

Penggunaan K-Nearest Neighbor (KNN) untuk Mengklasifikasi Citra Belimbing Berdasarkan Fitur Warna

Duwen Imantata Muhammad¹, Ermatita², Noor Falih³
Program Studi Informatika / Fakultas Ilmu Komputer
Universitas Pembangunan Nasional Veteran Jakarta
Jl. RS. Fatmawati, Pondok Labu, Jakarta Selatan, DKI Jakarta, 12460, Indonesia.
dwnimnt12@gmail.com¹, ermatita@upnvj.ac.id², falih@upnvj.ac.id³

Abstrak - Masih banyak yang belum mengetahui pasti tingkat kematangan buah. Akibatnya penjual maupun pembeli menjadi sulit untuk memperkirakan tingkat kematangan buah tersebut, khususnya buah belimbing. Berawal dari masalah tersebut dibutuhkan suatu sistem yang dapat membedakan tingkat kematangan dari buah. Berdasarkan hal tersebut tujuan penelitian ini dilakukan guna mengidentifikasi tingkat kematangan buah belimbing berdasarkan citra dengan algoritma K-Nearest Neighbor dan dengan menggunakan metode ekstraksi ciri Hue saturation Value (HSV) guna merubah nilai RGB kedalam nilai HSV dari suatu citra dan dengan menggunakan program Matlab guna membantu proses pengolahan citra digital. Dengan menggunakan algoritma KNN didapatkan akurasi sebesar 93.33% pada percobaan dengan menggunakan nilai K=7.

Kata kunci: Citra, Matlab, *K-Nearest Neighbor*, *Hue saturation Value* (HSV), Buah Belimbing

1. PENDAHULUAN

Belimbing merupakan salah satu buah yang tidak jarang ditemukan di suatu pasar tradisional maupun pinggir jalan sekalipun lebih tepatnya di kota Depok, karena Depok merupakan kota belimbing. Belimbing juga memiliki nilai jual di Indonesia dan memiliki porsi pasar yang cukup luas. Dapat dikatakan bahwa buah belimbing banyak dikonsumsi dan digemari oleh penduduk di Indonesia, juga memiliki daya tarik dan daya saing.

Untuk menaikkan daya tarik dan daya saing ini maka buah belimbing yang dihasilkan harus sesuai dengan standar pasar agar dapat diterima oleh konsumen. Belimbing sudah terdaftar dalam Standar Nasional Indonesia dengan nomor SNI 4491:2009 oleh Badan Standarisasi Nasional.

Selama ini pengklasifikasian kualitas belimbing hanya dilakukan dengan cara pengamatan manual pada permukaan luar buah secara langsung. Pengklasifikasian dengan cara manual menghasilkan klasifikasi yang kurang akurat dan tidak stabil yang disebabkan karena adanya kelalaian manusia. Sehingga dibutuhkan cara untuk dapat membedakan tingkat kematangan belimbing.

Salah satu cara untuk membedakan tingkat kematangan belimbing dalam bidang informatika yaitu dengan mengenali ciri dari belimbing tersebut menggunakan pengolahan citra. Pada penelitian ini digunakan Algoritma klasifikasi K-Nearest Neighbor (KNN) dan dengan menggunakan ekstraksi ciri HSV (hue, saturation, value) untuk mendapatkan ciri dari tingkat kematangan belimbing.

Algoritma K-Nearest Neighbor (KNN) merupakan sebuah teknik klasifikasi terhadap objek berdasarkan nilai (K) tetangga terdekatnya. Dan ekstraksi ciri HSV (hue, saturation, value) digunakan untuk mengekstraksi warna untuk membantu proses klasifikasi tingkat kematangan belimbing. Fitur warna digunakan guna mempermudah dalam pengolahan citra.

Terdapat banyak penelitian terdahulu yang bersangkutan dengan dilakukannya penelitian ini seperti penelitian Fuzy Yustika Manik dan Kana Saputra Saragih, 2017. Penelitian ini memiliki tujuan mengklasifikasi citra belimbing berdasarkan tingkat kemanisannya. Nilai dari RGB digunakan sebagai cara ekstraksi fitur untuk memperoleh ciri warna pada citra. Selanjutnya menggunakan Naïve Bayes untuk klasifikasi hasil dari ekstraksi fitur ini. Data citra buah belimbing yang digunakan

sebanyak 120 yang terbagi menjadi dua yaitu data latih sebanyak 90 dan data uji sebanyak 30. Akurasi yang didapatkan dari klasifikasi dengan ekstraksi fitur RGB yaitu sebesar 80%.

