

Perbandingan Euclidean Distance Dan K-Nearest Neighbor Dalam Pengenalan Bunga

Solehatin, M.Kom, Teknik Informatika, Sekolah Tinggi Ilmu Komputer PGRI, Banyuwangi, Indonesia

ABSTRAK

Penerapan pengenalan pola yang tepat saat melakukan ekstraksi fitur pada daun diharapkan dapat digunakan untuk penelitian pada karakteristik botani jenis daun. Penelitian ini menggunakan pengolahan gambar digital yang bertujuan mengotomatisasi dan klasifikasi tanaman bunga berdasarkan dari gambar bunga.

Dalam penyelesaiannya banyak metode yang dapat digunakan, dalam penelitian ini menerapkan 2 (dua) metode yaitu metode Euclidean Distance dan metode K-NN. Pada metode Euclidean Distance menggunakan nilai RGB per region sedangkan metode K-NN perhitungan yang dilakukan dengan 11 (sebelas) kriteria yang ada pada bunga antar antara lain nilai minimal RGB, nilai maksimal RGB, jarak minimal ujung daun dan putik, jarak maksimal ujung daun dan putik, jarak minimal antar ujung daun, jarak maksimal antar ujung daun, dan jumlah daun.

Dari hasil penelitian menunjukkan nilai prosentase kemiripan terhadap bunga dengan metode K-NN lebih tinggi dari pada menggunakan metode Euclidean distance, terlihat dari 5 hasil uji coba yang telah dilakukan. Metode Euclidean distance menghasilkan prosentase terkecil 56,526 % dan terbesar 86,447 %, sedangkan pada metode K-NN nilai prosentase terkecil 84,466 % dan terbesar 93,993 %. Hal ini dikarenakan metode K-NN menggunakan 11 (sebelas) kriteria sedangkan pada metode Euclidean distance hanya menggunakan nilai dari RGB per region.

Kata kunci: *RGB, K-NN, Euclidean Distance*

1. Pendahuluan

Tanaman mempunyai peranan penting dalam kehidupan manusia dan bisa memberikan informasi yang dibutuhkan dalam peradaban pembangunan manusia. Situasi genting adalah terjadinya degradasi lingkungan dan kurang kesadaran manusia, beberapa spesies tanaman langka mempunyai resiko yang tinggi terhadap kepunahan sehingga dibutuhkan dokumen agar bisa melakukan perlindungan terhadap tanaman tersebut. Langkah pertama adalah dengan membangun sebuah database untuk dokumentasinya. Sehingga dibutuhkan aplikasi komputer untuk melakukan penggolongan atau klasifikasi terhadap tumbuhan tersebut.

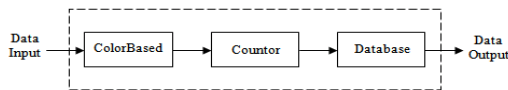
Pesatnya perkembangan basis data dalam menyimpan citra dalam ukuran besar memberikan pengaruh yang positif terhadap otomatisasi penyimpanan dan pemeliharaan data digital. Dalam penelitian tersebut dibentuk beberapa teknik untuk otomatisasi atau klasifikasi daun dalam kedalam bentuk citra digital. Penelitian ini bertujuan untuk memberikan kontribusi pada penelitian visi komputer dan penyimpanan citra berbasis konten.

P.E.Ajmire, dkk melakukan penelitian pada tahun 2012 pada pengenalan pola dengan Ekstraksi fitur yang merupakan langkah penting dalam pengenalan pola. Penerapan pengenalan pola untuk melakukan ekstraksi fitur yang tepat pada daun diharapkan dapat mendorong studi lebih lanjut pada karakteristik botani pada daun. Metodologi penelitian tersebut untuk menjelaskan pendekatan baru pada pengenalan pola untuk ekstraksi dan cita alam secara pendekatan seperti citra daun. Metode tersebut diajukan untuk pengenalan spesies tumbuhan hidup secara otomatis. Pengenalan spesies tumbuhan ini akan sangat berguna untuk penelitian botani dalam mengidentifikasi spesies tumbuhan.

