

## Penerapan K-Means Clustering dalam Pengelompokan Lipstik Matte Lip Cream Berdasarkan Warna RGB

Ananda Mulia Alhumaerah<sup>1</sup>, Arief Bramanto Wicaksono Putra<sup>2\*</sup>, Rihartanto<sup>3</sup>  
<sup>1,2,3</sup>Jurusan Teknologi Informasi/Politeknik Negeri Samarinda, Samarinda, Indonesia

\*E-mail: ariefbram@gmail.com

### Abstrak

Lipstik adalah jenis kosmetik yang paling banyak digunakan dan sangat familiar di masyarakat, hampir semua brand make-up memproduksi kosmetik jenis ini. Warna Lipstik pada umumnya dapat dikelompokkan menjadi seperti berikut : nude, pink, dan merah. Lipstik yang dijadikan obyek penelitian adalah 10 (sepuluh) sampel yang diperoleh dari Lipstik NYX Matte Lip Cream, dari masing-masing sampel dicuplik sebanyak 1 kali. Dari keberagaman warnanya, lipstik dapat dikelompokkan dalam 3 (tiga), 4 (empat) dan 5 (lima) kelompok. Berdasarkan hal tersebut, lipstik akan dikelompokkan dengan menggunakan metode K-Means Clustering, namun sebelum dikelompokkan diperlukan ciri yang didapatkan dengan metode Euclidean Distance dan Coefficient Correlation serta kombinasi pembentukan atribut yang beragam. Hasil pengelompokkan menggunakan metode K-Means Clustering menghasilkan 3 cluster, 4 cluster dan 5 cluster dimana pembentukan atribut dengan Menjumlahkan dan mencari Nilai Rata-Rata adalah pembentukan atribut yang menghasilkan clustering yang sama pada proses clustering 3 dan 4 cluster. Pembentukan nilai centroid juga berpengaruh pada hasil pengelompokan. Hasil perbandingan pengelompokan yaitu : pengelompokan 3 cluster menghasilkan kelompok A terdiri dari 8 member , kelompok B dan C memiliki 1 member yang berbeda. Dan pengelompokan 4 cluster menghasilkan pembagian kelompok A dan D memiliki 4 member yang berbeda serta kelompok B dan C memiliki masing masing cluster 1 member.

**Kata kunci:** Lipstik, Cluster, K-Means

### Abstract

*Lipstick is the most widely used type of cosmetics and very familiar in the community, almost all make-up brands produce this type of cosmetics. Lipstick colors can generally be grouped into the following: nude, pink, and red. The lipstick used as the object of the study were 10 (ten) samples obtained from NYX Matte Lip Cream Lipstick, from each sample sampled 1 time. From the diversity of colors, lipstick can be grouped into 3 (three), 4 (four) and 5 (five) groups. Based on this, the lipstick will be grouped using the K-Means Clustering method, but before grouping the characteristics obtained with the Euclidean Distance and Coefficient Correlation methods and the combination of various attribute formation are needed. The grouping results using the K-Means Clustering method produces 3 clusters, 4 clusters and 5 clusters where the formation of attributes by Summing up and searching for Average Values is the formation of attributes that produce the same clustering in clustering processes 3 and 4 clusters. The formation of centroid values also affects the results of grouping. The results of the grouping comparison are: grouping 3 clusters produces group A consisting of 8 members, group B and C have 1 different member. And grouping 4 clusters results in a division of groups A and D having 4 different members and groups B and C having each cluster 1 member.*

**Keywords:** Lipstick, Cluster, K-Means

### 1. Pendahuluan

Pewarna bibir (lipstik) merupakan sediaan kosmetika yang digunakan untuk mewarnai bibir dengan sentuhan artistik sehingga dapat meningkatkan estetika dalam tata rias wajah [1]. Sediaan pewarna bibir terdapat dalam berbagai bentuk, seperti cairan, krayon, dan krim.

