

## Perbandingan Klasifikasi Naive-Bayes dan KNN untuk Mengidentifikasi Jenis Buah Apel dengan Ekstraksi Ciri LBP dan HSV

Felicia Febriana<sup>1</sup>, Laurenza Setiana Riva<sup>2</sup>, Rafael Salomo Sahat Piero<sup>3</sup>, Muhammad Aldy Ikramsyah<sup>4</sup>,  
Mayanda Mega Santoni<sup>5</sup>  
Program Studi S1 Informatika / Fakultas Ilmu Komputer  
Universitas Pembangunan Nasional Veteran Jakarta  
Jl. RS. Fatmawati, Pondok Labu, Jakarta Selatan, DKI Jakarta, 12460, Indonesia.  
feliciaf@upnvj.ac.id<sup>1</sup>, laurenzasr@upnvj.ac.id<sup>2</sup>, rafaelssp@upnvj.ac.id<sup>3</sup>, muhammadaik@upnvj.ac.id<sup>4</sup>,  
megasantoni@upnvj.ac.id<sup>5</sup>

**Abstrak.** Apel memiliki ribuan jenisnya yang terkadang membuat sebagian besar orang bingung. Warna pada apel pun bervariasi, ada yang berwarna merah, oren, ataupun hijau. Pada paper ini akan dibahas bagaimana tahapan dalam mengidentifikasi buah apel sehingga diketahui masuk ke dalam jenis apakah buah apel yang ingin diketahui. Akan digunakan 5 jenis buah apel yaitu Fuji, Manalagi, Red-Delicious, Granny-Smith, dan Rome Beauty. Untuk mengidentifikasi buah apel maka pertama akan dilihat dari ciri tekstur masing-masing buah apel yaitu memakai teknik LBP untuk diketahui pola-polanya dan teknik HSV untuk mengekstraksi ciri berdasarkan warna. Untuk mendapatkan hasil yang maksimal akan dilakukan perbandingan antara 2 algoritma klasifikasi, yaitu KNN dan Naive Bayes. Setelah dilakukan penelitian, maka didapatkan hasil LBP dengan KNN sebesar 90%, dan LBP dengan Naive Bayes sebesar 90%, HSV dengan KNN sebesar 82%, dan HSV dengan Naive Bayes sebesar 97%.

**Kata Kunci:** LBP, HSV, KNN, Naive Bayes

### 1 Pendahuluan

Apel merupakan salah satu dari berbagai jenis buah-buahan yang dapat dibudidayakan di Indonesia. Tidak hanya di Indonesia saja namun apel juga dapat ditemui di berbagai negara. Buah Apel kaya akan kandungan gizi seperti Vitamin A, Vitamin B6, Vitamin C, Vitamin E, Vitamin K, Kalsium, Zat Besi, dan tentunya masih banyak manfaat lainnya. Banyak manfaat yang ditawarkan apel bagi tubuh manusia seperti diantaranya melawan Alzheimer, mencegah kanker, menstabilkan gula darah, menurunkan resiko stroke, mengurangi resiko terserang penyakit jantung, dan banyak lainnya. Buah apel ini dapat dimakan secara langsung ataupun disajikan menjadi jus ataupun dijadikan sebagai cemilan seperti kerupuk apel. Banyak orang menyukai buah apel karena rasanya, baik yang memiliki rasa yang manis maupun yang sedikit asam tergantung dari jenis apel itu sendiri. Apel memiliki ribuan jenisnya yang terkadang membuat sebagian besar orang bingung. Warna pada apel pun bervariasi, ada yang berwarna merah, oren, ataupun hijau.

Pada paper ini akan dibahas bagaimana tahapan dalam mengidentifikasi buah apel sehingga diketahui masuk ke dalam jenis apakah buah apel yang ingin diketahui. Disini peneliti memakai data dari 5 jenis buah apel yaitu Fuji, Manalagi, *Red-Delicious*, *Granny-Smith*, dan *Rome Beauty*. Apel yang akan diidentifikasi akan diselidiki dia akan masuk ke jenis apa dari kelima jenis data apel yang sudah diketahui sebelumnya. Untuk mengidentifikasi buah apel maka pertama akan dilihat dari ciri tekstur masing-masing buah apel yaitu memakai teknik LBP untuk diketahui pola-polanya. Digunakan juga teknik HSV untuk mengekstraksi ciri berdasarkan warna. Setelah didapatkan data tekstur/warnanya kemudian tahapan selanjutnya diperlukan metode yang tepat yang mana apel yang diuji tersebut akan masuk ke dalam jenis yang benar. Disini peneliti akan membandingkan 2 algoritma dalam pengklasifikasian jenis buah apel, dimana hasil dari masing-masing evaluasi menentukan seberapa baik algoritma yang digunakan sehingga sangat akurat dalam mengidentifikasikannya. Algoritma yang digunakan antara lain adalah metode KNN dan Naive Bayes yang mana kedua algoritma ini sering digunakan dalam pengidentifikasian.