Berbeda dengan disiplin ilmu pengolahan citra yang dibatasi oleh penggunaan citra sebagai masukan maupun keluarannya, suatu aplikasi pengenalan pola bertujuan untuk melakukan proses pengenalan terhadap suatu objek (misalnya citra) ke dalam salah satu kelas tertentu, berdasarkan pola yang dimilikinya. Secara umum dapat dilihat dimana gambar 1 merupakan proses pengenalan pola yang digambarkan dalam diagram blok sederhana.

* *Corresponding author.* Telepon: -.

E-mail: atin33@yahoo.co.id



Gambar 1 Diagram blok sistem pengenalan pola

Pola data input kemudian diolah dalam bagian-bagian yang akan dijelaskan yaitu :

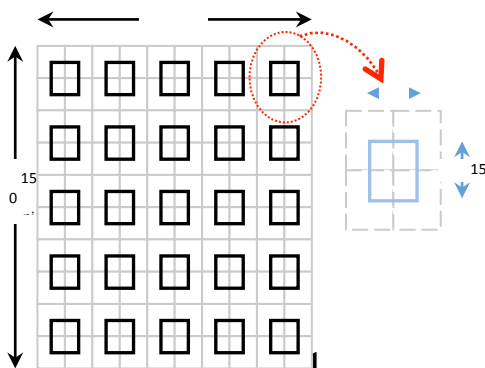
1. Proses penginputan data dengan cara menginputkan data gambar
2. *Color Based* meliputi tahapan-tahapan yang diperlukan untuk membawa input data ke dalam suatu bentuk yang dapat diterima melalui proses deteksi warna masing-masing lapisan pada bunga.
3. Tahap berikutnya adalah *Countour Based* meliputi langkah deteksi bentuk bunga dari setiap objek inputan data.

Dari hasil pencocokan warna dan bentuk bunga akan melakukan klasifikasi data yaitu pada saat melakukan proses pengenalan pola, di dalam sistem yang telah dirancang sehingga nantinya program akan mengklasifikasikan pola tersebut berdasarkan inputan data dengan mencocokkan pola-pola gambar yang ada dalam database.

Pada penelitian ini menggunakan pengolahan gambar digital bertujuann mengotomatisasi dan klasifikasi tanaman bunga berdasarkan dari gambar bunga. Kemudian dari ciri bunga yang ada diantaranya yaitu warna dan bentuk dari bunga yang digunakan sebagai pengenalan pola nantinya.

2. Euclidean Distance

Proses awal pada perhitungan region yaitu pembagian region per pixel dari gambar menjadi 25 region. Untuk ilustrasi pembagian region disajikan pada gambar 2.



Gambar 2 Pembagian Region Pada 150 x 150 Pixel

Citra input yang berukuran 150x150 dipecah menjadi 10x10 region sehingga menghasilkan 100 region. Kemudian dari 100 region yang ada dibentuk lagi ke dalam 25 region dari tiap 4 (empat) region yang berdekatan dan region yang terbentuk berada ada tiitk tengah 4 region tersebut dengan ukuran 15x15 pixel.

Pengukuran dengan metode Euclidean distance dilakukan dengan menghitung jumlah dari nilai masing-masing RGB per region kemudian mencari selisih masing-masing RGB dari data yang ada pada database dan data uji coba dari 25 region yang terbentuk. Kemudian hasil perhitungan yang didapat dicari prosentase per region, dari prosentase yang dihasilkan dijumlah seluruhnya. Jika hasil prosentase yang dihasilkan per region kurang dari 0 maka dianggap bernilai 0 dan tidak dijumlahkan untuk perhitungan prosentasenya. Hasil total prosentase yang ada kemudian di bagi dengan jumlah banyak region (25 region) dikalikan dengan 3 (RGB) kemudian dibagi 100 %.

Perhitungan Prosentase per regoin :

$$(r2 - (abs (r1 - r2))) / r2 \dots\dots (1)$$

Dimana :

r1 = nilai RGB pada Uji Coba

r 2 = nilai RGB pada Database

Perhitungan Prosentase Global :

$$(Total Prosentase / 75) * 100 \% \dots\dots (2)$$

3. K-Nearest Neighbour

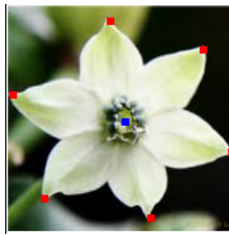
Penentuan nilai dengan metode K-NN dilakukan pada 11 kriteria yang dibutuhkan terlihat pada table 1.

