

IDENTIFIKASI JENIS TANAMAN BERDASARKAN EKSTRAKSI FITUR MORFOLOGI DAUN MENGGUNAKAN *K-NEAREST NEIGHBOR*

Kana Saputra S¹, Sri Wahyuni²

^{1,2}Program Studi Teknik Komputer

Fakultas Ilmu Komputer, Universitas Pembangunan Panca Budi

Jl. Jend. Gatot Subroto KM. 4,5 Medan, Sumatera Utara, Indonesia

¹kanasaputras@dosen.pancabudi.ac.id, ²sriwahyuni@dosen.pancabudi.ac.id

Abstrak – Indonesia merupakan salah satu negara dengan keanekaragaman tanaman yang cukup tinggi. Data yang diperoleh dari Indonesia Biodiversity Strategy and Action Plan 2003-2020 menunjukkan di Indonesia terdapat lebih dari 38.000 spesies tumbuhan yang tersebar di berbagai wilayah dan perhutanan Indonesia. Melihat jumlah spesies tanaman yang berlimpah maka dibutuhkan pendataan dan identifikasi jenis tanaman. Proses klasifikasi tanaman dapat dilakukan dengan cara mengidentifikasi gambar bentuk daun dari tanaman. Tujuan dari penelitian ini adalah menerapkan klasifier *k-Nearest Neighbors* untuk mengklasifikasi jenis tanaman secara cepat dan tepat. Ekstraksi fitur berdasarkan fitur morfologi daun, seperti: *area*, *perimeter*, *solidity*, dan *eccentricity* yang akan menjadi inputan untuk klasifier *k-Nearest Neighbors*. Model terbaik untuk klasifier *k-NN* yang dihasilkan adalah pada saat nilai $k = 5$ dengan hasil perhitungan akurasi sebesar 92%.

Kata kunci : Data Citra Daun, Ekstraksi Fitur, *k-Nearest Neighbors*, Morfologi Daun, anaman Indonesia

I. PENDAHULUAN

Indonesia merupakan salah satu negara dengan keanekaragaman tanaman yang cukup tinggi. Hampir setiap daerah di Indonesia memiliki tanaman yang khas dan mungkin saja tidak terdapat di negara lain. Tidak hanya keanekaragamannya saja, tanaman yang tumbuh di Indonesia juga memiliki kandungan vitamin, nilai gizi yang sangat tinggi, dan masih banyak manfaat lainnya bagi kesehatan. Selain itu, Indonesia juga merupakan negara agraris yang terkenal akan kekayaan rempah rempah dan tanaman herbal. Dari dulu hingga sekarang tanaman herbal ataupun tanaman obat dapat menyembuhkan berbagai macam penyakit. Dari data yang berhasil dirangkum ternyata di Indonesia sendiri terdapat lebih dari 38.000 spesies tumbuhan yang tersebar diberbagai wilayah dan perhutanan Indonesia [1].

Daun merupakan salah satu bagian tanaman yang sering digunakan untuk mengklasifikasikan jenis tanaman [2]. Daun digunakan untuk mengklasifikasikan tanaman, karena setiap jenis tanaman memiliki fitur daun yang berbeda [3]. Selain itu, daun lebih mudah diperoleh karena tidak tergantung pada musim dan posisi pada bagian tanaman. Posisi daun pada bagian tanaman tidak membutuhkan alat khusus pada saat pengambilan

sampel data di lapangan dan tidak mengganggu dan merusak tanaman.

Proses klasifikasi tanaman dapat dilakukan dengan cara mengidentifikasi gambar bentuk daun dari tanaman. Dengan cara tersebut dapat dilakukan langkah-langkah pengenalan pola daun dengan mengenali karakteristik daun seperti bentuk dan tekstur sebuah daun. Metode untuk melakukan pemrosesan terhadap citra masukan dengan pemanfaatan teknik pengolahan citra dilakukan untuk menganalisa karakteristik daun. Teknik pengolahan citra dilakukan pada tahapan pra-proses citra sampai diperoleh bentuk tepian dan ciri dari masing-masing daun. Metode yang digunakan pada ekstraksi fitur ini adalah pengenalan fitur morfologi daun [4]. Setelah dilakukan ekstraksi fitur dari citra daun tersebut maka diperoleh informasi penciri daun yang kemudian digunakan sebagai data pengelompokkan. Ekstraksi fitur daun yang tepat sangat mempengaruhi hasil klasifikasi tanaman.

