

Penentuan Jenis Tomat Menggunakan Ekstraksi Ciri Bentuk dan Ukuran dengan Metode K-Means

Devi Tiana Kartikasari ^a, Retno Wahyusari ^{b*}

^{ab}Sekolah Tinggi Teknologi Ronggolawe

*email: retnowahyusari@gmail.com

Intisari

Tentang naskah:

-diterima, 11 Jun 2022
-direview, 15 Jun 2022
-diterbitkan, 1 Jul 2022

Kata kunci:

Ekstraksi, K-means,
RGB, HSV, Segmentasi

Tanaman tomat dikembangkan dengan sangat pesat sehingga melahirkan jenis tomat baru. Mengakibatkan petani petani sulit membedakan jenis tomat satu dengan lainnya. Penentuan jenis tomat dapat dilihat berdasarkan ukuran, bentuk, warna dan keadaan kulit buahnya. Salah satu cara membedakan tomat satu dengan yang lainnya adalah dengan melihat ciri bentuk dan ukuran. Ekstraksi ciri digunakan agar dapat menjadi acuan pembeda pada jenis tomat. Tomat terlebih dahulu diubah dari citra asli atau RGB menjadi citra HSV kemudian mengambil kanal (S) dari HSV untuk segmentasi agar mendapatkan nilai ekstraksi ciri bentuk dan ukuran, pengelompokan jenis tomat sesuai dengan jenisnya masing-masing adalah dengan menggunakan metode K-means. Pemilihan metode K-Means sebab selain populer, juga merupakan metode yang sederhana dan efektif. Memastikan bahwa proses yang dilakukan mendapatkan hasil yang akurat maka perhitungan atau proses penentuan jenis tomat dapat di tambahkan tool bantuan seperti Rapid Miner dan Matlab. Ekstraksi ciri bentuk dan ukuran dengan metode K-means dinilai mampu membedakan antara jenis tomat satu dengan lainnya dan dapat dikelompokkan sesuai jenisnya. Jumlah data set 100 tomat dengan 5 jenis tomat terdapat 11 titik kesalahan dalam penentuan dengan metode K-means sehingga akurasi yang didapatkan adalah sebesar 89%.

Abstract

Tomato plants are developed very rapidly, giving birth to new types of tomatoes. As a result, it is difficult for farmers to distinguish different types of tomatoes from one another. Determination of the type of tomato can be seen based on the size, shape, color and state of the skin of the fruit. One way to distinguish tomatoes from one another is to look at the characteristics of shape and size. Feature extraction is used so that it can be used as a differentiating reference for the type of tomato. Tomatoes are first converted from the original image or RGB to an HSV image then take the channel (S) from HSV for segmentation in order to get the value of extracting shape and size characteristics, grouping the types of tomatoes according to their respective types is by using the K-means method. The choice of the K-Means method is because besides being popular, it is also a simple and effective method. Ensuring that the process carried out gets accurate results, the calculation or process of determining the type of tomato can be added to help tools such as Rapid Miner and Matlab. Extraction of shape and size features with the K-means method was considered capable of distinguishing between types of tomatoes and can be grouped according to type. The number of data sets of 100 tomatoes with 5 types of tomatoes contained 11 points of error in the determination by the K-means method so that the accuracy obtained was 89%.

Keyword:

Extraction, K-Means,
RGB, HSV, Segmentation

1. Pendahuluan

Tanaman tomat dikembangkan dengan sangat pesat bahkan dapat melahirkan jenis-jenis tomat baru. Tomat memiliki 10.000 jenis variasi (Aulia, 2019). Banyaknya variasi tomat menyulitkan para petani dalam membedakan jenis tomat.

Membedakan buah dapat dilakukan berdasarkan bentuk, ukuran, warna, dan tekstur (Raja Sekar L, Ambika N & T, 2018). Menurut Fereydoun Keshavarzpour dan Abdul Kabir Khan Achakzai

(Keshavarzpour & Achakzai, 2013) bentuk buah merupakan salah satu sifat fisik dan parameter yang paling penting dari semua produk pertanian.