Tabel 1 Hasil Perhitungan Pada 11 Kriteria

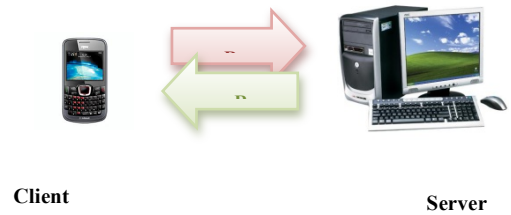
NO	Kriteria
1	Nilai Minimal Red
2	Nilai Minimal Green
3	Nilai Minimal Blue
4	Nilai Maksimal Red
5	Nilai Maksimal Green
6	Nilai Maksimal Blue
7	Jarak Minimal ujung daun dan putik
8	Jarak Maksimal ujung daun dan putik
9	Jarak Minimal antar ujung daun
10	Jarak Maksimal antar ujung daun
11	Jumlah Daun

Berbeda dengan metode Euclidean Distance, nilai minimal dan maksimal red, green dan blue didapat dari keseluruhan pixel gambar bunga. Untuk mendapatkan nilai tersebut dilakukan dengan cara scanning konvensional mulai dari kiri atas hingga kanan bawah. Alasan menggunakan cara scanning konvensional adalah untuk mendapatkan hasil yang akurat.

Penentuan titik putik dan ujung-ujung daun dilakukan secara manual dapat dilihat pada gambar 3. Dari hasil penentuan titik dapat dicari jarak minimal dan maksimal antar putik dan ujung daun, jarak minimal dan maksimal antar ujung daun, dan jumlah daun. Penentuan titik putik dan ujung daun dilakukan secara terurut searah jarum jam atau berlawanan.



Gambar 3 Contoh Titik Putik Dan Ujung Daun



Gambar 4 Blok Diagram Rancangan Sistem

Pengukuran pada metode K-NN dilakukan mencari nilai perkalian antara nilai bobot dan nilai kedekatan kemudian dibagi dengan jumlah bobot yang digunakan. Untuk penelitian ini nilai bobot di beri nilai 1 adapun contoh perhitungan yang dilakukan terlihat pada tabel 2 dan persamaan 3.

Tabel 2 Contoh Hasil Perhitungan Pada 11 Kriteria

Kriteria	Nilai Bobot	Nilai Kedekatan
Jarak Minimal ujung daun dan putik	1	0.4
Jarak Maksimal ujung daun dan putik	1	0.8
Jarak Minimal antar ujung daun	1	0.75
Jarak Maksimal antar ujung daun	1	0.5
Jumlah Daun	1	0.65

Perhitungan K-NN :

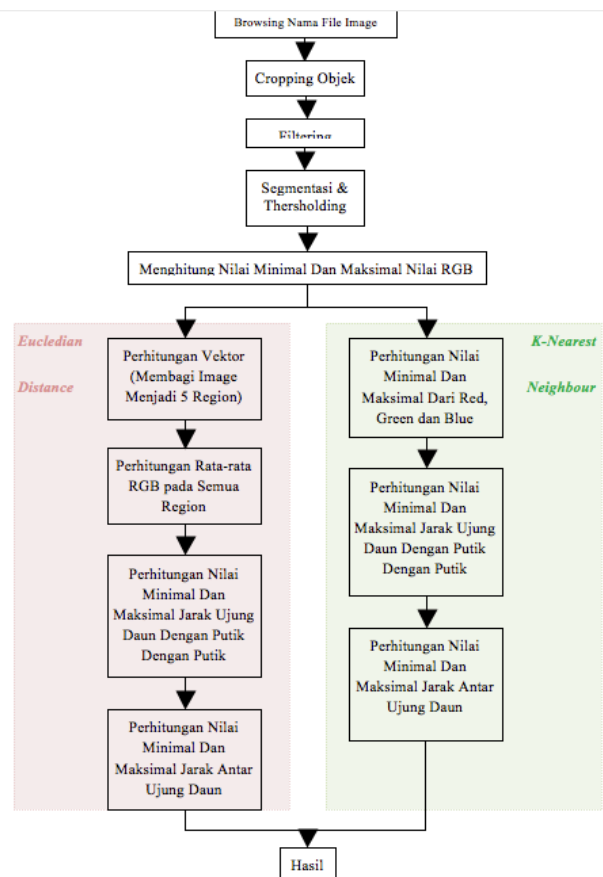
$$\frac{[(1*0.4)+(1*0.8)+(1*0.75)+(1*0.5)+(1*0.65)]}{5} = 0.62 = 62\% \dots\dots\dots (3)$$

