

Klasifikasi Daun Dengan Centroid Linked Clustering Berdasarkan Fitur Bentuk Tepi Daun

Febri Liantoni¹, Nana Ramadijanti², Nur Rosyid Mubtada'i³

Politeknik Elektronika Negeri Surabaya
Institut Teknologi Sepuluh Nopember
Kampus PENS-ITS Keputih Sukolilo Surabaya 60111
Telp (+62)31-5947280, 5946114, Fax. (+62)31-5946114
Email : luigie.liantonie@gmail.com

Abstrak

Ilmu tumbuhan pada waktu sekarang telah mengalami kemajuan yang sangat pesat, hingga bidang-bidang pengetahuan yang semula hanya merupakan cabang ilmu tumbuhan saja, sekarang ini telah menjadi ilmu yang berdiri sendiri. Salah satunya cabang ilmu tumbuhan yaitu morfologi tumbuhan, ilmu tumbuhan ini mempelajari bentuk dan susunan tubuh tumbuhan khususnya mengenai morfologi tumbuhan berdasarkan bentuk tepi daun. Pada proyek akhir ini akan dibangun sebuah program yang bertujuan untuk mengklasifikasikan jenis daun. Ciri yang digunakan dalam proses pengklasifikasian daun adalah bentuk dari tepi daun. Selanjutnya dilakukan pengolahan gambar meliputi proses gray-scale, thresholding, dan edge detection. Dari gambar hasil deteksi tepi ini akan dihitung menggunakan moment invariant. Dalam proyek akhir ini akan dilakukan klasifikasi menggunakan metode centroid linked dan metode k-means. Pada proses klasifikasi akan menghasilkan 10 (sepuluh) jumlah cluster yang berdasarkan bentuk tepi daun.

Kata kunci : Morfologi, gray-scale, thresholding, edge detection moment invariant, centroid linked, k-means

1. Pendahuluan

1.1 Latar Belakang

Ilmu tumbuhan pada waktu sekarang telah mengalami kemajuan yang sangat pesat, hingga bidang-bidang pengetahuan yang semula hanya merupakan cabang-cabang ilmu tumbuhan saja, sekarang ini telah menjadi ilmu yang berdiri sendiri-sendiri.

Dari berbagai cabang ilmu tumbuhan yang sekarang telah berdiri sendiri adalah Morfologi Tumbuhan. Morfologi Tumbuhan yang mempelajari bentuk dan susunan tubuh tumbuhan pun sudah demikian pesat perkembangannya hingga dipisahkan menjadi morfologi luar atau morfologi saja dan morfologi dalam atau anatomi tumbuhan. Salah satunya diambil pembahasan mengenai morfologi tumbuhan berdasarkan bentuk tepi daun.

Dalam proyek akhir ini akan dibahas mengenai pengklasifikasian tumbuhan berdasarkan bentuk tepi daun. Proses yang digunakan menggunakan teknik image processing, kemudian dilakukan proses klasifikasi. Untuk proses pengambilan gambar dilakukan secara load gambar yang nantinya akan digunakan sebagai sample yang akan digunakan untuk membandingkan dengan gambar-gambar

yang ingin diketahui bagian keanggotaannya dalam sistem klasifikasi yang dibuat.

1.2 Rumusan Permasalahan

Permasalahan utama pada proyek akhir bagaimana membangun sistem klasifikasi daun, dari permasalahan ini akan dibahas 4 permasalahan yang penting yaitu :

1. Mengenali motif atau bentuk dari tepi daun berdasarkan gambar.
2. Mengklasifikasikan berdasarkan keanggotaannya dalam sistem morfologi tumbuhan.
3. Merancang data klasifikasi guna sebagai sample acuan atau sebagai data training.
4. Membentuk sistem klasifikasi, dan menampilkan gambar sesuai dengan kelas klasifikasinya.

1.3 Penelitian Terkait

Adapun penelitian yang berkaitan dengan proyek akhir ini dan memiliki beberapa kesamaan, antara lain.

