

Implementasi Algoritma K-Nearest Neighbor untuk Identifikasi Citra Bunga Anggrek

Nofita Sari¹, Resty Wulanningrum²

^{1,2}Teknik Informatika, Fakultas Teknik Universitas Nusantara PGRI Kediri

E-mail: *¹nofitasari176@gmail.com, ²resty0601@gmail.com

Abstrak – *Phalaenopsis* adalah bahasa latin dari bunga anggrek yang merupakan salah satu bunga yang banyak digemari masyarakat untuk menghiasi rumah mereka. Bunga anggrek memiliki banyak jenis yang mungkin banyak masyarakat hanya mengetahui jenisnya dari warnanya saja. Banyak yang kurang mengamati tentang bunga anggrek itu sendiri. Terkadang 1 warna terdiri dari beberapa jenis yang sangat menonjol untuk membedakannya adalah dilihat dari kelopak bunga anggrek. Penelitian ini mengambil sampel 3 jenis bunga anggrek, yaitu jenis *Phalaenopsis Amabilis*, *Dendrobium Phalaenopsis*, dan *Phalaenopsis Violacea*. Tahapan implementasi yang dilakukan adalah dengan melakukan *preprocessing* yang meliputi *grayscale* dan deteksi tepi *Kirsch*, selanjutnya proses identifikasinya dengan menerapkan algoritma *K-Nearest Neighbor*. Hasil yang dari penelitian ini antara lain telah dihasilkan sebuah aplikasi untuk deteksi bunga anggrek berdasarkan kelopak bunga dan didapatkan akurasi sebesar 86,7%. Besarnya akurasi yang didapatkan berpengaruh dari banyaknya data *training* dan data *testing* yang digunakan saat ujicoba.

Kata Kunci — bunga anggrek, pengenalan pola, *K-Nearest Neighbor*.

Abstract – *Phalaenopsis* is the Latin word for orchid which is one of the most popular flowers for people to decorate their homes with. Orchid flowers have many types that maybe many people only know the type from the color. Many are less aware of the orchid itself. Sometimes 1 color consists of several types. What really stands out to distinguish them is seen from the petals of the orchid. This study took samples of 3 types of orchids, namely *Phalaenopsis Amabilis*, *Dendrobium Phalaenopsis*, and *Phalaenopsis Violacea*. The implementation stage is to do *pre-processing* which includes *grayscale* and *Kirsch* edge detection, then the identification process uses the *K-Nearest Neighbor* algorithm. The result of this research is that an application for the detection of orchids based on flower petals has been produced and an accuracy of 86.7% has been obtained. The amount of accuracy obtained affects the amount of training data and testing data used during testing.

Keywords — orchid, pattern recognition, *K-Nearest Neighbor*

1. PENDAHULUAN

Tanaman anggrek memiliki keistimewaan pada bagian bunganya yang sangat beranekaragam dalam hal bentuk, ukuran, warna maupun corak warnanya [1]. Bunga Anggrek mampu menarik perhatian, bukan hanya bagi penggemar di Indonesia, tetapi juga penggemar di luar negeri [2]. Cara yang dilakukan dalam mengklasifikasikan jenis bunga anggrek satu dengan yang lain adalah dengan mengamati tekstur, warna dan kelopak bunga anggrek. Dengan memahami perbedaan tersebut, jenis dari suatu bunga anggrek dapat dikenali. Akan tetapi pada umumnya jenis-jenis dari bunga anggrek mempunyai tekstur warna dan kelopak bunga yang mirip, hal tersebut yang mengakibatkan seseorang mengalami kendala dalam mengenali jenis bunga anggrek, pada khususnya orang awam yang belum memahami ciri-ciri dari beberapa jenis bunga anggrek [3].

Identifikasi Pola Bentuk Bunga dapat dilakukan dengan menerapkan metode *Principle Component Analysis* (PCA) Dan *K-NN*. Penerapan metode PCA pada masing-masing titik data sampel citra bunga ditransformasikan berdasarkan *principle component*-nya, sehingga keberagaman penyebaran data hasil transformasi yang ukurannya lebih kecil dapat mewakili keberagaman

