

# Identifikasi Citra Apel Menggunakan Fitur Warna Dengan Metode Klasifikasi *K-Nearest Neighbors*

Adla Anugrah Abbas<sup>1</sup>, Aini Cahyaning Putri<sup>2</sup>, Ananta Khayana Putra<sup>3</sup>, Mauludhanti Putri S<sup>4</sup>  
Informatika / Fakultas Ilmu Komputer

Universitas Pembangunan Nasional Veteran Jakarta

Jl. RS. Fatmawati Raya, Pd. Labu, Kec. Cilandak, DKI Jakarta 12450, Indonesia

[adlaaaa@upnvj.ac.id](mailto:adlaaaa@upnvj.ac.id)<sup>1</sup>, [ainic@upnvj.ac.id](mailto:ainic@upnvj.ac.id)<sup>2</sup>, [anantakp@upnvj.ac.id](mailto:anantakp@upnvj.ac.id)<sup>3</sup>, [mauludhantips@upnvj.ac.id](mailto:mauludhantips@upnvj.ac.id)<sup>4</sup>

**Abstrak.** Apel merupakan buah yang digemari oleh masyarakat Indonesia. Buah apel banyak digemari masyarakat karena rasanya yang bervariasi dan jenisnya yang beragam. Namun karena jenisnya yang beragam kita merasa sulit untuk membedakan jenisnya. Dalam penelitian ini bertujuan untuk mengklasifikasikan jenis buah apel. Metode yang digunakan dalam penelitian ini adalah *HSV (Hue Saturation Value)*. Nantinya akan diklasifikasi menggunakan algoritma *K-Nearest Neighbors (K-NN)* dan di akhir penelitian ini akan diuji tingkat keakuratannya guna mengetahui apakah penelitian ini sudah benar atau belum. Data yang dipakai di penelitian ini adalah sebanyak 160 data citra buah apel dan terbagi menjadi data training dan data testing. Hasil dari penelitian ini didapatkan nilai akurasi sebesar 93.75% dengan menggunakan nilai  $K = 3$ .

**Kata Kunci :** Apel, *HSV*, *K-Nearest Neighbors (K-NN)*, Klasifikasi

## 1. PENDAHULUAN

Apel (*Malus domestica*) merupakan tanaman buah tahunan berasal dari Asia Barat yang beriklim sub tropis. Tanaman apel ini dapat tumbuh di Indonesia setelah beradaptasi dengan iklim Indonesia, yaitu iklim tropis. Penanaman apel di Indonesia dimulai sejak tahun 1934 dan berkembang pesat pada tahun 1960 hingga sekarang. Apel merupakan buah yang digemari oleh masyarakat Indonesia, menurut Badan Pusat Statistik tahun 2006 rata-rata konsumsi apel di Indonesia hingga 1,1 kg perkapita pertahun[1].

Buah apel banyak digemari masyarakat karena rasanya yang bervariasi dan jenis nya yang beragam. Buah apel sendiri memiliki banyak nutrisi dan berbagai macam vitamin. Karena jenis nya yang beragam kita sebagai orang awam merasa sulit untuk membedakan jenis nya. Oleh karena itu dibutuhkan suatu teknologi yang dapat membedakan jenis apel secara akurat agar kita tidak merasa kebingungan.

Teknologi saat ini memungkinkan untuk melakukan klasifikasi citra digital. Secara umum tahapan dalam proses klasifikasi citra digital yaitu akusisi citra, pra pengolahan citra, ekstraksi ciri/ fitur, pelatihan, pengujian dan pengukuran akurasi. Tahapan mengekstrak ciri atau informasi dalam citra digital sangat mempengaruhi untuk mengenali objek yang ada dalam citra tersebut. Semakin banyak ciri yang diekstrak akan mempengaruhi tingkat akurasi klasifikasi citra.

Terdapat bermacam-macam ekstraksi dalam ekstraksi ciri citra yaitu ekstraksi ciri tekstur, bentuk, ukuran, geometri, dan warna[2].

