

Klasifikasi Kualitas Buah Melinjo Menggunakan K-NN

Cahaya Jatmoko¹, Daurat Sinaga², Heru Lestiawan³, Christy Atika Sari³

Program Studi Teknik Informatika, Universitas Dian Nuswantoro
Jl. Imam Bonjol 207 Semarang

¹jatmoko14@dsn.dinus.ac.id

²daurat.sinaga@dsn.dinus.ac.id

³heru.lestiawan@dsn.dinus.ac.id

⁴atika.sari@dsn.dinus.ac.id

Abstrak— Banyaknya industri pengolahan yang memanfaatkan melinjo sebagai bahan baku serta pasar yang membutuhkan melinjo untuk konsumsi segar juga turut mempengaruhi potensi untuk pengembangan budidaya melinjo. Di Indonesia penentuan kualitas buah melinjo kebanyakan menggunakan cara konvensional dengan mengamati buah secara manual. Pemilahan hasil panen terutama buah-buahan dapat dilakukan secara otomatis menggunakan high performance liquid chromatography, pencitraan nearinfrared, dan sensor gas, namun cara tersebut membutuhkan perangkat yang mahal dan operator professional. Penelitian ini mengusulkan analisis kualitas buah melinjo berdasarkan pengolahan citra. Dengan menggunakan algoritma K-Nearest Neighbor (KNN) dalam mengklasifikasikan buah melinjo ke dalam dua kelompok, yaitu baik dan buruk, serta atribut yang digunakan dalam pengklasifikasian berupa fitur ekstraksi RGB. Berdasarkan hasil eksperimen metode (KNN) mempunyai performa yang baik, dibuktikan dari hasil akurasi yang tinggi. Dari 100 dataset dibagi menjadi data training 70 citra dan data testing 30 citra. Hasil pengujian menggunakan data testing yaitu 28 citra masuk kategori benar dalam klasifikasi dan 2 salah sehingga menghasilkan akurasi 93 %, precision 94 persen dan recall 94 persen.

Kata kunci— Melinjo, K-NN, RGB, klasifikasi

I. PENDAHULUAN

Melinjo biasa dikenal oleh masyarakat Indonesia dengan berbagai nama, seperti Mlinjo, Tangkil, dan Bago. Tanaman Melinjo memiliki beberapa manfaat yang dapat digunakan oleh orang-orang, seperti daun dan buah yang dapat dikonsumsi, serta bagian batang yang dapat digunakan sebagai papan. Sebagai pohon, Tanaman Melinjo dapat bermanfaat sebagai peneduh dan pembatas pekarangan.

Melinjo sendiri dapat hidup hingga 100 tahun, dan setiap musim panen dapat menghasilkan buah atau biji melinjo sebanyak 80 kg hingga 100 kg. Pohon tanaman melinjo dapat tumbuh sampai 25 m. Kandungan nutrisi khususnya dari biji melinjo adalah senyawa antioksidan dengan prosentase sekitar 9-10% untuk setiap biji melinjo. Maka, dapat dikatakan bahwa kandungan biji melinjo efektif menangkal radikal bebas yang merupakan salah satu penyebab berbagai macam penyakit [1]. Disebutkan bahwa kulit biji melinjo mengandung asam askorbat, tokoferol, dan polifenol yang berfungsi sebagai inhibitor enzim xantin oksidase [2]. Kulit buah melinjo merupakan lapisan kulit yang dapat berubah warna dari hijau menjadi merah. Perubahan warna kulit melinjo diikuti dengan

tingkat kematangan dari bijinya. Perubahan warna kulit melinjo dimulai dari hijau, kemudian menjadi kuning, dan pada tingkat yang matang menjadi warna merah [3].