4. Desain Sistem

Langkah pertama yaitu citra bunga yang diambil secara langsung melalui Gated, kemudian data dikirim ke server. Image diperoleh dari pengambilan gambar langsung melalui gadget seperti handphone atau tablet, dan dikirim ke server melalui jalur internet. Jenis kompresi yang disupport dalam penelitian ini dibatasi beberapa tipe yaitu jpg, bmp, gif dimana tipe-tipe tersebut adalah tipe kompresi yang umum digunakan pada perangkat gadget yang beredar dipasaran utamanya wilayah Indonesia. Ada beberapa ketentuan atau aturan khusus dalam pengambilan gambar bunga, gambar bunga diambil dengan posisi tegak lurus dan sehingga posisi putik berada di tengah.

Selain itu diupayakan tidak ada bagian bunga yang terpotong dan diupayakan batas pengambilan gambar kiri, kanan, atas dan bawah adalah tepat pada batas dari ujung daun. Hal ini dimaksudkan untuk mempermudah user dalam proses cropping serta mempercepat kinerja sistem ditahap selanjutnya. Untuk ukuran optimal gambar inputan dibatasi hingga ukuran 1300 x 1300 pixel, sedangkan batas bawahnya 80 x 80 pixel.

Dari hasil gambar yang dikirim akan diproses dalam server dengan metode yang digunakan kemudian hasilnya akan dikirim lagi ke handphone berupa data link ke Wikipedia dan Google, link tersebut berisi informasi detail dari bunga yang diinputkan terkait dengan nama, genus, family dan ordo dari bunga tersebut.



Gambar 5 Blok Diagram Rancangan Algoritma

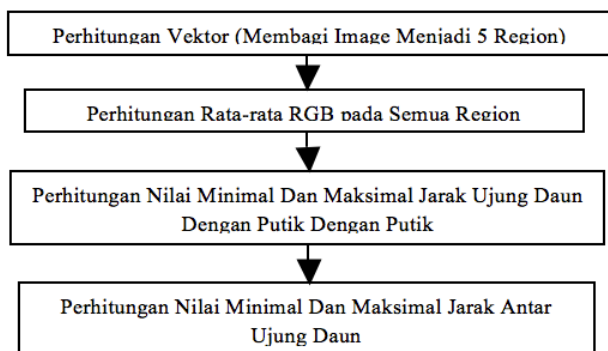
Rancangan algoritma yang digunakan untuk mendapatkan hasil yang sesuai dengan tujuan penelitian ini terlihat pada gambar 5. Untuk rancangan algoritma pada penelitian ini langkah awal yang dilakukan adalah mengolah gambar inputan agar dapat diproses pada proses lebih lanjut yang disebut sebagai preprocessing. Pada proses preprocessing dimulai dari browsing nama file kemudian dilakukan cropping terhadap gambar dan

difiltering selanjutnya baru dilakukan proses segmentasi dan thresholding.

Adapun tujuan utamanya proses preprocessing antara lain deteksi tepi, mencari titik-titik ujung daun dan mencari jumlah daun. Data tersebut nantinya akan digunakan pada proses recognition baik pada Euclidean distance ataupun k-nearest neighbour.

Proses kedua adalah menghitung nilai minimal dan maksimal untuk tiap-tiap komponen warna Red, Green dan Blue, dari hasil perhitungan ini nantinya digunakan sebagai kriteria yang akan diperhitungkan pada metode yang digunakan.

Proses pada ecludian distance ada beberapa tahapan yang dilakukan yaitu terlihat pada gambar 6.



Gambar 6 Blok Diagram Metode Ecludian Distance

Pada proses algoritma KNN terdapat beberapa perhitungan yang harus dicari terlebih dahulu antara lain terlihat pada gambar 7.

