

# Klasifikasi Kematangan Sayuran Pare dengan Metode KNN

I Gede Yogananda Adi Baskara<sup>a1</sup>, Luh Arida Ayu Rahning Putri<sup>a2</sup>

<sup>a</sup>Program Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,  
Universitas Udayana  
Jalan Raya Kampus UNUD, Bukit Jimbaran, Kuta Selatan, Badung, Bali, Indonesia  
<sup>1</sup>baskara.2208561061@student.unud.ac.id  
<sup>2</sup>rahningputri@unud.ac.id

## Abstract

*The bitter melon plant (*Momordica charantia* L) is a vegetable commodity that has commercial potential if cultivated on an agribusiness scale. The bitter melon plant products currently have quite a lot of consumers and have even entered supermarkets. However, the selection of bitter melon vegetables still uses human eye assessment which has the weakness of being subjective and inconsistent, so the level of accuracy is low. Based on these problems, researchers will create a system that is able to classify the level of maturity of bitter melon vegetables using HSV feature extraction with the KNN method at the classification stage and with the help of the Python programming language. In this research, 160 datasets will be used which are divided into 3 types of classes, namely category cooked bitter melon vegetables and raw bitter melon vegetables. The dataset is divided into two ories, namely 128 training data and 32 test data. The next stage is testing the data using the K-Nearest Neighbor method using the value  $k=3$ . From the test results, an accuracy rate of 88% was obtained.*

**Keywords:** Python, Kematangan, Pare, KNN, HSV

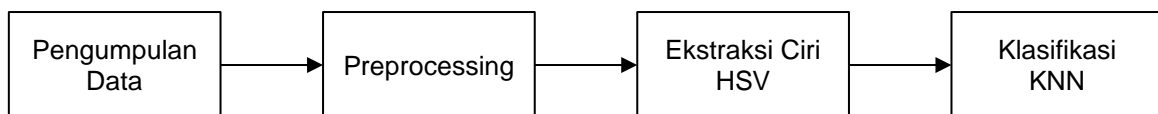
## 1. Pendahuluan

Tanaman pare merupakan sayuran yang berpotensi menguntungkan jika dibudidayakan secara besar-besaran untuk tujuan bisnis. Awalnya, tanaman merambat asal Asia ini hanya ditanam secara sederhana karena rendahnya permintaan pasar dan konsumen. Namun, seiring dengan banyaknya penelitian tentang tanaman pare dan munculnya varietas-varietas unggul dalam hal cita rasa dan penampilan, produk dari tanaman pare kini sudah cukup laku dan bahkan dijual di supermarket besar [1]. Biasanya pare diolah dengan cara digoreng, direbus atau dikukus, diasinkan, dijadikan jus dan dikeringkan menjadi teh. Pare mengandung senyawa fenolik yang berpotensi sebagai antioksidan. Kualitas produk olahan pare sangat dipengaruhi oleh tingkat kematangan pare itu sendiri. Pare berbentuk seperti mentimun dan berwarna hijau ketika masih muda, namun berubah menjadi kuning saat sudah matang. Masyarakat cenderung mengonsumsi pare yang belum matang. Menilai kematangan pare hanya dari penampilan luarnya tidaklah mudah, namun bisa dilakukan dengan melihat warna keseluruhan buah pare tersebut [2]. Namun, pemilihan sayuran pare masih menggunakan penilaian mata manusia memiliki kelemahan yaitu bersifat subjektif dan tidak konsisten sehingga tingkat akurasi rendah. Sehingga, dibutuhkan metode otomatis yang dapat meningkatkan akurasi dengan penilaian yang konsisten dalam klasifikasi tingkat kematangan sayuran pare berdasarkan fitur warna. Salah satu teknik pengolahan warna citra untuk mengklasifikasi kematangan suatu objek dengan menggunakan nilai HVS (*Hue, Saturation, Value*). Nilai HVS didapatkan dari mengkonversi nilai RGB citra. Sistem HVS dipandang lebih dekat daripada sistem RBG dalam mendeskripsikan sensasi warna oleh mata manusia sehingga perbedaan warna akan semakin terlihat dan meningkatkan akurasi dari metode klasifikasi. Salah satu metode yang dapat digunakan dalam proses klasifikasi adalah k-nearest neighbor (K-NN). Metode K-NN merupakan algoritma yang melakukan klasifikasi berdasarkan kedekatan posisi suatu data dengan data-data lainnya, artinya jarak terdekat antara data yang dievaluasi dengan k tetangga terdekatnya pada data pelatihan [3]. Penelitian dengan menggunakan metode klasifikasi K-NN sudah dilakukan