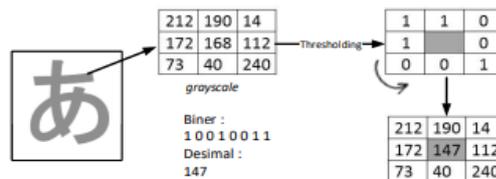
## 2 Tinjauan Pustaka

### 2.1 LBP

Metode LBP Local Binary Patterns merupakan metode analisis tekstur dengan menggunakan model statistik dan struktur [1]. LBP pertama kali diperkenalkan oleh Timo Ojala. Operator LBP memakai perbandingan nilai keabuan menurut piksel-piksel ketetanggaan. Operator dasar LBP memiliki ukuran 3 x 3 dan menggunakan 8 piksel ketetanggaan  $i_n$  dari sebuah piksel tengah  $i_c$ . Kemudian melakukan threshold pada piksel ketetanggaan ke- $n$  tersebut dengan menggunakan nilai keabuan dari piksel tengah seperti yang ditunjukkan pada persamaan (1) dan menggunakan fungsi thresholding  $s(x)$  seperti yang ditunjukkan pada persamaan (2). Kode binary output operator LBP piksel ketetanggaan akan dipakai untuk merepresentasikan fitur dari piksel tengah  $i_c$ . Gambar 1 menunjukkan diagram proses LBP.

$$LBP(x_c, y_c) = \sum_{n=0}^7 s(i_n - i_c)2^n \quad (1)$$

$$s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \quad (2)$$



**Gambar. 1.** Ilustrasi LBP

Pada Gambar 1, Proses pertama melibatkan penggunaan piksel tengah (1) untuk mengurangi piksel ketetanggaan. Setelah itu hasil pengurangan tersebut di-threshold menggunakan persamaan (2), apabila hasilnya  $\geq 0$  maka diberi nilai 1 & apabila hasilnya  $< 0$  maka diberi nilai 0. Setelah itu, nilai biner piksel ketetanggaan diurutkan berlawanan arah jarum jam, dan 8 bit biner diubah menjadi nilai desimal untuk menggantikan nilai piksel tengah  $i_c$  [2].

### 2.2 HSV

Fitur warna HSV mendefinisikan warna berdasarkan terminologi Hue, Saturation, dan Value. Hue (H) mewakili warna dasar, yang ditentukan oleh panjang gelombang dominan dalam distribusi spektrum panjang gelombang cahaya, dinyatakan dalam derajat, yaitu, 0 derajat merah, 120 derajat hijau, dan 240 derajat biru. Saturation (S) merupakan intensitas suatu warna atau bisa dikatakan kepekatan suatu warna. Semakin tinggi saturation suatu warna, maka semakin berwarna warna tersebut. Value (V) merupakan nilai yang menentukan tingkat kecerahan warna. Semakin cerah warna tersebut, maka semakin tinggi nilai value warna tersebut. Persamaan (6), Persamaan (7) dan Persamaan (8) di bawah ini digunakan untuk menghitung konversi RGB ke HSV:

$$H = \begin{cases} 0^\circ, & \text{jika } S = 0 \\ 60^\circ \times \left( \frac{G-B}{v_m} \bmod 6 \right), & \text{jika } V = R \\ 60^\circ \times \left( 2 + \frac{B-R}{v_m} \right), & \text{jika } V = G \\ 60^\circ \times \left( 4 + \frac{R-G}{v_m} \right), & \text{jika } V = B \end{cases} \quad (6)$$

$$S = \begin{cases} 0, & \text{jika } V = 0 \\ \frac{V_m}{V}, & \text{jika } V > 0 \end{cases} \quad (7)$$

$$V = \max (R, G, B) \quad (8)$$

dengan  $V_m$  yang diperoleh dengan menggunakan persamaan (9):

$$V_m = V - \min (R, G, B) \quad (9)$$

dimana pada RGB, lapisan pertama adalah  $R$  red, lapisan kedua adalah  $G$  green, dan lapisan ketiga adalah  $B$  blue [3]. Dari komponen warna Hue, Saturation, dan Value yang telah didefinisikan pada ruang warna HSV maka diperoleh nilai fitur HSV [4].

### 2.3 Naive Bayes

Klasifikasi Bayesian adalah metode pengklasifikasian statistik yang dapat digunakan untuk memprediksi probabilitas keanggotaan kelas. Dasar dari Klasifikasi Bayesian yaitu teorema Bayes yang mana kemampuan klasifikasinya mirip dengan Decision Tree dan Jaringan Saraf Tiruan. Klasifikasi Bayesian telah terbukti memiliki presisi dan kecepatan yang tinggi ketika diterapkan pada data yang besar [9]. Metode Bayes melakukan inferensi dengan penalaran dari yang khusus ke yang umum (induksi) terhadap masalah klasifikasi sebagai metode pendekatannya. Membahas tentang konsep dasar dan definisi pada Teorema Bayes terlebih dahulu, kemudian teorema ini digunakan untuk melakukan klasifikasi dalam Data Mining. Teorema Bayes mempunyai bentuk umum seperti berikut:

$$P(H | X) = \frac{P(X|H)P(H)}{P(X)} \quad (1)$$

Keterangan:

- X = Data dengan kelas yang belum diketahui
- H = Hipotesis data X merupakan suatu kelas spesifik
- $P(H|X)$  = Probabilitas hipotesis H berdasarkan kondisi x (posteriori prob.)
- $P(H)$  = Probabilitas hipotesis H (prior prob.)
- $P(X|H)$  = Probabilitas X berdasarkan kondisi tersebut
- $P(X)$  = Probabilitas dari X [5]