Dalam mengklasifikasikan jenis buah apel itu terdapat algoritma klasifikasi citra yang dapat digunakan salah satunya adalah *K-Nearest Neighbor (K-NN)*. Algoritma K-NN merupakan metode klasifikasi yang menentukan label dari sebuah objek baru berdasarkan mayoritas label dari jarak terdekat K dalam kelompok data latih[2].

## **2. TINJAUAN PUSTAKA**

### **2.1 Studi Literatur**

Pada tahap ini, kami mencari beberapa referensi jurnal terkait metode ini. Seperti contoh mengambil jurnal yang berjudul *Klasifikasi Jenis Buah Apel Dengan Metode K-Nearest Neighbors* yang disusun oleh Novan Wijaya dan Anugrah Ridwan dengan tingkat Akurasi pengujian data sebesar 82%.

Kemudian pada jurnal kedua yaitu jurnal dengan judul *Klasifikasi Kematangan Buah Jeruk Berdasarkan Fitur Warna Menggunakan Metode SVM* yang disusun oleh Zeni Dwi Lestari, Nur Nafi'iyah dan Purnomo Hadi Susilo dengan tingkat Akurasi pengujian data sebesar 80%.

### **2.2 Metode HSV (Hue Saturation Value)**

*HSV* mendefinisikan warna dalam terminologi *Hue*, *Saturation*, dan *Value* yang memiliki karakteristik sebagai berikut :

- a. *Hue* : menyatakan warna sebenarnya, seperti merah, violet, dan kuning dan digunakan menentukan kemerahan (redness), kehijauan (greeness), dsb.
- b. *Saturation* : kadang disebut chroma, adalah kemurnian atau kekuatan warna.
- c. *Value* : kecerahan dari warna. Nilainya berkisar antara 0-100 %. Apabila nilainya 0 maka warnanya akan menjadi hitam, semakin besar nilai maka semakin cerah dan muncul variasi-variasi baru dari warna tersebut.

Proses untuk mendapatkan nilai dari setiap warna yang ingin ditampilkan melalui proses perhitungan yaitu dengan melakukan konversi ruang warna RGB ke ruang warna HSV. Berikut adalah persamaan untuk mendapatkan setiap nilai HSV dengan menggunakan rumus yang dimana persamaan pertama digunakan untuk mencari nilai maksimum dari tiga parameter yang berasal dari ruang warna RGB dan akan dimasukkan kedalam variabel  $\min$ [3].

$$r = \frac{R}{(R+G+B)}, g = \frac{G}{(R+G+B)}, b = \frac{B}{(R+G+B)} \quad (1)$$

$$V = \max(r, g, b) \quad (2)$$

$$S = \begin{cases} 0, & \text{jika } V = 0 \\ 1 - \frac{\min(r, g, b)}{V}, & V > 0 \end{cases} \quad (3)$$

$$H = \begin{cases} 0, & \text{jika } S = 0 \\ \frac{60 * (g - b)}{S * V}, & \text{jika } V = r \\ 60 * \left[ 2 + \frac{b - r}{S * V} \right], & \text{jika } V = g \\ 60 * \left[ 4 + \frac{r - g}{S * V} \right], & \text{jika } V = b \end{cases} \quad (4)$$