2. TINJAUAN PUSTAKA

Buah Belimbing

Terdapat dua jenis belimbing, yaitu belimbing manis yang dapat dimakan langsung (*Averrhoa carambola L.*) serta belimbing wuluh atau belimbing sayur (*Averrhoa belimbi L.*). Seperti contohnya belimbing wuluh yang digunakan dalam suatu masakan sebagai bumbu dapur, terlebih untuk memberikan rasa dan aroma pada masakan yang dibuat [6].

Citra

Citra merupakan representasi dua dimensi dari dunia visual. Citra dapat berupa gambar dua dimensi seperti foto dan lukisan maupun tiga dimensi seperti patung [1]. Pada penelitian ini penulis menggunakan citra digital. Citra yang dapat diproses oleh computer merupakan citra digital, suatu citra digital dapat diartikan juga sebagai sebuah matriks yang terdiri dari M kolom dan N baris, dimana perpotongan antara kolom dengan baris disebut pixel.

Pengolahan Citra

Teknik yang dapat digunakan untuk memanipulasi atau memproses suatu citra dalam bentuk dua dimensi dikenal dengan pengolahan citra. Pengolahan citra dapat juga diartikan sebagai segala bentuk operasi baik itu untuk memperbaiki, menganalisa, atau juga mengubah suatu gambar [2].

Citra RGB

Jenis citra ini memiliki spesifikasi yang lebih kompleks jika dibandingkan dengan citra jenis citra *grayscale*, hasil yang diperoleh dari jenis citra ini sangat menyerupai dengan warna objek asli. Rentang warna yang sangat luas dalam pencampuran warna dapat ditemukan pada warna merah, hijau, dan biru. RGB juga dapat dikatakan sebagai warna primer (pokok). Intensitas pada suatu warna memiliki nilai tersendiri dengan maksimal 255 (8 bit) [7].

Model Citra Hue Saturation Value (HSV)

HSV menggambarkan nilai hue, saturation dan value pada citra. Nilai warna pada proses segmentasi digambarkan dengan hue dan digunakan untuk membedakan warna satu dengan yang lainnya serta dapat menentukan nilai kemerahan (redness), kehijauan (greenness) dan lainnya dari cahaya. Panjang gelombang cahaya juga bersatu bersama hue. Intensitas warna dari suatu citra dipresentasikan oleh saturation. Sedangkan tingkat kecerahan warna suatu citra dipresentasikan oleh value.

Berikut merupakan cara yang dapat dilakukan guna mengkonversi warna RGB menjadi warna HSV:

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

(Rumus mencari nilai R, G dan B)

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

(Rumus mencari nilai V atau Value)

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

(Rumus mencari nilai S atau Saturation)

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

$$H = H + 360 \text{ jika } H < 0$$

(Rumus mencari nilai H atau Hue)

Data Mining

Suatu tindakan mengekstrak informasi dan pengetahuan (knowledge) berharga dari suatu set data berukuran besar dengan memanfaatkan teknik tertentu disebut Data mining. Informasi yang di peroleh dari proses menambang data ini dapat dipakai untuk pengambilan keputusan [5].

Pre Processing

Tentunya dalam melakukan pengolahan data dibutuhkan berbagai macam langkah agar data tersebut dapat di proses dengan algoritma yang telah ditentukan dalam penelitian. Berikut tujuan dari *pre processing*:

- a. Merubah ukuran data
- b. Memperbaiki data asli untuk memperoleh data yang sesuai
- c. Dapat meminimalkan *noise*

K-Nearest Neighbor

Algoritma KNN ialah metode non-parametrik, yang dapat digunakan untuk melaksanakan klasifikasi pada sebuah objek, prinsip dari KNN adalah jika memiliki sekumpulan sampel data sebagai data training, kita berikan label untuk seluruh data tersebut maka kita akan mengetahui data tersebut masuk kedalam kelas mana. Jika diberikan data baru tanpa label, maka akan dibandingkan data tersebut dengan data yang sudah ada lalu akan melihat kesamaan dan mencari labelnya [4].