Penelitian tentang ekstraksi fitur citra daun berdasarkan morfologi daun pernah dilakukan untuk mengidentifikasi jenis penyakit pada tanaman Jabon. Hasilnya fitur morfologi daun seperti *rectangularity*, *roundness*, *compactness*, *solidity*, *convexity*, *elongation*, dan *eccentricity* dapat merepresentasikan karakteristik bentuk penyakit daun [4]. Penelitian lain menggunakan metode ekstraksi fitur citra daun untuk menerapkan pengenalan daun otomatis dengan tujuan mengklasifikasi jenis tanaman. Hasilnya metode tersebut cepat dalam eksekusi, efisien dalam pengenalan dan mudah dalam implementasikan [5]. Dari hasil penelitian tersebut menunjukkan bahwa ekstraksi fitur morfologi daun dapat digunakan sebagai suatu ciri khusus dalam proses identifikasi jenis tanaman.

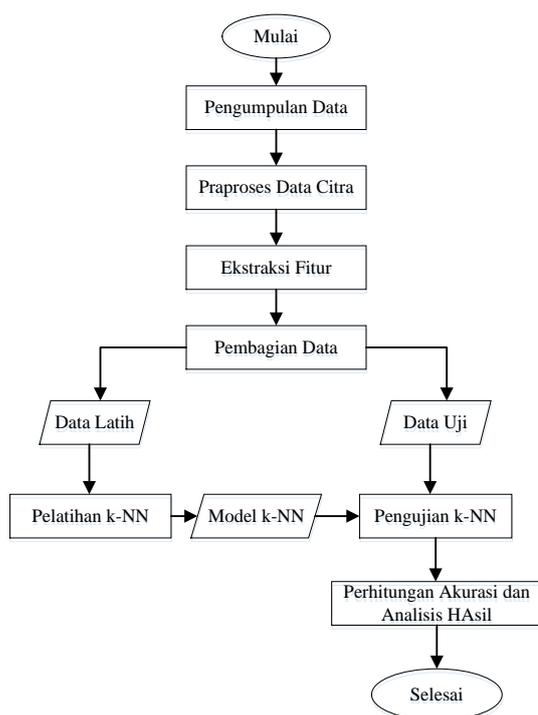
Penerapan *klasifier k-Nearest Neighbors* (*k-NN*) pernah dilakukan untuk mengklasifikasi daun dengan perbaikan fitur citra dengan hasil pengujian klasifikasi dengan menggunakan metode tersebut didapatkan nilai akurasi sebesar 86,67% [6]. Selain itu, penerapan *klasifier k-NN* juga pernah dilakukan untuk mengklasifikasi daun herbal dengan hasil didapatkan nilai akurasi sebesar 70,83% [7]. Dari beberapa penelitian tersebut menunjukkan bahwa *klasifier k-NN* mampu melakukan klasifikasi daun dengan baik.

Berdasarkan deskripsi diatas maka dilakukan penelitian untuk mengklasifikasikan jenis tanaman secara tepat. Penelitian ini akan dilakukan klasifikasi citra daun dengan *clasifier* k-NN. Proses klasifikasi berdasarkan fitur morfologi citra daun. *Classifier* k-NN melakukan klasifikasi terhadap objek berdasarkan data *training* yang terdekat. Sebelum tahapan klasifikasi terlebih dahulu dilakukan tahapan praproses citra dan ekstraksi fitur citra tepi daun agar didapatkan nilai masukan yang tepat untuk tahapan klasifikasi jenis tanaman berdasarkan ciri morfologi daun.

II. METODOLOGI PENELITIAN

A. Tahapan Penelitian

Penelitian ini akan melakukan akuisisi pola tekstur pada citra daun dan selanjutnya akan dilakukan praproses dan ekstraksi ciri. Hasil praproses dan ekstraksi ciri dari citra kemudian dijadikan sebagai input dalam pelatihan menggunakan *k-Nearest Neighbors* sehingga menghasilkan model yang dapat menggambarkan sifat dan mengklasifikasikan jenis tanaman. Adapun tahapan dalam penelitian ini seperti pada Gambar 1.