Pada tahun 2007, Helmy Hasniawati mengerjakan tugas akhir yang berkaitan dengan pengelompokan benda dengan objek yang dipilih yaitu buah. Dalam tugas akhir ini dijelaskan mengenai bentuk-bentuk clustering dan metode atau rumus perhitungannya. Salah satu metode

yang menjadi referensi yaitu mengenai centroid linkage clustering.

Pada tahun 2005, Kayhan Gulez melakukan penelitian menggunakan metode momen invarian. Pada metode momen invarian ini dapat digambarkan suatu obyek dapat dilakukan pengolahan melalui berbagai hal meliputi area, posisi, orientasi dan parameter terdefinisi lainnya. Dengan mendapatkan sejumlah informasi momen, baik momen tingkat ke nol (m_{00}) dan kesatu (m_{10} dan m_{01}) atau momen sentral, dan momen pada tingkat ≥ 2 atau momen invarian dari sebuah obyek, maka obyek tersebut dapat diidentifikasi sekalipun telah mengami pergeseran (tranlasi), perputaran (rotasi) maupun perubahan skala.

1.4 Tujuan Proyek

Tujuan utama dibuat proyek akhir

1. Mampu mengenali gambar-gambar yang diproses kemudian mampu memperoleh nilai fitur-fitur bentuk yang ada.
2. Mengklasifikasikan tumbuhan berdasarkan bentuk tepi daun.
3. Membantu memberikan kemudahan dalam pencarian letak keanggotaan suatu daun dalam morfologi / klasifikasi.

1.5 Kontribusi Proyek

Proyek Akhir ini nantinya diharapkan dapat dikembangkan untuk membantu menyelesaikan permasalahan klasifikasi daun secara cepat.

2. Teori Penunjang

2.1 Momen Invarian

Proses pengenalan sebuah obyek di dalam sebuah citra setelah proses segmentasi, sering terbentur pada permasalahan posisi obyek, rotasi sumbu obyek, dan perubahan skala dari obyek. Posisi obyek yang bergeser, berputar dan ukurannya yang lebih kecil atau lebih besar dari pada informasi yang sudah dimiliki sebagai knowledge dapat menyebabkan kesalahan dalam pengenalan / identifikasi obyek tersebut.

Momen dapat menggambarkan suatu obyek dalam hal area, posisi, orientasi dan parameter terdefinisi lainnya. Dengan mendapatkan sejumlah informasi momen, baik momen tingkat ke nol (m_{00}) dan kesatu (m_{10} dan m_{01}) atau momen sentral, dan momen pada tingkat ≥ 2 atau momen invarian dari sebuah obyek, maka obyek tersebut dapat diidentifikasi sekalipun telah mengami pergeseran (tranlasi), perputaran (rotasi) maupun perubahan skala.

Persamaan momen invarian yang digunakan adalah sebagai berikut.

$$Hu_1 = \eta_{20} + \eta_{02}$$

$$Hu_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$Hu_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$Hu_4 = (\eta_{30} + \eta_{12})^2 + (3\eta_{21} + \eta_{03})^2$$

$$Hu_5 = (\eta_{30} + 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

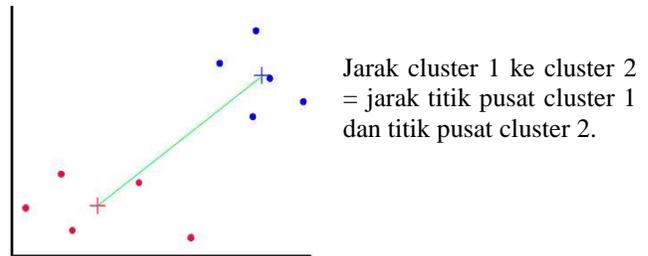
$$Hu_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$Hu_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