sebelumnya. Objek penelitian berupa buah pinang dengan mengklasifikasi menjadi tiga kelas yaitu mentah, matang, dan tua. Penelitian ini menggunakan metode K-Nearest Neighbor (K-NN) dengan data sebanyak 600 menggunakan parameter RGB. Didapatkan akurasi terbaik dengan akurasi 98,33% dengan K terbaik adalah  $K = 7$  [4]. Penelitian lain yang terkait juga pernah dilakukan untuk mendeteksi tingkat kematangan buah pisang berdasarkan ciri warna kulit pisangannya. Data yang digunakan berupa citra/gambar pisang dengan mengambil fitur warna yaitu nilai RGB lalu dikonversi ke Hue, Saturation, dan Intensitas (HSI) dengan tingkat akurasi mencapai 85% [5]. Penelitian serupa juga pernah dilakukan dengan menggunakan cabai sebagai objek, yang bertujuan untuk mengklasifikasikan tingkat kematangan cabai. Penelitian ini menggunakan ekstraksi ciri nilai HSV, sedangkan untuk proses klasifikasinya menggunakan metode K-Nearest Neighbor yang dikombinasikan dengan logika fuzzy. Pengujian dilakukan terhadap 60 data cabai rawit, terdiri dari 15 cabai matang, 15 cabai mentah, 15 cabai setengah matang, dan 7 cabai busuk. Hasil penelitian ini memperoleh akurasi sebesar 86,66% [6]. Dari pemaparan studi literatur sebelumnya pengolahan citra mampu memberikan solusi dari permasalahan klasifikasi kematangan buah dengan memanfaatkan pengolahan citra digital dan kecerdasan buatan dengan adanya penelitian ini diharapkan dapat membantu petani pare pasca panen mendapatkan kualitas pare terbaik berdasarkan kematangannya.

## 2. Metode Penelitian

Pada sistem klasifikasi kematangan sayuran pare akan diklasifikasikan kedalam dua kelas yaitu mentah dan matang. Data citra pare yang diambil dari website roboflow milik Kubo Data terdiri sebanyak 160 data citra pare akan dibagi kedalam dua label yang terdiri dari 80 citra pare mentah dan 80 citra pare matang. Penelitian ini menggunakan 80% dari dataset sebagai data latih dan 20% dari dataset sebagai data uji. Dibawah ini adalah gambar diagram alir dari sistem klasifikasi sayuran pare.



Gambar 1. Diagram Alir Sistem

### 2.1 Pengumpulan Data

Pada tahapan awal penelitian ini memperoleh dataset citra sayuran pare sebanyak 1100 data citra sayuran pare tetapi hanya akan diambil sebanyak 160 citra pare yang terdiri dari 80 buah citra pare mentah dan 80 buah citra pare matang. Dataset sayuran pare didapatkan dari dataset publik yang dipublikasi oleh KUBO DATA melalui website Roboflow.

### 2.2 Preprocessing

Pada tahapan preprocessing diawali dengan memotong citra sayuran pare dengan mengambil titik tengah citra lalu membuat titik penentuan yang membentuk persegi sehingga nanti akan menghasilkan citra berukuran 50x50 piksel. Pemotongan citra ini menggunakan matriks dengan memanfaatkan *library Numpy*.