Matlab

Matrix Laboratory yang dikenal sebagai Matlab merupakan suatu perangkat lunak yang menggunakan bahasa pemrograman tingkat atas yang diperuntukan untuk komputasi teknis. Pada aplikasi Matlab menggabungkan aspek komputasi, visualisasi dan pemrograman dalam suatu wilayah yang mudah dilakukan. Banyak cara yang dapat digunakan untuk menyelesaikan masalah melalui dukungan Matlab dan terlebih yang dapat di formulasikan kedalam bentuk matriks dan vektor diungkapkan oleh Etter & Kuncicky [3].

3. METODOLOGI PENELITIAN

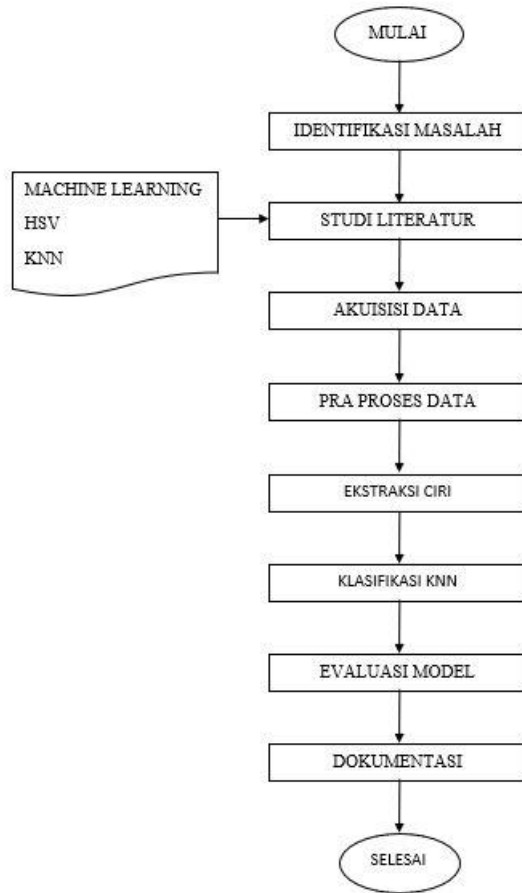
Identifikasi Masalah

Identifikasi masalah adalah bagian yang lebih fokus dalam menjelaskan poin penting yang dibahas dalam penelitian. Yaitu penerapan HSV (hue, saturation, value) dan algoritma K-Nearest Neighbor (KNN) untuk mengklasifikasi tingkat kematangan buah belimbing.

Studi Literatur

Studi Literatur umumnya digunakan oleh peneliti untuk dapat mengumpulkan berbagai macam data yang diperlukan dalam sebuah penelitian. Data didapat dari jurnal, buku dan sumber lainnya

Kerangka pikir



Gambar 1. Kerangka Pikir

Akuisisi Data

Akuisisi data diperlukan guna mendapatkan data penelitian, data yang dimaksud dalam penelitian ini adalah data citra buah belimbing dewi yang akan digunakan untuk data latih serta data uji. Dalam penelitian ini, proses pengumpulan data yang digunakan mencakup:

1. Observasi, di dalam penelitian ini, penulis akan melaksanakan observasi langsung ke lokasi penjualan belimbing dan juga kebun belimbing.
2. Dokumentasi, dalam penelitian ini, penulis akan mengambil beberapa dokumentasi citra yang berkaitan dengan hal penelitian yaitu buah belimbing dengan berbagai tingkat kematangan.

Pra Proses

Pra proses digunakan pada data citra sebelum diproses dengan metode HSV dan klasifikasi KNN, yaitu dengan cara resize citra menjadi ukuran 500x500 agar setiap ukuran buah belimbing lebih kecil dan sama rata sehingga memudahkan proses pengolahan citra pada tahap berikutnya. Selanjutnya melakukan segmentasi, proses ini dibutuhkan guna memisahkan objek dengan latar belakangnya. Pada tahap ekstraksi ciri HSV, data citra akan di olah agar dapat mengeluarkan ciri dari citra tersebut. Setelah itu data dibagi menjadi data latih yang berjumlah 80% serta data uji yang berjumlah 20%.

Klasifikasi KNN

Selanjutnya algoritma KNN digunakan untuk proses klasifikasi data citra. Terdapat beberapa tahapan dalam penggunaan KNN yaitu:

- a. Menetapkan nilai K yang digunakan
- b. Menghitung jarak euclidean antara data uji terhadap data training.
- c. Mengurutkan hasil dari terkecil sampai dengan terbesar
- d. Kelompokkan data dilandaskan nilai K yang sudah ditentukan
- e. Gunakan kelas yang paling banyak menjadi hasil klasifikasi data yang baru.

