

## SEGMENTASI OBJEK BUAH PADA CITRA DIGITAL MENGGUNAKAN METODE K-MEANS CLUSTERING

<sup>1)</sup>Wahyu Syaifullah J. S., <sup>2)</sup>Yisti Vita Via, <sup>3)</sup>Risma Hadi Setiawan

<sup>1,2,3)</sup>Program Studi Teknik Informatika Fakultas Teknologi Industri

Universitas Pembangunan Nasional “Veteran” Jawa Timur

Jl. Raya Rungkut Madya, Gunung Anyar, Surabaya, Jawa Timur 60294

<sup>1)</sup>wahyu.s.j.saputra@gmail.com, <sup>2)</sup>[yistivita@gmail.com](mailto:yistivita@gmail.com), <sup>3)</sup>risma1lhadi@gmail.com

**Abstrak.** Segmentasi merupakan langkah pertama dan menjadi kunci yang penting dalam suatu pengenalan objek (*object recognition*). Proses segmentasi merupakan suatu proses untuk memisahkan antara satu objek dengan objek lainnya. Dengan proses segmentasi masing-masing objek pada citra dapat diambil secara terpisah sehingga dapat digunakan sebagai masukan proses yang lain dan mempermudah untuk dikembangkan oleh proses selanjutnya. Perancangan program clustering ini bertujuan untuk mendapatkan cluster yang memiliki kemiripan atau kesamaan karakteristik antar objek. Sehingga mempermudah dalam hal memisahkan objek berdasarkan clusternya. Pada hasil uji coba yang dilakukan dengan membandingkan 20 contoh gambar hasil segmentasi dengan metode K-means Clustering dan segmentasi melalui aplikasi Adobe Photoshop cs5 dengan level threshold 200 sampai 250 sesuai dengan hasil yang paling maksimal didapatkan tingkat akurasi maksimal 99.50% dan minimal 85.08%.

**Kata Kunci:** Segmentasi, Clustering, K-Means

Pemrosesan citra digital memerlukan satu proses *preprocessing* yang selanjutnya akan digunakan untuk proses yang lain. Proses tersebut adalah segmentasi. Segmentasi merupakan langkah pertama dan menjadi kunci yang penting dalam suatu pengenalan objek (*object recognition*). Proses segmentasi merupakan suatu proses untuk memisahkan antara satu objek dengan objek lainnya. Dengan proses segmentasi masing-masing objek pada citra dapat diambil secara terpisah sehingga dapat digunakan sebagai masukan proses yang lain. Segmentasi citra merupakan suatu teknik pengelompokan (*clustering*) untuk citra. Dengan kata lain, merupakan suatu proses pembagian citra ke dalam wilayah (*region*) yang mempunyai kesamaan fitur antara lain : tingkat keabuan (*grayscale*), tekstur (*texture*), warna (*color*), gerakan (*motion*).

Pada tahap segmentasi tersebut hasil dari segmentasi juga harus akurat karena jika hasil segmentasi tidak akurat, maka akan mempengaruhi hasil selanjutnya. Untuk mengatasi hal itu agar mendapatkan hasil yang maksimal penulis mengutamakan untuk melakukan segmentasi. Penulis mencoba untuk membuat “SEGMENTASI OBJEK BUAH PADA CITRA DIGITAL MENGGUNAKAN METODE K-MEANS CLUSTERING” metode ini berbasis jarak yang membagi data-data ke dalam sejumlah *cluster*. Diharapkan dengan

metode tersebut segmentasi bisa diterapkan dengan maksimal.

### I. Metodologi Penelitian Pengolahan Citra

Secara harfiah, citra (*image*) adalah gambar pada bidang dwimatra (dua dimensi). Ditinjau dari sudut pandang matematis, citra merupakan fungsi menerus (*continue*) dari intensitas cahaya pada bidang dwimatra. Sumber cahaya menerangi objek, objek memantulkan kembali sebagian dari berkas cahaya tersebut. Pantulan cahaya ini ditangkap oleh alat-alat optik, misalnya mata pada manusia, kamera, pemindai (*scanner*), dan sebagainya, sehingga bayangan objek yang disebut citra tersebut terekam.