### 2.3 Ekstraksi Ciri HSV

Pada tahapan ekstraksi fitur akan dilakukan proses transformasi ruang warna dari RGB menjadi HSV menggunakan fungsi *cv2.cvtColor* pada *library OpenCV*. Berikut parafrasa dari penjelasan tentang HSV (Hue, Saturation, Value). HSV (Hue, Saturation, Value) merupakan representasi dari warna dasar dan distribusi dari panjang gelombang cahaya yang menentukan panjang gelombang dominan dari warna dasar tersebut. Dalam penentuan Hue digunakan satuan derajat, dengan ketentuan derajat pada warna dasar yaitu biru 240 derajat, hijau 120 derajat, dan merah 0 derajat. Saturation (S) menunjukkan intensitas warna, dimana nilai semakin tinggi membuat warna semakin cerah. Value (V) merupakan nilai kerapatan warna, pada warna-warna gelap nilai

kerapatannya rendah, sedangkan pada warna-warna cerah nilai kerapatannya tinggi [5]. Proses ini bertujuan untuk mendapatkan perbedaan warna yang lebih akurat dengan mendapatkan kanal warna *Hue*, *Saturation*, dan *Value* yang nanti nilai dari ketiga kanal warna tersebut akan dijadikan sebagai parameter untuk melakukan klasifikasi.

$$H = \tan\left(\frac{3(G-B)}{(R-G)+(R-B)}\right)$$
$$S = 1 - \frac{\min(R,G,B)}{V}$$
$$V = \frac{R+G+B}{3}$$

**Gambar 2.** Rumus Mengubah HSV

## 2.4 Klasifikasi KNN

KNN (K-Nearest Neighbor) merupakan salah satu metode untuk melakukan pengelompokan atau klasifikasi dengan mengacu pada ciri-ciri dari tetangga terdekatnya. Kelebihan dari KNN adalah kemampuannya untuk mengeksekusi data pelatihan dalam jumlah besar. Metode klasifikasi ini cukup sederhana namun memberikan hasil yang cukup akurat [7]. Pada penelitian ini memanfaatkan algoritma KNN pada *library Scikit*. Nilai rerata H, S, V akan dijadikan sebagai data inputan. KNN menggunakan pendekatan *euclidean* mendapatkan hasil yang lebih baik, berikut adalah rumus *euclidean*:

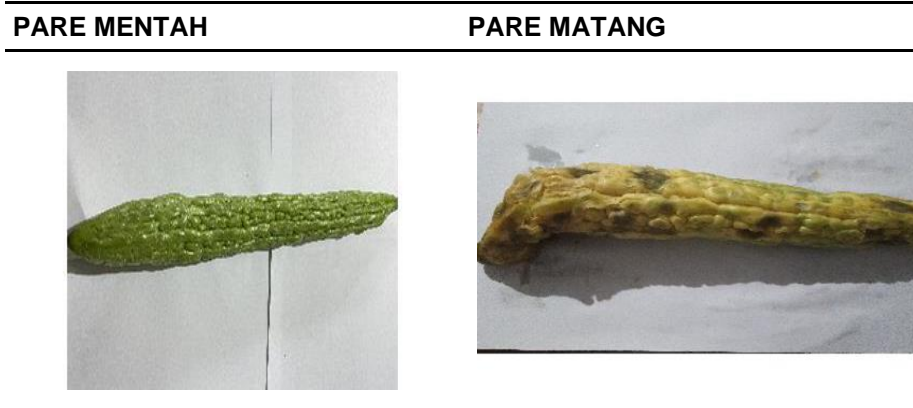
$$d(xy) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

Keterangan:

- d = jarak kedekatan
- x = data training
- y = data testing
- n = jumlah atribut antara 1 s.d n
- i = atribut individu antara 1 s.d n

## 3. Hasil dan Diskusi

Pada Sistem klasifikasi kematangan sayuran pare, penelitian ini menggunakan algoritma K-Nearest Neighbor untuk mengklasifikasi data dan menggunakan bahasa pemrograman python beserta *library seperti cv2, numpy, pyplot, pandas, dan scikit-learn* sebagai pendukung untuk mempermudah dalam penelitian kali ini. Tingkat kematangan sayuran pare yang digunakan adalah pare mentah dan pare matang. Berikut ini merupakan contoh tingkat kematangan sayuran pare yang digunakan dalam penelitian kali ini.



