

Penggunaan Algoritma K- Nearest Neighbor untuk Identifikasi Citra Kamboja

Resty Wulanningrum

Teknik Informatika, Fakultas Teknik, Universitas Nusantara PGRI Kediri

Abstrak– Teknologi saat ini sangat berkembang dengan pesat, terutama dari sector pertanian. Dalam hal pertanian saja, tidak hanya dalam bidang tanaman khusus makanan pokok, tetapi sudah merambah ke dunia tanaman hias. Tanaman hias saat ini juga sangat tinggi permintaan, salah satunya adalah tanaman hiaskamboja. Kamboja merupakan salah satu tanama hias dengan banyak jenis. Jenis yang berbeda- beda ini juga memiliki ciri serta aroma yang berbeda- beda. Tak jarang sebagai orang awam sering salah menyebutkan jenis bunga kamboja ini. Karena secara umum terlihat sama bentuknya, ternyata bunga kamboja memiliki ciri yang unik dari bentuk serta warnanya.

Dari permasalahan di atas, maka dibuatlah sebuah rumusan masalah bagaimana membuat sebuah system untuk mengidentifikasi citra bunga kamboja dengan menggunakan algoritma K-Nearest Neighbor?

Pada penelitian ini yang digunakan adalah jenis *plumeria obtusa*, *plumeria rubra*, dan *plumeria cendana*. Jeni tanaman yang digunakan ini memiliki jumlah kelopak bunga yang sama, yaitu 5, tetapi bentuknya berbeda- beda. Data yang digunakan sebanyak 20 citra bunga setiap jenisnya. Dari 20 data tersebut akan dibuat 5 sekenario ujicoba untuk mendapatkan hasil yang terbaik. Hasil terbaik pada ujicoba yang dilakukan adalah pada sekenario pertama dengan akurasi sebesar 88,9% dan yang terkecil pada sekenario ke- 5 yaitu sebesar 75,9 %. Hasil ujicoba yang bervariasi tersebut dipengaruhi oleh data training dan data testing. Semakin banyak data trainingnya maka semakin tinggi pula hasil akurasi identifikasinya, begitu sebaliknya. Disarankan menggunakan metode ekstraksi ciri yang lain untuk mendapatkan hasil yang maksimal.

Kata Kunci— Kamboja , K-NN, Identifikasi

Abstract – Today's technology is growing rapidly, especially from the agricultural sector. In the case of agriculture alone, not only in the field of special staple food crops, but has penetrated into the world of ornamental plants. Ornamental plants today are also very high demand, one of which is hiaskamboja plant. Cambodia is one of many ornamental plants. Different types also have different features and smells. Not infrequently as a layman often wrong mention this type of frangipani flowers. Because in general looks the same shape, it turns out frangipani flowers have unique characteristics of the shape and color.

From the above problem, then made a problem formulation how to make a system to identify image of frangipani flower by using K-Nearest Neighbor algorithm?

In this study used is *plumeria obtusa*, *plumeria rubra*, and *plumeria sandalwood*. The plant used has the same number of petals, 5, but the shape is different. The data used as many as 20 images of flowers of each type. Of the 20 data will be made 5 test scenarios to get the best results. The best results in the trial were in the first scenario with an accuracy of 88.9% and the smallest in the 5th scenario of 75.9%. The varied test results are influenced by training data and data testing. The more training data the higher the accuracy of the identification, and vice versa. It is recommended to use another feature extraction method to get maximum results.

Keywords— *Plumeria*, K-NN Identification

1. PENDAHULUAN

Tanaman bunga Kamboja merupakan salah satu tanaman hias outdoor yang eksklusif karena aroma dan warna bunganya yang khas serta berperan penting sebagai penghias tanaman tropis karena Kamboja memiliki ciri khas yang ever green. Terlebih bunganya tumbuh sepanjang tahun dan mudah beradaptasi dengan berbagai iklim [1].