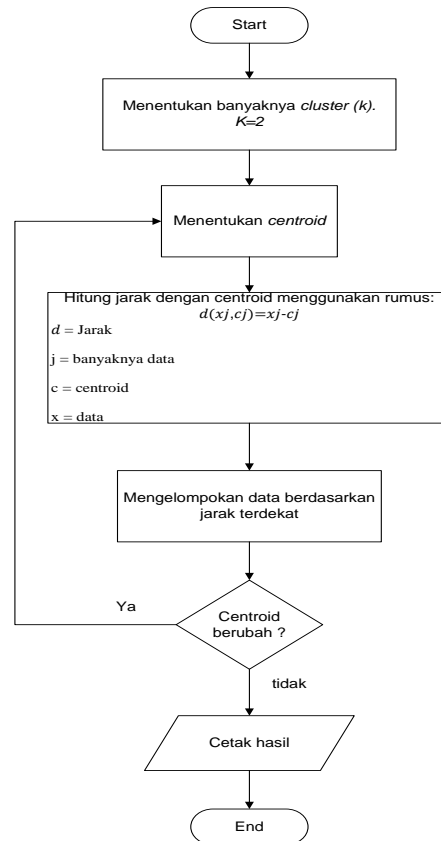
Pengenalan citra mengelompokkan data numerik dan simbolik (termasuk citra) secara otomatis oleh mesin (dalam hal ini komputer). Tujuan pengelompokan adalah untuk mengenali suatu objek didalam citra. Manusia bisa mengenali objek yang dilihatnya karena otak manusia telah belajar mengklasifikasi objek-objek di alam sehingga mampu membedakan suatu objek dengan objek lainnya. Kemampuan sistem visual manusia inilah yang dicoba ditiru oleh mesin. Komputer menerima masukan berupa citra objek yang akan diidentifikasi, memproses citra tersebut, dan memberikan keluaran berupa deskripsi objek didalam citra. [1].

### Metode K-Means Clustering

K-means ditemukan oleh beberapa orang yaitu Lloyd (1957, 1982), Forgey (1965), Friedman and Rubin (1967), and McQueen (1967). Ide dari *clustering* pertama kali ditemukan oleh Lloyd pada tahun 1957, namun hal tersebut baru dipublikasi pada tahun 1982. Pada tahun 1965, Forgey juga mempublikasi teknik yang sama sehingga terkadang dikenal sebagai Lloyd-Forgey pada beberapa sumber. [2]. K-Means Clustering merupakan metode untuk mengklasifikasikan atau mengelompokkan objek-objek (data) ke dalam K-group (*cluster*) berdasarkan atribut tertentu. Pengelompokan data dilakukan dengan memperhitungkan jarak terdekat antara data-data (objek observasi) dengan pusat *cluster* (*centroid*). Prinsip utama dari metode ini adalah menyusun K buah *centroid* atau rata-rata (*mean*) dari sekumpulan data berdimensi N, dimana metode ini mensyaratkan nilai K sudah diketahui sebelumnya (*apriori*). Algoritma K-means dimulai dengan pembentukan prototipe *cluster* diawal kemudian secara iteratif prototipe *cluster* tersebut diperbaiki sehingga tercapai kondisi konvergen, yaitu kondisi dimana tidak terjadi perubahan yang signifikan pada prototipe *cluster*. Perubahan ini diukur dengan menggunakan fungsi objektif D yang umumnya didefinisikan sebagai jumlah atau rata-rata jarak tiap item data dengan *centroid* groupnya. [3].

### Algoritma K-Means Clustering

Algoritma K-Means Clustering merupakan metode clustering berbasis jarak yang membagi data-data ke dalam sejumlah *cluster* dimana proses clustering tersebut dilakukan dengan memperhatikan kumpulan dari data-data yang akan dikelompokkan. Pada algoritma ini, pusat *cluster* atau *centroid* dipilih pada tahap awal secara acak dari sekumpulan koleksi (populasi) data. Kemudian K-Means menguji masing-masing komponen didalam populasi data dan menandai komponen tersebut ke salah satu *centroid* yang telah didefinisikan sebelumnya berdasarkan jarak minimum antara komponen (data) dengan masing-masing *centroid*. Posisi *centroid* akan dihitung kembali sampai semua komponen data dikelompokkan ke setiap *centroid* dan terakhir akan terbentuk posisi *centroid* baru. Iterasi ini akan terus dilakukan sampai tercipta kondisi konvergen.