penyebaran data asli yang ukurannya jauh lebih besar. Metode K-NN dipergunakan pada identifikasi pola untuk melakukan pengenalan terhadap citra bunga yang didasarkan dari jarak yang paling dekat antara data citra bunga yang akan dilakukan evaluasi dengan K tetangga (*neighbor*) terdekatnya didalam data *training*. Hasil dari evaluasi identifikasi pola pada citra bunga dengan menerapkan metode klasifikasi K-NN berdasarkan ekstraksi ciri menggunakan metode PCA diperoleh akurasi sebesar 97,70% [4]. Metode SVM untuk identifikasi secara otomatis terhadap citra tanaman anggrek yang sedang berbunga dan tanaman anggrek tanpa bunga. Penelitian ini menerapkan penggabungan dua ukuran *circular neighborhood* yang dilakukan oleh *fuzzy* sehingga diperoleh sebuah histogram dengan panjang *bin* yang merupakan hasil dari penjumlahan setiap *bin* pada operator [5]. Pengenalan jenis bunga anggrek berdasarkan warna dan bentuk dapat menerapkan metode ekstraksi fitur bentuk dan warna HSV, Prewit dan *Learning Vector Quantization* (LVQ) untuk proses klasifikasi, sehingga menghasilkan akurasi sebesar 73,33 % [6]. Selain menerapkan metode ekstraksi fitur, proses identifikasi bunga anggrek juga dapat dilakukan dengan menerapkan metode deteksi tepi. Dengan menerapkan metode deteksi tepi *Canny*, *Sobel* akurasi yang dihasilkan dari proses klasifikasi bunga anggrek diantara 40,7% hingga 55,5%, sedangkan dengan menggunakan metode HSV diperoleh akurasi sebesar 70% [7]. Dari beberapa penelitian sebelumnya, maka penulis membuat sebuah penelitian identifikasi pengenalan citra anggrek menggunakan *K-Nearest Neighbor* untuk mengelompokkan jenis anggrek berdasarkan kelopak bunganya.

2. METODE PENELITIAN

2.1. KNN (*K-Nearest Neighbor*)

Algoritma *K-Nearest Neighbor* (Pencarian tetanga terdekat) adalah metode klasifikasi yang sangat terkenal. Algoritma KNN diperkenalkan oleh *Fix* dan *Hodges* (1951), dan telah terbukti menjadi algoritma klasifikasi sederhana yang baik. *K-Nearest Neighbor* adalah salah satu teknik yang diterapkan untuk melakukan klasifikasi dengan memanfaatkan algoritma *supervised* [8]. Tujuan dari algoritma ini yaitu melakukan klasifikasi obyek baru berdasarkan jarak suatu obyek yang akan diklasifikasikan terhadap data contoh. *Classifier* hanya memanfaatkan fungsi jarak dari data baru ke data *training* [9]. Algoritma *K-Nearest Neighbor* adalah algoritma yang memilih nilai jarak pada pengujian data testing dengan data *training* yang diambil dari nilai paling kecil dari nilai ketetanggaan terdekat. Dalam penelitian ini data yang dipakai adalah citra bunga anggrek dengan warna latar belakang putih. *Dataset* yang dipakai adalah citra bunga anggrek yang terdiri dari 3 jenis. Data citra yang dipakai sejumlah 60 foto citra bunga anggrek. Proses klasifikasi citra bunga anggrek dengan menerapkan algoritma *K-Nearest Neighbor*. Jarak dari data yang ada masing-masing dihitung menggunakan persamaan *Euclidean distance* (Jarak *Euclidean*)[10]:

$$dv = \sqrt{\sum_{k=1}^m (fd_{i,k} - k_j)^2} \dots\dots\dots (1)$$

Dimana:

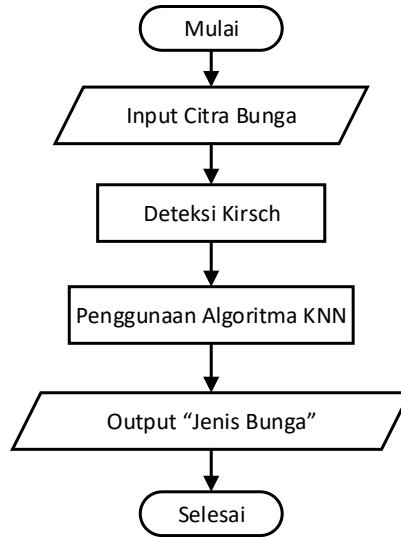
- dv = *Euclidean distance*
- m = jumlah data uji
- $fd_{i,k}$ = data *training*
- k_j = data uji

Proses pengujian dengan menerapkan deteksi tepi *Kirsch* untuk mendeteksi semua tepi obyek yang ada didalam citra dari berbagai arah. Operator ini mirip dengan model matriks 3x3 atau jendela dengan ukuran 3x3 piksel.