Pada penelitian sebelumnya yaitu tentang bunga adenium menggunakan metode Learning Vector Quantization. Learning Vector Quantization (LVQ) merupakan salah satu algoritma yang digunakan untuk clustering. Berdasarkan skenario pengujian yang dilakukan, aplikasi identifikasi citra kelopak bunga adenium menghasilkan akurasi sebesar 86,66%. [2]

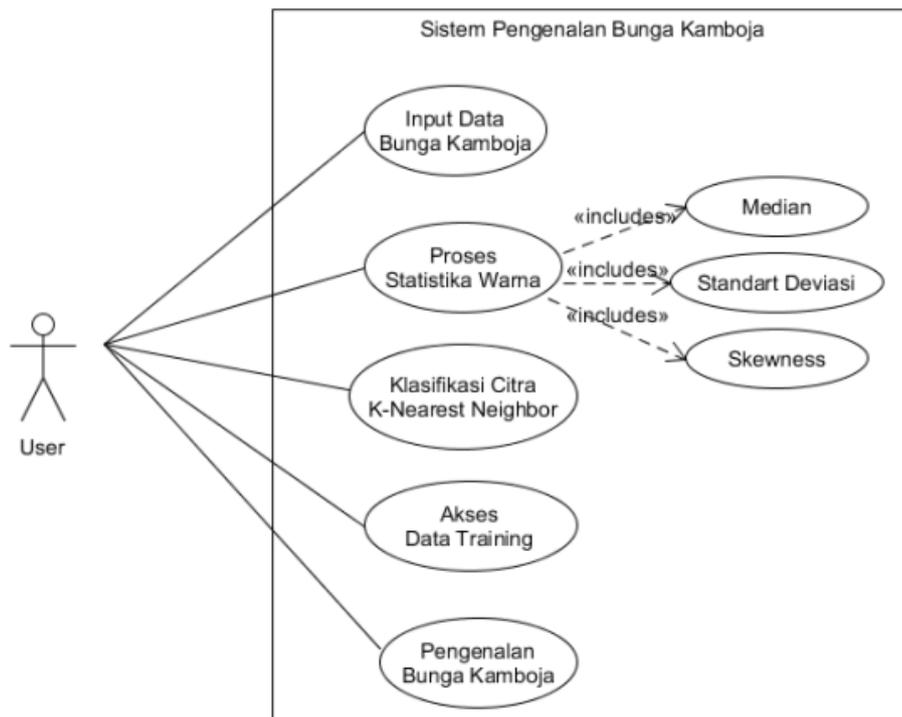
Pengenalan suatu objek merupakan suatu hasil yang mudah dilakukan oleh manusia, namun tidak demikian bagi sebuah mesin atau komputer. Seseorang dapat dengan mudah mengenali orang yang pernah dikenal sebelumnya walaupun hanya dengan mendengar suaranya dari kejauhan tanpa melihat langsung orang tersebut [3].

Penelitian ini mencoba memberikan alternatif baru untuk perhitungan nilai klasifikasi warna statistis yaitu fitur warna seperti rerata, deviasi standar, dan skewness. Fitur-fitur tersebut dapat digunakan untuk kepentingan identifikasi tanaman hias [4].

Berdasarkan latar belakang di atas, maka dibuatlah sebuah penelitian dengan judul Penggunaan Algoritma K- Nearest Neighbor untuk Identifikasi Citra Kamboja.

