

# Klasifikasi Tingkat Kematangan pada Buah Rambutan Berdasarkan Fitur Warna Menggunakan KNN dan Ekstraksi Warna HSV

Annisyao Octaviani<sup>1</sup>, Desta Sandya Prasvita<sup>2</sup>, Kiana Rizki Tsaniyah Zulkarnain<sup>3</sup>, Siti Hinggit<sup>4</sup>  
Informatika / Fakultas Ilmu Komputer

Universitas Pembangunan Nasional Veteran Jakarta

Jl. RS. Fatmawati, Pondok Labu, Jakarta Selatan, DKI Jakarta, 12460, Indonesia.

email: annisyao@upnvj.ac.id<sup>1</sup>, desta.sandya@upnvj.ac.id<sup>2</sup>, kianartz@upnvj.ac.id<sup>3</sup>, sitih@upnvj.ac.id<sup>4</sup>

**Abstrak.** Klasifikasi tingkat kematangan buah secara manual dalam penglihatan saja dirasa kurang tentu dipengaruhi beberapa faktor subyektivitas dari manusia. Akibat dari hal itu ialah tidak mengetahui dengan pasti buah tersebut belum matang, setengah matang, maupun sudah matang. Salah satunya pada buah rambutan. Berawal dari masalah tersebut, maka penelitian ini pun membahas mengenai perbedaan tingkat kematangan pada buah rambutan menggunakan ekstraksi ciri warna yang salah satunya ialah pada warna *Hue, Saturation, Value* (HSV) dengan menerapkan metode klasifikasi yaitu *K-Nearest Neighbors* (KNN) serta menerapkan metode hasil evaluasi menggunakan *Cross Validation* dengan pembagian secara *K-Fold*. Pengambilan sampel citra sebanyak 21 citra yang diproses, serta penggunaan nilai  $K = 7$  didapatkan hasil akurasi sebesar 76,6% yang merupakan hasil ketepatan dari citra dalam penelitian.

**Kata Kunci:** Buah Rambutan, *K-Nearest Neighbor* (KNN), *Hue Saturation Value* (HSV), *Cross Validation*.

## 1 Pendahuluan

### 1.1 Latar Belakang

Rambutan yang memiliki nama latin *Nephelium lappaceum L.* adalah salah satu buah tropis asli dari Indonesia yang tumbuh di dataran rendah hingga 600 mdpl dengan iklim basah. Di Indonesia, produksi buah rambutan berpusat di bagian barat wilayah Indonesia yaitu Pulau Sumatera, Jawa dan Kalimantan. Buah ini sangat terkenal karena warnanya yang memikat dan rasanya yang manis. Selain di Indonesia, rambutan juga digemari di berbagai negara lainnya sehingga Indonesia menjadi pengeksport rambutan ke macam negara seperti Uni Emirat Arab, Belanda, Arab Saudi, Taiwan, Singapura, Jerman, Prancis, dan Filipina. Musim panen utama rambutan di Indonesia adalah pada bulan September hingga April. Produksinya berfluktuasi dari 263.000 menjadi 350.000 MT/tahun.

Selama ini untuk menentukan tingkat kematangan buah rambutan hanya dilakukan secara manual atau dengan kata lain hanya menggunakan penglihatan manusia yang mana dibutuhkan tenaga dan juga waktu sehingga terkadang hasil dari klasifikasi tersebut kurang akurat dan tidak stabil. Dibutuhkan cara lain yang dapat membantu manusia untuk menentukan tingkat kematangan buah rambutan. Untuk menentukan tingkat kematangan rambutan dalam bidang informatika salah satunya yaitu dengan memanfaatkan ekstraksi ciri *Hue, Saturation, Value* (HSV) dan metode *K-Nearest Neighbor* (KNN).

Ekstraksi warna dilakukan untuk membantu pengklasifikasian tingkat kematangan. HSV menjelaskan warna berdasarkan terminologi *Hue, Saturation, Value*. Hue akan membedakan warna-warna dan menentukan kemerahan, kehijauan dsb. dari cahaya, *Saturation* menjelaskan tingkat kemurnian dari suatu warna. *Value* menjelaskan jumlah dari cahaya yang diterima oleh mata kita.