Gambar 1. Kerangka Kerja Penelitian

B. Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari <https://archive.ics.uci.edu/ml/datasets/Folio>. Data yang diperoleh berjumlah 5 jenis tanaman yang terdiri dari 75 data citra daun. Ekstraksi Fitur Citra

Pada tahap praproses, terdapat 2 langkah untuk mengolah citra, yaitu memotong (*cropping*) citra dan melakukan segmentasi. Teknik *cropping* dilakukan untuk memotong dan mengambil bagian dari daun yang dibutuhkan. *Cropping* dilakukan untuk melihat bentuk daun secara utuh. Hal ini dilakukan karena data yang diperoleh berjumlah sedikit. Selain itu melihat bentuk daun memerlukan komputasi yang besar saat melakukan proses segmentasi dan kesulitan dalam menentukan titik pusat dari tiap gejala yang terbentuk.

Berdasarkan fitur area dan perimeter dapat juga dihitung nilai-nilai fitur morfologi lainnya. Berikut adalah beberapa formula yang digunakan untuk mengekstraksi fitur morfologi [4]:

- *Area* adalah luas daerah yang dihitung berdasarkan jumlah pixel yang menempati objek gambar.
- *Perimeter* adalah daerah yang dihitung berdasarkan jumlah pixel disekitar gambar objek.
- *Solidity* adalah mengukur kepadatan dari sebuah objek, *solidity* merupakan rasio dari luas wilayah objek ke daerah *convex full* objek.

$$solidity = \frac{area}{convex_area}$$

- *Eccentricity* adalah rasio jarak antar fokus elips dengan panjang sumbu utama (*major axis*) sebuah objek. *Eccentricity* bernilai antara 0-1. *Eccentricity* merupakan teknik untuk menggambarkan sebuah objek dengan bentuk elips.

$$eccentricity = \frac{\sqrt{(major\ axis^2 - minor\ axis^2)}}{major\ axis}$$

Pembagian Data

Seluruh data hasil ekstraksi masing-masing ciri dibagi menjadi data latih dan data uji. Data *training* berjumlah 10 citra daun dan data *testing* berjumlah 4 citra daun untuk masing-masing jenis tanaman.

Implementasi k-Nearest Neighbor

Proses klasifikasi dengan k-NN dilakukan menggunakan data latih yang sebelumnya sudah dibagi. Dalam melakukan pelatihan dan pengujian data, karakter akan diambil satu per satu dari fitur yang ada. Dalam proses klasifikasi sebelumnya harus ditentukan dahulu nilai *k*, yaitu jumlah tetangga terdekat yang akan dilihat kelasnya untuk menentukan kelas terbanyak yang merupakan kelas dari titik baru. Nilai *k* akan sangat berpengaruh pada akurasi hasil klasifikasi. Nilai *k* yang terbaik untuk algoritma ini tergantung pada data. Secara umum, nilai *k* yang tinggi akan mengurangi efek *noise* pada klasifikasi, tetapi membuat batasan antara setiap klasifikasi menjadi kabur. Nilai *k* yang akan dicobakan pada penelitian ini adalah 3, 5, dan 7.

Evaluasi Model k-Nearest Neighbor

Evaluasi dilakukan dengan cara menghitung akurasi. Akurasi yang dihasilkan dihitung menggunakan *confusion matrix* [8]. Perhitungan *confusion matrix* dihitung berdasarkan prediksi positif yang benar (*True Positive*), prediksi positif yang salah (*False Positive*), prediksi negatif yang benar (*True Negative*), dan prediksi negatif yang salah (*False Negative*). Rumus menghitung akurasi adalah sebagai berikut:

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN}$$

Semakin tinggi hasil perhitungan akurasi yang dihasilkan, maka semakin baik metode yang digunakan.

III. HASIL DAN PEMBAHASAN

1. Pengumpulan Data

Data yang digunakan merupakan data citra daun dengan berbagai jenis tanaman. Detail data citra daun yang digunakan dapat dilihat pada Tabel 1.

Tabel 1. Detail Data Citra Daun Untuk Masing-Masing Jenis Tanaman

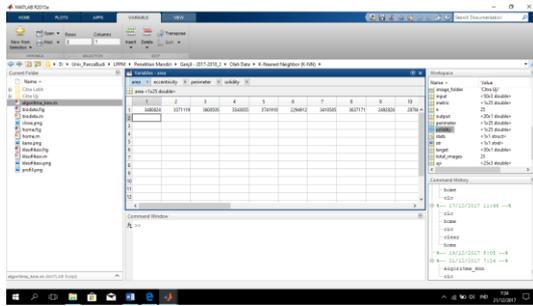
No	Jenis Tanaman	Jumlah Data	Citra Daun
1	<i>ashanti blood</i>	15	
2	<i>egg plant</i>	15	