Selama ini, kematangan buah melinjo dilakukan secara manual dengan mengamati perubahan warna pada kulit melinjo. Dimana dengan cara tersebut, akan lebih sering menghasilkan klasifikasi yang kurang akurat dan tidak stabil. Maka dari itu, dibutuhkan cara supaya dapat mendapatkan hasil klasifikasi kematangan buah melinjo dengan akurat. Salah satu cara untuk melakukan proses klasifikasi tingkat kematangan buah melinjo dalam bidang informatika ialah dengan mengenali ciri-ciri dari buah melinjo menggunakan pengolahan citra. Salah satu metode pengolahan citra yang akan digunakan pada penelitian ini ialah menggunakan K-nearest Neighbors.

II. METODE

A. State of The Art

Penelitian yang dilakukan oleh Krisna dkk, pada tahun 2022 menjelaskan tentang pengklasifikasian kematangan buah kersen berdasarkan warna kulitnya. Digunakan data training sebanyak 18 data, yang terdiri dari 6 data buah matang, 6 data buah setengah matang, dan 6 data buah mentah. Didapatkan hasil bahwa dengan menggunakan metode K-nearest Neighbor dan diuji dengan citra uji sebanyak 9 data, hasil akurasi yang didapatkan cukup tinggi yaitu sekitar 89% [4].

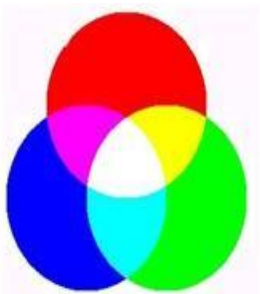
Penelitian yang dilakukan Cut Putri dkk, pada tahun 2017 membahas mengenai klasifikasi jenis pisang menggunakan Metode K-nearest Neighbor. Metode yang digunakan ialah Gray Level Co-occurrence Matrices sebagai ekstraksi ciri, dan K-nearest Neighbor sebagai metode untuk menentukan kedekatan antara citra uji dan latih berdasarkan tekstur jenis pisang yang dimiliki. Pada penelitian ini, didapatkan hasil akurasi sebesar 80% [5].

Penelitian yang dilakukan oleh Akmal dkk, pada tahun 2021 membahas tentang pengklasifikasian kematangan buah apel berdasarkan fitur warna yang menggunakan Metode Algoritma K-nearest Neighbor dan Ekstraksi Warna HSV. Dengan menggunakan 5 data testing, didapatkan hasil akurasi sebesar 95% dengan sensitivity sebesar 95% dan spesificitynya 100% yang dievaluasi menggunakan Model Hold out Estimation [6]. Penelitian yang dilakukan oleh Duwen Imantata Muhammad dkk, pada tahun 2021 membahas tentang mengklasifikasi citra belimbing berdasarkan fitur warna menggunakan Metode K-

nearest Neighbor. Digunakan sebanyak 15 buah belimbing berjenis belimbing dewi sebagai data yang diteliti. Dari ke-15 buah tersebut, 5 diantaranya merupakan belimbing mentah, 5 belimbing setengah matang, dan 5 belimbing matang. Proses pengambilan citra dilakukan 5 kali untuk setiap buah, sehingga dihasilkan 75 citra yang terdiri dari 25 citra belimbing mentah, 25 citra belimbing setengah matang, dan 25 citra belimbing matang. Didapatkan hasil akurasi 93.33% dengan jarak K=7 [7].