Sedangkan berdasarkan jenisnya lipstik dibedakan menjadi Matte Lipstik, Satin Lipstik, Glossy Lipstik, Liquid Lipstik, Lip Tint, Lip Balm. Matte Lipstik adalah jenis lipstik yang tidak bersinar, teksturnya membuat kontur bibir terlihat dan warnanya lebih tahan lama. Matte lipstik dapat membuat efek bibir terlihat lebih kecil dalam

warna baik terang maupun gelap. Kebanyakan lipstik matte memberikan kesan kering pada bibir.

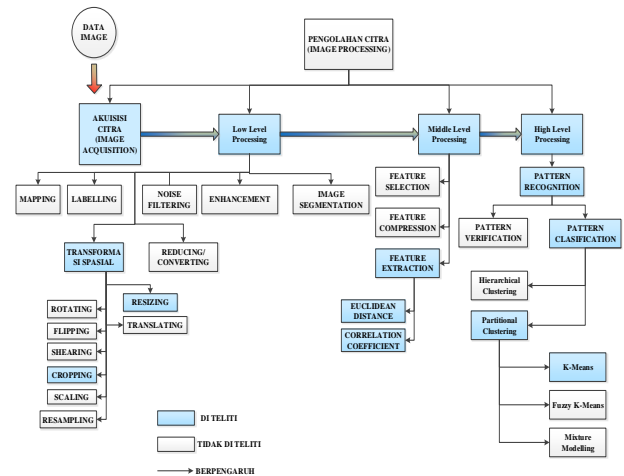
Tujuan pengelompokan adalah untuk mengenali suatu objek didalam citra. Manusia bisa mengenali objek yang dilihatnya karena otak manusia telah belajar mengklasifikasi objek-objek di alam sehingga mampu membedakan suatu objek dengan objek lainnya. Kemampuan sistem visual manusia inilah yang dicoba ditiru oleh mesin. Komputer menerima masukan berupa citra objek yang akan diidentifikasi, memproses citra tersebut, dan memberikan keluaran berupa deskripsi objek didalam citra [2].

Mengenali suatu obyek apakah berbeda dengan obyek lainnya dapat dilakukan dari perbedaan warnanya. Demikian pula pada citra, obyek yang berbeda memiliki warna yang berbeda pula. Ia juga menjelaskan bahwa Model warna RGB didasarkan pada tri-stimulus vision theory yang mengatakan bahwa manusia melihat warna dengan cara membandingkan cahaya yang datang dengan sensor-sensor peka cahaya pada retina (yang berbentuk kerucut). Sensor-sensor tersebut peka terhadap cahaya dengan panjang gelombang 630 nm (merah), 530 nm (hijau), 450 nm (biru)[3]. Ketika indeks warna RGB diekstraksi dan dijadikan sebagai acuan parameter dalam melakukan klasifikasi diketahui bahwa semakin tinggi indeks warnanya, maka citra tersebut semakin terang. Begitu pula sebaliknya, semakin kecil nilai indeks warnanya, maka citra akan semakin gelap [4].

Ketika metode K-Means Clustering dan metode Agglomerative Hierarchical Clustering (AHC) dibandingkan, terbukti bahwa meskipun harus menentukan nilai centroid terlebih dahulu hasil pengelompokan dengan K-Means menghasilkan data pengelompokan yang lebih baik, sedangkan metode AHC meskipun lebih efisien karena tidak menggunakan titik pusat pada awal pengelompokan dan hanya menghitung tingkat kemiripan tetapi memiliki hasil yang kurang baik [5].

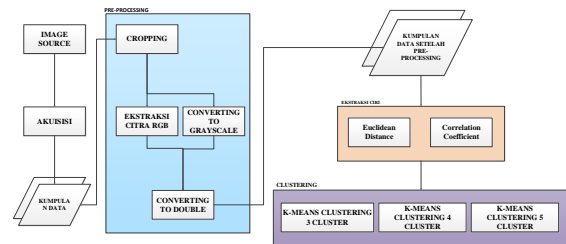
Dari uraian diatas yang menggambarkan keberagaman warna lipstik serta kemampuan K-mean Clustering dalam membedakan dan mengelompokkan objek berdasarkan warna, maka penelitian ini bertujuan untuk membentuk kelompok warna pada lipstik dengan K-Mean Clustering. Kelompok yang diinginkan terbentuk dari variasi ekstraksi ciri serta kombinasi pembentukan atribut keanggotaan, dimana pembentukan titik tengah (centroid) di tentukan

dengan aturan tertentu. Adapun Kerangka Konsep penelitian ini disajikan pada Gambar 1 dibawah ini.