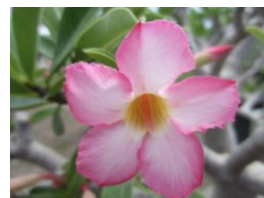


Gambar 7 Blok Diagram Metode K-NN

5. UJI COBA

Uji Coba pertama dilakukan pada jenis bunga kamboja merah dengan detail yaitu kelas *asteridae*, ordo *gentianales*, famili *apocynaceae* dan termasuk spesies *plumeria rubra l.* Adapun gambar bunga terlihat pada gambar 8.

Pada perhitungan lainnya berupa perhitungan pada 25 region terhadap nilai red, green dan blue. Perhitungan ini dilakukan dengan mencari nilai masing RGB kemudian dicari nilai rata-rata dari masing-masing nilai RGB. Contoh hasil perhitungan terlihat pada tabel 4.



Gambar 8 Contoh Bunga Kamboja Merah

Tabel 4 Hasil Perhitungan Pada 25 Region

	Rata-Rata Merah	Rata-Rata Hijau	Rata-Rata Biru
Region 1.1	95	99	91
Region 1.2	106	88	81
Region 1.3	195	138	178
Region 1.4	79	95	48
Region 1.5	67	89	65
Region 2.1	78	75	74
Region 2.2	153	110	126
Region 2.3	198	168	191
Region 2.4	193	116	157
Region 2.5	165	94	131
Region 3.1	174	116	157
Region 3.2	183	151	173
Region 3.3	136	73	34
Region 3.4	166	122	138
Region 3.5	110	79	85
Region 4.1	103	103	104
Region 4.2	142	88	118
Region 4.3	162	133	158
Region 4.4	159	130	153
Region 4.5	133	80	108
Region 5.1	59	57	59
Region 5.2	114	75	83
Region 5.3	133	75	110
Region 5.4	71	87	57
Region 5.5	51	60	35

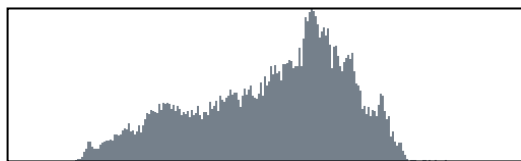
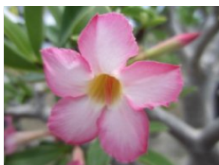
Sedangkan pada proses uji coba setelah melalui beberapa langkah sampai ditemukan hasil perhitungan 11 kriteria diantaranya yaitu nilai minimal red, nilai minimal green, nilai minimal blue, nilai maksimal red, nilai maksimal green, nilai mkasimal blu, jarak minimal ujung daun dan putik, jarak maksimal ujung daun dan putik, jarak minimal antar ujuang daun, jarak maksimal antar ujung daun dan jumlah daun. Dan perhitungan, hasil perhitungannya terlihat pada tabel 3.

Selain perhitungan yang dilakukan terhadap beberapa kriteria dari bunga yang dilakukan, juga dilakukan pencarian bentuk dari histogram dari warna bunga, terlihat ada gambar 9.

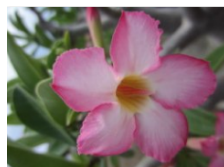
Dari data yang dihasilkan dan disimpan pada database, kemudian dilakukan uji coba dengan data yang sama dengan perbedaan cara pengambilan gambar bunganya dan meski dari data yang berbeda. Contoh dilakukan uji coba dari 2 (dua) gambar bunga pada gambar 10.

Tabel 3 Hasil Perhitungan Pada 11 Kriteria

Kriteria		Nilai
Nilai Minimal	Red	15
	Green	15
	Blue	10
Nilai Maksimal	Red	232
	Green	193
	Blue	205
Nilai Minimal	Ujung Daun Dan Putik	74,16
	Antar ujung Daun	89,56
Nilai Maksimal	Ujung Daun Dan Putik	78,791
	Antar ujung Daun	101,356
Jumlah Daun		5

**Gambar 9 Histogram Bunga Kamboja Merah**

Citra (a)



Citra (b)

Gambar 10 Uji Coba Pertama (a) Citra 1 (b) Citra 2

Uji coba yang dilakukan menggunakan dua metode dalam perhitungannya yaitu metode *ecludian distance* dan *k-nn*. Pada metode *ecludian distance* perhitungan dilakukan dengan membagi gambar bunga dalam 25 region, dari masing-masing region yang ada dicari nilai red, green dan blue kemudian dihitung rata-rata dari masing-masing red, green dan blue, ditunjukkan pada tabel 5.