Algoritma KNN akan mengklasifikasikan objek berdasarkan data *training* yang jaraknya paling dekat dengan objek tersebut. Metode ini merupakan salah satu jenis dari pembelajaran berdasarkan perwujudan. KNN digunakan oleh banyak peneliti sebagai metode klasifikasi yang hasil akurasi nya tinggi. Salah satu penelitian yang dilakukan Muhammad Duwen dkk. [6], mengklasifikasikan buah belimbing berdasarkan fitur warna yang dibagi menjadi 3 kelas klasifikasi dan memiliki hasil akurasi rata-rata mencapai 82,66%.

Pada kesempatan kali ini akan dibahas bagaimana cara membedakan tingkat kematangan buah rambutan berdasarkan citra buah rambutan menggunakan klasifikasi KNN dan ekstraksi ciri HSV sehingga dapat membantu dan memudahkan dalam menentukan apakah suatu buah rambutan sudah layak untuk dimakan atau belum dengan akurasi yang baik.

## 1.2 Studi Literatur

Sebelum melakukan penelitian, dilakukan studi literatur terhadap 2 jurnal penelitian yang telah ada sebelumnya.

Jurnal pertama disusun oleh Aminudin, A. [1] yang melakukan penelitian untuk mengklasifikasi tingkat kematangan dari buah Pepaya California menggunakan metode klasifikasi KNN berdasarkan warna kulit buah. Penulis menggunakan 3 kategori pengklasifikasian buah papaya Bangkok yaitu Matang, Setengah Matang dan Mentah. Terdapat 12 citra latih dan 12 citra uji yang digunakan. Didapat hasil akurasi kesesuaian dengan data tanpa digunakan rata-rata yaitu sebesar 75% untuk  $K = 1$ ,  $K = 5$ , dan  $K = 7$ . Jika digunakan rata-rata dalam evaluasi maka didapat hasil akurasi sebesar 66,67% pada  $K = 5$  dan 75% pada  $K = 7$  dengan data uji yang berbeda pula.

Jurnal kedua disusun oleh Areni dkk. [2] yang melakukan penelitian untuk mengklasifikasi tingkat kematangan dari buah Stroberi menggunakan metode klasifikasi SVM dan segmentasi warna HSV. Penulis menggunakan 3 jenis kelas yaitu matang, setengah matang dan belum matang serta menggunakan 180 citra yang dibagi 80% untuk dilatih dan 20% untuk diuji. Citra tersebut didapat dengan mengambil langsung menggunakan kamera *Logitech C920*. Citra yang diperoleh tersebut ditransformasi ke HSV lalu diubah menjadi hitam putih melalui proses *masking* untuk memisahkan antara *foreground* dan *background* citra sehingga dapat dilakukan ekstraksi fitur. Klasifikasi untuk menentukan tingkat kematangan buah stroberi menggunakan algoritma SVM menghasilkan tingkat akurasi sebesar 97% dengan kernel RBF.

## 1.3 Usulan Metode

Untuk proses ekstraksi citra digunakan fitur warna HSV berdasarkan pada tingkat kecerahan warna dari buah rambutan dan melakukan klasifikasi pengelompokkan berdasarkan metode *K-Nearest Neighbor* serta untuk evaluasi digunakan metode *Cross Validation*.

## 1.4 Data yang Digunakan

Pada penelitian ini digunakan dataset citra "*Fruits-262*" yang dapat diakses melalui <https://www.kaggle.com/aelchimminut/fruits262>. Dari dataset tersebut dilakukan penyaringan dan dipilih kembali sehingga hanya tersisa kumpulan citra buah rambutan yang akan digunakan.

# 2 Metodologi Penelitian

## 2.1 Tingkat Kematangan Buah Rambutan

Berdasarkan Studi Literatur yang dijadikan acuan dalam penelitian kali ini, tingkat kematangan pada buah rambutan berdasarkan warna kulitnya dapat dibedakan sebagai berikut:



**Gambar. 1.** Buah Rambutan Belum Matang.



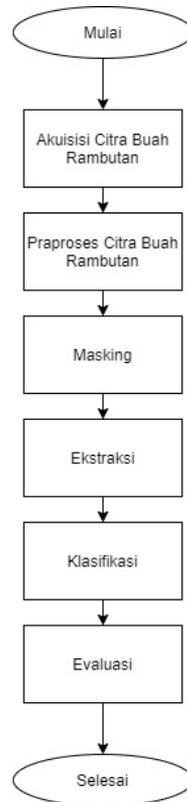
**Gambar. 2.** Buah Rambutan Setengah Matang.