Gambar 3. Tingkat Kematangan Sayuran Pare

### 3.1 Dataset Sayuran Pare

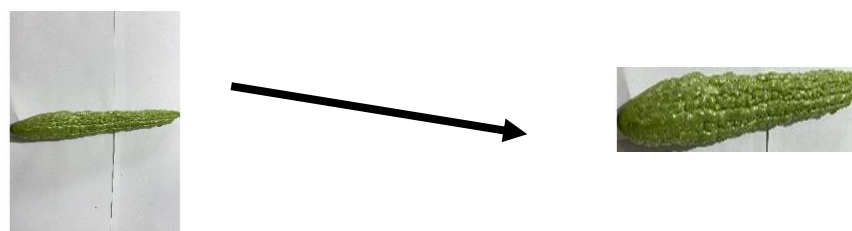
Penelitian ini memakai 160 data citra sayuran pare yang diambil dari Kubo Data melalui website Roboflow. Dari 160 data tersebut telah dikelompokkan kedalam dua label yaitu mentah dan matang dengan jumlah data yang seimbang. 80% dari dataset akan dipakai sebagai data latih dan 20% dari dataset akan digunakan sebagai data uji.



Gambar 4. Dataset Sayuran Pare

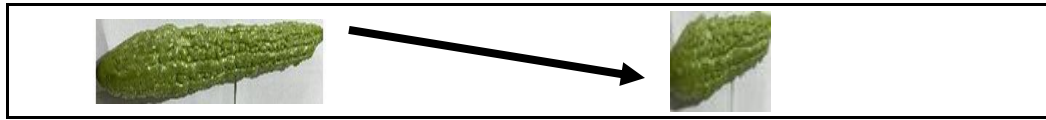
### 3.2 Tahap Preprocessing

Data citra yang telah dikumpulkan akan dilakukan cropping untuk menghilangkan bagian yang kurang penting pada image setelah itu dilakukan resize untuk mengubah rasio ukuran gambar sehingga mempermudah analisa gambar dan mengurangi ukuran penyimpanan yang digunakan. Berikut ini merupakan hasil dari cropping dataset.



**Gambar 5.** Hasil Proses Cropping

Setelah proses cropping dilakukan proses resize citra dari citra awal hasil cropping diubah ke dalam citra yang berukuran 50x50 pixel dengan menggunakan library OpenCV dalam melakukan resize citra.



**Gambar 6.** Hasil Proses Resize

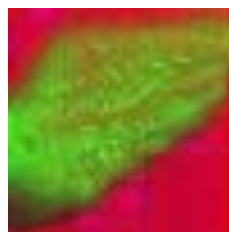
### 3.3 Ekstraksi Ciri HSV

Pada tahapan ini gambar akan diubah kedalam hsv untuk mendapatkan ekstraksi nilai objek gambar. Diawali dengan mengambil nilai gambar rgb yang diubah kedalam hsv lalu nilai dari setiap parameter yaitu hue, saturation, value dipisahkan untuk mendapatkan nilai dari setiap parameter dan dimasukkan kedalam tabel berbentuk csv.



**Gambar 7.** Gambar data dalam RGB

Dari gambar rgb tersebut akan diubah menjadi gambar hsv untuk mendapatkan informasi fitur dari data latih dan mendapatkan informasi apakah buah pare tersebut sudah matang atau masih mentah. Berikut ini merupakan gambar hasil hsv dari gambar diatas.