2. METODE PENELITIAN

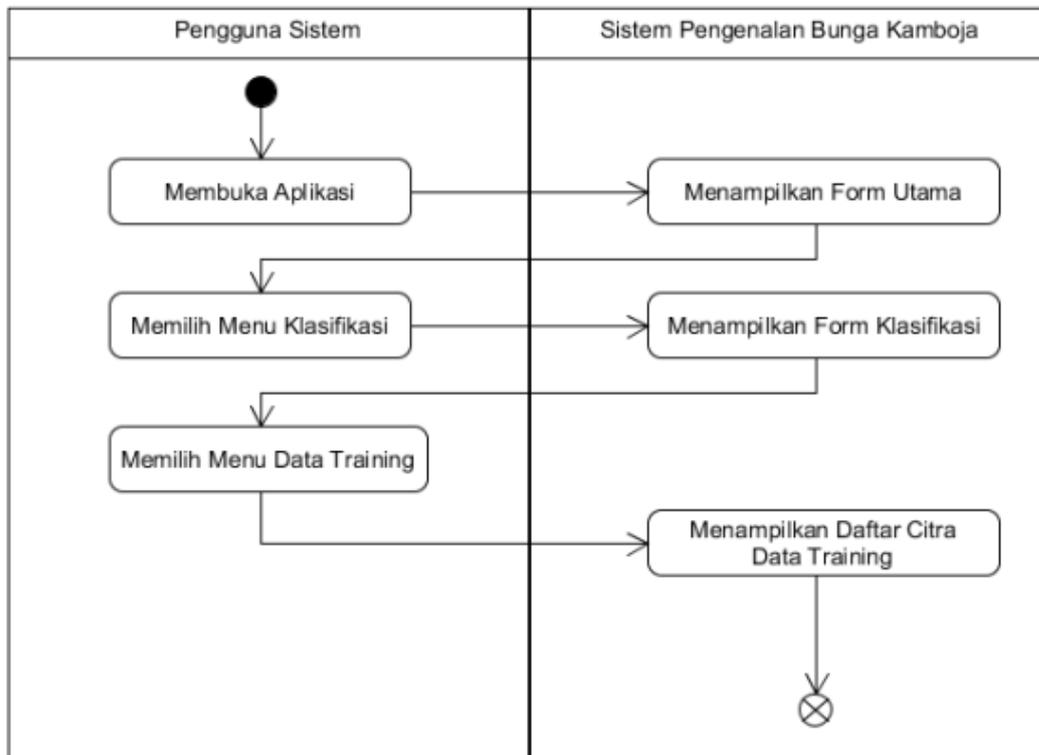
3. a. Use case Diagram



Gambar 1. Use Case Diagram

Penjelasan dari Use Case diagram pada gambar 1 adalah : Seorang user masuk aplikasi di menu utama. Kemudian Input Data Bunga Kamboja. User menginputkan data bunga Kamboja yang ada di media Penyimpanan. Setelah itu Proses Statistika Warna. User melakukan perhitungan proses statistika warna yaitu median, standart deviasi dan skewness. Proses selanjutnya Klasifikasi Citra K-Nearest Neighbor. Sistem akan melakukan klasifikasi menggunakan metode K-NN (K-Nearst Neighbor) dan program selesai. Terakhir Hasil klasifikasi ditampilkan ke aplikasi dan User dapat melihat hasil pencocokan kemudian keluar setelah selesai.

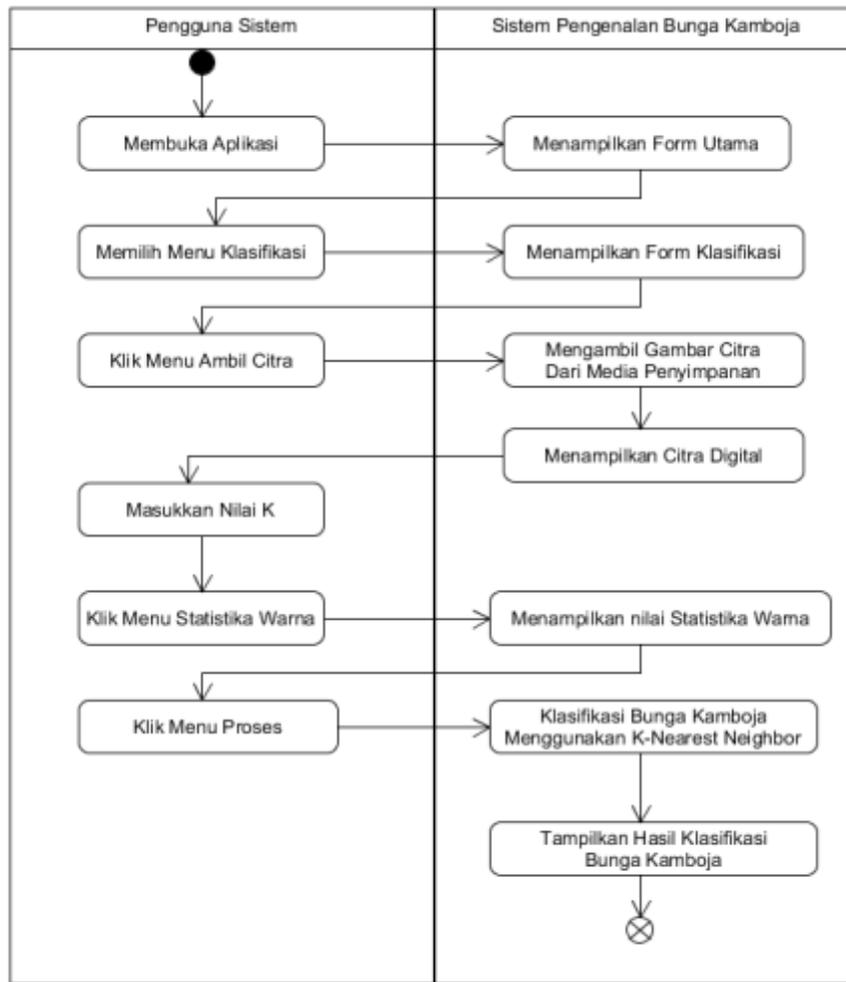
b. Activity Diagram



Gambar 2. Activity Diagram data *Training*

Pada Activity Diagram Data Training terlihat pada gambar 2 terdapat beberapa langkah-langkah untuk menjalankan Data Training yaitu :