$$H = H + 360 \text{ jika } H < 0 \quad (5)$$

Dimana  $r$  sebagai persamaan (1) merupakan persamaan untuk mencari nilai maksimum dari tiga parameter yang berasal dari ruang warna RGB dan akan dimasukkan kedalam persamaan (2) yaitu variabel  $V$  dimana  $\max$  adalah sebuah fungsi untuk mencari nilai maksimum dari beberapa angka  $R$  yang merupakan nilai warna merah,  $G$  merupakan nilai warna hijau, dan  $B$  merupakan nilai warna biru, persamaan berikutnya akan menunjukkan persamaan untuk mencari nilai minimum dari tiga parameter ruang warna RGB yang akan dimasukkan kedalam persamaan (3) yaitu variabel  $S$  untuk mendapatkan nilai saturation parameter ruang HSV dari hasil konversi tiga parameter ruang warna yang akan dimasukkan kedalam persamaan (4) yaitu variabel  $H$  untuk mendapatkan nilai hue parameter ruang warna HSV dari hasil konversi tiga parameter ruang warna RGB. Berdasarkan persamaan tersebut maka didapatkan persamaan (5) sebagai hasil nilai dari setiap warna HSV dan cara ini merupakan cara yang terbilang sangat efektif dipakai jika mendapatkan saturation bernilai 0 sehingga nilai dari hue akan terdefinisi[3].

### 2.3 Metode K-Nearest Neighbors (K-NN)

*K-Nearest Neighbors (K-NN)* merupakan algoritma dengan metode untuk melakukan klasifikasi terhadap objek yang berdasarkan dari data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Dimana kelas yang paling banyak muncul yang nantinya akan menjadi kelas hasil dari klasifikasi. Prinsip kerja *K-NN* adalah mencari jarak terdekat antara data yang akan dievaluasi dengan  $k$  tetangga (*neighbor*) terdekatnya dalam data pelatihan dengan menggunakan persamaan yang menghitung jarak tetangga terdekat.

$$D = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

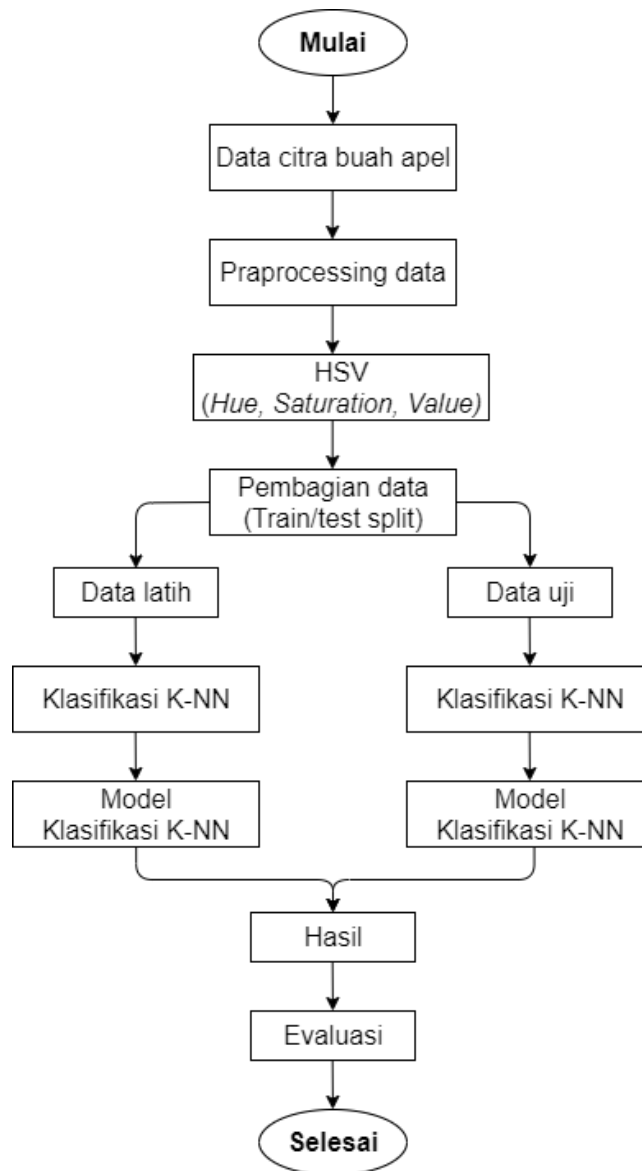
Keterangan :