2.2. Perencanaan dan Tahap Implementasi

2.2.1. Flowchart Sistem

Flowchart sistem menjelaskan bagaimana aplikasi ini bekerja dalam klasifikasi bunga *Phalaenopsis* sehingga kelopak bunga *Phalaenopsis* dapat dikenali dan diidentifikasi.

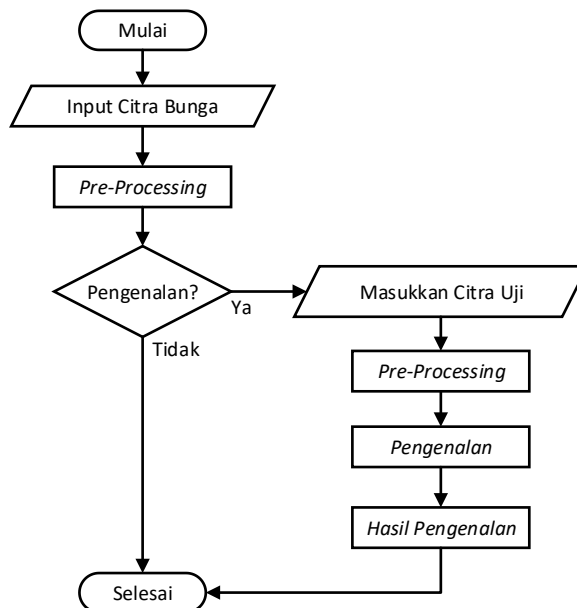


Gambar 1. Diagram alir sistem

Pada gambar 1 dapat diketahui bahwa langkah awal yang harus dilakukan yaitu memasukkan gambar. Selanjutnya penerapan deteksi tepi *Kirsch* serta perhitungan dengan Algoritma KNN untuk melakukan pencocokan dan terakhir *output*-nya hasil dari jenis bunga anggrek *Phalaenopsis*.

2.2.2. Flowchart sistem pengenalan bunga

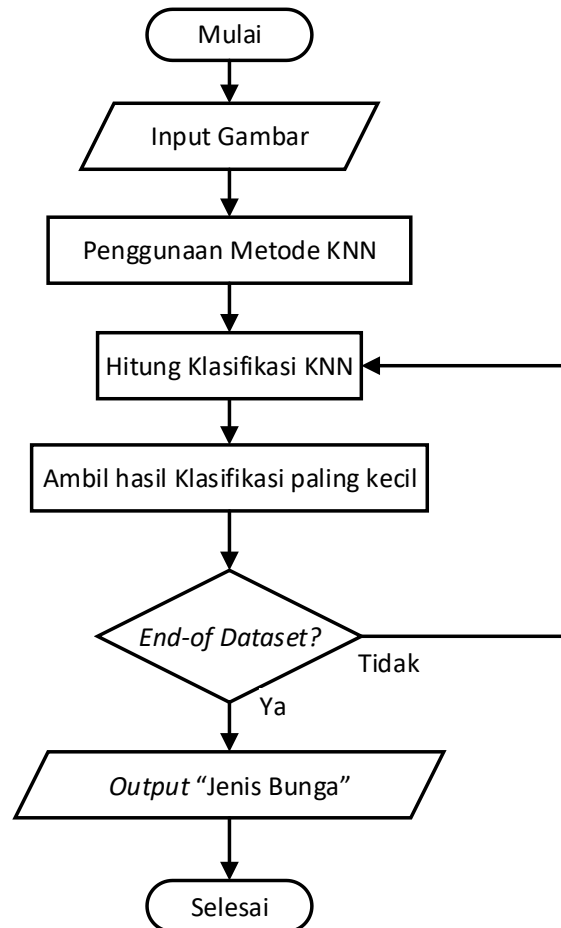
Pada tahap ini, *flowchart* digunakan untuk menjelaskan lebih detail mengenai alur aplikasi mulai dari proses memuat citra sampai pengenalan bunga *phalaenopsis*.



Gambar 2. Diagram alir sistem pengenalan bunga

Pada gambar 2 dapat dijelaskan bahwa langkah awal yang harus dilakukan memasukkan gambar, kemudian dilakukan *pre-processing*. Selanjutnya penerapan deteksi tepi *Kirsch* dan perhitungan dengan algoritma KNN untuk melakukan pencocokan dan terakhir *output*-nya hasil dari jenis bunga anggrek *Phalaenopsis*.

