaya dapat dimanfaatkan sebagai jus yang berkhasiat menurunkan kadar kolesterol tubuh dan sebagainya antioksidan melalui zat fitokimia yang dikandungnya yaitu meliputi flavonoid, saponin, dan tannin[1]. Bersama dengan Filipina, Thailand, dan Vietnam, Indonesia merupakan salah satu produsen buah-buahan tropis

terkemuka di ASEAN. Pisang, nanas, pepaya, dan mangga adalah beberapa ekspor Indonesia yang paling bernilai. Selain hasil ekspor Indonesia yang berkualitas tinggi, pepaya merupakan buah tropis yang bernilai ekonomi tinggi dan nilai gizi yang baik[1].

Penentuan tingkat kematangan buah pepaya didasarkan pada tekstur warna dari kulit buah pepaya tersebut. Analisis warna dalam pengolahan citra (image processing) memiliki beberapa model, termasuk Hue, Saturation, Value (HSV). Dalam bidang pengolahan citra, digunakan format citra berwarna yang dikenal dengan Hue, Saturation, dan Value (HSV). Hue, Saturation, dan Value merupakan elemen dari ruang warna HSV, dengan Hue mewakili warna asli seperti merah, ungu, kuning dan saturasi, atau membuat lebih hijau, lebih merah, dan seterusnya. Saturation juga dikenal sebagai kroma (chroma), mengacu pada tingkat kemurnian atau intensitas warna. Nilai (Value) suatu warna adalah tingkat pencahayaannya, dan dapat berkisar dari 0% hingga 100%. Misalnya, jika nilainya disetel ke 0, warnanya akan menjadi hitam, sedangkan peningkatan nilainya akan membuat warnanya lebih cerah dan menghadirkan nuansa baru[2].

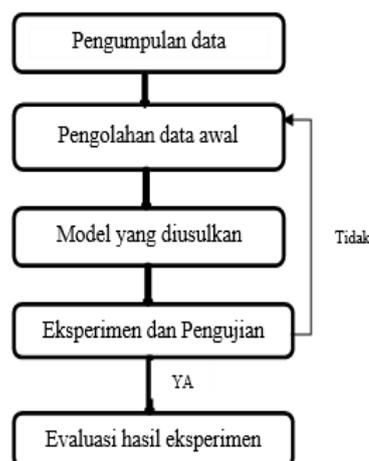
Selain warna, menganalisis tingkat kematangan buah pepaya adalah dengan cara mengidentifikasi tekstur. Tekstur adalah karakteristik penting untuk menganalisis permukaan berbagai jenis gambar dan pengelompokan gambar dalam kelas tertentu. Setiap gambar memiliki tekstur (unik) yang berbeda dengan gambar lain yang dapat menentukan komposisi struktur regional dan hubungan area sekitarnya. Model warna Hue / Saturation / Value (HSV) telah banyak digunakan dalam sistem pelacakan seperti sistem pelacakan mata dan sistem pelacakan jalur untuk mengemudi secara otonom di jalan yang tidak ditandai. Selain itu, kinerja model warna HSV lebih baik dalam klasifikasi tekstur warna jika dibandingkan dengan RGB[3].

Dalam mengklasifikasikan tingkat kematangan buah pepaya, metode K-Nearest Neighbor (K-NN) mengambil input dari hasil format citra warna menggunakan Hue, Saturation, Value (HSV). Mengklasifikasikan objek baru berdasarkan nilai (k) tetangga terdekatnya adalah tujuan dari algoritma K-Nearest Neighbor (K-NN). Sebagian besar kategori di K-NN digunakan untuk mengklasifikasikan hasil instans kueri baru. Ini karena K-NN adalah algoritma pembelajaran yang diawasi. Kategori kelas yang paling banyak muncul merupakan kelas yang dihasilkan dari proses klasifikasi. Metode Nearest Neighbor membandingkan bobot dari sekumpulan fitur yang sudah ada sebelumnya untuk menentukan seberapa mirip dua kasus [4].