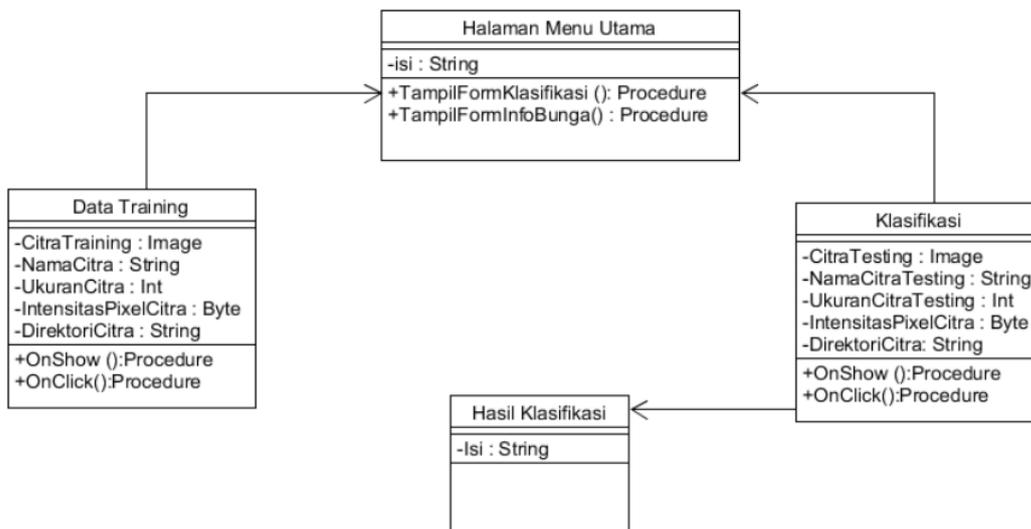
- 1) User membuka aplikasi maka sistem aplikasi akan menampilkan Form Utama.
- 2) Selanjutnya user memilih Menu Data Training maka sistem akan menampilkan daftar citra training.



Gambar 3. Activity Diagram Klasifikasi

Langkah-langkah pada menu Activity Diagram Klasifikasi pada gambar 3 adalah:

- 1) User membuka aplikasi, maka sistem aplikasi akan menampilkan Form Utama.
- 2) Selanjutnya user memilih menu Klasifikasi, maka sistem akan menampilkan Form Klasifikasi
- 3) Klik menu ambil citra dari data penyimpanan.
- 4) Menampilkan citra Digital Gambar.
- 5) Memasukkan nilai K .
- 6) Selanjutnya user memilih Statistka warna untuk menampilkan nilai statistika warna dari 3 nilai statistika warna yang telah.
- 7) Selanjutnya user akan memilih proses untuk mendapatkan nilai dari citra bunga agar dapat melakukan pencocokan data testing dengan data training, maka sistem akan menampilkan hasil pencocokan citra bunga.



Gambar 4. Class Diagram

Penjelasan gambar 4 Class Diagram :

Pada class Diagram mempunyai nilai disetiap masing-masing relasi, hal ini bertujuan untuk membedakan fungsi dari setiap relasi yang dibuat, Diantaranya adalah : Pada Halaman Utama terdapat 2 relasi yaitu “Klasifikasi” dan Data “Training”.

- 1) Pada “Data Training” berisi Citra training dan Spesifikasi dari citra training.
- 2) Pada “Klasifikasi” digunakan untuk menginputkan citra gambar dari kelopak bunga Kamboja untuk di analisis

A. Pengertian Bunga Kamboja Plumeria

Plumeria merupakan salah satu genus yang termasuk di dalam famili Apocynaceae. Famili ini merupakan salah satu dari sepuluh famili kelompok angiosperma terbesar menurut banyak jenisnya, selain Leguminosae, Orchidaceae, Rubiaceae, Cyperaceae, Begoniaceae, Solanaceae, Asteraceae, Myrtaceae, dan Euphorbiaceae. Tanamantanaman yang termasuk di dalam famili Apocynaceae dikenal menghasilkan berbagai metabolit sekunder seperti alkaloid dan cardenolid yang digunakan dalam obat-obatan tradisional untuk sakit maag, demam, asma, dan batuk rejan [5]. Kamboja bunga tropis yang cantik dan eksotis termasuk tanaman sukulen, yakni tumbuhan yang dapat menyimpan air pada seluruh bagian tubuhnya, dari akar, batang, daun, bunga, sampai buah [6].