2.2.3. Proses perhitungan KNN



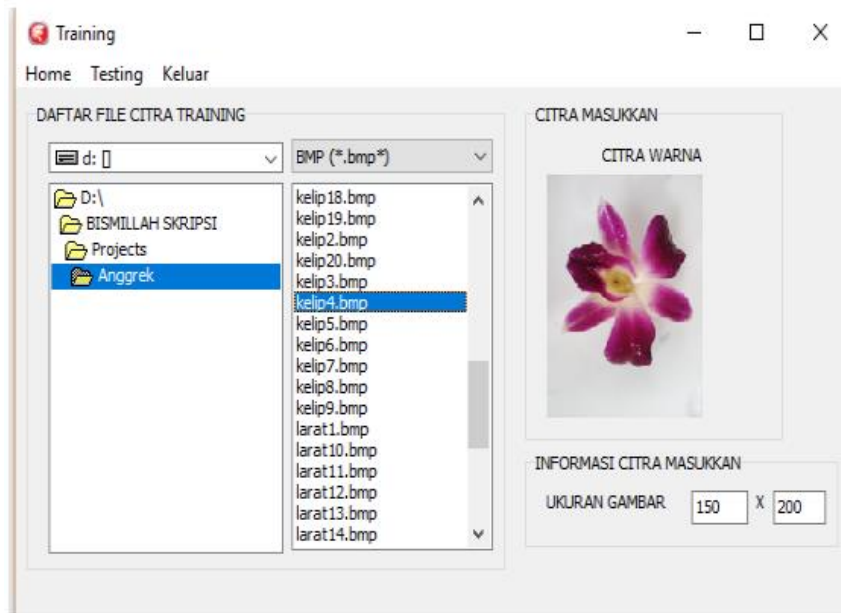
Gambar 3. Diagram alir perhitungan KNN

Pada gambar 3 dapat dijelaskan bahwa langkah awal yang harus dilakukan memasukkan gambar, hitung dengan algoritma KNN, setelah dihitung maka dikelompokkan jarak nilainya yang paling kecil/ terdekat. Jika jarak kemiripannya sudah mendekati proses akan menghasilkan output jenis bunga anggrek.

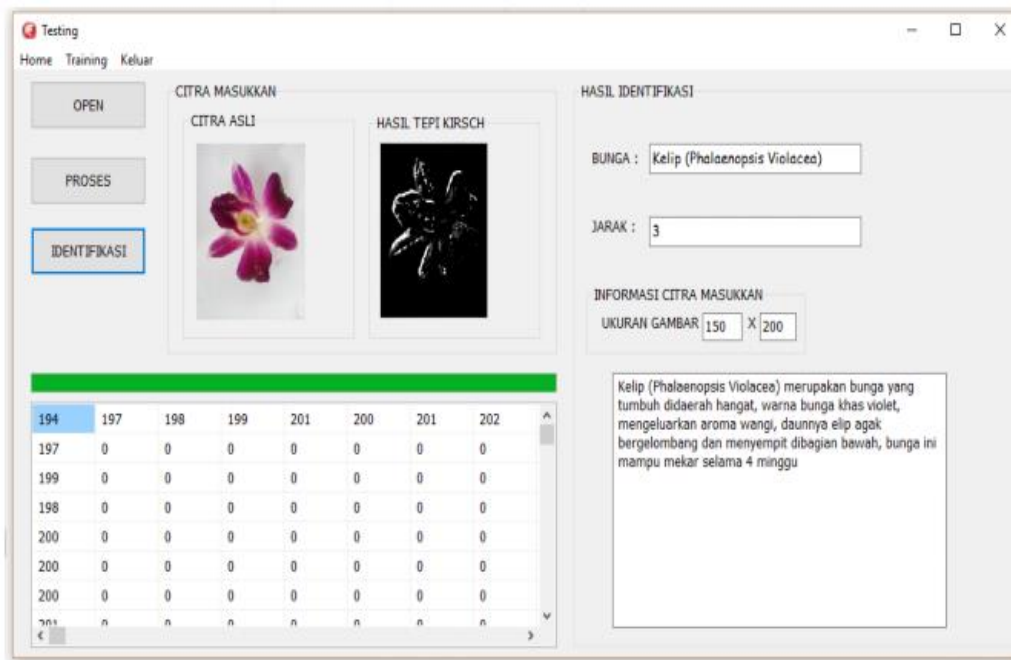
3. HASIL DAN PEMBAHASAN

3.1. Implementasi

Tampilan data *training* dari aplikasi pengenalan citra bunga anggrek terlihat pada gambar 4:



Gambar 4. Tampilan *training*



Gambar 5. Tampilan *Testing*

Gambar 5 menunjukkan tampilan *testing* yang terdiri dari *button open*, *proses* dan *identifikasi*. *Button open* berfungsi untuk menampilkan gambar citra asli yang dipilih, sedangkan *button proses* berfungsi menampilkan gambar citra asli yang telah diproses menjadi *grayscale* dan deteksi tepi *Kirsch*. Karena citra asli dari pengambilan citra digital berupa citra RGB [11]. Sedangkan *button identifikasi* berfungsi untuk menampilkan nilai piksel dari deteksi tepi *Kirsch* dan juga menampilkan jenis bunga serta nilai jarak dari metode KNN.

3.2. Hasil Pengujian Hardware

Tabel 1. Skenario Ujicoba *Data Training*

No	<i>Data Training</i>			Jumlah
	<i>Phalaenopsis Amabilis</i>	<i>Dendrobium Phalaenopsis</i>	<i>Phalaenopsis Violacea</i>	
1.	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15	45
2.	1,2,3,4,5,6,7,8,9,10,11,12,13	1,2,3,4,5,6,7,8,9,10,11,12,13	1,2,3,4,5,6,7,8,9,10,11,12,13	39
3.	1,2,3,4,5,6,7,8,9,10,11	1,2,3,4,5,6,7,8,9,10,11	1,2,3,4,5,6,7,8,9,10,11	33
4.	1,2,3,4,5,6,7,8,9	1,2,3,4,5,6,7,8,9	1,2,3,4,5,6,7,8,9	
5.	1,2,3,4,5	1,2,3,4,5	1,2,3,4,5	15

Tabel 1 menunjukkan banyaknya jumlah data *training* yang akan dilakukan untuk proses pembelajarannya.

Tabel 2. Skenario Ujicoba *data testing*

No	<i>Data Testing</i>			Jumlah
	<i>Phalaenopsis Amabilis</i>	<i>Dendrobium Phalaenopsis</i>	<i>Phalaenopsis Violacea</i>	
1.	16,17,18,19,20	16,17,18,19,20	16,17,18,19,20	15
2.	14,15,16,17,18,19,20	14,15,16,17,18,19,20	14,15,16,17,18,19,20	21
3.	12,13,14,15,16,17,18,19,20	12,13,14,15,16,17,18,19,20	12,13,14,15,16,17,18,19,20	27
4.	10,11,12,13,14,15,16,17,18,19,20	10,11,12,13,14,15,16,17,18,19,20	10,11,12,13,14,15,16,17,18,19,20	33
5.	6,7,8,9,10,11,12,13,14,15,16,17,18,19,20	6,7,8,9,10,11,12,13,14,15,16,17,18,19,20	6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,24,25	15

3.3. Hasil Uji coba

Pada uji coba aplikasi ini menggunakan 3 sampel citra bunga anggrek atau *Phalaenopsis* yaitu *Phalaenopsis Violacea*, *Dendrobium Phalaenopsis*, dan *Phalaenopsis Amabilis* dengan masing-masing citra memiliki ukuran 160x120 piksel. Tabel 3 adalah tabel hasil uji coba terhadap masing-masing sampel citra.

Tabel 3. Hasil ujicoba skenario

Skenario	Data		Hasil		Prosentase
	<i>Training</i>	<i>Testing</i>	Benar	Salah	
1	45	15	13	2	86,7%
2	39	21	18	3	85%
3	33	27	22	5	81,5%
4	27	33	25	8	75,8%
5	15	45	35	13	72%

Hasil yang diperoleh dari skenario uji coba aplikasi dapat dilihat bahwa hasil dari proses *training* dan proses *testing* pada masing-masing skenario yaitu:

Skenario 1: Diperoleh nilai benar 13 dan nilai salah 2, sehingga prosentase akurasi yang diperoleh sebesar 86,7%.

Skenario 2: Diperoleh nilai benar 18 dan nilai salah 3, sehingga prosentase akurasi yang diperoleh sebesar 85%.

Skenario 3: Diperoleh nilai benar 22 dan nilai salah 5, sehingga prosentase akurasi yang diperoleh sebesar 81,5%.

Skenario 4: Diperoleh nilai benar 25 dan nilai salah 8, sehingga prosentase akurasi yang diperoleh sebesar 75,8%.