B. Citra Warna (RGB)

Citra warna model RGB (Red Green Blue) merupakan citra saling lepas dalam tiga bidang citra. Yang mana terdiri dari warna dasar: Red (R), Green (G), Blue (B). RGB (Red Green Blue) adalah warna utama yang secara langsung mata manusia mampu menerimanya. Tiap pixel dalam warna citra mewakili citra yang merupakan gabungan dari tiga warna utama RGB (Red Green Blue) (Eel & ETP, 2018). Data tiga byte dibutuhkan citra warna pada tiap titiknya. Tiap warna utama mempunyai intensitas sendiri dengan nilai minimum nilai (0) dan maksimum 255 (8 bit). RGB (Red Green Blue) didasarkan pada teori bahwa mata manusia peka terhadap panjang gelombang 630 nm. Gambar 1 dibawah ini melihat bentuk geometri dari model warna RGB (Red Green Blue) dalam mendefinisikan warna dengan sistem kartesius koordinat. Spektrum tingkat keabuan (Grayscale) adalah tiga warna utama yang membentuk warna dengan jumlah yang sama dan berada di garis titik hitam dan putih yang saling berhubungan. Warna didefinisikan kedalam salah satu sinar tambahan dalam pembentukan warna yang baru, dan membentuk sinar campuran yang saling berhubungan. Model warna RGB (Red Green Blue) kebanyakan digunakan pada vidio kamera maupun layar komputer.



Gambar 1. Representasi Warna RGB

Proses data citra menjadi nilai RGB seperti berikut :

Rata-rata fitur R (Red),
$$R = \frac{\sum_{i=1}^j R_{ij}}{j}$$

$R =$

$$\frac{\sum_{i=1}^j R_{ij}}{j}$$

(1)

Rata-rata fitur G (Green),
$$G = \frac{\sum_{i=1}^j G_{ij}}{j} \quad (2)$$

$G =$

$$\frac{\sum_{i=1}^j G_{ij}}{j}$$

Rata-rata fitur B (Blue),
$$B = \frac{\sum_{i=1}^j B_{ij}}{j} \quad (3)$$

$B =$

$$\frac{\sum_{i=1}^j B_{ij}}{j}$$

(3)

C. K-nearest Neighbors (KNN)

Algoritma KNN merupakan metode non-parametrik yang dapat digunakan untuk mengklasifikasi sebuah objek [8]. Algoritma KNN termasuk ke dalam Algoritma Supervised. Tujuan dari digunakannya Algoritma KNN adalah sebagai pengklasifikasi objek berdasarkan data training dan uji [9]. Algoritma K-Nearest Neighbor (KNN) adalah algoritma yang digunakan untuk mengklasifikasikan suatu objek, berdasarkan kereta data pelatihan yang lebih dekat dengan objek. Persyaratan K seharusnya tidak lebih besar dari jumlah data pelatihan, dan nilai K harus aneh dan lebih dari satu. Algoritma K-Nearest Neighbor menggunakan klasifikasi lapangan kerja sebagai nilai prediksi konsultasi konsultasi baru. Algoritma ini sederhana, berfungsi berdasarkan jarak terpendek dari kueri instance ke pelatihan sampling untuk menentukan ketetanggan. K-Nearest Neighbor (K-NN) memiliki kelebihan yang dapat menghasilkan data yang kuat atau jelas dan efektif jika digunakan dalam data besar. Tetangga terdekat juga memiliki kerugian, yang membutuhkan nilai K, jarak dari data uji tidak jelas dengan jarak yang digunakan, untuk mendapatkan hasil terbaik, Anda harus menggunakan semua atribut atau hanya atribut yang ditentukan [10]. Metode K-Nearest Neighbor (KNN) cukup sederhana, tidak ada asumsi sehubungan dengan distribusi data dan mudah dipublikasikan. Pemilihan nilai K (nomor data / tetangga) ditentukan oleh peneliti. Tahapan menghitung metode KNN sebagai berikut:

- Memastikan parameter K (jumlah tetangga terdekat)
- Menghitung kuadrat Euclidean distance (queri instance) masing-masing objek terhadap data sampel.
- 3. Kemudian mengurutkan objek- objek tersebut kedalam kelompok yang mempunyai jarak euclid terkecil.
- Menggabungkan kategori y (klasifikasi Nearest Neighbor)

Dengan menggunakan kategori Nearest Neighbor yang paling mayoritas, maka dapat diprediksi nilai query instance yang telah dihitung.