2.2 Centroid Linkage Clustering

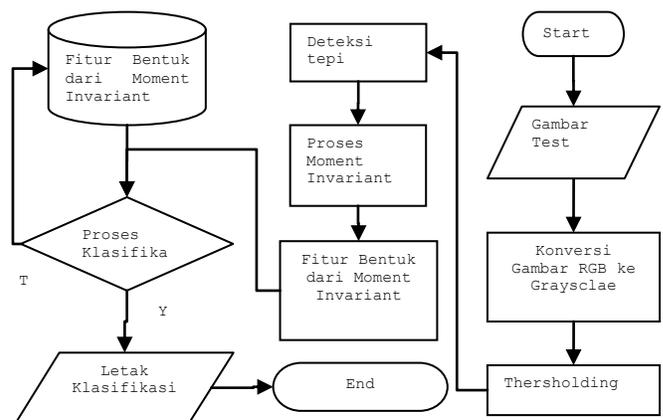
Centroid linkage clustering adalah proses pengklasteran yang didasarkan pada jarak antar centroidnya. Metode ini bagus untuk memperkecil *variance within cluster* karena melibatkan centroid pada saat penggabungan antar cluster. Metode ini juga baik untuk data yang mengandung outlier.

Berikut ini ilustrasi pada proses metode centroid linkage clustering.



Gambar 1 Centroid Linkage

3. Rancangan Sistem



Gambar 2 Desain pemrosesan klasifikasi

3.1 Proses Momen Invarian

Dari proses thresholding yang telah dilakukan sebelumnya, didapatkan nilai 0 dan 255 yang kemudian juga telah diproses dengan menggunakan deteksi tepi sehingga citra lebih terlihat garis tepinya, dari hasil deteksi tepi ini kemudian dapat dilakukan proses momen invarian.

Dari proses momen invarian ini akan dihasilkan nilai tujuh momen invarian yang akan digunakan untuk data proses klasifikasi.

3.2.1 Normalisasi Data

Dari data-data berupa tujuh momen invarian ini kemudian akan dilakukan proses normalisasi data. Proses normalisasi ini dilakukan sebab nilai tujuh momen invarian yang didapat berupa data yang acak tidak beraturan. Dimana proses normalisasi ini akan menghasilkan data dengan range antara 0 (nol) sampai dengan 1 (satu).

3.3 Centroid Linkage Clustering

Proses clustering dimaksudkan untuk tujuan mengelompokkan data-data yang memiliki kemiripan nilai. Metode ini bagus untuk memperkecil variance within cluster karena melibatkan centroid pada saat penggabungan antar cluster. Metode ini juga baik untuk data yang mengandung outlier.

Pada proses klasifikasi ini data akan dijadikan kedalam 10 jenis cluster.

3.4 K-Means

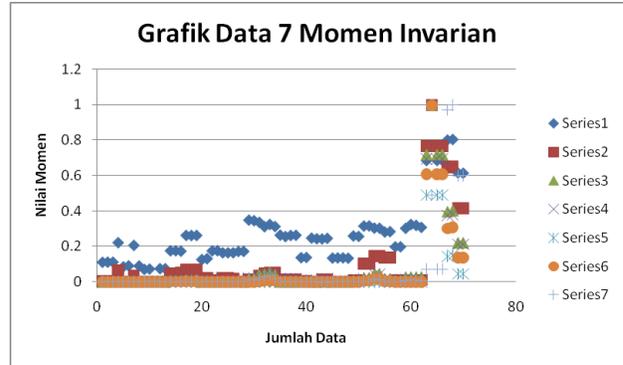
Algoritma k-means termasuk partitioning clustering yang memisahkan data ke k daerah bagian yang terpisah. Algoritma k-means sangat terkenal karena kemudahannya dan kemampuannya untuk mengklaster data besar dan data outlier dengan sangat cepat. Sesuai dengan karakteristik partitioning clustering. Setiap data harus termasuk ke cluster tertentu, dan Memungkinkan bagi setiap data yang termasuk cluster tertentu pada suatu tahapan proses, pada tahapan berikutnya berpindah ke cluster yang lain.