Gambar 1. Flowchart K-Means Clustering

Secara lebih detail, algoritma K-means Clustering dapat dijabarkan seperti mendefinisikan jumlah K *cluster*, Inisialisasi K pusat *cluster* (*centroid*) sebagai *cluster* awal. *Centroid* ini diperoleh secara acak atau dipilih dari K objek data pertama. Untuk setiap komponen data, hitung dan tandai jarak (*distance*) ke *centroid* awal kemudian masukan data tersebut ke *centroid* yang paling dekat jaraknya, hitung dan ubah kembali *centroid* tiap *cluster* sebagai rata-rata dari seluruh anggota kelompok *cluster* tersebut. Cek semua data kembali dan taruh setiap data yang terdekat dengan *centroid* baru. Jika anggota tiap *cluster* tidak berubah, maka langkah berhenti dan jika masih berubah, kembali ke langkah 2. [3].

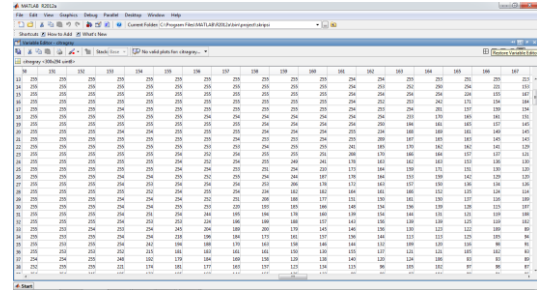
## II. Hasil dan Pembahasan

### Rancangan Uji Coba

Pada tahap ini terdapat beberapa tahapan uji coba untuk mencoba apakah metode yang digunakan berjalan dengan benar atau tidak dan mengecek apakah terdapat kekurangan dari program ini, beberapa tahapan uji coba yang dilakukan antara lain:

- a. Mencoba program dengan menginputkan 30 gambar dengan ukuran yang berbeda.

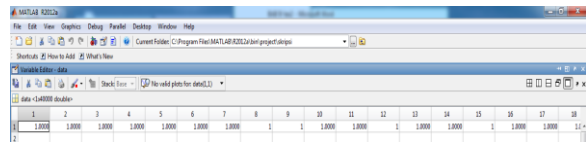
- b. Mencoba dengan mengubah ukuran resizeselain 200x200 yaitu dengan ukuran 150x150, 250x250, 280x280,dan 300x300.
- c. Mencoba dengan gambar objek yang sama tetapi berbeda posisi seperti 5 gambar pisang dengan ukuran dan posisi yang berbeda.
- d. Mencoba dengan background yang tidak berwarna putih.



Gambar 4. Hasil Grayscale

Kita dapat melihat nilai hasil *preprocessing grayscale* yang merupakan perubahan dari citra red, green, blue (RGB) pada Gambar 4 Untuk nilai 255 menampilkan warna putih dan 0 dengan warna hitam dengan nilai maksimal 255.

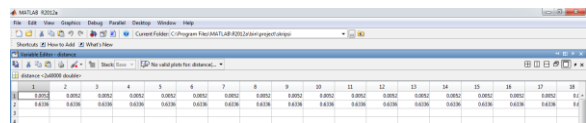
- c. Nilai Data Dalam 1 Dimensi  
Berikut adalah tampilan nilai data dalam wujud 1 dimensi.



Gambar 5. Data Dalam 1 Dimensi

Pada Gambar 5 dapat kita lihat bagaimana hasil mengubah citra ke dalam wujud satu dimensi, nilai tersebut mulai dari piksel pertama 1 sampai dengan piksel yang paling terakhir 40.000.

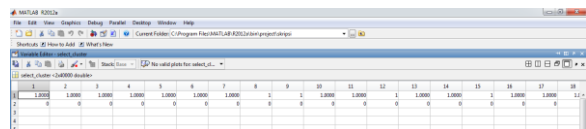
- d. Hasil Perhitungan Terhadap *Centroid*  
Berikut adalah tampilan hasil perhitungan data terhadap *centroid*.



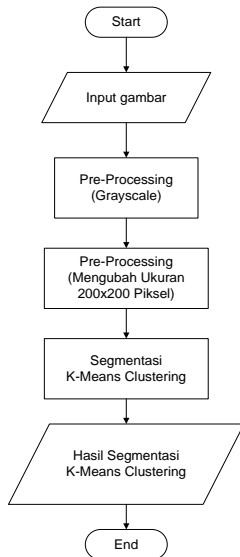
Gambar 6. Hasil Perhitungan Terhadap Centroid

Pada hasil perhitungan jarak data seperti yang tertera pada Gambar 6 dapat kita lihat baris pertama ialah hasil jarak data dengan cluster 1, baris ke dua jarak data dengan cluster 2.