### 2.4 KNN

Algoritma K-Nearest Neighbor adalah metode klasifikasi yang mengelompokkan suatu data berdasarkan jarak data tersebut ke beberapa data terdekat (tetangga/neighbors). Algoritma KNN menggunakan algoritma supervised learning untuk mengelompokkan hasil dari data yang baru lalu diklasifikasikan dengan berlandaskan mayoritas kategori dari KNN. Nearest Neighbor itu sendiri menghitung seberapa dekat jarak antara kasus baru dengan kasus lama dengan menggunakan perbandingan bobot dari fitur-fitur yang ada sebagai metode pendekatannya. Berikut adalah rumus penghitungan kemiripan antar dua kasus:

$$\text{Similarity} (T, S) = \frac{\sum_{i=1}^n f(T_i, S_i) * w_i}{w_i} \quad (1)$$

Keterangan:

- S : Data Kasus lama
- T : Data Kasus baru
- n : jumlah suku dalam kasus
- i : indeks batas 1 s/d n

f : fungsi similaritas atribut indeks ke-i antara data kasus T dan S  
w: bobot atribut pada indeks ke-i

## 2.5 Evaluasi

Dalam pengertian umum, evaluasi dapat dipahami sebagai proses sistematis dalam menentukan nilai sesuatu (kondisi, kegiatan, keputusan, kinerja, proses, orang, objek, dll.) menurut kriteria tertentu melalui penilaian. Evaluator dapat langsung menentukan nilai sesuatu dengan cara membandingkannya dengan kriteria umum, mereka juga dapat mengukur terhadap sesuatu yang dievaluasi kemudian membandingkan dengan kriteria tertentu [7]. Evaluasi adalah suatu metode atau proses yang bertujuan untuk memperoleh nilai yang terukur.

Cross-validasi atau dapat disebut estimasi rotasi adalah sebuah teknik validasi model yang digunakan untuk mengevaluasi bagaimana hasil statistik analisis akan menggeneralisasi kumpulan data independen. Metode ini digunakan untuk memprediksi model memperkirakan keakuratan model prediksi selama dijalankan dalam praktiknya. Salah satu teknik dari validasi silang yang membagi data menjadi k bagian set data dengan ukuran yang sama adalah k-fold cross validation. K-fold cross validation digunakan untuk menghilangkan bias pada data. Pelatihan dan pengujian dilakukan sebanyak k kali. (Bramer, 2007).

Untuk mengukur kinerja klasifikasi, yaitu membandingkan semua data uji yang diklasifikasi benar dengan banyaknya data uji. Persamaan dibawah ini adalah model yang digunakan untuk mengukur kinerja klasifikasi.

$$akurasi = \frac{\sum \text{klasifikasi benar}}{\sum \text{data uji}} \times 100\% \quad (1)$$

## 2.6 Penelitian Terdahulu

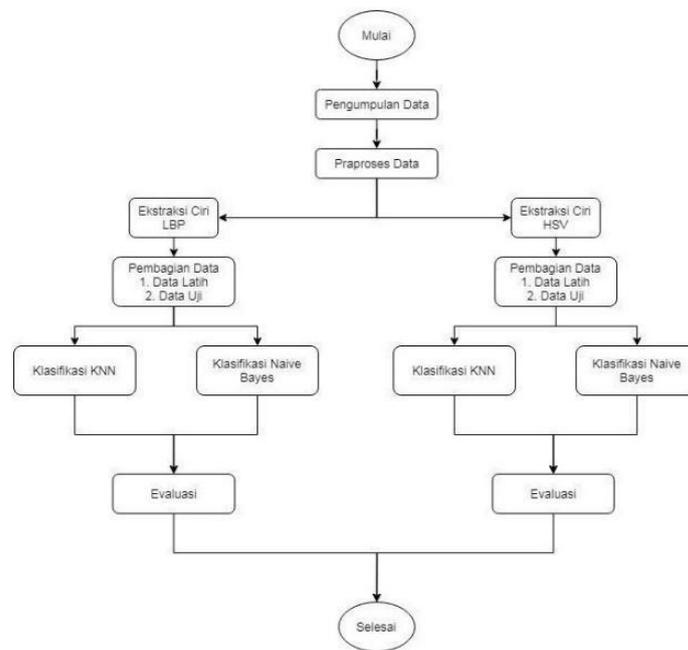
Terdapat penelitian terlebih dahulu yang relevan terhadap penelitian ini. Penelitian-penelitian tersebut memiliki kesamaan dan perbedaan terhadap penelitian ini. Peneliti menggunakan 2 penelitian yang relevan sebagai acuan bagi peneliti untuk membuat karya ilmiah ini. Penelitian yang dimaksud adalah penelitian dengan judul “Klasifikasi Jenis Buah Apel Dengan Metode K-Nearest Neighbors” karya Novan Wijaya dan Anugrah Ridwan [8]. Terdapat Perbedaan antara penelitian tersebut dengan penelitian ini yaitu adalah penelitian ini menggunakan 2 metode klasifikasi, yaitu KNN dan Naive Bayes dengan ekstraksi ciri LBP dan HSV yang digunakan secara terpisah, sedangkan penelitian tersebut hanya menggunakan metode KNN dengan ekstraksi ciri LBP dan HSV yang digunakan secara bersamaan.