Metode ini dengan cara memisahkan data ke k daerah bagian yang terpisah. Algoritma k-means sangat terkenal karena kemudahannya dan kemampuannya untuk mengklaster data besar dan data outlier dengan sangat cepat.

3.5 Pengolahan Data

3.5.1 Data

Data yang digunakan pada penelitian ini berjumlah 70 buah gambar daun, setiap gambar memiliki nilai tujuh momen invarian.



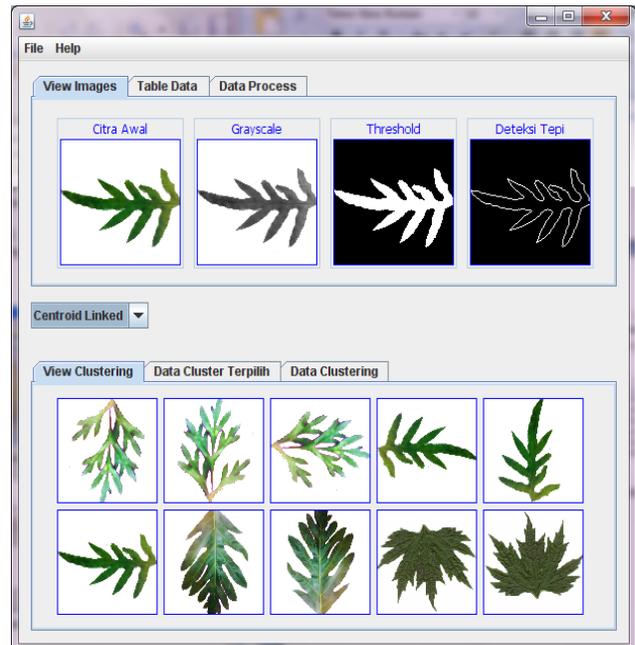
Gambar 3 Grafik data momen invarian

3.5.2 Proses Klasifikasi

Pada proses klasifikasi ini data dijadikan kedalam 10 kelas. Untuk hasil yang termasuk kedalam kelas yang sama maka gambar daun akan ditampilkan. Pada tampilan program ini disediakan 10 buah lokasi yang akan digunakan untuk menampilkan gambar daun tersebut.

3.5.2.1 Klasifikasi Dengan Centroid Linked

Berikut ini pengujian terhadap berbagai macam kelas dalam klasifikasi menggunakan metode centroid linkage.



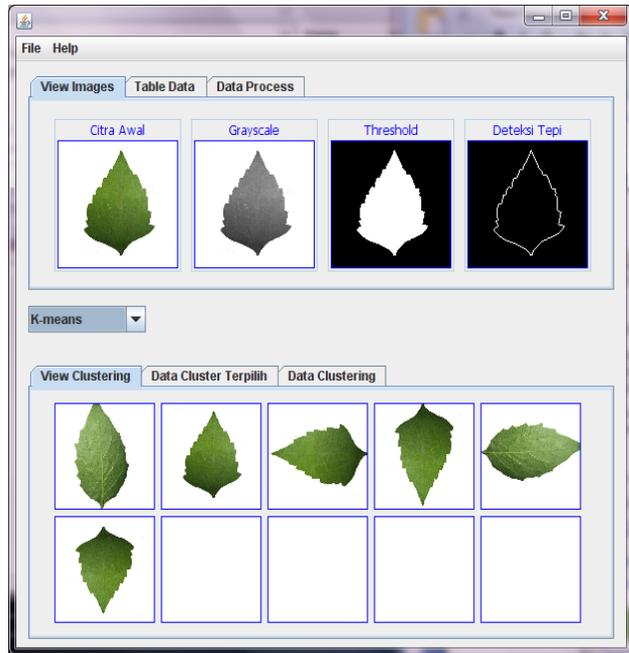
Gambar 4 Hasil pengujian centroid

Proses klasifikasi diatas dipilih kelas daun bertipe daun bertepi berbagi menyirip. Dari hasil yang ditampilkan tersebut dapat diketahui jumlah anggota dalam klasifikasi sebanyak 10 daun. Dari data yang ditampilkan terdapat 2 gambar yang tidak termasuk

kedalam kelas ini yaitu dari kelas bercanggap menjari (pada daun ke 9 dan 10).