D. Pengukuran Hasil

Pengukuran ketepatan klasifikasi sangat perlu dilakukan karena metode ini digunakan untuk menghitung kinerja akurasi hasil kalsifikasi yang digunakan pada penelitian ini. Sistem yang dilakukan kinerja klasifikasi diharapkan mampu menghasilkan

akurasi yang sangat tinggi, tetapi tidak dapat dipungkiri bahwa tidak mungkin sebuah kinerja sebuah sistem mencapai 100% benar. Maka mutlak harus diukur kinerja suatu sistem klasifikasinya. Umumnya, pengukuran kinerja klasifikasi dilakukan dengan matrik konfusi (confusion matrix). Rumus dalam penghitungan akurasi adalah sebagai berikut :

$$akurasi = \left(\frac{TP+TN}{TP+TN+FP+FN} \right) \quad (4)$$

Keterangan:

TP : keseluruhan data yang benar yang terklasifikasi dalam kelas

TN : keseluruhan data yang benar yang terklasifikasi sebaliknya dalam kelas

FP : keseluruhan data dianggap sistem benar berada dalam kelas namun seharusnya data tersebut berada dalam kelas sebaliknya

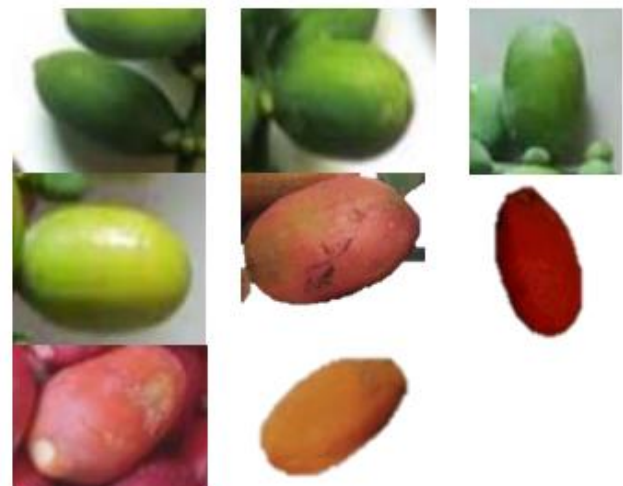
FN : keseluruhan data yang dianggap sistem berada dalam kelas sebaliknya namun seharusnya data tersebut berada dalam kelas yang benar

E. Dataset

Dataset tersebut memiliki 2 kelas yaitu kelas baik dan kelas buruk. Dataset akan dibagi menjadi data training dan data testing. Total data yang digunakan dalam percobaan adalah 200 citra dengan sebaran 100 citra melinjo kelas baik dan 100 citra melinjo pada kelas buruk. Dari 100 citra, 70 citra digunakan sebagai data training sedangkan 30 citra digunakan sebagai data testing. Contoh visual citra dataset yang digunakan dapat dilihat pada Gambar 3.