Evaluasi Model

Setelah citra selesai diklasifikasi dengan KNN dapat dihitung tingkat akurasi pengujian KNN dengan cara mengukur hasil keluarannya.

Dokumentasi

Ditahap ini kesimpulan akhir dapat diperoleh dari hasil pengujian algoritma KNN dan metode ekstraksi ciri HSV terhadap tingkat kematangan buah belimbing. Setelah semua tahapan selesai dilakukan maka tahap terakhir yaitu melakukan dokumentasi yang dimana semua kegiatanyang kita lakukan ditulis untuk dijadikan laporan penulisan yang disusun secara sistematis.

4. HASIL DAN PEMBAHASAN**Akuisisi Data**

Pada penelitian ini, menggunakan buah belimbing berjenis belimbing dewi sebanyak 15 buah dalam pengambilan data. Dari 15 buah belimbing terdapat 5 belimbing mentah, 5 belimbing setengah matang, dan 5 belimbing matang. Proses pengambilan citra dilakukan 5 kali setiap buah dengan cara memutar buah belimbing sesuai dengan sisi nya, jumlah citra yang diambil sebanyak 75 citra yang terdiri dari 25 citra belimbing mentah, 25 citra belimbing setengah matang, dan 25 citra belimbing matang. Pengambilan citra diperoleh dari kamera handphone di kejauhan 20 cm dari buah belimbing yang merupakan objek penelitian,

Pra Proses

Setelah data didapatkan, hal yang dilakukan yaitu pra proses data. Tahapan pra proses bertujuan untuk mendapatkan ciri dari citra belimbing, tahapan ini dilakukan untuk mempermudah proses perubahan RGB ke HSV. Tahapan pra proses yang dilakukan yaitu resize dan segmentasi background.

Proses Resize

Tahap pertama yang dilakukan dalam praproses ini yaitu proses resize. Proses ini disebut juga perubahan ukuran citra ini akan merubah ukuran awal menjadi citra berukuran 500x500 pixel menggunakan fungsi "imresize" [500 500]. Resize berguna mempersingkat proses pada sistem matlab agar mendapatkan hasil yang lebih sesuai.

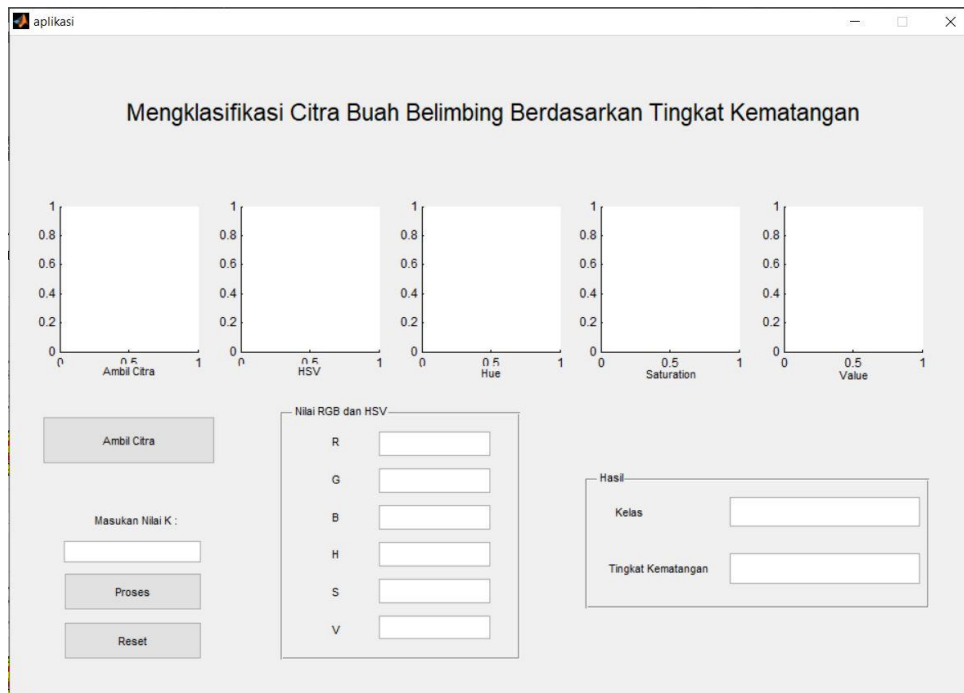
Proses Segmentasi

Tahap selanjutnya yang dilakukan yaitu proses segmentasi. Proses ini bertujuan untuk memisahkan foreground dengan background. Pemisahan ini dilakukan guna meringankan proses berikutnya yaitu proses ekstraksi HSV supaya warna terfokus pada objek yaitu buah belimbing.