**Gambar. 3.** Buah Rambutan Matang.

Dapat dilihat dari citra diatas bahwa rambutan yang memiliki tingkat kematangannya sudah mencapai tingkat siap dikonsumsi atau matang yang memiliki kecerahan warna dominan merah yang kuat dan segar. Dibandingkan rambutan yang belum matang tingkat kecerahan serta warnanya masih tidak menunjukkan warna kemerahan atau dominan hijau. Sedangkan, rambutan yang memiliki tingkat kematangan setengah matang akan memiliki warna dominan kuning yang masih tidak menunjukkan warna kemerahan.

## 2.2 Implementasi Penelitian



**Gambar. 4.** Rancangan Penelitian.

### 2.2.1 PraPengolahan Data

Penelitian ini menggunakan Kaggle.com sebagai dataset untuk memproses citra buah rambutan dengan judul “*Fruits-262*”. Dalam dataset tersebut berisikan satu folder 1117 citra dengan kualitas, pengambilan posisi yang berbeda-beda serta terdapat banyak salinan citra yang sama. Dari 1117 citra tersebut di *sorting* kembali untuk memilih citra terbaik yang sesuai dengan kebutuhan penelitian ini sehingga diperoleh 21 citra yang akan dijadikan dataset berisikan data *training* dan data *testing*.

### 2.2.2 Akuisisi Citra Buah Rambutan

Penelitian kali ini menggunakan 21 citra yang didapatkan dari dataset yang akan digunakan sebagai data percobaan dari keseluruhan total dataset. Untuk melakukan penelitian kali ini menggunakan 3 gambar buah rambutan yang belum matang, 8 buah rambutan yang setengah matang dan 10 gambar buah rambutan yang sudah matang.

### 2.2.3 Masking

*Masking* adalah salah satu metode yang digunakan dengan tujuan untuk dapat proses menyembunyikan atau menutupi suatu objek dengan objek lain, sehingga objek yang menutupi terlihat transparan dan menyatu dengan objek yang ditutupi. Dimana dilakukan 2 kali proses *masking* hal ini agar citra didapatkan tingkat kecerahan warna berwarna merah sesuai dengan tingkat kematangan dari masing-masing buah atau citra yang digunakan.

#### 2.2.4 Ekstraksi Menggunakan *Hue Saturation Value* (HSV)

Pada penelitian kali ini untuk mengekstraksi warna digunakan metode algoritma *Hue Saturation Value* yang dilakukan dengan menganalisis nilai warna tiap piksel citra sesuai fitur yang diinginkan. Pada penelitian ini image citra yang tertangkap lebih banyak memiliki warna RGB. Untuk dapat mengurangi efek ilusi pada sebuah citra, warna citra tersebut dapat dikonversikan ke *colour space* yang lain, salah satunya adalah ruang warna HSV. Tujuan penelitian kali ini melakukan ekstraksi fitur warna menggunakan metode HSV untuk menjadikan ciri warna sebagai batasan atau patokan apakah sebuah citra rambutan atau buah rambutan dapat dikatakan matang atau belum berdasarkan pada proses *masking* sebelumnya yang menggunakan warna merah sebagai acuan dari penentuan kali ini.

#### 2.2.5 Klasifikasi Menggunakan *K-Nearest Neighbor* (KNN)

Pada dataset yang digunakan setelah proses ekstraksi pada citra menggunakan metode *Hue Saturation Value* sebelumnya. Kemudian dilakukan klasifikasi pada citra rambutan tersebut menggunakan metode *K-Nearest Neighbor* (KNN) dengan menggunakan  $K = 7$ . Untuk pembagian klasifikasinya dibagi menjadi 3 kelas/label yaitu matang, setengah matang dan belum matang.

#### 2.2.6 Evaluasi Menggunakan *Cross Validation*

Pada tahap akhir, perlu dilakukan evaluasi terhadap klasifikasi yang dilakukan agar dapat menentukan akurasi dari klasifikasi ini, maka diperlukan sebuah model evaluasi yang digunakan. Dari hasil evaluasi tersebut diputuskan bahwa metode yang akan digunakan dalam evaluasi model kali ini adalah *Cross Validation*. Selanjutnya, digunakan *K-fold Cross-Validation* karena untuk meningkatkan tingkat performansi dari model tersebut dan dapat mengolah dataset dengan kelas yang seimbang. Metode standar yang digunakan untuk evaluasi adalah 4-fold cross-validation, agar dapat melihat hasil akurasi dalam klasifikasi kali ini.