Penelitian lain yang relevan dengan penelitian ini berjudul “Analisis Tekstur Dan Ekstraksi Fitur Warna Untuk Klasifikasi Apel Berbasis Citra” karya Arie Qur’ania, Lita Karlitasar, dan Sufiatul Maryana [9]. Sedangkan perbedaannya adalah penelitian ini menggunakan 2 metode klasifikasi, yaitu KNN dan Naive Bayes, sedangkan penelitian tersebut hanya menggunakan metode KNN. Selain itu, ekstraksi ciri yang digunakan berbeda. Penelitian ini menggunakan ekstraksi ciri LBP dan HSV, penelitian tersebut menggunakan analisis tekstur dan ekstraksi ciri warna.

Walaupun terdapat perbedaan, tetapi 2 penelitian diatas memiliki kesamaan dengan penelitian ini. Persamaan tersebut adalah objek yang digunakan sama-sama buah apel, dan juga tujuan penelitiannya adalah untuk mengklasifikasikan jenis buah apel dengan banyaknya kelas adalah 5.

## 3 Metodologi Penelitian

Pada bab ini akan dijelaskan bagaimana tahapan-tahapan penelitian yang mana bisa dilihat pada Gambar 2.



**Gambar. 2.** Tahap Penelitian

Dari flowchart diatas, kami memulai dengan mengumpulkan data citra yang akan diolah. Setelah data dikumpulkan, dilakukan praproses data agar data lebih bersih dan dapat digunakan. Setelah itu data citra diekstrak ciri LBP dan HSV nya. Hasil dari ekstraksi ciri itu akan dibagi lagi ke data latih dan data uji. Kemudian data latih dan data uji tersebut digunakan untuk melakukan klasifikasi KNN dan Naive Bayes. Terakhir hasil dari klasifikasi tersebut dievaluasi.

### 3.1 Mengumpulkan Data

Dikarenakan penelitian ini dilakukan di masa pandemi, maka pengumpulan data penelitian ini dilakukan secara daring atau online. Peneliti menggunakan dataset Apel yang sudah ada sebelumnya di internet untuk meminimalisir mobilitas dalam pengumpulan data. Adapun dataset yang peneliti pakai adalah dataset 5 jenis buah apel hingga nantinya akan membentuk 5 kelas data. 5 jenis buah apel yang dimaksud adalah manalagi, fuji, *red-delicious*, *granny-smith*, dan *rome-beauty*.

### 3.2 Pra Proses Data

Setelah semua data telah terkumpul, maka proses selanjutnya adalah praproses data. Pra proses data yang dilakukan adalah dengan menghilangkan background, melakukan resize citra, dan pada proses LBP dilakukan perubahan warna citra menjadi Grayscale.

### 3.3 Ekstraksi Ciri Tekstur Menggunakan LBP

Dalam penelitian ini, untuk mengekstraksi ciri apel berdasarkan tekstur, maka dilakukan ekstraksi ciri LBP (Local Binary Pattern). LBP terbilang efektif dalam mendeteksi tekstur, karena LBP tidak terpengaruh pencahayaan yang tidak merata.

### 3.4 Ekstraksi Ciri Warna Menggunakan HSV

Dalam penelitian ini, untuk mengekstraksi ciri apel berdasarkan warna, maka dilakukan ekstraksi ciri HSV (*Hue Saturation Value*). Dalam proses HSV, data citra yang awalnya RGB akan diubah terlebih dahulu ke dalam bentuk HSV. Setelah itu akan didefinisikan masing-masing komponen, yaitu *Hue*, *Saturation*, dan *Value*.

### 3.5 Pembagian Data

Data testing sebagai data yang menguji model untuk simulasi penggunaan model di dunia nyata, data training sebagai data yang melatih model, dan data validasi sebagai data yang digunakan untuk proses validasi model dan menghindari terjadinya overfitting.

### 3.6 Klasifikasi KNN

Klasifikasi KNN adalah metode klasifikasi sekumpulan data yang telah dilatih sebelumnya untuk menentukan label pada data uji menggunakan mayoritas jarak terdekat dan untuk mengukur kinerja klasifikasi KNN pada penelitian ini menggunakan confusion matrix.

### 3.7 Klasifikasi Naive Bayes

Klasifikasi Naive Bayes merupakan metode klasifikasi yang menggunakan probabilitas dan teorema bayesian. Dimana terdapat data yang sudah diklasifikasikan, sehingga dapat menentukan label (class) data uji. Dan untuk mengukur kinerja klasifikasi Naive bayes pada penelitian ini menggunakan confusion matrix.