Gambar 1. Kerangka Konsep Penelitian

## 2. Metode Penelitian



Gambar 2. Metodologi Penelitian

Dari Gambar 2, tahap-tahap yang dilakukan dalam penelitian ini adalah Pre-processing, ekstraksi ciri dan klasifikasi. Berikut adalah penjelasan dari gambar diatas :

### 1. Data

Tahap awal sebelum pemrosesan citra adalah tahap akuisisi. Data diakuisisi melalui suatu proses tertentu untuk kemudian dikumpulkan dalam suatu kumpulan data. Adapun data citra lipstik yang digunakan dalam penelitian ini adalah citra lipstik NYX Soft Matte Lip Cream.

### 2. Pre-processing

Tahap Pre-processing yang dilakukan antara lain melakukan Cropping dengan Photoshop CC, Ekstraksi Ciri Citra RGB, Converting RGB to Grayscale, Converting to Double. Tahap-tahap tersebut dilakukan agar citra yang akan diolah

memiliki kualitas yang baik dan dapat diproses pada tahap selanjutnya.

### 3. Ekstraksi Ciri

Ekstraksi ciri atau ekstraksi fitur merupakan bagian fundamental dari analisis citra[6]. Karakteristik fitur yang baik sebisa mungkin memenuhi persyaratan berikut :

- Dapat membedakan suatu objek dengan yang lainnya
- Memperhatikan kompleksitas komputasi dalam memperoleh fitur. Kompleksitas komputasi yang tinggi tentu akan menjadi beban tersendiri dalam menemukan suatu fitur.
- Tidak terikat (independence) dalam arti bersifat invariant terhadap berbagai transformasi (rotasi, penskalaan, penggeseran, dan lain sebagainya).
- Jumlahnya sedikit, karena fitur yang jumlahnya sedikit akan dapat menghemat waktu komputasi dan ruang penyimpanan untuk proses selanjutnya (proses pemanfaatan fitur)

Metode yang digunakan pada tahapan ekstraksi ciri yang digunakan adalah Euclidean Distance dan Correlation Coefficient. Hasil dari ekstraksi ciri tersebut akan dikelompokkan menjadi 3 (tiga), 4 (empat) dan 5 (lima) kelompok.

#### a. Euclidean Distance

Euclidean Distance adalah perhitungan jarak dari 2 buah titik dalam Euclidean space yang merupakan selisih antara 2 buah vector yang akan dibandingkan untuk pengenalan sebuah objek yang akan diuji [7].

Euclidean Distance adalah metrika yang paling sering digunakan untuk menghitung kesamaan dua vector[8]. Euclidean distance menghitung akar dari kuadrat perbedaan dua vektor yang dapat dirumuskan seperti pada persamaan (1) dibawah ini :

$$d = \sqrt{\sum_{k=1}^n (x_{ij} - x_{jk})^2} \quad (1)$$

Keterangan :

d = Euclidean Distance

x<sub>ij</sub> = Citra Uji

x<sub>jk</sub> = Citra Acuan

#### b. Correlation Coefficient

Ekstraksi ciri menggunakan *Coefficient Correlation* adalah menghitung nilai kesamaan

dibandingkan ketidaksamaan, hasil dari perhitungan tersebut akan dirata-rata berdasarkan baris dan kolom untuk merepresentasikan ciri pada citra menggunakan *Coefficient Correlation*[9].

### 4. Pengelompokan

Pada tahapan Pengelompokan, hasil ekstraksi ciri pada tahap sebelumnya akan dikelompokkan menggunakan metode K-Means Clustering. Dimana kelompok tersebut akan terbentuk dari kombinasi pembentukan atribut keanggotaan dan aturan pembentukan titik tengah (centroid).