**Gambar 8.** Gambar data dalam HSV

Hasil dari ekstraksi gambar diatas maka diperoleh nilai dari objek gambar berupa hue, saturation, value yang termasuk kedalam kelompok data latih seperti tabel berikut ini seperti tabel berikut.

**Tabel 1.** Nilai Hue, Saturation, Value

H	S	V
414.184	960.504	1.433.056

Berikut ini merupakan 15 data ekstraksi ciri dari 32 data uji dimana data ini menampilkan nilai hue, saturation, value serta label yang bernilai 1 untuk pare yang matang dan bernilai 0 untuk pare yang mentah. Setelah itu data akan diolah menggunakan metode selanjutnya yaitu K-

Nearest Neighbor agar diketahui akurasi sistem klasifikasi ini menggunakan metode K-Nearest Neighbor.

**Tabel 2.** Contoh Ekstraksi Tabel Uji

H	S	V	Label
59.263.175	64.235.875	1.247.065	1
47.831.875	63.885.725	1.264.319	1
5.440.765	583.777	123.001.125	1
536.868	62.521.475	12.265.665	1
5.506.665	29.139.925	1.672.659	1
48.185.825	210.978	200.690.325	1
60.706.725	36.410.775	1.858.143	1
32.841.125	35.307.225	175.629	1
7.891.975	46.457.375	1.347.658	1
392.883	563.666	136.901.475	0
73.900.325	595.007	1.412.405	0
64.666	528.793	12.942.055	0
8.546.995	48.753.675	14.175.405	0
6.773.985	478.426	12.980.355	0
89.883.225	5.114.305	133.840.275	0
4.243.255	5.603.525	14.593.875	0

### 3.4 Klasifikasi KNN

Pada metode klasifikasi ini akan menggunakan 32 data uji dengan 16 data pare matang dan 16 pare mentah. Akurasi yang didapatkan dari masing-masing nilai k mempunyai rentang nilai yang sangat beragam. Berikut ini merupakan akurasi yang didapatkan berdasarkan masing-masing nilai k yang digunakan.

**Table 3.** Hasil Uji Coba Akurasi

Nilai K	Data yang Sesuai	Data Keseluruhan	Akurasi
3	28	32	88%
4	26	32	82%
5	25	32	78%
6	26	32	82%
7	23	32	72%
8	22	32	69%
9	22	32	69%
10	20	32	63%

Akurasi yang didapatkan dari hasil uji coba tersebut menunjukkan bahwa nilai tertinggi yang dapat diperoleh adalah 88% dengan menggunakan nilai  $k = 3$  sehingga acuan dalam melakukan klasifikasi sistem dengan metode K-Nearest Neighbor adalah  $k$  dengan nilai terbaik yaitu 3.

### 3.5 Hasil Akhir

Setelah diketahui bahwa nilai  $k = 3$  merupakan nilai terbaik dari hasil uji coba akurasi. Berikut akan ditampilkan hasil data uji sebanyak 32 data yang terdiri dari 16 data pare matang dan 16 data pare mentah dengan menggunakan metode K-Nearest Neighbor memakai nilai  $k = 3$ .

**Tabel 4.** Hasil 32 Data Uji

Data Tes	Kelas Asli	Hasil KNN	Informasi
2	Mentah	Mentah	Tepat
6	Mentah	Mentah	Tepat
8	Mentah	Mentah	Tepat
10	Mentah	Mentah	Tepat
11	Mentah	Mentah	Tepat
17	Matang	Matang	Tepat
19	Matang	Matang	Tepat
25	Matang	Matang	Tepat
27	Mentah	Matang	Tidak Akurat
29	Mentah	Mentah	Tepat
30	Mentah	Mentah	Tepat
33	Matang	Matang	Tepat
35	Matang	Matang	Tepat
36	Matang	Matang	Tepat
40	Matang	Matang	Tepat
43	Matang	Matang	Tepat
45	Matang	Matang	Tepat
59	Mentah	Mentah	Tepat
60	Mentah	Mentah	Tepat
68	Mentah	Mentah	Tepat
69	Mentah	Mentah	Tepat
85	Matang	Mentah	Tidak Akurat
88	Matang	Matang	Tepat
97	Mentah	Mentah	Tepat
101	Mentah	Mentah	Tepat
102	Mentah	Mentah	Tepat
103	Mentah	Mentah	Tepat
104	Matang	Mentah	Tidak Akurat