3.5.2.1 Klasifikasi Dengan K-Means

Berikut ini pengujian terhadap berbagai macam kelas dalam klasifikasi menggunakan metode k-means.



Gambar 5 Hasil pengujian k-means

Proses klasifikasi diatas dipilih kelas daun bertipe daun bertepi bergerigi. Dari hasil yang ditampilkan tersebut dapat diketahui jumlah anggota dalam klasifikasi sebanyak 6 daun.

3.5.3 Penghitungan Kesalahan

Dari hasil pengujian yang telah dilakukan pada masing-masing metode akan dihitung nilai kesalahan/error yang terjadi.

Tabel berikut ini merupakan persentase kelasahan/error yang dihasilkan dalam proses klasifikasi menggunakan metode centroid linkage.

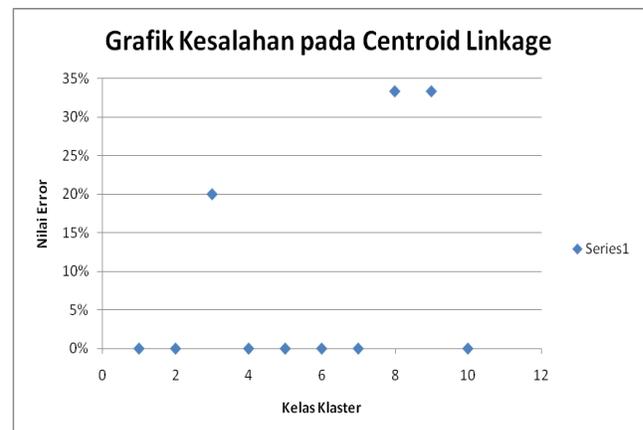
Tabel 1 Persentase error pada centroid linkage

Kelas klaster	Jumlah anggota	Jumlah error	Persentase error
Klaster 1	8	0	0%
Klaster 2	5	0	0%
Klaster 3	10	2	20%
Klaster 4	5	0	0%

Klaster 5	6	0	0%
Klaster 6	6	0	0%
Klaster 7	7	0	0%
Klaster 8	9	3	33.33%
Klaster 9	6	2	33.33%
Klaster 10	8	0	0%
Total	70	7	10%
Erorr rata-rata	$(20\%+33.33\%+33.33\%)/10 = 8.666\%$		

Dari data tabel tersebut kita dapat menyimpulkan bahwa dengan menggunakan metode centroid linkage jumlah kesalahan error tertinggi mencapai 33.33%, sedangkan untuk jumlah error rata-rata yang dihasilkan sebesar 8.666%.

Berikut ini grafik yang dihasilkan dari data kesalahan /error yang dihasilkan dari proses klasifikasi menggunakan metode centroid linkage.



Gambar 6 Grafik data kesalahan pada centroid linkage

3.5.3.2 Menggunakan K-Means

Tabel berikut ini merupakan persentase kelasahan/error yang dihasilkan dalam proses klasifikasi menggunakan metode k-means.