Beberapa penelitian yang berkaitan dengan usulan metode yang dapat membantu dalam membedakan jenis tanaman berdasarkan ciri fisiknya adalah: Tahta Herdian Andika dan Nuari Sivi Anisa (Sivi Anisa & Herdian Andika, 2020) melakukan identifikasi citra daun menggunakan ekstraksi ciri bentuk dan

metode K-Means. Hasil yang diperoleh dari 5 kali percobaan terdapat satu kesalahan identifikasi.

Rahma Ferika shaumi (Shaumi, 2018) penyortiran tomat berdasarkan ukuran dilakukan, citra tomat yang digunakan di tangkap menggunakan webcam dan diolah dengan Raspberry pi 3 pada proses sortir digunakan deteksi tepi canny deteksi ini digunakan karena mempunyai kemampuan dalam mendeteksi tepi sebuah objek. Penyortiran tomat dibedakan berdasarkan ukuran yang sudah ditentukan, menurut penelitian metode canny adalah metode yang dapat di aplikasikan untuk proses penyortiran buah tomat berbasis Raspberry pi 3. Tingkat kesalahan mencapai 16% hal itu di pengaruhi oleh hasil citra yang blur sehingga hasil deteksi tepi kurang maksimal, serta hasil perhitungan jarak dari euclidean distance.

Sarlita Hartiningtyas, dkk (Hartiningtyas et al., 2018) melakukan klasifikasi jenis mangga berdasarkan bentuk dan metode K-Nearest Neighbor. Hasil penelitian menghasilkan akurasi sebesar 88% dari 25 citra uji mangga.

Metode yang digunakan dalam penelitian adalah metode K-Means. Metode K-Means adalah metode yang populer dalam analisa data dengan tujuan yang penting pada statistik (Deshpande & Mandi, 2020). Selain itu metode K-Means merupakan metode clustering yang sederhana dan efektif (Huang et al., 2021). Berdasarkan latar belakang diatas maka penelitian bertujuan untuk menentukan jenis tomat menggunakan ekstraksi ciri bentuk dan ukuran serta metode K-means dalam mengklaster jenis tomat.

2. Kerangka Teori

2.1 Segmentasi

Segmentasi warna adalah proses pemisahan segmen dalam suatu citra berdasarkan warna, dalam beberapa perkembangan telah di lakukan berbagai macam metode untuk melakukan segmentasi warna seperti metode clustering dan indeks. Segmentasi ruang warna HSV memiliki 3 ruang lingkup atau 3 kanal yaitu *Saturation*, *Hue*, *Value*. Proses segmentasi warna HSV

menggunakan RGB sebagai acuan warna dasar dan mengkonversi menjadi ruang warna HSV. Kanal (S) atau Saturation yang mewakili tingkat dominasi warna di segmentasi untuk menghasilkan citra yang di inginkan.

2.2 Konversi Citra RGB ke HSV

Ruang warna HSV diperkenalkan oleh A.R Smith pada tahun 1978. Adapun perhitungan untuk konversi RGB ke HSV adalah sebagai berikut (Acharya & Ray, 2005):

$$H = \tan \left[\frac{3(G-B)}{(R-G)+(R-G)} \right] \quad (1)$$

$$S = 1 - \frac{\min(R,G,B)}{V} \quad (2)$$

$$V = \frac{R+G+B}{3} \quad (3)$$

Akan tetapi ketika $S = 0$ maka hue tidak dapat di tentukan maka perlu adanya normalisasi nilai RGB terlebih dahulu.