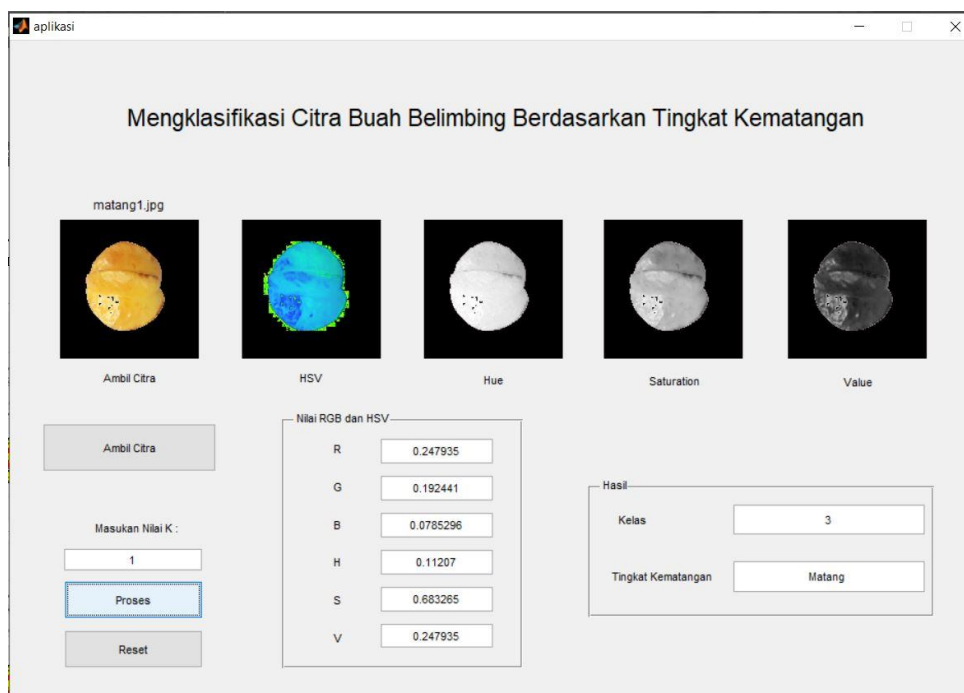
Ekstraksi Ciri HSV

Tahap selanjutnya yang dilakukan yaitu ekstraksi ciri dengan menggunakan metode HSV. Proses ini dilakukan untuk mengeluarkan nilai citra HSV. Pada proses ini akan mengambil citra testing1.jpg yang telah melalui tahap praproses resize dan segmentasi.

Tampilan Aplikasi



Gambar 2. Tampilan Awal



Gambar 3. Tampilan Input dan Ekstraksi Ciri Data Training

Tampilan dari program “mengklasifikasi citra belimbing” ini dibuat menggunakan GUI yang gunanya mempermudah pengguna untuk mengolah data. Tahapannya sebagai berikut :

- a. Pertama user klik button Ambil Citra untuk memilih citra yang akan diuji dengan data training yang sudah dimiliki dan memasukkan nilai K yang diinginkan. Dalam penulisan ini menggunakan K=1, K=3, K=5, K=7, dan K=9
- b. Kedua user klik button Proses dan akan muncul semua nilai RGB dan HSV serta hasil klasifikasi yang dihasilkan.

Hasil Ekstraksi Ciri

Sebanyak 60 data training di kelompokkan menjadi 20 data citra belimbing mentah, 20 data citra belimbing setengah matang, dan 20 data citra belimbing matang.

Tabel 1. Hasil Ekstraksi Ciri Data Training

CITRA KE	R	G	B	H	S	V	KELAS
1	0,2033	0,1939	0,063	0,1556	0,6898	0,2033	1
2	0,1766	0,1488	0,0602	0,1269	0,6594	0,1766	1
3	0,174	0,1475	0,0606	0,1278	0,6516	0,174	1
4	0,1677	0,1432	0,0599	0,1288	0,6431	0,1677	1
5	0,1601	0,138	0,0443	0,1348	0,7232	0,1601	1
6	0,159	0,1388	0,0443	0,1374	0,7213	0,159	1
7	0,1633	0,1425	0,0517	0,1356	0,6835	0,1633	1
8	0,1717	0,1477	0,0513	0,1334	0,7012	0,1717	1
9	0,1799	0,1617	0,0597	0,1414	0,6682	0,1799	1
10	0,1961	0,1756	0,0599	0,1415	0,6947	0,1961	1
..
60	0,1793	0,1561	0,065	0,1329	0,6372	0,1793	3

Data testing yang digunakan sebanyak 15 data yang dibagi menjadi 5 data citra belimbing mentah, 5 data citra belimbing setengah matang, dan 5 data citra belimbing matang.