Penelitian ini diuji berdasarkan pada tekstur warna kulit buah pepaya yang akan diekstrak menjadi fitur warna HSV (Hue, Saturation, Value) untuk dijadikan sebagai *input* dalam mengklasifikasikan tingkat kematangan buah pepaya menggunakan metode klasifikasi *K-Nearest Neighbor (K-NN)*.

## 2. METODE PENELITIAN

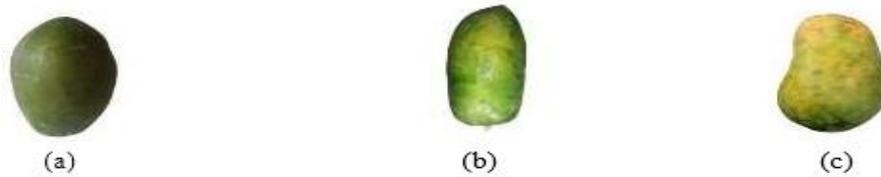
Pada penelitian ini penulis mengusulkan metode klasifikasi K-Nearest Neighbor untuk membedakan tingkat kematangan dari buah pepaya. Ini adalah situasi di mana algoritma K-Nearest Neighbor akan digunakan untuk memproses hasil pemrosesan baik data latih maupun data uji baru, dengan beberapa uji coba menggunakan nilai k yang berbeda untuk menentukan nilai akurasi terbaik. Pada Gambar 1 dapat dilihat langkah-langkah prosedur penelitian:



**Gambar 1.** Metode penelitian

### A. Pengumpulan Data

Data yang digunakan dalam penelitian ini adalah data citra dari buah pepaya california (Carica papaya L). Data diperoleh sebanyak 65 data citra yang terdiri dari 50 data training dan 15 data testing. Gambar pepaya diambil menggunakan Hp Oppo A37f dengan kamera belakang berukuran 8 Mp. Dataset ini akan diklasifikasikan ke dalam 3 kelas yang terdiri dari mentah, mengkal (setengah matang), dan matang yang ditunjukkan pada Gambar 2 sebagai berikut :



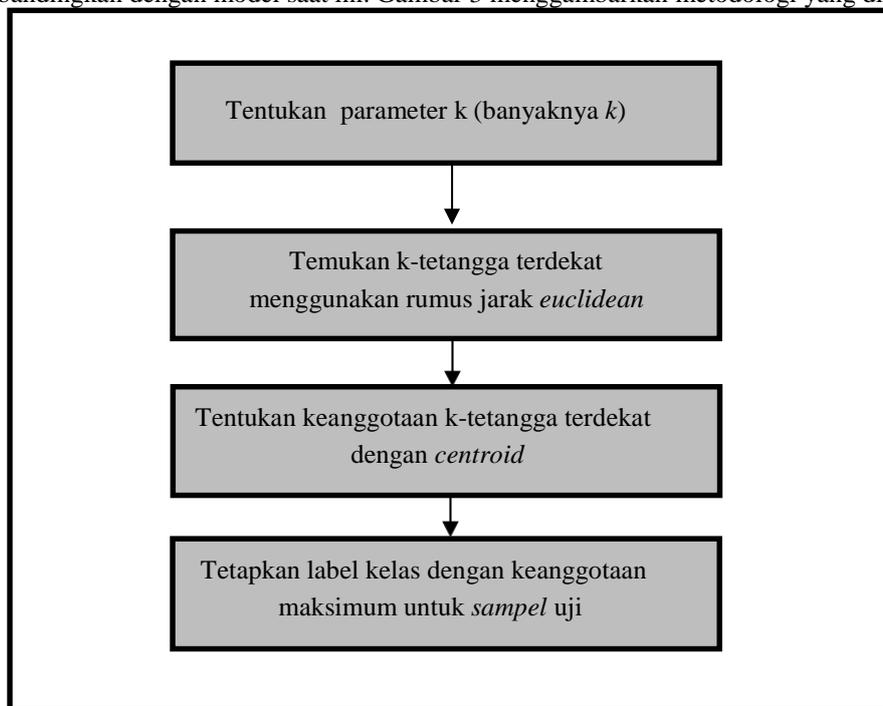
**Gambar 2.** Klasifikasi buah pepaya (a) Pepaya mentah, (b) Pepaya mengkal (setengah matang), (c) Pepaya matang