Data clustering menggunakan metode K-Means Clustering ini secara umum dilakukan dengan algoritma dasar sebagai berikut[10] :

- Pilih jumlah cluster k.
- Inisialisasi k pusat cluster ini bisa dilakukan dengan berbagai cara. Namun yang paling sering dilakukan adalah dengan cara random. Pusat-pusat cluster diberi nilai awal dengan angka-angka random,
- Alokasikan semua data/ objek ke cluster terdekat. Kedekatan dua objek ditentukan berdasarkan jarak kedua objek tersebut. Demikian juga kedekatan suatu data ke cluster tertentu ditentukan jarak antara data dengan pusat cluster. Dalam tahap ini perlu dihitung jarak tiap data ke tiap pusat cluster. Jarak paling antara satu data dengan satu cluster tertentu akan menentukan suatu data masuk dalam cluster mana. Untuk menghitung jarak semua data ke setiap titik pusat cluster dapat menggunakan teori jarak Euclidean.
- Hitung kembali pusat cluster dengan keanggotaan cluster yang sekarang. Pusat cluster adalah rata-rata dari semua data/ objek dalam cluster tertentu. Jika dikehendaki bisa juga menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang bisa dipakai.
- Tugaskan lagi setiap objek memakai pusat cluster yang baru. Jika pusat cluster tidak berubah lagi maka proses clustering selesai. Atau, kembali ke langkah c sampai pusat cluster tidak berubah lagi.

Pembentukan centroid awal pada penelitian ini ditentukan sebagai berikut :

- Centroid 1 didapatkan dengan merata-ratakan setiap kriteria.
- Centroid 2 didapatkan dengan mencari nilai maksimum setiap kriteria.

- c. Centroid 3 didapatkan dengan mencari nilai minimum setiap kriteria.
- d. Centroid 4 didapatkan dengan mencari nilai tengah setiap kriteria.
- e. Setiap centroid terakhir didapatkan dengan cara merata-ratakan nilai centroid sebelumnya.

### 3. Hasil Penelitian

Pada hasil dan pembahasan akan dijelaskan tentang tahapan-tahapan yang telah dilakukan untuk memperoleh kelompok data dari objek yang diteliti. Tahapan yang dilakukan adalah : akuisisi data, pre-processing, analisis ciri menggunakan ekstraksi ciri dengan metode Euclidean Distance dan metode Correlation Coefficient, dan tahap clustering data dengan metode K-Means Clustering.

#### 3.1. Akuisisi

Pada penelitian ini akusitor yang digunakan untuk proses akuisisi citra adalah kamera smartphone Xiaomi Redmi Note 3 Pro dengan resolusi kamera 16 MegaPiksel atau 4608 x 3456 pixels (4:3). Objek yang menjadi penelitian kali ini adalah lipstik dengan 10 warna yang berbeda yang sudah dipilih yaitu 4 warna lipstik nude, 3 warna lipstik pink dan 3 warna lipstik red. Format gambar yang dipilih adalah JPG. Adapun kondisi pencahayaan saat melakukan akuisisi adalah gelap gulita dengan flash kamera menyala dan kamera tegak lurus dengan objek yang dituju.

Data citra yang telah diakuisisi akan diolah kedalam proses pre-processing. Akan diambil 1 (satu) sample warna per lipstik untuk diambil ciri dari citra warna lipstik tersebut.

#### 3.2. Pre-Processing

Pada tahap ini input berupa citra digital hasil akuisisi akan melalui tahapan cropping, ekstraksi komponen RGB, konversi tipe warna dan konversi tipe data sebagai berikut :

##### a. Cropping

Untuk mempermudah peneliti dalam mengekstraksi ciri citra maka digunakan aplikasi tambahan untuk melakukan pemotongan citra. Proses cropping pada tahap ini menggunakan aplikasi Adobe Photoshop CC.