Skenario 5: Diperoleh nilai benar 35 dan nilai salah 13, sehingga prosentase akurasi yang diperoleh sebesar 72%.

Kesimpulan dari data prosentasi akurasi yang paling akurat ini terjadi pada skenario 1, dimana diperoleh nilai benar 13 dan nilai salah 2, sehingga prosentase akurasi yang diperoleh sebesar 86,7%.

4. KESIMPULAN

Setelah melakukan analisis, perancangan dan implementasi Algoritma *K-Nearest Neighbor* dalam melakukan klasifikasi citra bunga anggrek diperoleh kesimpulan bahwa:

1. Deteksi tepi menggunakan metode *Kirsch* dapat diterapkan untuk melakukan klasifikasi jenis bunga anggrek berdasarkan kelopaknya dan telah berhasil dibuat dan diujicobakan.
2. Penerapan metode deteksi tepi *Kirsch* dan *K-Nearest Neighbor* untuk klasifikasi bunga anggrek menunjukkan hasil berupa identifikasi dari setiap kelopak bunga anggrek. Sehingga prosentase akurasi yang diperoleh sebesar 86,7% dengan rasio data *training* dan *testing* 45:15.
3. Ukuran piksel citra memiliki dampak terhadap nilai matriks. Jika ukuran piksel citra semakin besar, maka proses pengambilan nilai matriks akan semakin lama.
4. Penggunaan perangkat lunak pendukung berupa *software* aplikasi *Delphi Xe5* dan citra objek diimplementasikan dengan menggunakan metode deteksi tepi *Kirsch* dan penerapan algoritma *K-Nearest Neighbor* pada proses klasifikasi mampu bekerja dengan baik serta memperoleh hasil yang maksimal.

DAFTAR PUSTAKA

- [1] Henuhili, V. (2007). Pewarisan warna bunga pada anggrek. In Seminar Nasional MIPA (Vol. 25).
- [2] Gunawan, L. W. (1986). Budidaya anggrek (Vol. 41). Niaga Swadaya.
- [3] Pamungkas, D. P. (2019). Ekstraksi Citra menggunakan Metode GLCM dan KNN untuk Identifikasi Jenis Anggrek (Orchidaceae). *Jurnal INNOVATICS: Innovation in Research of Informatics*, 1(2), 51-56.
- [4] Herfina, H. (2013). PENGENALAN POLA BENTUK BUNGA MENGGUNAKAN *PRINCIPLE COMPONENT ANALYSIS* DAN K-NN. *SEMNAS TEKNOMEDIA ONLINE*, 1(1), 07-25.
- [5] Widyawati, D. K., & Zuriati, Z. (2014). *Detecting Resemblance Of Orchid Plant Image Through Support Vector Machine (SVM) Of Kernel Linear Method*. *Jurnal Ilmiah ESAI*, 8(3), 244-257.
- [6] Awanda, M., Rismawan, T., & Midyanti, D. M. (2018). Aplikasi Klasifikasi Anggrek Berdasarkan warna dan bentuk bunga dengan metode LVQ berbasis *web*. *Coding Jurnal Komputer dan Aplikasi*, 6(2).
- [7] Pratyaswara, E. C., Wirdiani, N. K. A., & Sasmita, G. M. A. Analisis Perbandingan Metode *Canny*, *Sobel* dan HSV dalam Proses Identifikasi Bunga Anggrek Hibrida. *Jurnal Ilmiah Merpati (Menara Penelitian Akademika Teknologi Informasi)*, 11-11.
- [8] Chan, L. H., Salleh, S. H., & Ting, C. M. (2010). *Face biometrics based on principal component analysis and linear discriminant analysis*. *Journal of Computer Science*, 6(7), 693.
- [9] Whidhiasih, R. N., Wahanani, N. A., & Supriyanto, S. (2013). Klasifikasi Buah Belimbing Berdasarkan Citra *Red-Green-Blue* Menggunakan Knn Dan Lda. *Penelitian Ilmu Komputer Sistem Embedded Dan Logic*, 1(1), 155397.
- [10] Backer, D. (2005). *Learning from Data Nearest neighborhood Classification*.

-
- [11] D. Erwanto, S. A. D. Prasetyowati, dan E. N. Budi Susila, "Utilization of Digital Image Processing in Process of Quality Control of the Primary Packaging of Drug Using Color Normalization Method," in *IOP Conference Series: Materials Science and Engineering*, 2017
-