3	<i>guava</i>	15	
4	<i>jackfruit</i>	15	
5	<i>sweet potato</i>	15	
Total		75	-

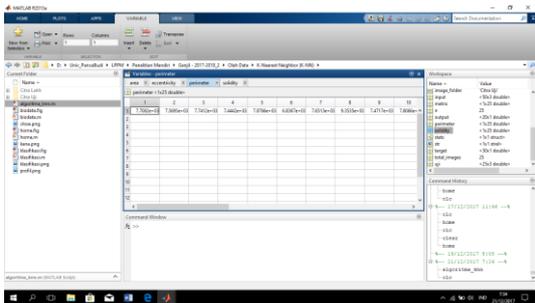
Dari data citra daun yang berjumlah 75 dibagi menjadi 50 data *training* dan 25 data *testing*. Masing-masing citra daun untuk data *training* berjumlah 10 citra daun dan untuk data *testing* berjumlah 5 citra daun. Data citra daun tersebut memiliki latar belakang berwarna putih. Latar belakang tersebut tidak akan mempengaruhi hasil ekstraksi ciri karena ekstraksi ciri yang dilakukan berdasarkan ciri morfologi daun bukan dari segi warna daun.

2. Hasil Ekstraksi Fitur Citra

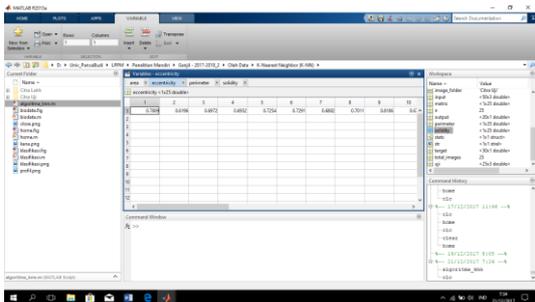
Data citra daun yang digunakan dalam penelitian ini tidak memiliki ukuran yang seragam. Data citra daun yang digunakan memiliki titik fokus yang berbeda dengan menampilkan bentuk daun secara keseluruhan. Ini dilakukan untuk melihat kemampuan model yang terbentuk dari *klasifier k-NN*. Contoh hasil ekstraksi ciri morfologi (*area, perimeter, eccentricity, solidity*) dapat dilihat pada Gambar 2, Gambar 3, Gambar 4, dan Gambar 5.



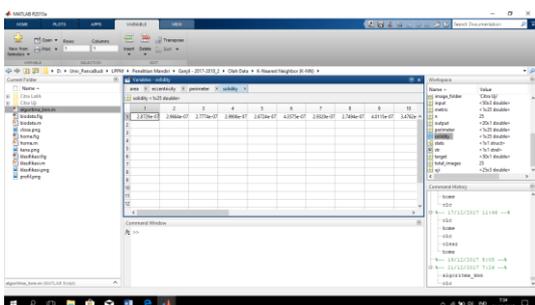
Gambar 2. Tampilan Contoh Hasil Perhitungan Area Untuk Setiap Citra Daun



Gambar 3. Tampilan Contoh Hasil Perhitungan Perimeter Untuk Setiap Citra Daun



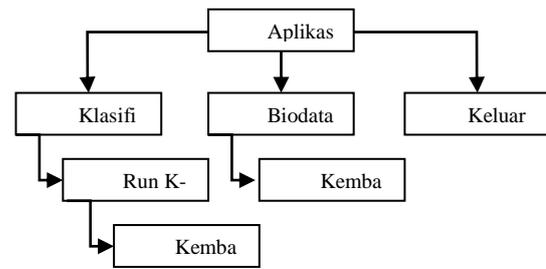
Gambar 4. Tampilan Contoh Hasil Perhitungan Eccentricity Untuk Setiap Citra Daun



Gambar 5. Tampilan Contoh Hasil Perhitungan Solidity Untuk Setiap Citra Daun

3. Hasil Klasifikasi *k-Nearest Neighbor*

Hasil klasifikasi menggunakan bantuan software *Matlab R2013a* dengan memanfaatkan GUI *Matlab*. Struktur menu aplikasi yang dibuat dapat dilihat pada Gambar 6.