Pada *K-fold cross-validation*, data awal dibagi secara random ke dalam k bagian,  $D_1, D_2, \dots, D_k$ . Data *training* dan data *testing* dilakukan sebanyak k kali. Kemudian, pada pengulangan ke-i, partisi data  $D_i$  digunakan sebagai *testing*, sisanya digunakan *training*. Untuk mendapatkan hasil dari evaluasi tersebut maka dapat dilakukan dengan perhitungan seperti ini :

Akurasi :

$$\frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (1)$$

Dimana Akronim yang memiliki arti sebagai berikut:

*True Positive* (TP) : Jumlah tuple positif yang diprediksi juga sebagai positif.

*True Negative* (TN) : Jumlah tuple negatif yang diprediksi juga sebagai negatif.

*False Positive* (FP) : Jumlah tuple negatif yang salah diprediksi sebagai positif.

*False Negative* (FN) : Jumlah tuple positif yang salah diprediksi sebagai negatif.

### 3 Hasil dan Pembahasan

#### 3.1 Implementasi

Pengambilan citra didapatkan dari situs *open source*, yaitu pada situs *kaggle.com* dengan judul datasetnya "*Fruit-262*" dengan folder didalamnya yang bernama "Rambutan". Kemudian citra yang telah didapatkan akan diproses yaitu dengan memilih citra yang sesuai dengan kebutuhan klasifikasi, dengan pembagian tiga kelas, yaitu kelas 0 untuk belum matang, kelas 1 untuk setengah matang, dan kelas 2 untuk matang. Sebelum dilakukannya proses klasifikasi, tentu citra yang telah dipilih tadi dilakukan proses *masking* dengan menggunakan *Hue Saturation Value*

(HSV) untuk mendapatkan warna merah dari citra tersebut. Pada proses *masking* pun dilakukan operasi *open* dan *close*. Setelah didapatkan citra yang telah di-*masking*, langkah selanjutnya ialah mendapatkan nilai dari ekstraksi ciri warna yaitu HSV dari histogram setiap citra, lalu diubah menjadi vektor matriks dan mentranspose matriks dari setiap citra.

Kemudian dilakukan penggabungan data yang diperoleh dari setiap citra menjadi satu matriks yang digunakan sebagai data training maupun data testing. Untuk pemilihan data training dan data testing, menggunakan operasi *K-Fold*, dimana perbandingannya dibagi 1:3,. Dilakukan perulangan sebanyak 4 kali.

### 3.2 Hasil Klasifikasi

Untuk mengetahui serta mengukur akurasi dari klasifikasi yang menggunakan metode KNN dengan  $K = 7$ , maka digunakan *Cross Validation* dengan pembagian data dengan metode *K-Fold* dimana 25% digunakan sebagai data *testing*, dan 75% digunakan sebagai data *training*. Hasilnya akan disajikan menggunakan *Confusion Matrix* seperti pada tabel di bawah ini dimana hasil diperoleh dengan perhitungan menggunakan rumus akurasi yang terdapat pada bagian 2.2.6.

**Tabel. 1.** *Confusion Matrix* Pertama

		Kelas Prediksi		
		0	1	2
Kelas Sebenarnya	0	0	1	0
	1	0	2	0
	2	0	1	2

Akurasi : 66,6%

**Tabel. 2.** *Confusion Matrix* Kedua

		Kelas Prediksi		
		0	1	2
Kelas Sebenarnya	0	0	1	0
	1	0	2	0
	2	0	1	1

Akurasi : 60%

**Tabel. 3.** *Confusion Matrix* Ketiga

		Kelas Prediksi		
		0	1	2
Kelas Sebenarnya	0	0	1	0
	1	0	2	0
	2	0	0	2

Akurasi : 80%

**Tabel. 4.** *Confusion Matrix* Keempat

		Kelas Prediksi		
		0	1	2
Kelas Sebenarnya	0	0	0	0
	1	0	2	0
	2	0	0	3

Akurasi : 100%

Hasil Evaluasi :

- Akurasi rata-rata : 76,6%.

### 3.3 Pembahasan

Berdasarkan dari hasil percobaan pada data di tabel 1 memberikan akurasi sebesar 66,6%, lalu untuk tabel 2 akurasi yang didapatkannya sebesar 60%, serta untuk tabel 3 memberikan akurasi sebesar 80%, dan pada tabel terakhir yaitu dengan akurasi sebesar 100%. Berdasarkan pada hasil tersebut bahwa terdapat beberapa data dengan prediksi yang salah, tidak sesuai dengan label aslinya.