### B. Pengolahan Data Awal

Memisahkan data citra yang terkumpul menjadi data latih dan data uji merupakan langkah awal dalam pengolahan data. Sebelum mengklasifikasikan dengan metode K-NN, data yang telah dibagi akan diekstraksi untuk memisahkan background dari objek, dengan tujuan agar karakteristik warna yang dihasilkan dari setiap citra dapat digunakan sebagai tolak ukur untuk mempermudah proses klasifikasi.

### C. Metode yang Diusulkan

Bagian ini menguraikan fungsi model yang diusulkan dan menjelaskan alur metode yang diusulkan. Model berbasis data akan dijelaskan secara garis besar, dan hasil pemrosesan model di masa mendatang akan dievaluasi dibandingkan dengan model saat ini. Gambar 3 menggambarkan metodologi yang disarankan.



**Gambar 3.** Metode yang diusulkan

### D. Eksperimen dan Pengujian Model

Dalam tahapan eksperimen pada penelitian ini adalah:

1. Menyiapkan dataset citra buah pepaya untuk eksperimen.
2. Melakukan pengolahan data awal yaitu dengan membagi dataset yang ada kedalam data training dan data testing.
3. Ruang warna HSV (Hue, Saturation, Value) digunakan untuk mengekstraksi hasil dari data pelatihan dan data pengujian.
4. Algoritma K-NN (K-Nearest Neighbor) akan digunakan untuk mengklasifikasikan data hasil ekstraksi.

### E. Evaluasi dan Validasi

Setelah model dikembangkan, model tersebut akan dilakukan pengujian menggunakan *confusion matrix* untuk melihat seberapa baik kinerjanya. Akurasi direpresentasikan dalam *confusion matrix* sebagai rangkaian baris, yang masing-masing menggambarkan hasil dari prediksi positif yang benar, prediksi positif yang salah, prediksi negatif yang benar, dan prediksi negatif yang salah. Dalam hal ini akurasi akan ditentukan dengan menjumlahkan semua keberhasilan dan kegagalan prediksi yang dibuat, baik positif maupun negatif. Keunggulan model yang diusulkan meningkat sebanding dengan nilai akurasi yang dihasilkannya[5].

### 3. HASIL DAN PEMBAHASAN

#### A. Pengolahan Data

Penelitian ini digunakan data sebanyak 65 citra pepaya callifornia yang terdiri dari 50 data *training* dan 15 data *testing*. Citra pepaya yang didapatkan akan dibagi ke dalam 3 kelas klasifikasi yaitu kelas mentah, mengkal (setengah matang), dan matang. Data *training* dan data *testing* yang telah didapatkan akan diekstrak terlebih dahulu. Selanjutnya akan dilakukan uji data dengan menggunakan nilai  $k$  yang bervariasi sebelum akhirnya dilakukan proses klasifikasi data.

#### B. Klasifikasi dengan metode $K$ -NN

Hasil dari data yang telah diekstraksi akan di klasifikasi menggunakan metode  $K$ -NN dengan nilai ( $k$ ) = 3. Teknik  $K$ -NN digunakan untuk mengolah data dengan cara menentukan seberapa jauh perbedaan data uji dengan data latih. Baik label maupun atribut digunakan dalam penelitian ini, yang pertama berasal dari tingkat kematangan buah, sedangkan yang kedua diambil dari fitur yang diambil dari gambar. Tabel 1 digunakan pada tahap ini dalam proses untuk menentukan nilai ekstraksi fitur dari data pelatihan.