$x$  = sample data

y = data uji  
D = Jarak

### 3. METODOLOGI PENELITIAN

Pada tahap ini tim peneliti akan mempelajari metode *K-NN* dengan ekstraksi fitur warna *HSV* kemudian mengumpulkan informasi yang diperoleh dari internet dan beberapa jurnal penelitian orang lain yang pernah dibuat sebelumnya sebagai gambaran untuk penelitian yang kami lakukan saat ini. Berikut bagan keseluruhan perancangan sistem yang dapat dilihat pada gambar dibawah :



**Gambar 1.** Rancangan metode yang diusulkan

### 3.1 Data Citra Buah Apel

Data yang digunakan dalam penelitian ini adalah 5 kelas jenis apel meliputi kelas *Apple Granny Smith*, *Apple Braeburn*, *Apple Crimson Snow*, *Apple Pink lady*, *Apple Red Delicious*. Data tersebut didapatkan melalui situs online Kaggle.com. Citra apel tersebut berformat JPG.

### 3.2 Praprosesing Data

Sebelum masuk ke tahap pembagian data, citra harus dilakukan praprosesing terlebih dahulu. Untuk mengetahui suatu citra, diperlukan adanya ekstraksi ciri. Ekstraksi ciri dapat diambil dari beberapa citra yang bisa menunjukkan ciri khas citra tersebut. Misal warna, pola citra, dan lain lain.

### 3.3 HSV (Hue Saturation Value)

Pada proses ini objek didalam citra akan menghitung objek yang berkaitan sebagai ciri. Dalam melakukan ekstraksi citra warna apel ini akan menggunakan metode *HSV (Hue Saturation Value)*. *HSV* mendefinisikan warna dalam terminologi *Hue*, *Saturation* dan *Value*.

### 3.4 Pembagian Data

Identifikasi citra apel pada penelitian ini dibagi menjadi dua yaitu data latih dan data uji.

#### 3.4.1 Data Latih

Citra data latih yaitu data yang berisi nilai yang digunakan untuk menentukan kelas yang cocok.

#### 3.4.2 Data Uji.

Citra data uji adalah data yang berisi data baru yang akan diklasifikasikan oleh model yang telah dibuat dan diketahui nilai akurasi dari proses klasifikasi.

Proses pembagian data menggunakan metode *train/test split* yang membagi keseluruhan dataset menjadi data latih dan data uji, data latih berisi keluaran yang sudah diketahui nilainya dan data uji akan digeneralisasikan ke data lainnya sehingga membentuk model.

### 3.5 Klasifikasi K-NN

*K-Nearest Neighbor (K-NN)* merupakan metode klasifikasi yang menentukan label (*class*) dari sebuah objek baru berdasarkan mayoritas *class* dari jarak terdekat  $k$  dalam kelompok data latih. Nilai  $k$  yang digunakan bernilai 3 yang digunakan dalam menggunakan metode K-NN.

### 3.6 Evaluasi

Pada tahapan ini yaitu tahapan untuk menganalisis dan mengevaluasi model yang telah diujikan pada penelitian ini. Proses perhitungan akurasi menggunakan rumus Confusion Matrix. Confusion Matrix merupakan table yang terdiri atas banyaknya baris dan data uji yang diprediksi benar dan tidak benar sesuai hasil dari klasifikasi.

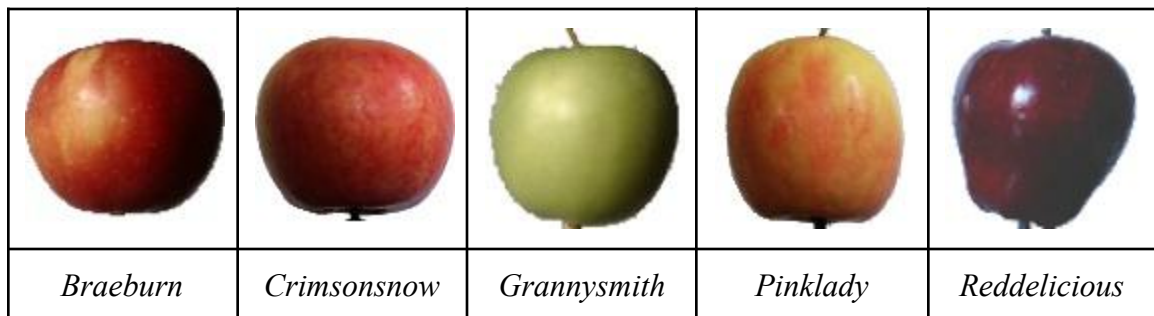
		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