Melinjo Kualitas Baik



Melinjo Kualitas Buruk

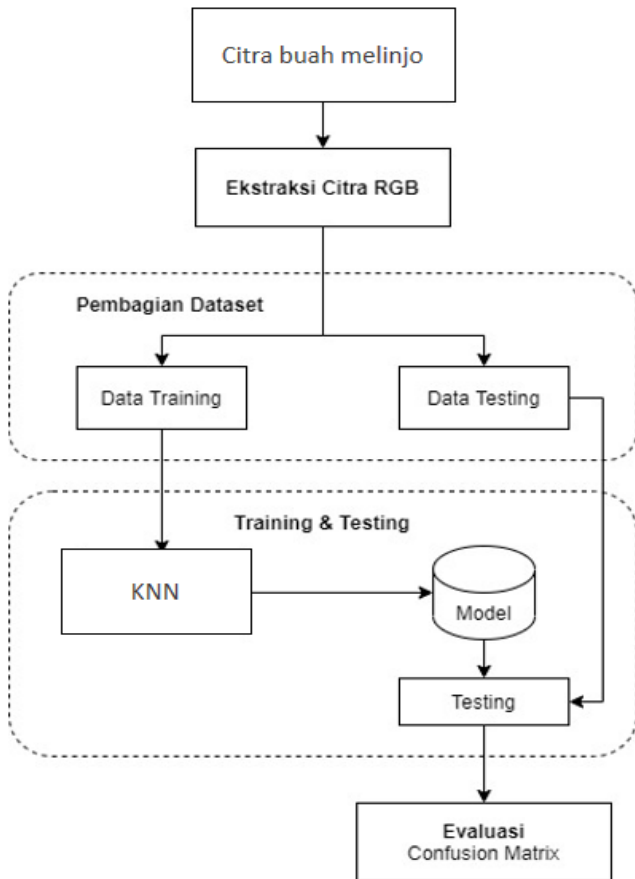
Gambar 3. Contoh Dataset Citra Melinjo

F. Proses Klasifikasi Citra

Terdapat tahap-tahap utama yaitu ekstraksi citra menggunakan RGB, pembagian dataset, klasifikasi dengan KNN dan evaluasi seperti pada Gambar 4. Langkah-langkah pada metode yang diusulkan dijelaskan sebagai berikut:

- Persiapan dataset yaitu menyiapkan dataset citra melinjo dengan citra ukuran 256x256 pixel sebanyak 100 citra.
- Dataset kemudian di proses Ekstraksi citra menggunakan Red, Green, Blue
- Red yaitu mengambil nilai pada lapisan Red pada citra dengan formula (1)
- Green yaitu mengambil nilai pada lapisan Green pada citra dengan formula (2)
- Blue yaitu mengambil nilai pada lapisan Blue pada citra dengan formula (3)

- Setelah dataset melalui proses ekstraksi kemudian dibagi menjadi data training dan data testing. Data training digunakan untuk proses pelatihan KNN. Sedangkan data testing digunakan untuk menguji model yang telah dilatih.
- Tahap pelatihan atau training, merupakan proses pelatihan KNN pada nilai K=1
- Output dari hasil klasifikasi adalah confusion matrix yang akan dihitung akurasi untuk mengetahui performa dari model klasifikasi.

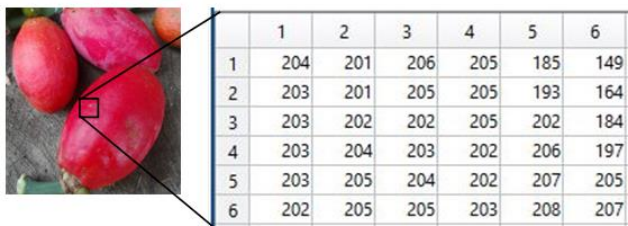


Gambar 4. Tahapan Klasifikasi Buah Melinjo

III. HASIL DAN PEMBAHASAN

A. Normalisasi Citra

Pada tahap ini data berupa citra melinjo di ekstraksi fitur RGB. Tahap ekstraksi fitur mengambil nilai rata-rata RGB untuk setiap citra seperti pada Gambar 5.



Gambar 5. Proses Normalisasi Citra

Dari Gambar 5, diperoleh nilai Rata-rata fitur R (Red)

$$R = \frac{204+201+206+\dots+207}{65536} = 188$$

$$R = \frac{204+201+206+\dots+207}{65536} = 188$$

$$R = \frac{204+201+206+\dots+207}{65536} = 188$$

188

Rata-rata fitur G (Green) $G = \frac{210+209+208+\dots+209}{65536} = 180$

$$G = \frac{210+209+208+\dots+209}{65536} = 180$$

$$B = \frac{208+207+206+\dots+207}{65536} = 174$$

$B = \frac{208+207+206+\dots+207}{65536} = 174$. Setelah dilakukan proses ke semua data hasil fitur ekstraksi RGB ditampilkan pada Tabel 1.