##### b. Read Image

Proses ini berguna untuk membaca citra dalam bentuk file. Software komputasi yang digunakan adalah Matlab. Format image yang digunakan

adalah JPG, proses baca menggunakan program dari M-File, hasil pembacaan data dari gambar menggunakan syntax sebagai berikut :

```
>> LIP1 = imread('nude1.jpg');
```

Pada perintah diatas variabel yang digunakan adalah variabel LIP1 didalam fungsi imread untuk membaca data citra 'nude1.jpg' yang terdapat pada data collection, kemudian image yang sudah berupa array atau matriks akan tersimpan di workspace matlab. Pada proses ini citra yang terbaca dalam format truecolor (RGB).

##### c. Ekstraksi Citra RGB

Setelah dilakukan proses data citra dibaca (Read Image) maka proses selanjutnya yang dilakukan Ekstraksi Ciri Citra RGB. Default citra berjenis truecolor RGB (Red, Green, Blue). Proses Ekstraksi yang dilakukan memisahkan ketiga komponen utama citra yaitu R, G, dan B. Berikut adalah syntax yang digunakan untuk melakukan ekstraksi citra :

```
>> R = LIP1 (:,:,1);
```

```
>> G = LIP1 (:,:,2);
```

```
>> B = LIP1 (:,:,3);
```

LIP1 (:,:,1) berarti semua baris dan semua kolom dalam bidang gambar pertama. LIP1 (:,:,2) berarti semua baris dan semua kolom dalam bidang gambar kedua. LIP1 (:,:,3) berarti semua baris dan semua kolom dalam bidang gambar ketiga.

##### d. Converting to Grayscale

Selain mengekstraksi ciri citra RGB, penelitian ini juga mengkonversi citra RGB menjadi menjadi citra Grayscale. Proses konversi yang dilakukan adalah merubah true color menjadi grayscale, warna yang dimiliki adalah warna dari hitam,keabuan dan putih, warna keabuan yang disebut disini merupakan berbagai tingkat dari warna hingga mendekati warna putih. Pada matlab telah disediakan syntax untuk mengubah citra RGB menjadi citra Grayscale. Berikut adalah syntax yang digunakan :

```
>> Gray = rgb2gray( LIP1 );
```

##### e. Converting to Double

Agar proses aritmatika pada matriks dapat dilakukan, maka perlu dilakukan proses konversi dari tipe data uint8 ke tipe data double dengan menggunakan fungsi im2double seperti berikut ini.

```
>> RDb = im2double ( R );
```

```
>> GDb = im2double ( G );
>> BDb = im2double ( B );
>> GrayDb = im2double ( Gray );
```

### 3.3. Ekstraksi Ciri

Setelah melalui tahap pre-processing, tahap selanjutnya adalah tahap Ekstraksi Ciri. Adapun metode yang digunakan adalah metode *Euclidean Distance* dan metode *Correlation Coefficient*.

#### a. Euclidean Distance

Pada proses ekstraksi ciri menggunakan *Euclidean Distance* semua citra telah melalui tahap pre-processing. Sintak untuk memperoleh nilai *Euclidean Distance* yang kemudian disimpan pada variabel R\_ED telah disediakan oleh MATLAB seperti contoh dibawah ini :

```
>> R_ED = dist (RDb);
```

Pada sintak diatas, dist merupakan fungsi dari *Euclidean Distance* yang tersedia pada MATLAB.

Untuk memudahkan komputasi, pada penelitian ini akan dihitung rata-rata baris dan rata-rata kolom dari setiap piksel dengan sintak sebagai berikut :

```
>> R_ED = mean (dist(Rdb));
```

Fungsi mean berfungsi untuk merata-rata baris dan kolom sehingga menghasilkan matriks berukuran 1 x 670.