Sedangkan pada perhitungan dengan metode *k-nn* dilakukan dengan mencari perhitungan pada 11 kriteria yang ada yaitu nilai minimal red, nilai minimal green, nilai minimal blue, nilai maksimal red, nilai maksimal green, nilai maksimal blue, jarak minimal ujung daun dan putik, jarak maksimal ujung daun dan putik, jarak minimal antar ujung daun, jarak maksimal antar ujung daun dan jumlah daun. Adapun perhitungan yang dihasilkan dari gambar uji coba tabel 6.

Dari hasil uji coba yang dilakukan akan dibandingkan dengan 4 (empat) gambar bunga lainnya dengan menampilkan beberapa kriteria pembandingan mulai dari nama bunga, kategori, dari gambar asli, gambar hasil segmentasi, tresholding dan perhitungan deteksi tepi, adapun tampilan dari histogram masing-masing gambar

bunga dan prosentase yang dihasilkan dari perhitungan menggunakan metode *ecludian distance* serta data link yang disediakan dari hasil pencarian yang diperoleh. Link yang ditampilkan ada 2 (dua) yaitu link ke jalur Wikipedia dan link melalui penelusuran Google.

Tabel 5 Hasil Perhitungan 25 Region Pada Uji Coba Citra (a)

	Rata-Rata	Rata-Rata	Rata-Rata
Region 1.1	138	162	143
Region 1.2	186	122	156
Region 1.3	151	120	127
Region 1.4	143	111	129
Region 1.5	62	79	51
Region 2.1	94	113	68
Region 2.2	203	122	165
Region 2.3	192	156	172
Region 2.4	182	154	177
Region 2.5	128	90	93
Region 3.1	220	153	203
Region 3.2	221	201	220
Region 3.3	160	87	34
Region 3.4	187	153	177
Region 3.5	163	104	137
Region 4.1	93	107	73
Region 4.2	181	142	131
Region 4.3	209	175	198
Region 4.4	169	102	136
Region 4.5	160	90	121
Region 5.1	114	120	115
Region 5.2	108	118	85
Region 5.3	202	140	180
Region 5.4	142	142	140
Region 5.5	127	128	125

Tabel 6 Hasil Perhitungan 11 Kriteria Pada Uji Coba Citra (a)






Kriteria		Nilai
Nilai Minimal	Red	18
	Green	34
	Blue	6
Nilai Maksimal	Red	250
	Green	243
	Blue	254
Nilai Minimal	Ujung Daun Dan Putik	55.946
	Antar ujung Daun	93.005
Nilai Maksimal	Ujung Daun Dan Putik	77.026
	Antar ujung Daun	102.859
Jumlah Daun		5

Dalam laporan ini, presentasi citra dari setiap jenis bunga diwakili oleh 3 (tiga) gambar bunga yang jenis dengan perbedaan terletak pada sisi pengambilan gambar yang berbeda, satu sebagai data dan 2 (dua) gambar bunga sebagai testing. Percobaan ini dilakukan untuk 5 (lima) jenis bunga yang berbeda.

Urutan gambar bunga hasil pencarian yang ditampilkan berdasarkan dari prosentase yang dihasilkan, bunga yang memiliki nilai kemiripan yang tinggi memiliki prosentase

yang tinggi dan berada pada urutan deret teratas dari tampilan dan diikuti dengan prosentase dibawahnya atau ditampilkan urutan secara ascending. Sedangkan dengan metode k-nn data yang ditampilkan urutannya berdasarkan data yang ada di database dengan 4 (empat) prosentase kemiripan bunga yang mendekatinya.