- e. Hasil Pengelompokan Data  
Berikut adalah tampilan hasil dari pengelompokan data.



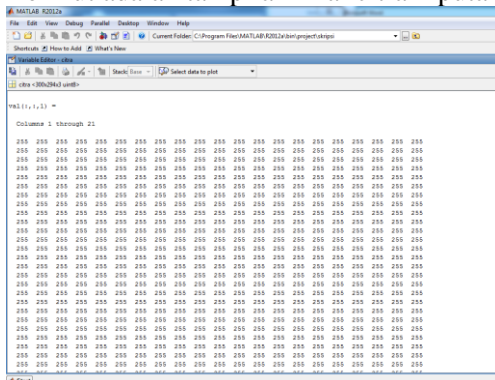
Gambar 7. Hasil Pengelompokan Data



Gambar 2. Flowchart Alur Sistem

Sebagaimana design *flowchart* seperti pada Gambar 2, berikut adalah hasil implementasi dari program sesuai dengan metode yang digunakan.

- a. Nilai Piksel Citra Inputan  
Berikut adalah tampilan nilai citra inputan.



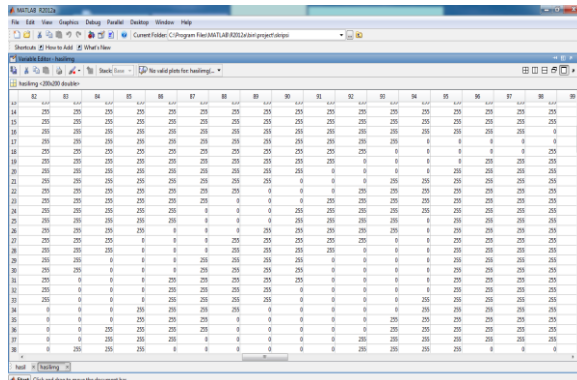
Gambar 3. Nilai Citra Inputan

Data yang awalnya berupa gambar (*image*) di representasikan ke dalam nilai. Dapat dilihat pada Gambar 3 ialah nilai masing-masing piksel dari citra inputan yang akan diolah atau digunakan untuk tahap uji coba.

- b. Hasil *Grayscale*  
Berikut adalah tampilan nilai dari *grayscale*.

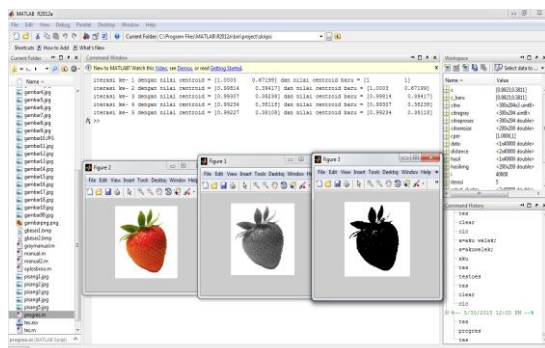
Pada Gambar 7 dapat dilihat bagaimana hasil pengkelompokan data sesuai dengan clusternya berdasarkan jarak yang terdekat. Jika jarak data1 kurang dari jarak data 2, maka data 1 masuk ke *cluster* 1. Jika jarak data1 lebih dari jarak data 2, maka data 1 masuk ke *cluster* 2.

- f. Hasil Akhir Perhitungan  
Berikut adalah hasil akhi dari perhitungan K-Means Clustering.



Gambar 8. Hasil Akhir Perhitungan

Hasil akhir yang dapat dilihat pada Gambar 8 ialah hasil dalam ukuran citra 200x200 piksel.



Gambar 9. Hasil Segmentasi

Pada Gambar 9 menjelaskan hasil segmentasi menggunakan algoritma K-Means Clustering terhadap citra. Setelah program di *run* maka akan terlihat gambar inputan, gambar *grayscale*, *output* hasil pengolahan gambar segmentasi dan hasil perubahan *centroid*. Pada *workspace* Matlab dapat dilihat jalannya perhitungan algoritma K-means Clustering mulai dari variabel, matriks, ukuran citra, nilai citra, hasil *centroid*. fungsi dari *Workspace window* ialah sebagai navigator bagi pemakai dalam penyediaan informasi mengenai variabel yang sedang aktif dalam *workspace* pada saat pemakaian hingga hasil akhir perhitungan citra.