B Algoritma K-Nearest Neighbor (KNN)

K-Nearest Neighbor Classifier didasari pada pembelajaran dengan persamaan sampel yang diujikan untuk diuraikan oleh ndimensi atribut angka. Masing-masing sampel mempresentasikan sebuah titik pada ruang n dimensi. Ketika ada sebuah sampel yang tidak diketahui, K-Nearest Neighbor Classifier mencari pola ruang dari K sampel yang diuji yang sama sampel tersebut dicari yang terdekat dengan sampel yang tidak diketahui. Terdekat artinya jarak kedekatan data berdasarkan aturan Euclidean Distance. Aturan pada Euclidean Distance dihitung diantara dua titik, dan setiap atribut telah mendapat bobot nilainya [7]. K-Nearest Neighbor merupakan metode klasifikasi yang tangguh terhadap training data yang memiliki banyak noise dan metode ini memiliki tingkat akurasi yang cukup tinggi serta efektif apanila training datanya besar [8]. Sesuai dengan prinsip kerja K-Nearest Neighbor yaitu mencari jarak terdekat antara data yang akan dievaluasi dengan K tetangga (Neighbor) terdekatnya dalam data pelatihan. Persamaan dibawah

ini menunjukkan rumus perhitungan untuk mencari jarak terdekat dengan d adalah jarak dan p adalah dimensi data [9]:

$$d_i = \sqrt{\sum_{i=1}^p (x_{2i} - x_{1i})^2}$$

Dengan keterangan indeks sebagai berikut :

x_1 : sampel data
 x_2 : data uji
 i : variabel data
 d : jarak
 p : dimensi data

Pada fase training, algoritma ini hanya melakukan penyimpanan vektor-vektor fitur dan klasifikasi data training sample. Pada fase klasifikasi, fitur-fitur yang sama dihitung untuk testing data (yang klasifikasinya tidak diketahui). Jarak dari vektor baru yang ini terhadap seluruh vektor training sample dihitung dan sejumlah k buah yang paling dekat diambil. Titik yang baru klasifikasinya diprediksikan termasuk pada klasifikasi terbanyak dari titik-titik tersebut. Adapun algoritma dalam melakukan prediksi dengan menggunakan metode k-nearest neighbor adalah sebagai berikut ini [9]:

- 1) Tentukan nilai K dimana nilai tersebut akan digunakan sebagai parameter pembatas jumlah titik sampel yang akan diuji
- 2) Hitung beda antara semua titik sampel dengan titik uji. Kemudian urutkan dari yang paling dekat dengan titik uji.
- 3) Dimulai dari titik sampel yang paling dekat, sampai dengan ke- K lakukan pencarian nilai mayoritas yang dinyatakan oleh titik sampel.
- 4) Nilai mayoritas tersebut adalah merupakan hasil prediksi