$$r = \frac{R}{R+G+B} \quad (4)$$

$$g = \frac{G}{R+G+B} \quad (5)$$

$$b = \frac{B}{R+G+B} \quad (6)$$

Setelah dilakukan normalisasi R,G, dan B maka rumus konversi RGB ke HSV sebagai berikut:

$$V = \max(r, g, b) \quad (7)$$

$$S = \begin{cases} 0 & \text{Jika } V = 0 \\ V - \frac{\min(r, g, b)}{v} & \text{Jika } V > 0 \end{cases} \quad (8)$$

$$H = \begin{cases} 0, & \text{jika } S = 0 \\ \frac{60 * (g - b)}{S * V}, & \text{jika } V = r \\ 60 * \left[2 + \frac{b - r}{S * V} \right], & \text{jika } V = g \\ 60 * \left[4 + \frac{r - g}{S * V} \right], & \text{jika } V = b \end{cases} \quad (9)$$

2.3 Ekstraksi Ciri

Ekstraksi ciri proses mengambil ciri-ciri yang terdapat pada objek di dalam citra untuk mengenali objek tersebut. Biasanya

ciri-ciri umum yang yang dapat mengenali satu atau lebih objek di dalam citra adalah ukuran, posisi atau lokasi, dan orientasi atau sudut kemiringan objek terhadap garis acuan yang digunakan.

Ciri bentuk (*Eccentricity*) diperoleh berdasarkan nilai perbandingan antara jarak foci elips minor dengan foci elips major suatu benda atau objek . Sehingga dapat diketahui seberapa dekat benda tersebut dengan bentuk lingkaran. *Eccentricity* dirumuskan dengan rumus.

$$e = \sqrt{1 - \frac{b^2}{a^2}} \quad (10)$$

Keterangan :

e = eccentricity

b = minor axis

a = mayor axis

Ciri ukuran biasa juga di sebut *Parimeter*, *Parimeter* adalah perbandingan antara luas (Area) dan keliling benda. Benda yang metriknya mendekati 0 maka bentuk benda mendekati bentuk garis yang panjang, sedangkan benda yang metriknya mendekati angka 1 maka bentuk benda mendekati bentuk bulat. Perhitungan metrik diilustrasikan dalam rumus:

$$M = \frac{4\pi A}{P^2} \quad (11)$$

Keterangan :

M = *Roundness*

A = Area

P = *Parimeter*

2.4 K-Means Clustering

Metode K-means sangat terkenal karena kemudahan dan kemampuannya untuk mengcluster data yang besar dan data *outlier* dengan sangat cepat. Dalam metode K-Means, setiap data harus termasuk ke *cluster* tertentu dan bisa dimungkinkan bagi setiap data yang termasuk *cluster* tertentu pada suatu tahapan proses, pada tahapan berikutnya berpindah ke *cluster* lainnya. Metode K-Means dihitung dengan menentukan K sebagai jumlah *cluster* yang ingin dibentuk, bangkitkan K *centroid* awal secara random, hitung jarak setiap data pada masing-masing *centroid* dengan rumus:

$$d(x_j, c_j) = \sqrt{\sum_{j=1}^n (x_j - c_j)^2} \quad (12)$$

Dimana:

d = jarak

j = banyaknya data

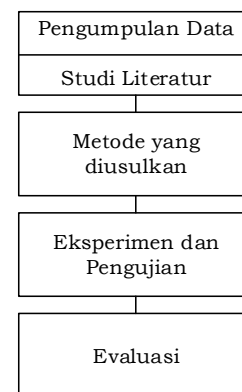
c = *centroid*

x = data

Setiap data memilih *centroid* yang terdekat, kelompokan data sesuai dengan *clusternya*, untuk tahap selanjutnya hitung jarak pada setiap baris data, kembali ke posisi *centroid* baru denan *centoroid* yang lama tidak sama, jika tidak terjadi perubahan maka peoses *clustering* selesai.