**Gambar 2.** Tabel *Confussion Matrix*

#### 4. HASIL DAN PEMBAHASAN

##### 4.1 Data Citra Apel

Data citra apel yang digunakan dalam penelitian ini dibagi menjadi lima kelas yang terdiri dari *Apple Granny Smith*, *Apple Braeburn*, *Apple Crimson Snow*, *Apple Pink lady*, *Apple Red Delicious*. Tiap kelas terdiri dari 32 gambar dengan format JPG, dengan total semua gambar adalah 160 gambar.



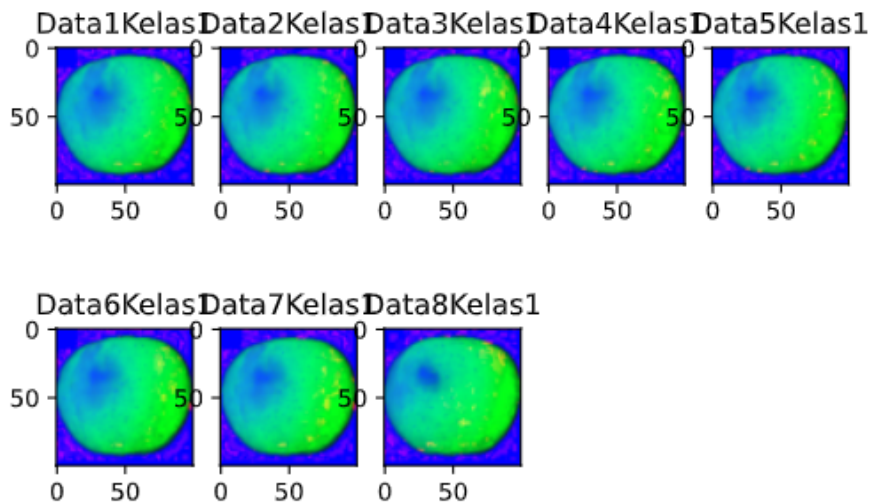
**Tabel 1.** Gambar apel dari tiap kelas

##### 4.2 Praprosessing Data

Pada tahap prosessing data dilakukan ekstraksi citra. Ekstraksi citra pada penelitian ini dengan cara mengambil warna pada buah apel untuk mengetahui jenis apel yang nantinya akan diuji.

##### 4.3 HSV (Hue Saturation Value)

Pada tahap ini gambar apel tiap kelas dilakukan ekstraksi citra apel dengan metode HSV (*Hue Saturation Value*). Pada tahap ini kami menggunakan kelas 1 dan hanya 8 data yang ditampilkan, kemudian didapat hasilnya sebagai berikut :



**Gambar 3.** Hasil ekstraksi HSV citra apel

#### 4.4 Pembagian Data

##### 4.4.1 Data Latih (Data Training)

Data yang digunakan untuk data latih adalah sebanyak 80% dari keseluruhan dataset maka data yang dipakai sebanyak 25 atau 26 gambar berformat JPG.

##### 4.4.2 Data Uji (Data Testing)

Data yang digunakan untuk data uji adalah sebanyak 20% dari keseluruhan dataset maka data yang dipakai sebanyak 6 gambar berformat JPG.

#### 4.5 Model Klasifikasi K-NN

Model pada gambar dibawah ini adalah hasil dari eksekusi data yaitu :

```
# Inisialisasi KNN (K = 3)
knn = KNeighborsClassifier(n_neighbors=3)
knn.fit(x_train, y_train) # Melatih Data (Train)
res = knn.predict(x_test) # Prediksi Data Test
```

**Gambar 4.** Model Klasifikasi KNN dengan k=3

Dengan mengklaster k=3 dari data 160 gambar yang terdiri dari 5 kelas buah apel.