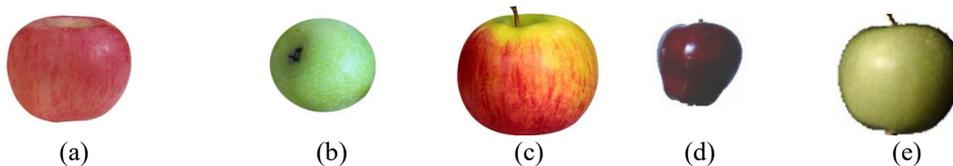
### 3.8 Evaluasi

Setelah hasil dari klasifikasi KNN dan Naive Bayes telah didapat, maka langkah selanjutnya adalah mengevaluasi hasil klasifikasi tersebut. Manakah diantara klasifikasi KNN dan Naive Bayes yang lebih baik berdasarkan ekstraksi ciri LBP/HSV. Klasifikasi yang memiliki angka yang lebih tinggi berarti dalam percobaan ini lebih baik daripada klasifikasi lawannya, karena memiliki tingkat keakuratan yang lebih tinggi.

## 4 Hasil

### 4.1 Pengumpulan Data

Pada penelitian kali ini digunakan dataset Apel yang sudah ada sebelumnya. Sample dataset yang telah kami kumpulkan, dapat dilihat pada Gambar 3.



**Gambar. 3.** (a) Apel Fuji, (b) Apel Manalagi, (c) Apel Rome-Beauty, (d) Apel Red-Delicious, dan (e) Apel Granny-Smith

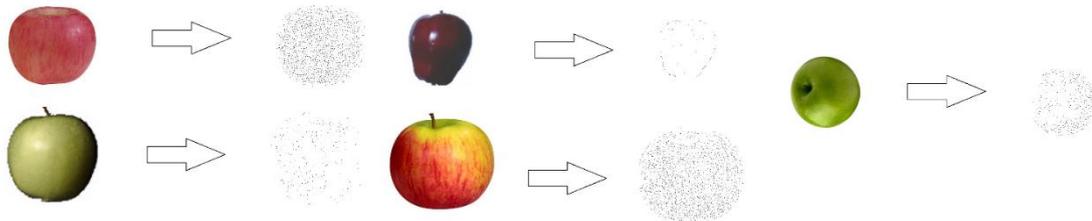
Dataset yang peneliti pakai sebanyak 20 buah apel pada setiap jenisnya. Karena terdapat 5 jenis apel maka didapat keseluruhan data adalah 100 buah.

### 4.2 Pra Proses

Untuk pra proses, peneliti melakukan penghilangan background dan resize data. Untuk melakukan pra proses digunakan software pendukung, yaitu Photoshop. Untuk menyelaraskan ukuran dari setiap data, dalam penelitian ini dilakukan resize pada setiap data apel sehingga berukuran 640 x 480 agar mengurangi terjadinya kekeliruan dalam pengidentifikasiannya. Dan untuk proses ekstraksi citra LBP dilakukan perubahan warna citra yang sebelumnya RGB menjadi Grayscale.

#### 4.3 Ekstraksi Ciri Tekstur Menggunakan LBP

Peneliti menggunakan metode LBP untuk ekstraksi ciri berdasarkan tekstur. Awal tahap ini adalah dengan mengubah citra RGB yang sudah dipisahkan dari backgroundnya menjadi citra grayscale, selanjutnya akan masuk ke tahap klasifikasi sehingga diubah dahulu ke histogram matriks kemudian diubah ke bentuk histogram vektor dan siap di klasifikasi. Hasil dari ekstraksi ciri tekstur LBP dapat dilihat pada Gambar 4.

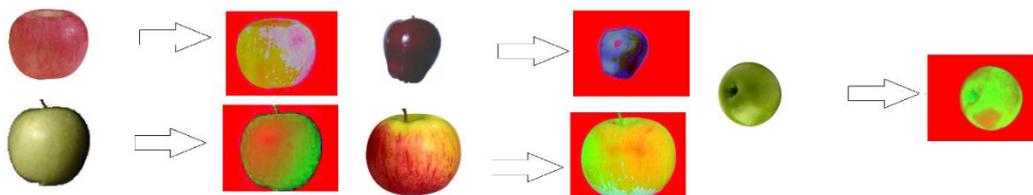


**Gambar. 4.** Hasil LBP

Dataset tersebut ditempatkan secara berurut dari 20 dataset Fuji, 20 dataset Manalagi, 20 dataset Red Delicious, 20 dataset Rome Beauty, dan 20 dataset Granny Smith.

#### 4.4 Ekstraksi Ciri Warna Menggunakan HSV

Peneliti menggunakan metode HSV untuk ekstraksi ciri berdasarkan warna. Awal tahap ini adalah dengan mengubah citra RGB yang sudah dipisahkan dari backgroundnya menjadi citra HSV, selanjutnya sama seperti LBP untuk bisa masuk ke dalam tahap klasifikasi maka diubah dahulu ke histogram matriks baru diubah lagi menjadi histogram vektor yang kemudian siap di klasifikasi. Hasil dari ekstraksi ciri tekstur HSV dapat dilihat pada Gambar 5.



**Gambar. 5.** Nilai HSV

Dataset tersebut ditempatkan secara berurut dari 20 dataset Fuji, 20 dataset Manalagi, 20 dataset Red Delicious, 20 dataset Rome Beauty, dan 20 dataset Granny Smith.

#### 4.5 Pembagian Data

Dari keseluruhan data yang berjumlah 100, data tersebut akan dibagi kedalam 3 bagian yaitu data latih, data evaluasi, dan data uji. Banyaknya masing-masing total data latih, data evaluasi, dan data uji secara berturut-turut adalah 80, 10, dan 10.