## 4 Kesimpulan dan Saran

### 4.1 Kesimpulan

Penggunaan dataset dalam penelitian ini yaitu sebanyak 21 citra dengan judul “*Fruit-262*” yang dapat diakses melalui <https://www.kaggle.com/aelchimminut/fruits262>. Citra pun diproses menggunakan masking dengan operasi *open* dan *close* dengan tujuan untuk mendapatkan warna merah saja menggunakan warna *Hue Saturation Value* (HSV) yang kemudian dilakukan ekstraksi ciri warna dengan metode *Hue Saturation Value* (HSV) dan diklasifikasi dengan *K-Nearest Neighbor* (KNN) dengan K sebanyak 7. Serta dilakukan evaluasi model dengan *Cross Validation* dengan metode *K-Fold* yang dilakukan sebanyak 4 kali mendapatkan hasil akurasi sebesar 76,6%.

### 4.2 Saran

Pada penelitian ini didapatkan nilai akurasi tidak terlalu besar, karena citra yang digunakan tidak cukup banyak. Diharapkan pada penelitian selanjutnya akan lebih baik jika dataset yang digunakan lebih banyak, karena penulis hanya menggunakan 21 data citra dan untuk pembagian citra ke dalam kelasnya tidak merata.

## Referensi

- [1] Aminudin, A. (2019). *KLASIFIKASI TINGKAT KEMATANGAN BUAH PEPAYA MENGGUNAKAN METODE K-NEAREST NEIGHBOR BERDASARKAN WARNA KULIT BUAH* (Doctoral dissertation, University of Technology Yogyakarta).
- [2] Areni, I. S., Amirullah, I., & Arifin, N. (2019). Klasifikasi Kematangan Stroberi Berbasis Segmentasi Warna dengan Metode HSV. *Jurnal Penelitian Enjiniring*, 23(2), 113-116.
- [3] Handayani, L. (2016). Rancang Bangun sistem Temu Kembali Gambar Ikan Berdasarkan Ekstraksi Ciri Warna HSV, Bentuk Canny dan tekstur Orde Dua. *Jurnal Sains, Teknologi dan Industri*, 13(2), 174-180.
- [4] Khotimah, H., Nafi'iyah, N., & Masruroh, M. (2020). Klasifikasi Kematangan Buah Mangga Berdasarkan Citra HSV dengan KNN. *Jurnal Elektronika Listrik dan Teknologi Informasi Terapan*, 1(2), 1-4.
- [5] Kusuma, S. F., Pawening, R. E., & Dijaya, R. (2017). Otomatisasi klasifikasi kematangan buah mengkudu berdasarkan warna dan tekstur. *Register: Jurnal Ilmiah Teknologi Sistem Informasi*, 3(1), 17-23.
- [6] Muhammad, D. I., Ermatita, E., & Falih, N. (2021). Penggunaan K-Nearest Neighbor (KNN) untuk Mengklasifikasi Citra Belimbing Berdasarkan Fitur Warna. *Informatik: Jurnal Ilmu Komputer*, 17(1), 9-16.
- [7] Paramita, C., Rachmawanto, E. H., Sari, C. A., & Setiadi, D. R. I. M. (2019). Klasifikasi Jeruk Nipis Terhadap Tingkat Kematangan Buah Berdasarkan Fitur Warna Menggunakan K-Nearest Neighbor. *Jurnal Informatika*, 4(1).
- [8] Rachmawanto, E. H., & Salam, A. (2019). Pengukuran Tingkat Kematangan Kopi Robusta Menggunakan Algoritma K-Nearest Neighbor.
- [9] Rianto, P., & Harjoko, A. (2017). Penentuan Kematangan Buah Salak Pondoh Di Pohon Berbasis Pengolahan Citra Digital. *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, 11(2), 143-154.
- [10] Rohmi, G. F. (2017). *Implementasi citra digital berdasarkan nilai HSV untuk mengidentifikasi jenis tanaman mangga menggunakan algoritma K-Nearest Neighbor* (Doctoral dissertation, UIN Sunan Gunung Djati Bandung).
- [11] Solehatin, S., Kom, M., & Anam, C. (2020). *E-Deteksi Kematangan Buah Jeruk Banyuwangi Menggunakan Metode KNN Berbasis Android*. Deepublish.