Tabel 7 Perbandingan Hasil Uji Coba

Gambar	Nama Bunga	Prosentase Kemiripan	
		Metode Ecludian Diatance	Metode K-NN
	Kamboja Merah	73,793 %	85,676 %
	Tapak Dara	56,526 %	93,993 %
	Angrek Bulan Unggu	76,624 %	84,737 %
	Gardenia Merah	86,447 %	94,003 %
	Angrek Unggu	80,273 %	84,466 %

Dari hasil penelitian menunjukkan nilai prosentase kemiripan terhadap bunga dengan metode K-NN lebih tinggi dari pada menggunakan metode Eclidean Distance. Uji coba pertama dilakukan pada bunga Kamboja Merah prosentase kemiripan dengan menggunakan metode Eclidean Distance dengan nilai kemiripan sebesar 73,793% dan pada metode K-NN nilai kemiripan sebesar 85,676%.

Uji coba kedua dilakukan pada bunga Tapak Dara prosentase kemiripan dengan menggunakan metode Eclidean Distance dengan nilai kemiripan sebesar 56,526% dan pada metode K-NN nilai kemiripan sebesar 93,993%. Berikutnya uji coba ketiga dilakukan pada bunga Angrek Bulan Unggu prosentase kemiripan dengan menggunakan metode Eclidean Distance dengan nilai kemiripan sebesar 76,624% dan pada metode K-NN nilai kemiripan sebesar 84,737%.

Pada uji coba keempat dilakukan pada bunga Gardenia Merah prosentase kemiripan dengan menggunakan metode Eclidean Distance dengan nilai kemiripan sebesar 86,447% dan pada metode K-NN nilai kemiripan sebesar 94,003%. Untuk Uji coba kelima dilakukan pada bunga Angrek Unggu prosentase kemiripan dengan menggunakan metode Eclidean Distance dengan nilai kemiripan sebesar 80,273% dan pada metode K-NN nilai kemiripan sebesar 84,466%. Contoh hasil dari uji coba terlihat pada tabel 7.

6. Kesimpulan dan Saran

6.1. Kesimpulan

Perbandingan prosentase kemiripan yang dihasilkan dari penelitian pengenalan bunga dengan metode K-NN lebih tinggi dari pada menggunakan metode Euclidean Distance.

Pada Metode Euclidean Distance dengan menggunakan nilai dari color (RGB) per region menghasilkan prosentase terkecil 56,526% dan terbesar 86,447%.

Sedangkan pada metode K-NN dengan menggunakan 11 (sebelas) criteria pada struktur bunga menghasilkan prosentase nilai terkecil sebesar 84,466% dan terbesar 93,993%.

6.2. Saran

Pada penelitian ini hanya saja dalam proses pencarian data bunga masih dilakukan secara manual dengan menginputkan titik putik dan ujung daun bunga sehingga untuk penelitian lebih lanjut dapat dilakukan secara komputerisasi.

DAFTAR PUSTAKA

- [1] Ashit Gandhi, "Content-Based Image Retrieval: Plant Species Identification", Oregon State University
- [2] D S Guru, Y. H. Sharath, S. Manjunath (2010) "Texture Features and KNN in Classification of Flower Images" *IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition" RTIPPR*,
- [3] Gurpreet kaur, Himanshu Monga, (2012) "Classification of Biological Species Based on Leaf Architecture-A review" IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555 Vol. 2
- [4] Jie Zou; Nagy, G.(2004) Evaluation of model-based interactive ower recognition, *Pattern Recognition*, 2004. ICPR 2004. Proceedings of the 17th International Conference on , vol.2, no., pp. 311- 314 Vol.2, 23-26 Aug.
- [5] Jyotismita Chaki, Ranjan Parekh, (2011) "Plant Leaf Recognition using Shape based Features and Network classifiers" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 10,
- [6] Keita Fukuda, Tetsuya Takiguchi, Yasuo Arika, "Multiple Classifier Based on Fuzzy C-Means for a Flower Image Retrieval"
- [7] P.E.Ajmire, Rajeev Dharaskar, V M Thakare, (2012) "Pattern Recognition Method for Study of botanical characteristics of Leaf" "Recent Trends in Computing" *Proceedings published by International Journal of Computer Applications® (IJCA)ISSN: 0975 - 8887* 1 MPGI National Multi Conference 2012 (MPGINMC-2012) 7-8 April.
- [8] Vuarnoz Vincent (2010) "Flower Recognition, Semester Project Final Report"