

# Klasifikasi Beras Menggunakan Metode K-Means Clustering Berbasis Pengolahan Citra Digital

Atriyani Trisnawan<sup>1)</sup>, Wahyudi Harianto<sup>2)</sup>, Syahminan<sup>3)</sup>

<sup>1,2,3)</sup> Teknik Informatika Fakultas Sains Dan Teknologi Universitas Kanjuruhan Malang  
atriyan1994@gmail.com, wahyoudi@unikama.ac.id, syahminan@unikama.ac.id

**Abstract.** Rice plants (*Oryza Sativa L.*) are important food crops that have become a staple food for more than half of the world's population. In Indonesia rice is the main commodity in supporting community food. The process of processing rice into rice is done in two ways, namely the processing of pulverized and modern processing using a grinding tool. Rice is an important component in daily food. There are several types on the market, namely: fragrant pandan rice, IR 64, IR 42, C 4, and others. With the variety of forms and types of rice on the market, there are many weaknesses that humans have in perceiving the classification of rice using the senses of vision. Therefore, digital image processing techniques are needed to help analyze the type of rice. This study aims to analyze the type of rice using the K-Means Clustering method based on RGB colors. Before the K-Means calculation, the RGB color feature extraction process must be carried out to get the red value, green value, blue value in each image. The results of this study found that image processing to determine the type of rice using the k-means clustering method can help users to know the type of rice.

**Keyword :** Analysis; Rice; K-Mean Clustering; Digital Image Processing.

**Abstract** Tanaman padi (*Oryza Sativa L.*) merupakan tanaman pangan yang penting yang telah menjadi makanan pokok lebih dari setengah penduduk dunia. Di Indonesia padi merupakan komoditas utama dalam menyokong pangan masyarakat. Proses pengolahan padi menjadi beras dilakukan dengan dua cara yaitu proses pengolahan ditumbuk dan proses pengolahan modern memakai alat penggilingan. Beras merupakan komponen penting dalam makanan sehari-hari. Ada beberapa jenis yang beredar di pasaran yaitu : beras pandan wangi, IR 64, IR 42, C 4, dan lain-lain. Dengan beraneka ragam bentuk dan jenis beras di pasaran banyak kelemahan yang dimiliki manusia di dalam mempersepsi klasifikasi jenis beras menggunakan indera penglihatan. Oleh karena itu, teknik pengolahan citra digital diperlukan untuk membantu menganalisa jenis beras. Penelitian ini bertujuan untuk menganalisa jenis beras menggunakan metode K-Means Clustering berdasarkan warna RGB. Sebelum dilakukan perhitungan K-Means, harus dilakukan proses ekstraksi fitur warna RGB untuk mendapatkan nilai red, nilai green, nilai blue pada tiap citra. Hasil dari penelitian ini didapatkan bahwa image processing untuk menentukan jenis beras menggunakan metode k-means clustering dapat membantu pengguna untuk mengetahui jenis beras.

**Kata Kunci :** Analisa; Beras; K-Mean Clustering; Digital Image Processing.

## Pendahuluan

padi (*Oryza Sativa L.*) merupakan tanaman pangan yang penting yang telah menjadi makanan pokok lebih dari setengah penduduk dunia. Di Indonesia padi merupakan komoditas utama dalam menyokong pangan masyarakat. Makanan lokal seperti umbi-umbian tidak mampu menggeser keberadaan tanaman padi sebagai pangan pokok. Terdapat 18 juta petani padi di Indonesia yang menyumbang 66% terhadap produk domestik bruto (PDB) tanaman pangan.

Di Indonesia proses pengolahan padi menjadi beras dilakukan dengan dua cara yaitu proses pengolahan ditumbuk dan proses pengolahan modern memakai alat penggilingan. Proses pengolahan paska panen di Indonesia identik menggunakan penggilingan modern. Beras merupakan komponen penting dalam makanan sehari-hari. Ada beberapa jenis yang beredar di pasaran yaitu : beras pandan wangi, rojo lele, IR 64, IR 42, C 4, dan lain-lain. Dengan beraneka ragam bentuk, warna dan jenis beras di pasaran banyak kelemahan yang dimiliki manusia di dalam mempersepsi klasifikasi jenis beras menggunakan indera penglihatan

Pada penelitian kali ini, penelitian mencoba menggunakan teknik dengan pendekatan metode untuk menentukan analisa dan pengelompokan jenis beras dengan pendekatan menggunakan pencitraan dengan judul "Klasifikasi beras menggunakan Metode K-Means Clustering Berbasis pengolahan citra digital".