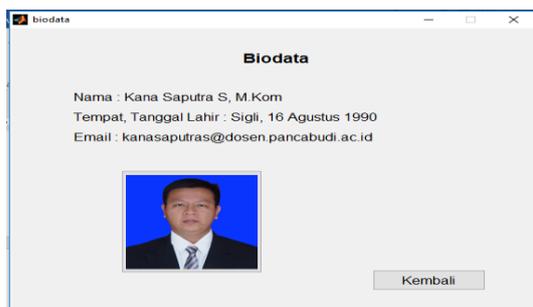
Gambar 6. Struktur Menu Aplikasi Yang Dibuat

Aplikasi yang dibuat dapat digunakan untuk mengklasifikasikan data citra daun dengan ekstraksi ciri morfologi daun (*area, perimeter, eccentricity, solidity*) menggunakan *klasifier k-NN*. Untuk parameter *k* yang digunakan dapat bervariasi sesuai dengan kebutuhan. Parameter *k* yang digunakan adalah 3, 5, dan 7. Selain itu aplikasi akan menampilkan *confusion matrix* agar dapat digunakan untuk melihat data citra daun yang salah terklasifikasi. Dari *confusion matrix* juga dapat dilakukan perhitungan akurasi untuk melihat kemampuan model *klasifier k-NN* dalam mengklasifikasikan data citra daun. Tampilan aplikasi yang telah dibuat dapat dilihat pada Gambar 7, Gambar 8, dan Gambar 9.



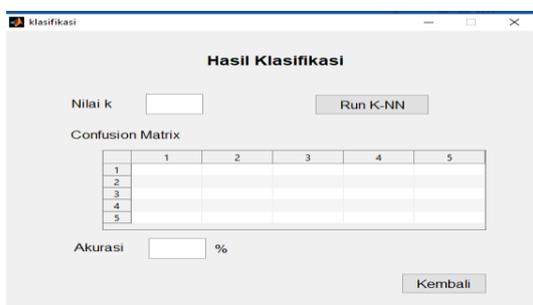
Gambar 7. Tampilan Utama Dari Aplikasi Identifikasi Jenis Tanaman

Gambar 7 merupakan tampilan utama dari aplikasi identifikasi jenis tanaman berdasarkan ciri morfologi daun menggunakan algoritma *k-NN*. Menu yang terdapat dalam tampilan utama aplikasi identifikasi jenis tanaman berdasarkan ciri morfologi daun menggunakan algoritma *k-NN* adalah menu klasifikasi, menu biodata, dan menu keluar. Untuk setiap menu tersebut memiliki tampilannya masing-masing, kecuali menu keluar. Menu klasifikasi untuk melakukan proses klasifikasi menggunakan *klasifier k-NN*, sedangkan menu biodata untuk melihat biodata peneliti.



Gambar 8. Tampilan Biodata Dari Aplikasi Identifikasi Jenis Tanaman

Gambar 8 merupakan tampilan menu biodata dari peneliti. Menu biodata menampilkan biodata peneliti yang terdiri dari nama lengkap, tempat dan tanggal lahir, e-mail, dan foto peneliti. Selain itu terdapat tombol untuk kembali. Tombol tersebut berfungsi untuk kembali ke menu utama.



Gambar 9. Tampilan Hasil Klasifikasi Dari Aplikasi Identifikasi Jenis Tanaman

Gambar 9 merupakan tampilan halaman klasifikasi dari aplikasi identifikasi jenis tanaman berdasarkan ciri morfologi daun menggunakan algoritma k-NN. Menu klasifikasi tersebut menampilkan nilai parameter k yang digunakan, *confusion matrix*, dan akurasi dalam bentuk persen. Nilai k dapat diinputkan sesuai dengan kebutuhan. Nilai k biasanya merupakan bilangan ganjil dengan nilai yang digunakan adalah 3, 5, dan 7. Setelah menginputkan nilai k , klik tombol Run K-NN untuk melakukan klasifikasi.

Hasil klasifikasi untuk setiap nilai k diperoleh dalam bentuk *confusion matrix* seperti yang terlihat pada Tabel 2, Tabel 3, dan Tabel 4.

Tabel 2. *Confusion Matrix* Untuk $k = 3$

Kelas	a b	e p	g	j	s p
ab	5	0	0	0	0
ep	1	3	0	1	0
g	0	0	5	0	0
j	0	0	0	5	0
sp	0	0	0	0	5

Untuk nilai $k = 3$ terdapat kesalahan dalam pengklasifikasian pada data citra daun *egg plant* sebanyak 2 data citra daun.