#### 4.6 Klasifikasi dan Evaluasi

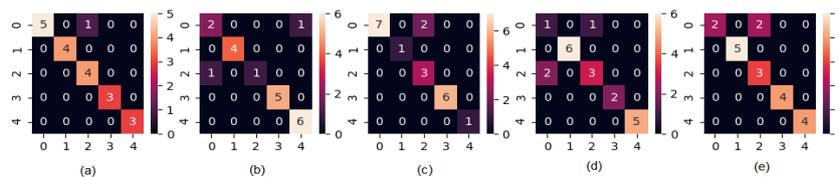
#### 4.6.1 LBP dengan Klasifikasi KNN dan Naïve Bayes

Hasil Klasifikasi berdasarkan atribut LBP nya yaitu memakai radius 3 dan n point =  $8 \times 3$  yaitu 24. Pada Tabel 1 merupakan hasil dari membandingkan nilai k dari 1 sampai 10 dengan fold 1 sampai 5.

**Tabel 1.** Perbandingan nilai k

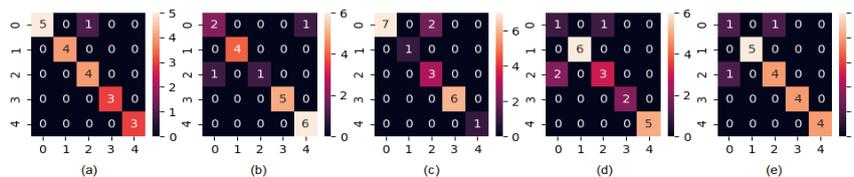
Nilai K	Fold 1					Rata2
	1	2	3	4	5	
1	0.95	0.90	0.80	0.85	0.90	0.88
2	0.95	0.90	0.90	0.85	0.90	0.9
3	0.90	0.90	0.85	0.90	0.80	0.87
4	0.80	0.95	0.80	0.85	0.90	0.86
5	0.80	0.90	0.80	0.90	0.90	0.86
6	0.85	0.85	0.80	0.90	0.90	0.86
7	0.90	0.85	0.80	0.85	0.90	0.86
8	0.90	0.85	0.85	0.85	0.90	0.87
9	0.90	0.85	0.80	0.85	0.90	0.86
10	0.85	0.85	0.85	0.85	0.85	0.85

Dari Tabel 1 diatas, didapati bahwa hasil k dari 1-10 disetiap fold dirata-ratakan kemudian diambil nilai tertinggi dari rata-rata tersebut yaitu 0.9 pada k yaitu 2. Kemudian tampilkan matriks confusion menggunakan model k=2, Matriks confusion dapat dilihat pada Gambar 6.



**Gambar 6.** Confusion Matrix KNN pada LBP. (a) Fold 1, (b) Fold 2, (c) Fold 3, (d) Fold 4, dan (e) Fold 5

Untuk Naïve bayes, rata-rata dari akurasi adalah 0.90 atau 90%. Untuk matriks confusionnya dapat dilihat pada Gambar 7.



**Gambar 7.** confusion matrix NB pada LBP. (a) Fold 1, (b) Fold 2, (c) Fold 3, (d) Fold 4, dan (e) Fold 5

Selanjutnya peneliti membandingkan hasil setiap fold dan rata-rata akurasi antara KNN dengan k adalah 2 dan Naïve bayesnya. Pada Tabel 2 merupakan perbandingan dari KNN dan Naive Bayes.

**Tabel 2.** Perbandingan KNN dan NB pada LBP

Fold	KNN	NB
1	0.95	0.95
2	0.90	0.90
3	0.90	0.90
4	0.85	0.85
5	0.90	0.90
Rata-rata	0.90	0.90

Dari Tabel 2 diatas terlihat bahwa nilai rata-rata dari akurasi menggunakan metode KNN dengan k=2 adalah 0.9 atau 90% dan jika dengan menggunakan metode Naïve bayes adalah sebesar 0.90 atau 90%. Dari hasil akhir

tersebut yang mana adalahimbang maka baik penggunaan dengan metode naïve bayes maupun dengan KNN sama-sama memberikan hasil yang optimal.

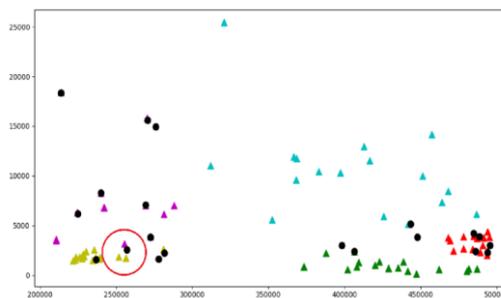
#### 4.6.2 HSV dengan Klasifikasi KNN dan Naïve Bayes

Pada Tabel 3 merupakan hasil dari membandingkan nilai k dari 1 sampai 10 dengan fold 1 sampai 5.