Karena hasil dari rata-rata baris dan kolom masih memiliki banyak nilai, maka proses selanjutnya adalah mencari rata-rata dari hasil rata-rata baris dan kolom dengan menggunakan sintak berikut ini :

```
>> R_ED = mean (mean (dist(Rdb)));
```

Pada sintak diatas, fungsi mean yang tidak berada dalam tanda kurung bertujuan merata-ratakan hasil dari rata-rata baris dan kolom, sehingga diperoleh satu nilai ciri menggunakan *Euclidean Distance*. Sehingga ciri citra menggunakan *Euclidean Distance* dapat dilihat pada tabel 1 berikut :

Tabel 1 Ciri Citra menggunakan *Euclidean Distance*

Lipstik	<i>Euclidean Distance</i>			
	R	G	B	Gray
nude1	1.237	2.699	2.849	1.675
nude2	1.170	3.345	3.128	2.526
nude3	1.491	3.053	3.128	2.522
nude4	0.990	3.201	3.026	2.400
pink1	1.220	3.907	2.121	2.440
pink2	1.041	3.452	2.490	2.200
pink3	1.251	2.213	1.792	1.351
red1	0.900	3.467	2.381	2.338
red2	0.936	3.926	2.759	2.684
red3	1.068	5.367	2.759	3.499

#### b. Correlation Coefficient

Pada proses ekstraksi ciri menggunakan *Correlation Coefficient* semua citra telah melalui tahap pre-processing. Sintak untuk memperoleh nilai *Correlation Coefficient* yang kemudian disimpan pada variabel R\_CC telah disediakan oleh MATLAB seperti contoh dibawah ini :

```
>> R_CC = corrcoef (RDb);
```

Pada sintak diatas, corrcoef merupakan fungsi dari *Correlation Coefficient* yang tersedia pada MATLAB.

Untuk memudahkan komputasi, pada penelitian ini akan dihitung rata-rata baris dan rata-rata kolom dari setiap piksel dengan sintak sebagai berikut :

```
>> R_CC = mean (rot90(corrcoef(Rdb)));
```

Fungsi rot90 berfungsi untuk merotasi matriks sebesar 900 sedangkan fungsi mean berfungsi untuk merata-rata baris dan kolom sehingga menghasilkan matriks berukuran 1 x 670.

Karena hasil dari rotasi serta rata-rata baris dan kolom masih memiliki banyak nilai, maka proses selanjutnya adalah mencari rata-rata dari hasil rata-rata baris dan kolom dengan menggunakan sintak berikut ini :

```
>>R_CC=mean(mean(rot90(corrcoef(Rdb))););
```

Pada sintak diatas, fungsi mean yang tidak berada dalam tanda kurung bertujuan merata-ratakan hasil dari rotasi serta rata-rata baris dan kolom, sehingga diperoleh satu nilai ciri menggunakan *Correlation Coefficient* dapat dilihat pada tabel 2.

Tabel 2 Ciri Citra menggunakan *Correlation Coefficient*

Lipstik	<i>Correlation Coefficient</i>			
	R	G	B	Gray
nude1	0.376	0.496	0.512	0.464
nude2	0.198	0.451	0.449	0.448
nude3	0.341	0.461	0.477	0.454
nude4	0.145	0.458	0.465	0.443
pink1	0.163	0.363	0.326	0.335
pink2	0.227	0.464	0.470	0.448
pink3	0.438	0.521	0.470	0.420
red1	0.111	0.504	0.483	0.478
red2	0.252	0.401	0.411	0.403
red3	0.240	0.415	0.414	0.401

### 3.4. K-Means Clustering

Setelah melalui tahap Ekstraksi Ciri, tahap selanjutnya adalah tahap Clustering. Adapun metode yang digunakan adalah metode K-Means Clustering. Clustering pada penelitian ini dilakukan sebanyak 3 (tiga) kali. Dengan jumlah cluster yang berbeda pada tiap clustering. Clustering pertama dengan jumlah 3 (tiga) cluster, clustering kedua dengan jumlah 4 (empat) cluster dan clustering ketiga dengan jumlah 5 (lima) cluster.

Untuk clustering 3 (tiga) dan 5 (lima) cluster digunakan ciri citra dengan komponen *Red*, *Green*, dan *Blue*. Sedangkan untuk clustering 4 (empat) cluster digunakan ciri citra dengan komponen *Red*, *Green*, *Blue* dan *Grayscale*. Kombinasi kriteria juga bervariasi dengan menggunakan Jumlah, Rata-Rata dan Nilai Minimum berbagai kriteria.