**Tabel 1.** Ekstraksi ciri data latih

No	Hasil Ekstraksi Ciri							Label
	Red	Green	Blue	Hue	Saturation	Value	Area	
1	119.7159	109.9773	25.46690	0.151975	0.793577	0.478046	2436649	Matang
2	193.5432	168.3648	38.57719	0.141509	0.801803	0.761489	2140936	Matang
3	169.8985	158.3684	64.73596	0.148794	0.640752	0.669310	2240659	Matang
4	145.2874	122.1724	56.1618	0.130479	0.636261	0.577902	2117500	Matang
5	125.5376	145.2929	68.2265	0.209844	0.553356	0.570794	2035052	Mengkal
6	99.92119	127.0789	51.2405	0.228803	0.625215	0.498693	2261690	Mengkal
7	115.9060	128.5615	74.3541	0.207862	0.432956	0.508744	1976325	Mengkal
8	123.3725	145.3880	46.6520	0.196375	0.695748	0.579120	1946699	Mengkal
9	91.4718	145.0510	78.8528	0.308263	0.487479	0.568851	2704240	Mentah
10	119.9828	162.5459	88.2319	0.267326	0.487373	0.637435	2838090	Mentah

Selanjutnya akan dihitung jarak terdekat dari data uji terhadap 50 data latih. Sebagai contoh akan diambil 1 data uji yang akan digunakan sebagai contoh perhitungan menggunakan rumus euclidean distance.. Hasil dari ekstraksi data uji ditampilkan pada Tabel 2.

**Tabel 2.** Ekstraksi ciri data uji 1

No	Hasil Ekstraksi Ciri							Label
	Red	Green	Blue	Hue	Saturation	Value	Area	
1	179.8245	131.9713	35.343	0.1157	0.8086	0.706	2221825	?

$K$ -NN menggunakan jarak terpendek yang ditentukan dengan rumus *Euclidean distance* untuk mengklasifikasikan data yang ditunjukkan pada Tabel 1 dan 2 setelah dilakukan perhitungan ekstraksi fitur. Pada penelitian ini terdapat 15 data uji dan dilakukan klasifikasi dengan nilai  $k = 3$ . Adapun hasil akhir yang didapatkan berdasarkan pengujian dari data uji terhadap data latih ditunjukkan pada Tabel 3.

**Tabel 3.** Hasil akhir pengujian

No	Data Uji	Kelas Asli	Hasil $K$ -NN	Keterangan
1	1	Matang	Matang	Akurat
2	2	Matang	Matang	Akurat
3	3	Matang	Matang	Akurat
4	4	Matang	Matang	Akurat
5	5	Matang	Matang	Akurat
6	6	Mengkal	Mengkal	Akurat
7	7	Mengkal	Mengkal	Akurat
8	8	Mengkal	Mengkal	Akurat
9	9	Mengkal	Mentah	Tidak Akurat
10	10	Mengkal	Mengkal	Akurat
11	11	Mentah	Mengkal	Tidak Akurat
12	12	Mentah	Mentah	Akurat
13	13	13	Mentah	Mentah

14	14	14	Mentah	Mentah
15	15	15	Mentah	Mentah

**C. Pengujian**

Tujuan dari pengujian ini adalah untuk mengevaluasi keefektifan algoritma K-NN dalam melakukan klasifikasi terhadap kategori yang telah ditetapkan sebelumnya. Pengujian akurasi dilakukan dengan menggunakan *confussion matrix*. Berikut perhitungan akurasinya menggunakan *confussion matrix*[6].

**Tabel 4.** Perhitungan *confussion matrix*

Aktual	Prediksi		
	Mentah	Mengkal	Matang
Mentah	3	1	0
Mengkal	1	4	0
Matang	0	0	6

$$A = \left( \frac{\text{Jumlah data benar}}{\text{jumlah seluruh data}} \right) \times 100\% = \dots\% \dots\dots\dots(1)$$

Maka didapatkan tingkat akurasi dari hasil pengujian sebagai berikut:

$$A = \left( \frac{13}{15} \right) \times 100\% = 86,6667\%$$

Jadi, berdasarkan pengujian *K-NN* di atas maka dapat diketahui bahwa pada *confussion matrix* menunjukkan hasil akurasi sebesar 86,6667 %.