Tabel 1. Hasil Ekstraksi RGB

Normalisasi Warna			
R	G	B	Label
0.869914	0.822458	0.777111	1
0.943817	0.872548	0.777795	1
-0.65488	-0.52792	-0.51429	-1
-0.81755	-0.81981	-0.78112	-1
0.943817	0.872548	0.777795	1
-0.65919	-0.66154	-0.5801	-1
-0.87887	-0.75598	-0.69006	-1
-0.77692	-0.80057	-0.76953	-1
1	1	1	1
0.263304	0.329403	0.217625	-1
-0.89281	-0.72686	-0.68182	-1
0.853464	0.769385	0.73289	1
0.92034	0.960431	0.979151	1
-0.40169	-0.5348	-0.59455	-1
0.896242	0.942881	0.889491	1
-0.02627	-0.1151	-0.19997	-1
0.048732	-0.01044	-0.07632	-1
-0.64056	-0.65899	-0.57135	-1
0.867026	0.764627	0.733607	1
0.97516	0.915795	0.91032	1
-0.62762	-0.80011	-0.84093	-1
0.477193	0.382106	0.317806	1
-0.69746	-0.60695	-0.53382	-1
-0.88404	-0.90094	-0.82328	-1
-0.63186	-0.7378	-0.7326	-1
0.896242	0.942881	0.889491	1
-0.75743	-0.8866	-0.91947	-1
0.794978	0.672932	0.679289	1
-0.76821	-0.87255	-0.92329	-1
-0.98509	-1	-1	-1

Tahap akhir dalam penelitian ini adalah pengujian metode. Pengujian dilakukan dengan mengklasifikasi 30 data testing. Berdasarkan 30 data citra yang telah dilakukan pengujian klasifikasi ditampilkan pada Tabel 2.

Tabel 2. Hasil Prediksi Kelas

Normalisasi warna			Actual Class	Predicted Class
R	G	B		

-0.8789	-0.756	-0.6901	1	1
0.86703	0.76463	0.73361	2	2
-0.0263	-0.1151	-0.2	1	1
-0.6549	-0.5279	-0.5143	1	1
-0.6647	-0.5806	-0.5153	1	1
-0.791	-0.6417	-0.6199	1	1
0.92382	0.92746	0.86565	2	2
-0.7609	-0.6126	-0.5675	1	1
0.92382	0.92746	0.86565	2	2
-0.5191	-0.5037	-0.5207	1	1
0.47719	0.38211	0.31781	2	2
0.83748	0.8097	0.76893	2	2
-0.7063	-0.7108	-0.6301	1	1
0.94382	0.87255	0.7778	2	2
0.89624	0.94288	0.88949	2	2
0.94382	0.87255	0.7778	2	2
0.69889	0.59434	0.53944	2	2
-0.8175	-0.8198	-0.7811	1	1
0.89624	0.94288	0.88949	2	2
-0.8928	-0.7269	-0.6818	1	2
0.9493	0.94843	0.86383	2	2
-0.7769	-0.8006	-0.7695	1	1
-0.8784	-0.7296	-0.7452	1	1
-0.8113	-0.8269	-0.7732	1	1
0.47719	0.38211	0.31781	2	1
-0.8651	-0.7846	-0.7251	1	1
0.81198	0.73859	0.70907	2	2
-0.4017	-0.5348	-0.5945	1	1
0.03744	0.01425	-0.0684	1	1
-0.7862	-0.6564	-0.6115	1	1

Berdasarkan hasil pengujian pada Tabel 2, dibuat confusion matrix untuk perhitungan performa klasifikasi.