Maka dengan majunya teknologi dan informasi, dibidang *digital image processing*. Untuk mengatasi permasalahan jenis beras, maka disini penulis akan berusaha membuat suatu aplikasi yang menggunakan metode *k-means clustering* yang nantinya dapat digunakan untuk menentukan klasifikasi beras. Berdasarkan permasalahan di atas

## Metode

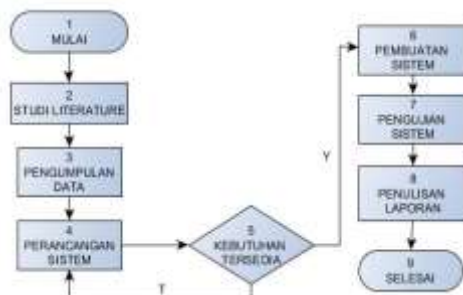
### 1. Alat Pendukung

Alat pendukung yang digunakan dalam penelitian ini adalah sebagai berikut:

1. Perangkat Keras:
  - Laptop processor core i3
  - Ram 6GB
  - Hardisk 500GB
  - LCD 14"
2. Perangkat Lunak:
  - Sistem Operasi Windows 7 Ultimate
  - Borland Delphi 7
  - Microsoft Access 2013
  - Microsoft Office Word 2013
  - Alpha Skins

### 2. Tahapan Penelitian

Tahapan penelitian merupakan langkah-langkah yang akan dilakukan peneliti dalam melakukan penelitian. Tahapan yang dilakukan dalam penelitian ini meliputi studi literatur,



pengumpulan data, perancangan sistem, pembuatan sistem, pengujian sistem, dan pembuatan laporan. Bagan tahapan penelitian disajikan pada Gambar 1.

### 3. Perhitungan *K-means Clustering*

*K-means clustering* adalah salah satu algoritma dalam menentukan klasifikasi terhadap objek berdasarkan fitur dari objek tersebut kedalam **K** kluster. **K** adalah angka positif yang menyatakan jumlah grup/kluster/partisi terhadap objek. Pemartisian data dilakukan dengan mencari nilai jarak minimum antara data dan nilai *centroid* yang telah di set baik secara random atau pun dengan *initial set of centroid*, kita juga dapat menentukan nilai *centroid* berdasarkan **K** objek yang berurutan.

*Centroid* adalah nilai rata-rata aritmetik dari sebuah objek dari seluruh titik dalam objek tersebut. Penerapan *K-means clustering* ini dapat dilakukan dengan prosedur *step by step* berikut:

No	Data	R	G	B	Pixel
1	B1	4497322	287410	280391	255
2	B2	4519442	316529	296287	255
3	B3	4464346	255928	250218	255
4	B4	4499680	290553	276541	255
5	B5	4483034	275847	264708	255
6	B6	4484693	277393	265664	255
7	B7	4471565	261805	250235	255
8	B8	4466688	258150	242576	255
9	B9	4492792	286712	284523	255
10	B10	4485676	273983	265215	255

Akan dilakukan pemartisian data terhadap data diatas sebanyak lima (5) partisi, maka tahapannya adalah sebagai berikut:

1.  $K=5$ , ( $K$ =Jumlah Cluster)
2. Nilai awal centroid (4519442, 316529, 296287) untuk partisi 1, centroid (4499680, 290553, 276541) untuk partisi 2, centroid (4484693, 277393, 265664) untuk partisi 3, centroid (4471565, 261805, 250235) untuk partisi 4, centroid (4485676, 273983, 265215) untuk partisi 5

3. Maka perhitungan *Euclidean Distance* adalah

$$D(p, c)_0 = \sqrt{\sum_{i=0}^n (p_i - c_i)^2}$$

dimana  $p$  adalah data,  $c$  adalah centroid,  $n$  adalah jumlah data,  $I$  adalah iterasi.

4. Hasil perhitungan jarak minimum B1 adalah

$$D(p, c)_1 = \sqrt{(4497322 - 4519442)^2 + (287410 - 316529)^2 + (280391 - 296287)^2} = 70886281111584$$

$$D(p, c)_1 = \sqrt{(4497322 - 4499680)^2 + (287410 - 290553)^2 + (280391 - 276541)^2} = 67247626671623$$

$$D(p, c)_1 = \sqrt{(4497322 - 4484693)^2 + (287410 - 277393)^2 + (280391 - 265664)^2} = 69886281111584$$

$$D(p, c)_1 = \sqrt{(4497322 - 4471565)^2 + (287410 - 261805)^2 + (280391 - 250235)^2} = 69886281111584$$

$$D(p, c)_1 = \sqrt{(4497322 - 4485676)^2 + (287410 - 273983)^2 + (280391 - 265215)^2} = 100886281111584$$