Tabel 2. Hasil Ekstraksi Ciri Data Testing

CITRA KE	R	G	B	H	S	V	KELAS
1	0,1973	0,1953	0,0559	0,1644	0,7167	0,1973	1
2	0,1838	0,1604	0,0888	0,1256	0,5167	0,1838	1
3	0,2479	0,1924	0,0785	0,1121	0,6833	0,2479	1
4	0,2739	0,226	0,088	0,1238	0,6788	0,2739	1
5	0,2629	0,2051	0,0832	0,113	0,6835	0,2629	1
6	0,2369	0,1755	0,0658	0,1069	0,7223	0,2369	2
7	0,2502	0,2009	0,0862	0,1166	0,6556	0,2502	2
8	0,1602	0,1495	0,0544	0,1498	0,6603	0,1602	2
9	0,1714	0,1456	0,0631	0,127	0,6316	0,1714	2
10	0,1691	0,1488	0,0485	0,1387	0,7132	0,1691	2
..
15	0,1972	0,1782	0,0995	0,1342	0,4954	0,1972	3

Hasil Klasifikasi KNN

Berikut merupakan rumus akurasi untuk mendapatkan hasil ketepatan dari citra dalam penelitian :

Tabel 3. Hasil Akurasi

K	AKURASI
1	66.66%
3	80%
5	86.66%
7	93.33%
9	86.66%
RATA-RATA	82.66%

Table ini mempresentasikan bahwa akurasi dari nilai $k=1$, $k=3$, $k=5$, $k=7$, $k=9$, Didapatkan hasil optimal yaitu $k=7$, dengan akurasi = 93.33% dan dengan nilai rata-rata akurasi yaitu 82.66%, Hasil evaluasi sistem diatas menunjukkan kenaikan akurasi terhadap peningkatan nilai K. Tetapi ada penurunan pada $K=9$ yang nilainya sama dengan $K=5$

5. KESIMPULAN DAN SARAN

Kesimpulan

- Kesimpulan yang baik dari ekstraksi ciri HSV dapat klasifikasikan oleh algoritma KNN, dari penelitian yang dibuat ini algoritma KNN dengan jarak $K=7$ dapat mengklasifikasikan ciri HSV pada citra buah belimbing dengan keakuratan sebesar 93.33%
- Bisa dibilang pada penelitian ini semakin tinggi nilai K maka semakin meningkat nilai akurasi, walaupun terjadi penurunan akurasi pada nilai $K=9$

Saran

- Menggunakan data citra buah belimbing dengan jumlah yang lebih banyak
- Menggunakan dan mengaplikasikannya dengan metode lainnya
- Menggunakan keseluruhan sisi dari objek yang digunakan yaitu buah belimbing

DAFTAR PUSTAKA

- [1.] Qurania dan Arie. 2012. Analisis Tekstur Dan Estraksi Fitur Warna Untuk Klasifikasi Apel Berbasis Citra. Pakuan Bogor: Lokakarya Komputasi dalam sains dan teknologi Nuklir, 2012. 296-304.
- [2.] Gonzales, R. C., & Woods, R. E. 2002. Digital Image Processing. New Jersey: Prentice-Hall, Inc.
- [3.] Santosa, B. 2007. Data Mining Terapan dengan MATLAB: Graha Ilmu.
- [4.] Budiharto, W. 2016. Machine Learning & Computational Intelligence: Andi
- [5.] Santosa, B. & Umam, A. 2018. Data Mining dan Big Data Analytics: Penebar Media Pustaka
- [6.] Manik, F. Y., & Saragih K. S. 2017. Klasifikasi Belimbing Menggunakan Naïve Bayes Berdasarkan Fitur Warna RGB, 11(1), 99-108.
- [7.] Sari, Yuslena & Pramunendar, Ricardus, Pengolahan Citra dalam *Soft Computing*, Sukabumi : CV jejak, 2017.