**Tabel 3.** Perbandingan nilai k

Nilai K	Fold 1					Rata2
	1	2	3	4	5	
1	0.85	0.90	0.85	0.80	0.95	0.87
2	0.75	0.85	0.80	0.80	0.90	0.82
3	0.70	0.75	0.75	0.80	0.95	0.79
4	0.75	0.75	0.75	0.75	0.90	0.78
5	0.75	0.75	0.80	0.80	0.80	0.78
6	0.70	0.75	0.70	0.75	0.85	0.75
7	0.65	0.75	0.70	0.85	0.85	0.76
8	0.65	0.65	0.70	0.85	0.85	0.74
9	0.65	0.65	0.75	0.85	0.85	0.75
10	0.65	0.65	0.75	0.80	0.75	0.72

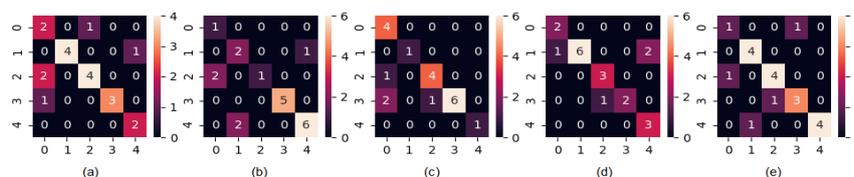
Dari Tabel 3 diatas, didapati bahwa hasil k dari 1-10 disetiap fold dirata-ratakan kemudian diambil nilai tertinggi dari rata-rata tersebut yaitu 0.87 pada k yaitu 1. Namun penggunaan k=1 itu kurang optimum, dapat dilihat scatternya pada Gambar 8.



**Gambar 8.** Scatter

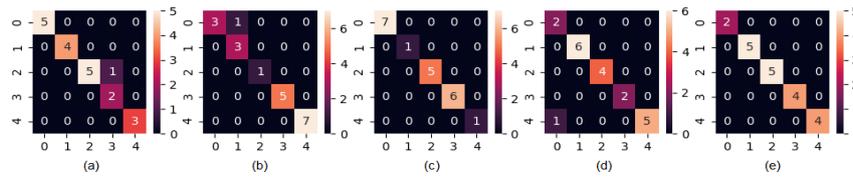
Gambar 8 diatas merupakan salah satu fold yang menampilkan scatter dari data apel dimana untuk data training dilambangkan dengan segitiga dan untuk ujiinya dilambangkan dengan lingkaran hitam. Jika digunakan k=1 terkadang kurang optimal karena yang dibandingkan hanya dengan 1 data dan belum tentu yang didekatnya adalah cluster yang sama. Perhatikan bulat merah, karena mengambil 1 tetangga terdekat maka sistem membacanya dia sebagai kelas ungu padahal jika tetangganya diperluas sedikit maka didapatkan kelas kuning yang mana lebih menyebar di sekelilingnya. Sehingga dari Gambar 8 diatas dapat disimpulkan bahwa penggunaan KNN = 1 kurang optimal jika digunakan sehingga sebaiknya dihindari.

Peneliti menggunakan k optimal kedua yaitu 2 yang kemudian ditampilkan matriks confusionnya pada Gambar 9.



**Gambar 9.** Confusion Matrix KNN pada HSV. (a) Fold 1, (b) Fold 2, (c) Fold 3, (d) Fold 4, dan (e) Fold 5

Untuk Naïve bayes, rata-rata dari akurasi adalah 0.97 atau 97%. Matriks confusionnya dilihat pada Gambar 10.



**Gambar. 10.** Confusion Matrix NB pada HSV. (a) Fold 1, (b) Fold 2, (c) Fold 3, (d) Fold 4, dan (e) Fold 5

Selanjutnya peneliti membandingkan hasil setiap fold dan rata-rata akurasi antara KNN dengan k adalah 2 dan Naïve bayesnya. Pada Tabel 4 merupakan perbandingan dari metode KNN dan Naive Bayes.

**Tabel. 4.** Perbandingan KNN dan NB pada HSV

Fold	KNN	NB
1	0.75	0.95
2	0.85	0.95
3	0.80	1.00
4	0.80	0.95
5	0.90	1.00
Rata-rata	0.82	0.97

Dari Tabel 4 diatas terlihat bahwa nilai rata-rata dari akurasi menggunakan metode KNN dengan k=2 adalah 0.82 atau 82% dan jika dengan menggunakan metode Naïve bayes adalah sebesar 0.97 atau 97%. Terdapat perbedaan yang cukup signifikan pada saat menggunakan dengan metode naïve bayes dimana rata-rata akurasinya lebih optimal digunakan jika dibandingkan dengan menggunakan metode KNN.

## 5 Kesimpulan Dan Saran

Dari hasil penelitian ini, maka dapat disimpulkan penggunaan metode klasifikasi Naive bayes lebih baik jika dibandingkan dengan menggunakan KNN baik dengan ekstraksi ciri tekstur menggunakan LBP ataupun ekstraksi ciri warna menggunakan HSV. Pada LBP, klasifikasi menggunakan Naive Bayes dan KNN memberikan hasil yang sama pada akurasinya yaitu sebesar 0.90 atau 90%. Pada HSV, klasifikasi menggunakan Naive Bayes dan KNN memberikan hasil yang lumayan signifikan dengan hasil akurasi menggunakan dari Naive Bayes lebih besar yaitu sebesar 0.97 atau 97% dibandingkan dengan KNN yaitu sebesar 0.82 atau 82%.