5. Hasil perhitungan jarak keseluruhan

Data	C1	C2	C3	C4	C5
B1	7088 6281 1115 84	67247 62667 1623	69886 28111 1584	90886 28111 1584	100886 281111 584
B2	0	21833 .8374 7764	24162 .0860 4404	39862 .3754 5606	261242 474970 5
B3	6062 5.88	34002 5.886	63058 .2659	93958 .5000	088628 111158

	6104 86	10486	8472	8381	4
B4	3336 6.53 3383 01	0	30918 .7322 6702	38316 .6445 2990	385488 048771 94
B5	2271 7.94 2204 34	25166 .9353 38300	38136 .6445 2990	54947 .6464 9484	672476 266716 23
B6	3854 8804 8771 94	26833 82948 8166	0	41838 36807 8834	126124 247470 5
B7	3757 6576 1384 66	26124 24747 052	70886 28111 1584	0	378967 178190 93
B8	3819 6104 1487 21	12612 42474 705	37576 57613 8466	56798 .3939 494	418383 680788 34
B9	4183 8368 0788 34	57597 04453 3552	38191 .0414 8721	26124 24747 050	418383 680788 34
B10	3789 6.17 8190 93	15692 .4308 9568	67869 .2430 89568	38548 .0487 7194	0

Hasil perhitungan partisi diambil dari jarak minimum, sehingga di dapat sebagai berikut:

Data	C1	C2	C3	C4	C5
B1	0	1	0	0	0
B2	1	0	0	0	0
B3	0	1	0	0	0
B4	0	1	0	0	0
B5	0	1	0	0	0
B6	0	0	1	0	0
B7	0	0	0	1	0
B8	0	1	0	0	0
B9	0	0	0	1	0
B10	0	0	0	0	1

Pada C2 67247626671623 lebih kecil dari C1 70886281111584, C3 69886281111584, C4 90886281111584, C5 100886281111584 maka T1 termasuk ke dalam Cluster 2.

6. Setelah data di partisi, maka selanjutnya nilai centroid harus dihitung ulang untuk menentukan centroid baru:

$$c_i^{(t+1)} = \frac{1}{|S_i^{(t)}|} \sum_{P_j \in S_i^{(t)}} P_j$$

$$c_1 = \left( \frac{(4497322 + 4519422)}{2}, \frac{(287410 + 316529)}{2}, \frac{(280391 + 296287)}{2} \right) = 5098036,7554512,7312517$$

$$c_2 = \left( \frac{(464346 + 4499680)}{2}, \frac{(255928 + 290553)}{2}, \frac{(250218 + 276541)}{2} \right) = 5760220,5569626,5224930$$

$$c_3 = \left( \frac{(4483034 + 4484693)}{2}, \frac{(275847 + 277393)}{2}, \frac{(264708 + 265664)}{2} \right) = 5775815,581726,5276042$$

$$c_4 = \left( \frac{(4471565 + 4466688)}{2}, \frac{(261805 + 258150)}{2}, \frac{(250235 + 242576)}{2} \right) = 5647232,5659580,5430346$$

$$c_5 = \left( \frac{(4492792 + 4485676)}{2}, \frac{(2686712 + 273983)}{2}, \frac{(284523 + 265215)}{2} \right) = 5870871,5661924,5439344$$

7. jarak minimumnya kembali dengan menggunakan centroid yang baru, sehingga didapat hasilnya sebagai berikut:

$$D(p, c)_0 = \sqrt{(9622437 - 6511822)^2 + (9536920 - 63492876)^2 + (9416651 - 60412542)^2} = 80494838245916$$

$$D(p, c)_1 = \sqrt{(9622347 - 5861628)^2 + (9536920 - 5661660)^2 + (9416651 - 5432017.8)^2} = 67113286495484$$

$$D(p, c)_0 = \sqrt{(9622437 - 5775815)^2 + (9536920 - 581726)^2 + (9416651 - 5276042)^2} = 45604838245916$$

$$D(p, c)_0 = \sqrt{(9622437 - 5647232)^2 + (9536920 - 5659580)^2 + (9416651 - 5430346)^2} = 65394838245916$$

$$D(p, c)_0 = \sqrt{(9622437 - 5870871)^2 + (9536920 - 5661924)^2 + (9416651 - 5439344)^2} = 87644838245916$$