#### 4.6 Hasil

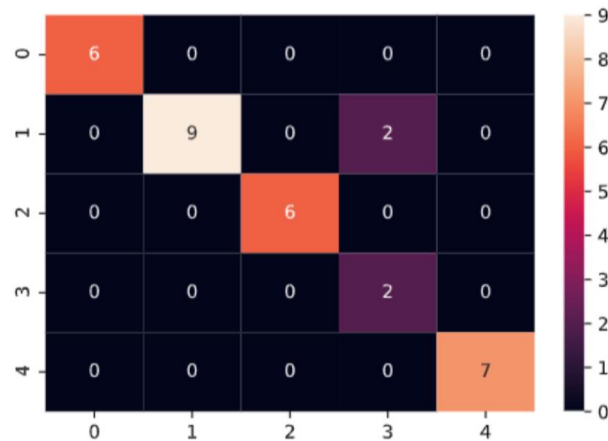
Klasifikasi apel menggunakan KNN dimana K adalah 3 didapat hasil sebagai berikut :

```
Hasil Klasifikasi Data Testing (KNN) = [1. 3. 3. 5. 5. 2. 1. 5. 1. 1. 4. 2. 2. 2. 4. 2. 5. 3. 3. 2. 2. 3. 2. 2.
2. 5. 1. 3. 1. 5. 5. 2.]
Kelas yang seharusnya benar (KNN)   = [1. 3. 3. 5. 5. 2. 1. 5. 1. 1. 4. 2. 2. 2. 4. 4. 2. 5. 3. 3. 2. 2. 3. 2. 2.
2. 5. 1. 3. 1. 5. 5. 4.]
```

**Gambar 5.** Hasil pengujian model

#### 4.7 Evaluasi

Tahap ini menampilkan total perhitungan dataset dengan *confusion matrix* dari hasil eksekusi data yaitu :



**Gambar 6.** Hasil *Confusion matrix*

Berdasarkan hasil output dari gambar 5, diketahui bahwa model dapat membedakan semua data untuk kelas 1 dengan sempurna. Tetapi pada kelas lain seperti kelas 3 masih terdapat sedikit kesalahan klasifikasi yang ditandai dengan *precision* dan *recall* yang hanya mencapai angka 2.

#### 5. KESIMPULAN

Ekstraksi fitur buah apel dilakukan dengan *HSV (Hue Saturation Value)*. Hasil ekstraksi ciri tersebut digunakan untuk mengklasifikasi jenis buah apel dengan metode *K-Nearest Neighbor (K-NN)*. Berdasarkan penelitian yang dilakukan, maka didapatkan hasil klasifikasi yang menunjukkan tingkat akurasi sebesar 93.75%. Hasil ini sudah maksimal karena data yang di uji sama dengan kelas yang seharusnya dan sudah sangat baik melihat juga pada nilai score *confusion matrix* yang tinggi menandakan bahwa sebaran hasil klasifikasi yang seimbang untuk kelas-kelas lainnya.

#### DAFTAR PUSTAKA

- [1] Arifah, Fina A'Yuni. Ivana. (2019). Potensi Buah Apel (*Malus domestica*) Dalam Mengatasi Penyakit Asma. *Proceeding of Biology Education*, 3(1), 208-212.
- [2] Ciputra, Antoni. dkk. (2018). Klasifikasi Tingkat Kematangan Buah Apel Manalagi Dengan Algoritma Naive Bayes Dan Ekstraksi Fitur Citra Digital. *Jurnal SIMETRIS*, Vol. 9, 465-472.
- [3] Wijaya, Novan., Anugrah R. (2019). Klasifikasi Jenis Buah Apel Dengan Metode K-Nearest Neighbors. *Jurnal SISFOKOM*, Volume 08, Nomor 01, 74-78.