Data Tes	Kelas Asli	Hasil KNN	Informasi
111	Matang	Matang	Tepat
115	Matang	Matang	Tepat
125	Matang	Matang	Tepat
127	Matang	Mentah	Tidak Akurat

Setelah itu akan dilakukan perhitungan manual untuk menentukan akurasi dari hasil 32 data uji yang menggunakan k=3 dengan confusion matrix euclidean distance pada sebagai berikut.

**Tabel 5.** Confusion Matrix k = 3

Kelas Asli	Kelas Prediksi	
	Matang	Mentah
Matang	13	3
Mentah	1	15

$$\text{Akurasi} = \frac{TP}{Data} \times 100\%$$

$$= \frac{(13 + 15)}{32} \times 100\% = 88\%$$

Jadi akurasi yang didapatkan dari perhitungan confusion matrix k=3 adalah sebesar 88%

#### 4. Kesimpulan

Berdasarkan hasil pengujian yang sudah dilakukan untuk mengklasifikasi kematangan sayuran pare dengan menggunakan metode K-Nearest Neighbor mendapatkan akurasi terbaik yaitu 88% untuk nilai k=3 dengan menggunakan data latih sebanyak 128 data dan data uji sebanyak 32 data dengan masing-masing kelas yakni kelas matang dan kelas mentah. Jadi berdasarkan ekstraksi ciri HSV dan metode K-Nearest Neighbor dapat diterapkan untuk mengklasifikasi sayuran pare dari tingkat kematangannya berdasarkan nilai dari Hue, Saturation, Value.

#### Daftar Pustaka

- [1] W. Gunadi, K. Sulandjari, and M. Syafi'i, "Respon Pertumbuhan dan Hasil Tanaman Pare (Momordica charantia L) Varietas Lipa F1 pada Berbagai Taraf Ketebalan Mulsa Jerami Padi," *Jurnal Agrotek Indonesia (Indonesian Journal of Agrotech)*, vol. 6, no. 2, pp. 62-69, 2021.
- [2] T. Saras, Pare: Sayuran Ajaib untuk Kesehatan Anda. Tiram Media, 2023.
- [3] F. Liantoni and F. N. Annisa, "Fuzzy K-Nearest Neighbor Pada Klasifikasi Kematangan Cabai Berdasarkan Fitur Hsv Citra," *JIPI (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 3, no. 2, 2018.
- [4] D. R. Febrinamas, R. Hidayati, and I. Nirmala, "Klasifikasi Buah Pinang Berdasarkan Data Sensor Menggunakan Metode K-Nearest Neighbor Berbasis Web," *Journal of Computer System and Informatics (JoSYC)*, vol. 4, no. 4, pp. 1046-1055, 2023.
- [5] M. Ismail, I. Indrabayu, and I. S. Areni, "Support Vector Machine Method to Reduce the Execution Time of Vehicle Plate Recognition System," *EPI International Journal of Engineering*, vol. 1, no. 1, pp. 69-75, 2018.
- [6] F. M. Sarimole and A. Rosiana, "Classification of Maturity Levels in Areca Fruit Based on HSV Image Using the KNN Method," *Journal of Applied Engineering and Technological Science (JAETS)*, vol. 4, no. 1, pp. 64-73, 2022.
- [7] E. Arisawati, "Penerapan K-Nearest Neighbor Berbasis Genetic Algorithm Untuk Penentuan Pemberian Kredit," *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, vol. 1, no. 1, pp. 1-11, 2017.