Data	C1	C2	C3	C4	C5
B1	80494	67113	45604	6539	8764
	83824	28649	83824	4838	4838
	5916	5484	5916	2459	2459
B2	80389	49135	47890	8494	7463
	19001	79054	53263	8736	8275
	9152	5766	7284	3552	3536
B3	80386	25811	69863	4739	7493
	58992	79891	42638	3923	7634

	4976	5418	4353	5532 76	2748 44
B4	80329 65886 7861	51099 67801 9793	74638 30735 6282	3629 2635 3426 80	8892 6537 5274 84
B5	58848 88155 4157	34702 65464 2836	53427 39202 7354	4648 3625 2640 17	4926 4527 2547 34
B6	80056 71282 5809	15725 07329 0767	90362 36453 8284	8940 4735 4745 3	6038 4548 4638 33
B7	80078 28332 4574	15181 85696 8105	67946 42794 7352	7464 8392 6542 2	6583 3527 4957 46
B8	55683 82093 1942	16576 01363 4164	78534 27365 4909	1204 8463 8363 5	9074 5375 3735 52
B9	80018 83880 7614	44240 26340 3375	20183 74527 3840	6449 3832 6353 6	5756 3836 4846 83
B10	90076 89112 2713	10037 20912 1065	60383 64673 5363	4639 4746 3792 0	5836 4835 4892 92

8. Klasifikasi kembali data berdasarkan jarak minimum di atas:

Data	C1	C2	C3	C4	C5
B1	0	0	1	0	0
B2	0	0	1	0	0
B3	0	1	0	0	0
B4	0	0	0	1	0
B5	0	1	0	0	0
B6	0	1	0	0	0
B7	0	1	0	0	0
B8	0	0	0	1	0
B9	0	0	1	0	0
B10	0	1	0	0	0

Karena ada data yang berpindah ke cluster yang berbeda, sehingga iterasi harus di hitung ulang sampai dengan nilai centroid tidak ada yang berpindah ke cluster yang berbeda.

#### 4. Pengujian Sistem

Pengujian sistem merupakan rangkaian uji coba yang terdapt di dalam sistem untuk klasifikasi telur ayam. Pada proses pengujian ini dimaksudkan untuk mengetahui jalannya sistem, apakah sistem

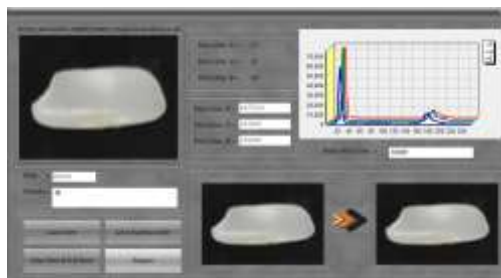
yang dibuat telah sesuai harapan atau belum. Berikut proses pengujian pada sistem penentuan klasifikasi telur ayam dapat dilihat pada gambar berikut.

#### 4.1 Halaman Utama



Halaman utama ini dirancang untuk menampilkan judul dan deskripsi pembuat aplikasi diantaranya nama dan nim pembuat aplikasi. Halaman utama berfungsi untuk menampilkan tombol *home*, *input* data master, analisa, rekap laporan dan keluar.

#### 4.2 Halaman Load Citra



Halaman *Load Citra* ini digunakan untuk menginput data aturan / data yang digunakan sebagai pembandingan nilai yang similar (mirip) satu sama lain dalam *cluster* dan disimilar (tidak mirip) terhadap objek-objek yang berbeda *cluster*, sehingga sebelum melakukan analisa harus menginputkan data master terlebih dahulu. Data master sangat berperan penting terhadap hasil analisa, karena nilai RGB telur ayam yang akan dianalisa akan di cocokkan dengan nilai RGB pada data master. *Input* data master dapat dilakukan oleh pakar dengan *login* terlebih dahulu.

#### 4.3 Halaman Data



Pada halaman Data ini pengguna aplikasi dapat melakukan klasifikasi. klasifikasi yang dimaksud disini yaitu mengetahui klasifikasi Beras. Pada berbagai menu analisa ini, pengguna dapat melihat gambar Beras yang ingin diketahui nilai RGBnya, dan mengklik tombol perintah secara urut sesuai pada *form* di atas.