Tabel 3. *Confusion Matrix* Untuk $k = 5$

Kelas	a b	e p	g	j	s p
ab	5	0	0	0	0
ep	1	3	0	1	0
g	0	0	5	0	0
j	0	0	0	5	0
sp	0	0	0	0	5

Untuk nilai $k = 5$ terdapat kesalahan dalam pengklasifikasian pada data citra daun *egg plant* sebanyak 2 data citra daun.

Tabel 4. *Confusion Matrix* Untuk $k = 7$

Kelas	a b	e p	g	j	s p
ab	5	0	0	0	0
ep	1	3	1	0	0
g	0	0	5	0	0
j	0	0	1	4	0
sp	0	0	0	0	5

Untuk nilai $k = 7$ terdapat kesalahan dalam pengklasifikasian pada data citra daun *egg plant* sebanyak 2 data citra daun dan pada citra daun *jackfruit* sebanyak 1 data citra daun.

4. Evaluasi Model k-Nearest Neighbor

Evaluasi model *k-Nearest Neighbor* digunakan dengan cara menghitung dan membandingkan hasil perhitungan akurasi dan waktu eksekusi program. Hasil perhitungan akurasi untuk setiap nilai k dapat dilihat pada Tabel 5.

Tabel 5. Tingkat Akurasi Untuk Setiap Nilai k

Nilai k	Akurasi (%)
3	92
5	92
7	88

Dari Tabel 5 dapat dilihat bahwa akurasi tertinggi terjadi pada saat nilai $k = 3$ dan $k = 5$, yaitu sebesar 92 %. Sebagai perbandingan lain dapat digunakan hasil perhitungan waktu eksekusi yang dibutuhkan oleh setiap nilai k . Perbandingan waktu eksekusi program dapat dilihat pada Gambar 10.



Gambar 10. Perbandingan *Execution Time* Untuk Setiap Nilai k

Hasil akurasi tertinggi terjadi saat nilai $k = 3$ dan $k = 5$, sehingga waktu eksekusi program yang dibandingkan hanya pada nilai k tersebut. Semakin cepat waktu eksekusi program, maka semakin baik untuk dipilih. Dari Gambar 10 terlihat bahwa waktu eksekusi program tercepat antara nilai $k = 3$ dan $k = 5$ adalah saat nilai $k = 5$, yaitu 140 detik. Model terbaik untuk *classifier* k-NN yang dihasilkan adalah pada saat nilai $k = 5$.

IV. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, maka dapat disimpulkan sebagai berikut:

1. Parameter nilai k yang digunakan adalah 3, 5, dan 7.
2. Hasil perhitungan akurasi menggunakan *clasifier* k-NN terjadi pada saat nilai $k = 3$ dan $k = 5$ sebesar 92 %.
3. Data citra daun yang salah klasifikasi adalah data citra daun *egg plant* dan *jackfruit*.
4. Berdasarkan hasil perbandingan model terbaik untuk *classifier* k-NN yang dihasilkan adalah pada saat nilai $k = 5$.

DAFTAR PUSTAKA

- [1] Bappenas. 2013. Indonesia Biodiversity Strategy and Action Plan 2003-2020. Jakarta: Bappenas.
- [2] Hati S, Sajeevan G. 2013. Plant Recognition from Leaf Image through Artificial Neural Network. *International Journal of Computer Applications*. 62(17): 15-18.
- [3] Wu Q, Zhou C, Wang C. 2006. Feature Extraction and Automatic Recognition of Plant Leaf Using Artificial Neural Network. *Advances in Artificial Intelligence Research in Computing Science*. 20: 3-10.
- [4] Manik FY, Herdiyeni Y, Herliyana EN. 2016. Leaf Morphological Feature Extraction of Digital Image Anthocephalus Cadamba. *TELKOMNIKA*. 14(2): 630-637.
- [5] Wu SG, Bao FS, Xu EY, Wang YX, Chang YF, Xiang QL. 2007. A Leaf Recognition Algorithm for Plant Classification Using K-Nearest Neighbors. *IEEE International Symposium on Signal Processing and Information Technology*. 1-6.
- [6] Liantoni F. 2015. Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor. *ULTIMATICS*. 7(2): 98-104.
- [7] Liantoni F, Nugroho H. 2015. Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan K-Nearest Neighbor. *Jurnal Simantec*. 5(1): 9-16.
- [8] Andriani Anik. 2012. Penerapan Algoritma C4.5 Pada Program Klasifikasi Mahasiswa Dropout. *Prosiding Seminar Nasional Matematika*.