**Hasil Uji Coba Akurasi Sistem**

Hasil akurasi didapatkan dari menyamakan ke dua gambar segmentasi yang dilakukan berdasarkan algoritma K-means Clustering dan Photoshop, dua gambar tersebut dibandingkan pikselnya satu persatu. Setiap perbedaan diberi nilai 1, dari semua perbedaan yang diperoleh selanjutnya ialah menjumlahkan perbedaan nilai tersebut hingga mendapatkan hasil yang diinginkan. Sebelum membandingkan piksel kita harus memastikan ukuran dan nilai piksel dilakukanlah proses normalisasi. Nilai akurasi didapatkan dengan cara

$$x = \frac{\text{jumlah total kesamaan piksel}}{\text{banyaknya piksel}} \times 100\%$$

K-means Clustering	Adobe Photoshop	Jumlah Kesamaan (Piksel)	Hasil Perhitungan Akurasi
		39801	99.50%
		34034	85.08%

Gambar 9. Perbandingan Segmentasi

Tabel 1. Hasil Akurasi

No	Nama	Jumlah Kesamaan Piksel	Akurasi
1	Gambar 1	39568	98.92%
2	Gambar 2	37935	94.84%
3	Gambar 3	39801	99.50%
4	Gambar 4	38007	95.02%
5	Gambar 5	37675	94.19%
6	Gambar 6	39517	98.79%
7	Gambar 7	38202	95.50%
8	Gambar 8	39197	97.99%
9	Gambar 9	37308	93.27%
10	Gambar 10	38360	95.9%
11	Gambar 11	34055	85.14%
12	Gambar 12	36102	90.26%
13	Gambar 13	38815	97.04%
14	Gambar 14	39065	97.66%
15	Gambar 15	37837	94.59%
16	Gambar 16	39754	99.38%
17	Gambar 17	39474	98.68%
18	Gambar 18	38598	96.49%
19	Gambar 19	39143	97.86%
20	Gambar 20	39310	98.27%
21	Gambar 21	38712	96.78%
22	Gambar 22	37123	92.81%
23	Gambar 23	35586	88.96%
24	Gambar 24	37582	93.95%

25	Gambar 25	36726	91.81%
26	Gambar 26	38450	96.12%
27	Gambar 27	37038	92.59%
28	Gambar 28	34034	85.08%
29	Gambar 29	39134	97.83%
30	Gambar 30	37564	93.91%

### III. Simpulan

- Dari 20 uji coba citra pencahayaan dan *background* sangat berpengaruh pada hasil segmentasi.
- Hasil tidak maksimal apabila *resize* dilakukan dengan ukuran yang berbeda seperti 160x120 piksel, tapi jika pada tahap *resize* dilakukan dengan ukuran yang sama maka hasil akan terlihat baik seperti ukuran 200x200 piksel.
- Dalam proses segmentasi kita perlu memberi batas toleransi eror untuk mengatasi looping yang terus menerus tanpa henti atau hasil tidak terbatas (*NaN*).

### IV. Daftar pustaka

- [1] Eka Widya Wardani, 2013. Pengnalan Motif Batik Menggunakan Metode Transformasi Paket Wavelet. Informatika: Bandung.
- [2] X. Wu and V. Kumar, eds., The Top Ten

Algorithms in Data Mining, Chapman and Hall, 2009. Mengenal MATLAB, diakses pada 10 Maret 2015 dari : <http://www.mikron123.com/index.php/Tutorial-Matlab/Tutorial-1-Mengenal-Matlab.html>

- [3] (Aris, Maria, Novianto 2013). ANALISIS ALGORITMA K-MEANS CLUSTERING DENGAN MENGGUNAKAN OPENMP.
- [4] Edi Satriyanto, M.SI, 2012. Diakses pada 25 Maret dari : kangedi.lecturer.pens.ac.id
- [5] Nur Wakhidah, 2011. Perbaikan Kualitas Citra Menggunakan Metode Contrast Stretching. JURNAL TRANSFORMATIKA, Volume 8, No.2.
- [6] Saptono Widodo 2004. SEGMENTASI CITRA MENGGUNAKAN TEKNIK PEMETAAN WARNA (COLOR MAPPING) DENGAN BAHASA PEMROGRAMAN DELPHI.
- [7] Saluky, 2014. Dasar Pemrograman Matlab. Diakses pada 29 Maret dari : <http://www.etunas.com/web/dasar-pemrograman-matlab.htm>

*Halaman ini sengaja dikosongkan.*