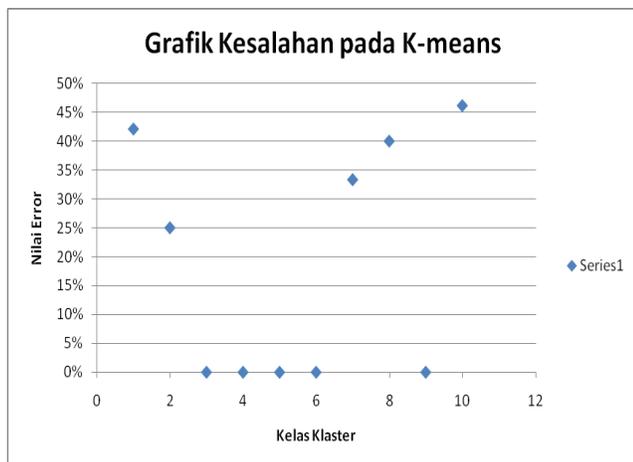
Tabel 2 Persentase error pada k-means

Kelas klaster	Jumlah anggota	Jumlah error	Persentase error
Klaster 1	19	8	42.10%
Klaster 2	4	1	25%
Klaster 3	4	0	0%
Klaster 4	2	0	0%

Klaster 5	6	0	0%
Klaster 6	4	0	0%
Klaster 7	6	2	33.33%
Klaster 8	10	4	40%
Klaster 9	2	0	0%
Klaster 10	13	6	46.15
Total	70	21	30%
Erorr rata-rata	$(42.10+25\%+33.33\%+40\%+46.15\%) / 10 = 18.658\%$		

Dari data tabel tersebut kita dapat menyimpulkan bahwa dengan menggunakan metode k-measn didapatkan jumlah kesalahan error tertinggi mencapai 46.15 %, sedangkan jumlah error rata-rata yang dihasilkan sebesar 18.658%.

Berikut ini grafik yang dihasilkan dari data kesalahan error yang dihasilkan dari proses klasifiaksi menggunakan metode k-means.



Gambar 7 Grafik data kesalahan pada k-means

4. Hasil Dan Kesimpulan

Pada penggunaan metode seven momen invarian dapat dibuktiakn bahwa perubahan posisi, rotasi dan skala tidak begitu berpengaruh pada nilai seven momen invarian yang dihasilkannya.

Pada tugas akhir ini penggunaan metode centroid linkage lebih baik dari pada menggunakan metode k-means. Pada metode centroid linkage didapatkan nilai error rata-rata yang relatif lebih kecil yaitu sebesar 8,666%, sedangkan jika dengan menggunakan metode k-means didapatkan nilai error rata-rata sebesar 18,658%.

Daftar Pustaka

- [1] Gembong, Tjitrosoepomo. 2005. Morfologi Tumbuhan. Gadjah Mada University. Yogyakarta.
- [2] Muharrem Mercimek, Kayhan Gulez And Tarik Veli Mumcu. December 2005. Real object recognition using moment invariants. Besiktas-Istanbul, Turkey.
- [3] V.devendra, Amitabh Wahi, and Hemalatha Thiagarajan. 2008. Invariant Moments to Scene Categorization Using Support Vector Machines. Department of Computer Applications, Tamil Nadu, India.
- [4] Raicu, Daniela. 2004. Image Feature Extraction. DePaul University.
- [5] Artikel dan Tutorial pada www.java2s.com
- [6] Raicu, Daniela. 2004. Image Feature Extraction. DePaul University.
- [7] Bagus, Bayu. 2007. Image Database Menggunakan Sistem Content Based Image Retrieval dengan Ekstraksi Fitur Terstruktur. Jurusan Teknologi Informasi Politeknik Elektronika Negeri Surabaya Institut Teknologi Sepuluh Nopember, Surabaya.
- [8] Nizar Grira, Michael E.Houle. 2007. A Hybridized Centroid-Medoid Clustering Heuristic. National Institute of Informatics, Tokyo, Japan.
- [9] Hasniawati, Helmy. 2007. Image Clustering Berdasarkan Warna untuk Identifikasi Buah dengan Metode Valley Tracing. Jurusan Teknologi Informasi Politeknik Elektronika Negeri Surabaya Institut Teknologi Sepuluh Nopember, Surabaya.

[CV Penulis]

Febri Liantoni, menjalankan studi D4 bidang Teknik Informatika pada Politeknik Elektronika Negeri Surabaya – Institut Teknologi Sepuluh Nopember(PENS-ITS) semester 8.