Berikut adalah algoritma pemrograman untuk K-Means Clustering 3 cluster

- Baca data member dan Inisialisasi dalam Variabel X
- Pembentukan centroid awal dengan aturan  $C1 = \text{Mean}(X)$ ,  $C2 = \text{Max}(X)$ ,  $C3 = \text{Min}(X)$
- $c = [C1, C2, C3]$
- $\text{err} = 1$
- $\text{iter} = 0$
- Hitung jarak Euclidean dengan persamaan 2.1
- Periksa jarak terdekat
- Update anggota cluster
- Jika posisi  $j = 1$  maka cluster  $(1,j) = 1$
- Kecuali jika posisi  $j = 2$  maka cluster  $(2,j) = 1$
- Kecuali yang lain cluster  $(3,j) = 1$
- Update centroid
- Cek perubahan anggota cluster
- $\text{err} = \text{mean}(\text{sum}(\text{abs}(G-G_{\text{new}})))$ ;
- Jika  $\text{err} > 0$  maka  $c = c_{\text{new}}$ ,  $G = G_{\text{new}}$ .
- Ulang proses 6 – 14 jika  $\text{err} > 0$

Algoritma diatas juga berlaku untuk kondisi 4 dan 5 cluster, dengan menambahkan kondisi pemilihan cluster pada tahap h sampai dengan k.

### 3.5. Perbandingan Hasil Pengelompokan

Hasil pengelompokan menggunakan K-Means Clustering dapat dilihat dari tabel 3,4 dan 5. Yang dapat dibandingkan sebagai berikut :

Tabel 3. Hasil Perbandingan Iterasi Terakhir Dengan 3 Cluster

Cluster	Iterasi ke 3		
	Variasi Penyederhanaan		
	Jumlah	Rata-Rata	Nilai Minimum
C1	nude1, nude2, nude3, nude4, pink1, pink2, red1, red2 (8)	nude1, nude2, nude3, nude4, pink1, pink2, red1, red2 (8)	nude2, nude4, pink1, pink2, red1, red2 (6)
C2	red3 (1)	red3 (1)	nude1, nude3, red3 (3)
C3	pink3 (1)	pink3 (1)	pink3 (1)

Tabel 4. Hasil Perbandingan Iterasi Terakhir Dengan 4 Cluster

Cluster	Iterasi ke 5		
	Variasi Penyederhanaan		
	Jumlah	Rata-Rata	Nilai Minimum
C1	nude1, nude2, nude3, nude4 (4)	nude1, nude2, nude3, nude4 (4)	nude2, nude4, pink2, red1 (4)
C2	red3 (1)	red3 (1)	nude1, nude3, pink3 (3)
C3	pink3 (1)	pink3 (1)	pink1 (1)
C4	pink1, pink2, red1, red2 (4)	pink1, pink2, red1, red2 (4)	red2, red3 (2)

Tabel 5. Hasil Perbandingan Iterasi Terakhir Dengan 5 Cluster

Cluster			
Iterasi ke 4			
Variasi Penyederhanaan			
Cluster	Jumlah	Rata-Rata	Nilai Minimum
C1	nude3 (1)	pink2, red1 (2)	- (0)
C2	red3 (1)	red3 (1)	pink3 (1)
C3	pink3 (1)	pink3 (1)	pink1, pink2, red1, red2 (4)
C4	nude1, nude2, nude4 (3)	nude1, nude2, nude3, nude4 (4)	nude1, nude2, nude3, nude4 (4)
C5	pink1, pink2, red1, red2 (4)	pink1, red2 (2)	red3 (1)

Dalam tabel 3 dan tabel 4 dapat dilihat bahwa penyederhanaan atribut dengan cara Menjumlahkan dan mencari Nilai Rata-Rata cocok digunakan dalam penelitian ini karena keduanya menghasilkan pengelompokan yang sama. Penentuan nilai centroid juga berpengaruh dalam proses clustering karena apabila centroid yang diinisialisasikan secara random kurang baik, maka hasil pengelompokan menjadi kurang optimal karena pengelompokan data yang dihasilkan berbeda-beda.