**D. Implementasi**

Berdasarkan pengolahan data yang telah dilakukan dengan menggunakan metode K-NN, selanjutnya akan diimplementasikan untuk membuat program klasifikasi tingkat kematangan buah pepaya. Adapun desain tampilan dan tampilan programnya sebagai berikut.

**1) Tampilan Halaman Utama**



**Gambar 4.** Tampilan halaman utama program

Gambar 4. merupakan tampilan halaman utama yang terdapat tombol klasifikasi buah pepaya yang ketika diklik akan menampilkan halaman klasifikasi.

**2) Tampilan Halaman Klasifikasi**



**Gambar 5.** Tampilan halaman klasifikasi

Gambar 5. terdapat kolom *processing*, hasil ekstraksi citra dan ciri citra. Dimana, pada kolom *processing* berisi beberapa tombol yang akan digunakan untuk memproses data, kolom hasil ekstraksi citra

berfungsi untuk menampilkan hasil citra yang telah di ekstraksi sedangkan, kolom ciri citra berfungsi untuk menampilkan nilai yang diperoleh dari hasil ekstraksi citra.

### 3) Tampilan program pada klasifikasi K-NN



Gambar 6. Tampilan Program klasifikasi K-NN

## 4. KESIMPULAN

Hasil percobaan dan pengujian sebanyak 65 data citra buah pepaya california yang diklasifikasikan tingkat kematangannya menggunakan algoritma K-NN diperoleh akurasi sebesar 86,6667% dengan kriteria good classification menggunakan confusion matrix. Sebagai sarana untuk meningkatkan kinerja dan menyempurnakan penelitian yang ada, penulis memberikan beberapa saran yaitu:

- Untuk penelitian lebih lanjut disarankan untuk menambah data yang akan digunakan untuk diuji agar hasil pengujian yang dilakukan dapat lebih sempurna.
- Penelitian ini dapat dikembangkan dengan teknik klasifikasi berbeda maupun ekstraksi ciri dengan metode yang lain untuk dapat menghasilkan hasil prediksi yang lebih baik.
- Ekstraksi fitur HSV pada penelitian ini dapat digabungkan dengan model lain seperti GLCM. Karena, GLCM memiliki tingkat pengenalan yang tinggi berdasarkan nilai kontras, korelasi, homogenitas dan energi.

## REFERENSI

- [1] S. Sugiyanto and F. Wibowo, "Klasifikasi Tingkat Kematangan Buah Pepaya (Carica Papaya L) California (Callina-Ibp 9) Dalam Ruang Warna Hsv Dan Algoritma K-Nearest Neighbors," *Pros. Senat.*, vol. 1, no. November, 2015.
- [2] E. D. Nurcahya, "EKSTRAKSI FITUR PERTUMBUHAN PADI BERDASAR WARNA DAUN MENGGUNAKAN ANALISA RUANG WARNA HUE SATURATION VALUE," *MULTITEK Indones.*, vol. 13, no. 1, 2019, doi: 10.24269/mtkind.v13i1.1515.
- [3] Q. Al-Jubouri, R. J. Al-Azawi, M. Al-Tae, and I. Young, "Efficient individual identification of zebrafish using Hue/Saturation/Value color model," *Egypt. J. Aquat. Res.*, vol. 44, no. 4, 2018, doi: 10.1016/j.ejar.2018.11.006.
- [4] P. Cunningham and S. J. Delany, "K-Nearest Neighbour Classifiers-A Tutorial," *ACM Computing Surveys*, vol. 54, no. 6. 2021. doi: 10.1145/3459665.
- [5] O. R. Indriani, E. J. Kusuma, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, "Tomatoes classification using K-NN based on GLCM and HSV color space," in *Proceedings - 2017 International Conference on Innovative and Creative Information Technology: Computational Intelligence and IoT, ICITech 2017*, 2018, vol. 2018-January. doi: 10.1109/INNOCIT.2017.8319133.
- [6] H. Annur, "Klasifikasi Masyarakat Miskin Menggunakan Metode Naive Bayes," *Ilk. J. Ilm.*, vol. 10, no. 2, 2018, doi: 10.33096/ilkom.v10i2.303.160-165.