Tabel 3. Confusion Matrix

		Actual	
		Baik	Buruk
Predicted	Baik	16	1
	Buruk	1	12

$$Accuracy = \left(\frac{True(bagus) + True(jelek)}{N} \right) = \left(\frac{16 + 12}{30} \right) = 0.93$$

$$Precision = \frac{TP}{TP+FP} = \frac{16}{16+1} = 0.94$$

$$Precision = \frac{TP}{TP+FP} = \frac{16}{16+1} = 0.94$$

$$Accuracy = \left(\frac{True(bagus) + True(jelek)}{N} \right) = \left(\frac{16+12}{30} \right) = 0.93$$

$$Recall = \frac{TP}{TP + FN} = \frac{16}{16 + 1} = 0.94$$

IV. KESIMPULAN

Dari hasil penelitian dan pengujian klasifikasi kualitas buah melinjo dengan metode KNN dan ekstraksi citra RGB. Metode yang diusulkan yaitu KNN telah berhasil dianalisis untuk menyelesaikan kasus klasifikasi kualitas buah melinjo baik atau buruk. Metode KNN mempunyai performa yang baik,

dibuktikan dari hasil akurasi yang tinggi. Dari 100 dataset dibagi menjadi data training 70 citra dan data testing 30 citra. Hasil pengujian menggunakan data testing yaitu 28 citra masuk kategori benar dalam klasifikasi dan 2 salah sehingga menghasilkan akurasi 93 %, precision 94 persen dan recall 94 persen.

REFERENSI

- [1] (2022) agrotek.id website : <https://agrotek.id/klasifikasi-dan-morfologi-tanaman-melinjo/>
- [2] Santoso, M., Naka, Y., Angkawidjaja, C., Tamaguchi, T., Matoba, T. & Takamura, H. 2010. *Antioxidant and damage Prevention Activities of Edible Parts of Gnetum Gnemon and Their Change upon Heat Treatment*. Journal Food Science and Technology 15 (6): 549-556.
- [3] Nisa, R. 2017. *Struktur Anatomis dan Profil Fitokimia Kulit Lar Biji Melinjo (Gnetum Gnemon L.) pada Empat Tingkat Kemasakan Biji*. Skripsi, Fakultas Biologi Universitas Gadjah Mada Yogyakarta.
- [4] Pratama, K. A., Atmaja W. P., Lusiana, V. 2022. *Klasifikasi Tingkat Kematangan Buah Kersen menggunakan Citra HIS dengan Metode K-nearest Neighbor (KNN)*. Smart Comp Vol. 11 (1): 105-108.
- [5] Iklima, C. P., Nasir, M., HariTohaHidayat. 2017. *Klasifikasi Jenis Pisang menggunakan Metode K-nearest Neighbors (KNN)*. Jurnal Teknologi Rekayasa Informasi dan Komputer Vol. 1 (1): 11-14.
- [6] Ilimi, A., Razka, M. A., Wiratomo, D. S., Prasvita, D. S. 2021. *Klasifikasi Tingkat Kematangan Buah Apel berdasarkan Fitur Warna menggunakan Algoritma K-nearest Neighbor dan Ekstraksi Warna HSV*. Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA): 176-182.
- [7] Muhammad, D. I., Ermatita., Falih, N. 2021. *Penggunaan K-nearest Neighbors (KNN) untuk Mengklasifikasi Citra Belimbing berdasarkan Fitur Warna*. Jurnal Informatik Vo. 17 (1): 9-16.
- [8] Budiharto, W. 2016. *Machine Learning & Computational Intelligence*: Andi.
- [9] Febri Liantoni, *Klasifikasi Daun dengan Perbaikan Fitur Citra menggunakan K-nearest Neighbor*. Institut Teknologi Adhi Tama, 2015.
- [10] Paramita C, Rachmawanto EH, Sari CA, Setiadi DR. 2019. *Klasifikasi Jeruk Nipis Terhadap Tingkat Kematangan Buah Berdasarkan Fitur Warna Menggunakan K-Nearest Neighbor*. Jurnal Informatika. Vol. 4 (1).