Terdapat berbagai masalah yang mungkin saja ditimbulkan sehingga menurunkan keakuratan adalah seperti dari segi kualitas gambar walau peneliti sudah menyamakan ukuran pixel data citra namun tetap ada data citra yang memiliki kualitas yang buruk seperti data yang blur, juga dari segi warna yaitu adanya perbedaan kontras yang mana saat pengambilan gambar memakai kamera yang berbeda-beda dan intensitas cahaya yang berbeda pula.

Sebaiknya menggunakan data citra dengan tingkat kecerahan dan kualitas gambar yang sama. Untuk memudahkan lebih baik mengambil langsung objek citra sehingga keseragamannya dapat langsung diatur. Namun jika harus mengambil dataset dari internet, diusahakan untuk mengambil dari 1 sumber saja.

## Referensi

- [1] Kurniawardhani, Arrie, Nanik Suciati,& Isye Arieshanti.2014.Jurnal."KLASIFIKASI CITRA BATIK MENGGUNAKAN METODE EKSTRAKSI CIRI YANG INVARIANT TERHADAP ROTASI",JUTI: Jurnal Ilmiah Teknologi Informasi - Volume 12, Nomer 2. Dalam <http://juti.if.its.ac.id/index.php/juti/article/view/322/269> diakses pada 27 Juni 2021.
- [2] Amat, Rabiuldien, Jayanti Yusmah Sari, Ika Purwanti Ningrum.2017.Jurnal."IMPLEMENTASI METODE LOCAL BINARY PATTERNS UNTUK PENGENALAN POLA HURUF HIRAGANA DAN KATAKANA PADA SMARTPHONE",JUTI: Jurnal Ilmiah Teknologi Informasi - Volume 15, Nomor 2.Dalam [uti.if.its.ac.id/index.php/juti/article/view/612/344](http://juti.if.its.ac.id/index.php/juti/article/view/612/344) diakses pada 27 Juni 2021.
- [3] R. H. Sianipar. Dasar Pemrosesan Citra Digital dengan Matlab. Yogyakarta: ANDI, 2018.

- [4] Auliasari, Rahma Nur, Ledy Novamizanti, & Nur Ibrahim. 2020. Jurnal. "Identifikasi Kematangan Daun Teh Berbasis Fitur Warna Hue Saturation Intensity (HSI) dan Hue Saturation Value (HSV)", JUITA: Jurnal Informatika e-ISSN: 2579-8901; Volume 8, Nomor 2. Dalam <http://jurnalnasional ump.ac.id/index.php/JUITA/article/view/7387/3566> diakses pada 27 Juni 2021.
- [5] Annur, Haditsah. 2018. Jurnal. "KLASIFIKASI MASYARAKAT MISKIN MENGGUNAKAN METODE NAÏVE BAYES". Dalam <https://jurnal.fikom.umi.ac.id/index.php/ILKOM/article/view/303/142> diakses pada 24 Juni 2021.
- [6] Rohman, Abdul. 2015. Jurnal. "MODEL ALGORITMA K-NEAREST NEIGHBOR (K-NN) UNTUK PREDIKSI KELULUSAN MAHASISWA". Dalam <http://jurnal.unpand.ac.id/index.php/NT/article/view/350> diakses pada 24 Juni 2021.
- [7] L., Idrus. 2019. Jurnal. "EVALUASI DALAM PROSES PEMBELAJARAN". Dalam <https://jurnal.iain-bone.ac.id/index.php/adara/article/download/427/352> diakses pada 24 Juni 2021.
- [8] Wijaya, N., & Ridwan, A. 2019. Jurnal. "Klasifikasi Jenis Buah Apel Dengan Metode K-Nearest Neighbors". Dalam [https://www.researchgate.net/publication/332276594\\_KLASIFIKASI\\_JENIS\\_BUAH\\_APEL\\_DENGAN\\_METODE\\_K-NEAREST\\_NEIGHBORS\\_DENGAN\\_EKSTRAKSI\\_FITUR\\_HSV\\_DAN\\_LBP](https://www.researchgate.net/publication/332276594_KLASIFIKASI_JENIS_BUAH_APEL_DENGAN_METODE_K-NEAREST_NEIGHBORS_DENGAN_EKSTRAKSI_FITUR_HSV_DAN_LBP) diakses pada 3 Juni 2021.
- [9] Qur'ania, A., Karlitasar, L., & Maryana, S. 2018. Jurnal. "ANALISIS TEKSTUR DAN EKSTRAKSI FITUR WARNA UNTUK KLASIFIKASI APEL BERBASIS CITRA". Dalam <https://docplayer.info/31373019-Analisis-tekstur-dan-ekstraksi-fitur-warna-untuk-klasifikasi-apel-berbasis-citra-arie-qur-ania-lita-karlitasar-sufiatul-maryana.html> diakses pada 3 Juni 2021.
- [10] Tempola, Firman, Miftah Muhammad, & Amal Khairan. 2018. Jurnal. "PERBANDINGAN KLASIFIKASI ANTARA KNN DAN NAIVE BAYES PADA PENENTUAN STATUS GUNUNG BERAPI DENGAN K-FOLD CROSS VALIDATION". Dalam <https://jtiik.ub.ac.id/index.php/jtiik/article/view/983> diakses pada 2 Juli 2021.