4. HASIL DAN PEMBAHASAN

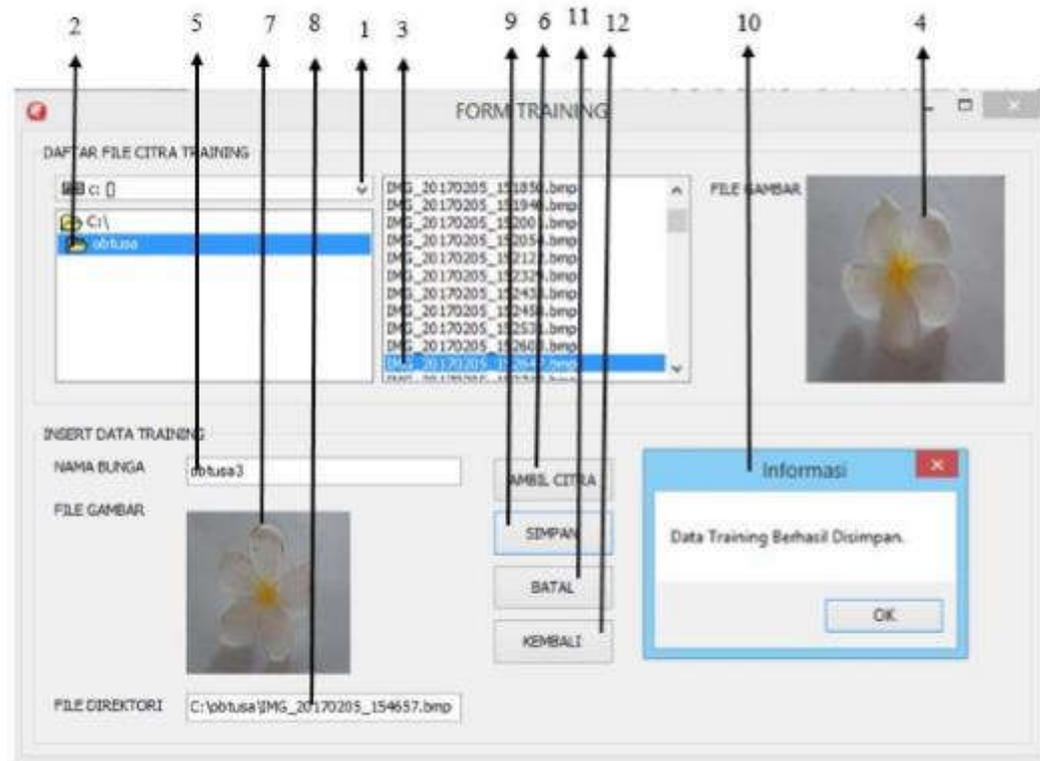
A. Implementasi

Keterangan dari gambar 5. adalah sebagai berikut :

1. Tombol Klasifikasi. Berfungsi untuk masuk ke form klasifikasi atau data testing.
2. Tombol Info Bunga. Berfungsi untuk menampilkan info bunga.
3. Tombol Keluar. Berfungsi untuk keluar dari aplikasi



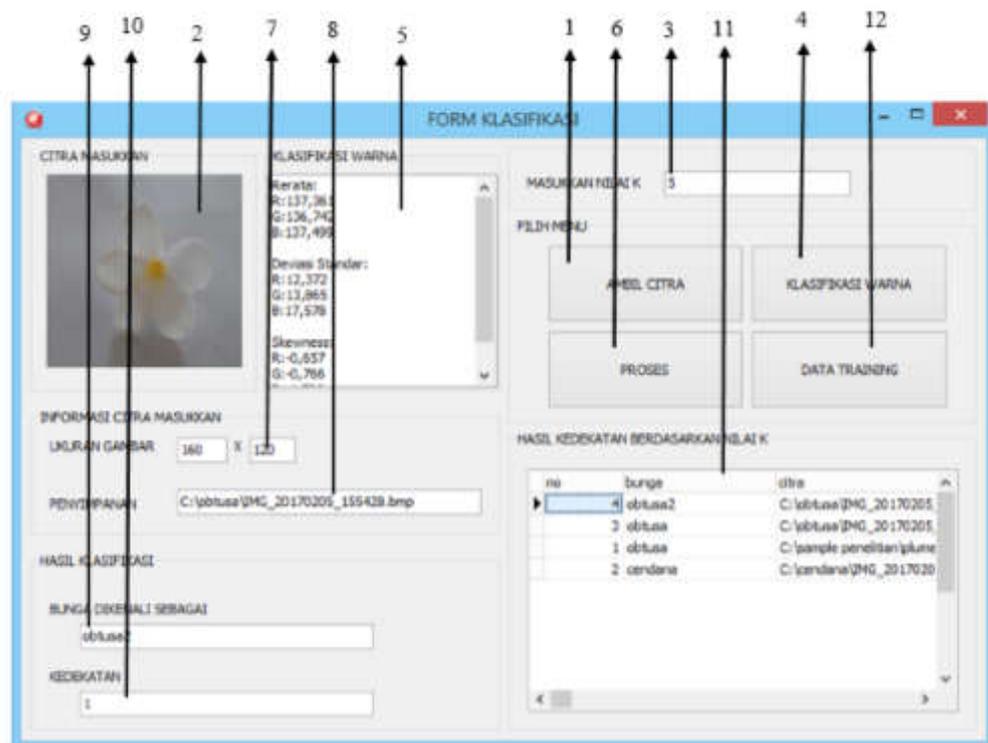
Gambar 5. Tampilan Utama



Gambar 6. Tampilan Form Training

Penjelasan Gambar 6 adalah sebagai berikut:

1. ComboBox. Berfungsi untuk menampilkan device komputer yang dipilih.
2. Directory ListBox. Berfungsi untuk menampilkan folder dimana citra disimpan.
3. File ListBox. Berfungsi menampilkan nama file yang disimpan dalam folder.
4. Image Preview. Berfungsi untuk menampilkan citra bunga yang ada dalam filelistbox.
5. Nama Bunga. Menambahkan nama bunga untuk diinputkan kedalam database.
6. Tombol Ambil Citra. Memasukkan gambar yang akan disimpan di database.
7. Image Preview. Berfungsi untuk menampilkan citra bunga yang ada dalam filelistbox.
8. File Direktori. Menampilkan folder dimana citra disimpan.
9. Tombol Simpan. Untuk menyimpan data yang sudah diinputkan kedalam database.
10. Form Informasi. Jika form informasi muncul maka data training berhasil disimpan di database.
11. Tombol Batal. Berfungsi jika tidak jadi menambahkan data di database.
12. Tombol Kembali. Berfungsi untuk kembali ke form klasifikasi



Gambar 7. Tampilan form Klasifikasi

Keterangan dari gambar 7 adalah sebagai berikut :

1. Tombol Ambil Citra. Berfungsi untuk memasukkan gambar yang akan diproses.
2. Image Preview Citra Masukkan. Berfungsi untuk menampilkan gambar yang dipilih dan akan diproses.
3. Masukkan Nilai K. Menginputkan nilai K untuk klasifikasi.
4. Tombol Klasifikasi Warna. Berfungsi untuk memproses gambar yang diinputkan perhitungan statistis seperti rerata, deviasi standar, dan skewness untuk setiap komponen R, G, dan B.
5. Tombol Proses. Berfungsi untuk menghasilkan hasil klasifikasi pengenalan jenis bunga yang telah diuji.
6. Ukuran Gambar. Berfungsi untuk menampilkan ukuran piksel gambar yang akan diproses.
7. Penyimpanan. Berfungsi untuk menampilkan nama file yang disimpan dalam folder.
8. Bunga dikenali sebagai. Menampilkan hasil klasifikasi dari citra yang telah diuji.
9. Kedekatan. Menampilkan hasil nilai kedekatan terkecil dari citra yang telah diuji.
10. Hasil Kedekatan Berdasarkan Nilai K. Menampilkan tabel hasil data citra yang diuji.
11. Tombol Data Training. Berfungsi untuk masuk ke form data training.

B. Skenario Uji Coba

Bunga kamboja proses pengujian dilakukan lima kali dengan memperhatikan jumlah citra bunga kamboja yang terdapat pada data training dan data testing. Pengujian dilakukan dengan ketentuan citra bunga kamboja yang digunakan sebagai data training. Berikut adalah tabel dari skenario uji coba data training :

Tabel 1. Skenario Ujicoba Data Training

No	Data Training			Jumlah
	Plumeria Obtusa	Plumeria Rubra	Plumeria Cendana	
1.	1,2,3,4,5,6,7,8,9,10,11,12,13,14	1,2,3,4,5,6,7,8,9,10,11,12,13,14	1,2,3,4,5,6,7,8,9,10,11,12,13,14	42
2.	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
3.	1,2,3,4,5,6,7,8	1,2,3,4,5,6,7,8	1,2,3,4,5,6,7,8	24
4.	1,2,3,4,5	1,2,3,4,5	1,2,3,4,5	15
5	1,2	1,2	1,2	6

Tabel 2. Skenario Ujicoba Data Testing

No	Data Testing			Jumlah
	Plumeria Obtusa	Plumeria Rubra	Plumeria Cendana	
1.	15,16,17,18,19,20	15,16,17,18,19,20	15,16,17,18,19,20	18
2.	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
3.	9,10,11, 12,13,14,15,16,17,18,19,20	9,10,11, 12,13,14,15,16,17,18,19,20	9,10,11, 12,13,14,15,16,17,18,19,20	36
4.	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	45
5	3,4,5,6,7,8,9,10,11, 12,13,14,15,16,17,18,19,20	3,4,5,6,7,8,9,10,11, 12,13,14,15,16,17,18,19,20	3,4,5,6,7,8,9,10,11, 12,13,14,15,16,17,18,19,20	54

1. Hasil Skenario Ujicoba

Tabel 3. Hasil Skenario Ujicoba

Skenario	Data		Hasil		Prosentase
	Training	Testing	Benar	Salah	
1.	42	28	16	2	88,9%
2.	33	27	23	4	85,2%
3.	24	36	31	5	86,1%
4.	15	45	37	8	82,2%
5.	6	54	41	13	75,9%

Dari hasil skenario uji coba dapat dijelaskan bahwa hasil dari proses training dan proses testing di masing-masing skenario adalah :

Skenario 1 : Menghasilkan nilai benar 16 dan nilai salah 2, maka prosentase akurasi didapat sebesar 88,9%.