#### 4. Kesimpulan

Dalam penelitian ini telah didapatkan hasil berdasarkan pengujian perbandingan bahwa penyederhanaan atribut dengan cara Menjumlahkan dan mencari Nilai Rata-Rata cocok digunakan dalam penelitian ini karena keduanya menghasilkan pengelompokan yang sama pada clustering 3 cluster dan clustering 4 cluster. Sehingga penyederhanaan dengan cara mencari nilai minimum dapat dihindari ketika akan melakukan proses penyederhanaan. Penentuan pembentukan titik tengah (centroid) juga berpengaruh pada hasil clustering. Jika centroid yang diinisialisasikan secara random kurang baik, maka hasil pengelompokan menjadi kurang optimal karena pengelompokan data yang dihasilkan berbeda-beda. Penelitian dengan konsep klasifikasi diharapkan dapat dilakukan sebagai kelanjutan dari tahap ini.

#### 5. Daftar Pustaka

- [1] N. Adliani, Nazliniwyaty, and D. Purba, "Formulasi Lipstik Menggunakan Zat Warna Dari Ekstrak Bunga Kecombrang (Etlingera elatior (Jack) R.M.Sm.)," *J. Pharm. Pharmacol.*, vol. 1, no. 2, pp. 87–94, 2012.
- [2] W. Syaifullah, Y. V. Via, and R. H. Setiawan, "Segmentasi Objek Buah Pada Citra Digital Menggunakan Metode K-Means Clustering," *SCAN - J. Teknol. Inf. DAN Komun.*, vol. X, no. 2, pp. 69–74, 2015.
- [3] Y. R. W. Utami, "K-Means Clustering Untuk Pengenalan Buah Berdasarkan Karakteristik Warna Citra," *J. Ilm. SINUS*, vol. 9, no. 2, pp. 11–20, 2011.
- [4] M. A. Bustomi and Z. Dzulfikar, "Analisis Distribusi Intensitas RGB Citra Digital untuk Klasifikasi Kualitas Biji Jagung menggunakan Jaringan Syaraf Tiruan," *J. Fis. dan Apl.*, vol. 10, no. 3, pp. 127–132, 2014.
- [5] L. Zahrotun, "Analisis Pengelompokan Jumlah Penumpang Bus Trans Jogja Menggunakan Metode Clustering K-Means Dan Agglomerative Hierarchical Clustering (AHC)," *J. Inform.*, vol. 9, no. 1, pp. 1039–1047, 2015.
- [6] A. B. W. Putra, D. S. B. Utomo, and M. D. Rahmawan, "Verifikasi Golongan Darah Manusia Berbasis Citra Dijital Menggunakan Logika Fuzzy," *J. Sains Terap.*, vol. 4, no. 1, 2018.
- [7] B. Aditya, A. Hidayatno, and A. A. Zahra, "Sistem Pengenalan Buah Menggunakan Metode Discrete Cosine Transform Dan Euclidean Distance," *J. Ilm. Tek. Elektro*, vol. 3, no. 2, pp. 134–138, 2014.
- [8] A. B. W. Putra and E. Subkhiana, "Ekstraksi Ciri Entropy Untuk Pengenalan Pola Wajah Menggunakan Fuzzy Rule Base," *J. SMARTICS*, vol. 2, no. 2, pp. 35–42, 2016.

- [9] I. H. Qasim and A. B. W. Putra, "Penggunaan Metode Euclidean Distance , Coefficient Correlation Dan Center Of Gravity Untuk Pembangunan Prototype Ciri Citra Tanda Tangan," *Semin. Nas. Serba Inform. 2017*, vol. 1, no. 1, pp. 52–58, 2017.
- [10] P. R. Nastiti and A. B. W. Putra, "Perbandingan Algoritma K-Means Dan Fuzzy C-Means Clustering Untuk Kualifikasi Data Kinerja Dosen Di Jurusan Teknologi Informasi POLNES," *Semin. Nas. Serba Inform. 2017*, vol. 1, no. 1, pp. 71–76, 2017.