#### 4.4 Halaman Proses Analisa



Pada halaman proses pengguna yang telah menginputkan gambar beras dan memilih tombol Proses, maka sistem secara otomatis melakukan perhitungan dengan menggunakan metode *k-means clustering* dan akan menampilkan hasil klasifikasi. Hasil klasifikasi diperoleh dari nilai RGB yang paling similar (mirip) antar nilai RGB beras yang akan di analisa dengan beras yang disimpan pada data master. Hasil analisa ini berupa informasi klasifikasi Beras berdasarkan inputan data yang dimasukkan oleh pengguna aplikasi

#### 4.5 Uji Coba Menggunakan *Black Box Testing*

Pengujian sistem dengan menggunakan *black box* bertujuan untuk mengetahui jalannya suatu sistem. Dengan melakukan pengujian seperti ini, dapat diketahui hasil dari proses program apakah telah sesuai dengan hasil yang diharapkan, sehingga dapat menghasilkan *output* yang sesuai dengan hasil tes-tes yang dijalankan. Hasil yang diharapkan disini adalah dengan memasukkan gambar telur ayam dapat diketahui nilai RGBnya, sehingga dengan perbandingan nilai kedekatan antara nilai RGB data master dan data telur yang di analisa dapat menghasilkan informasi berupa klasifikasi telur ayam terhadap pengguna. Berikut adalah hasil pengujian menggunakan *black box*.

Untuk mengetahui seberapa besar keakuratan sistem maka dilakukan uji coba menggunakan kuisisioner antara pendapat responden tentang sistem aplikasi image processing. berdasarkan hasil pengujian kuisisioner yang dihasilkan dari sistem aplikasi image processing sebanyak 10 responden maka hasil tingkat keakurasian sistem sebesar 100% dengan perhitungan sebagai berikut.

10 responden x 8 pertanyaan x bobot nilai tertinggi (3) =

Sangat Setuju

$$10 \times 8 \times 3 = 240$$

$$93/240 \times 100 = 38.75\% \text{ di bulatkan menjadi } 39\%$$

Setuju

$$10 \times 8 \times 3 = 240$$

$$76/240 \times 100 = 31.6\% \text{ di bulatkan menjadi } 32\%$$

Cukup Setuju

$$10 \times 8 \times 3 = 240$$

$$15/240 \times 100 = 6.25\% \text{ di bulatkan menjadi } 7\%$$

Kurang Setuju

$$10 \times 8 \times 3 = 240$$

$$0/240 \times 100 = 0\%$$

Total Keseluruhan 78%.

Dengan demikian tingkat keakurasian pada pengujian pada *black box testing* pada aplikasi adalah 78% yang artinya masih ada kegagalan pada sistem aplikasi yang harus diperbaiki

#### 4.6 Pengujian Sistem Dengan Pakar

Berdasarkan hasil pengujian yang dilakukan sebanyak 40 kali terdapat kesalahan sebesar 9 kali dan kesesuaian 31 kali, sehingga dapat disimpulkan bahwa tingkat keakuratan sistem sebanyak 77,5%

### Kesimpulan dan Saran

#### 1. Kesimpulan

Dari penelitian yang telah dilakukan, maka dapat diperoleh kesimpulan:



1. Dengan menggunakan aplikasi ini masyarakat dapat mengimplmentasikan data berupa gambar, sehingga bisa dimasukkan kedalam aplikasi *image processing* dengan menggunakan pendekatan metode *k-means clustering* agar masyarakat dapat mengetahui jenis beras.
2. Dari hasil uji coba dengan pakar maka dapat disimpulkan tingkat keakurasian dengan menggunakan metode *k-means clustering* mencapai 77.5% sehingga aplikasi ini dapat digunakan.

## 2. Saran

Saran untuk penelitian pengembangan selanjutnya:

1. Untuk mengembangkan aplikasi selanjutnya bisa digunakan untuk mengclustering semua citra digital bukan hanya yang berformat *bmp*.
2. Untuk meningkatkan hasil pengelompokkan citra digital yang lebih relavan sebaiknya algoritma *K-means* digabung dengan algoritma lain, seperti deteksi tekstur atau yang lain dalam pengolahan citra digital (*image procceng*).
3. Penambahan kamera yang langsung dihubungkan keprogram dan dapat mengambil gambar lalu diproses secara otomatis.

## Daftar Pustaka

- Abdul. 2012. *Pengujian Perangkat Lunak Dengan Menggunakan Metode White Box Dan Black Box*. Semarang
- Awaluddin, M dan B.D Yuwono. 2010. *Penajaman dan Segmentasi Citra Pada Pengolahan Citra Digital*. TEKNIK – Vol. 31 No. 1 Tahun 2010, ISSN 0852-1697, hal:63
- Desmanto S, Irwan, dan R. Angreni. 2015. *Penerapan Algoritma K-Means Clustering Untuk Pengelompokkan Citra Digital Dengan Ekstraksi Fitur Warna RGB*. STMIK GI MDP. Palembang.
- Nuryani. *Potensi substitusi beras putih dengan beras merah sebagai makanan pokok untuk perlindungan diabetes militus*. Makasar.