Skenario 2 : Menghasilkan nilai benar 23 dan nilai salah 4, maka prosentase akurasi didapat sebesar 85,2%.

Skenario 3 : Menghasilkan nilai benar 31 dan nilai salah 5, maka prosentase akurasi didapat sebesar 86,1%.

Skenario 4 : Menghasilkan nilai benar 37 dan nilai salah 8, maka prosentase akurasi didapat sebesar 82,2%.

Skenario 5 : Menghasilkan nilai benar 41 dan nilai salah 13, maka prosentase akurasi didapat sebesar 75,9%.

Jadi, kesimpulannya dari skenario uji coba semakin banyak gambar training akan membuat akurasi aplikasi semakin tinggi yaitu menghasilkan nilai prosentase 88,9%.

5. SIMPULAN

Berdasarkan hasil uji coba dapat disimpulkan bahwa

1. KNN dapat digunakan untuk identifikasi citra bunga Kamboja.
2. Penggunaan data ujicoba yang bervariasi menghasilkan hasil akurasi yang berbeda- beda pula.
3. Penggunaan data training yang banyak menghasilkan akurasi yang tinggi pula.
4. Hasil terbaik dari 5 skenario ujicoba adalah pada skenario 1, yaitu dengan hasil 88,9% dan terendah pada skenario 5, yaitu 75,9 %.

6. SARAN

Saran dari peneliti untuk pengembangan selanjutnya adalah

1. Mengembangkan aplikasi dengan penggunaan ekstraksi ciri.
2. Membuat aplikasi menjadi lebih interaktif sehingga akan mempermudah penggunaannya.
3. Mampu mengenali dengan jenis data bunga yang lebih banyak.

DAFTAR PUSTAKA

- [1] Utami, Prapti & Tim Lentera. 2003. Tanaman obat untuk mengatasi diabetes mellitus. Jakarta : Agromedia Pustaka. Hal:1,2,6,123,24.
- [2] Wulanningrum, Resty, Bagus F.R. 2017. Klasifikasi Citra Adenium Menggunakan Learning Vector Quantization. Prosiding Seminar Nasional Teknologi Informasi dan Multimedia 2017. STIMIK AMIKOM Yogyakarta. 4 Februari 2017.
- [3] Duda R, O., Hart P. E., and Strok D. G., 2000. Pattern Clasification. 2ed., Wiley, USA
- [4] Kadir, A.; Nugroho, L.E.;Susanto, A.,and Santoso, P.I.2011b. Leaf Classification Using Shape, Color , and Texture Features. On International Journal of Computer Trends and Technology,Vol.2, Issue 1. 225-230. Kadir, A.; Nugroho, L.E.;Susanto, A.,and Santoso, P.I.2011c. Neural Network Application on Foliage Plant Identification. On International Journal of Computer Applications,Vol.29,No.9. 15-22.
- [5] Van der Heijden, E.W., Kuyper, T.W., 2001. Does origin of mycorrhizal fungus or mycorrhizal plant influence effectiveness of the mycorrhizal symbiosis? Plant and Soil 230: 161-174
- [6] Handoko, Bagus Haryo. 2008. Pachypodium. Yogyakarta: Gramedia Pustaka Utama.
- [7] Mehmed, Kantardzic. 2002. Data Mining Concepts, Models, Methods and Algorithms. USA : a John Wiley & Sons.
- [8] Yofianto, Evan. 2010. K-Nearest Neighbor. (Online)
<http://kuliahinformatika.wordpress.com/2010/02/13/buku-ta-k-nearestneighbor-knn/>
diakses 28 November 2016.
- [9] Agusta, Y. 2007. K-Means, Penerapan, Permasalahan Dan Metode Terkait, Jurnal Sistem dan Informatika, Vol 3, (Februari 2007), 47-60.