3. Metodologi

Alur metode penelitian dari pengumpulan data hingga mengetahui nilai akurasi dapat digambarkan sebagai berikut:



Gambar 1 Alur Metode Penelitian

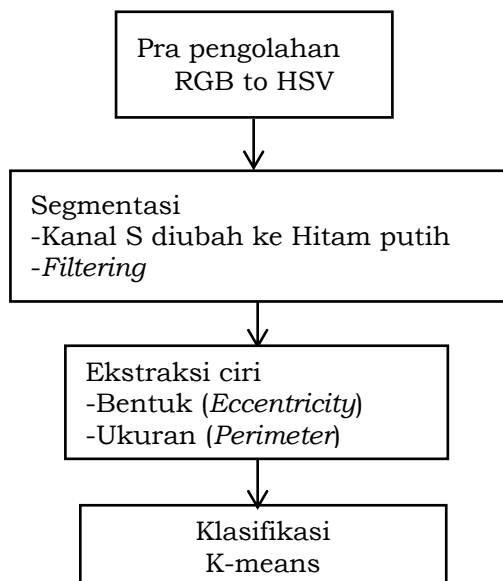
Pengumpulan data dengan mengumpulkan materi yang berhubungan dengan metode yang digunakan dalam penentuan jenis tomat, serta mengumpulkan citra tomat. Beberapa contoh citra yang digunakan terlihat pada Gambar 2.



Gambar 2 Citra Uji

Metode yang digunakan terlihat pada alur metode Gambar 3, dimana diawali dengan merubah citra kedalam ruang

warna HSV, segmentasi dilanjutkan dengan melakukan *clustering*.



Gambar 3 Metode yang Diusulkan

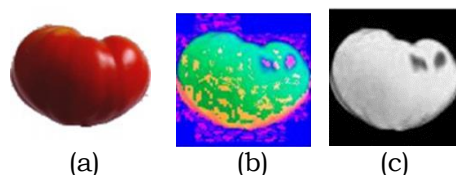
Eksperimen dan pengujian dengan membandingkan hasil *cluster* secara manual/hasil perhitungan, aplikasi Rapid Miner dan Matlab. Hasil cluster dari ketiga cara memberikan data yang dijadikan bahan evaluasi.

4. Hasil dan Pembahasan

4.1 Hasil

a Konversi Citra RGB to HSV

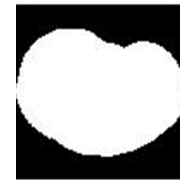
Citra asli pada ruang RGB dirubah keddalam citra ruang HSV. Hasil konversi nantinya digunakan untuk proses segmentasi. Gambar 4 merupakan hasil konversi pada ruang RGB ke HSV.



Gambar 4 (a) citra asli atau RGB, (b) Citra HSV, (c) Citra kanal Saturation dari warna HSV

b Segmentasi Warna

Proses segmentasi dilakukan pada kanal S (Saturation), sebab pada kanal ini mewakili tingkat dominasi warna. Hasil segmentasi terlihat pada Gambar 5.



Gambar 5 Citra Segmentasi

c Ekstraksi Ciri

Ekstraksi ciri yang digunakan adalah ekstraksi bentuk dan ekstraksi ukuran. Pada ekstraksi bentuk terdapat nilai *eccentricity*, dimana nilai *eccentricity* merupakan nilai bentuk untuk menentukan bagaimana objek yang terbentuk tidak mendekati lingkaran atau mendekati lingkaran. Ekstraksi ciri ukuran terdapat nilai area dan parimeter. Area atau keliling adalah piksel yang mengelilingi objek, sedangkan parimeter atau luas adalah piksel yang menyusun objek. Beberapa nilai hasil pengambilan ekstraksi ciri bentuk dan ukuran terlihat pada Tabel 1.

Tabel 1 Nilai Ekstraksi Ciri Bentuk dan Ukuran

NO	Area	Parimeter	Eccentricity
1	7323	319,3208512	0,324485176
2	7298	320,7350647	0,302910323
3	7411	322,5929291	0,365333015
4	7151	329,9066376	0,16537504
5	7239	324,1492783	0,210726614
6	7323	319,3208512	0,324485176
7	7337	319,1787156	0,392723166
8	7036	315,1787156	0,474387243
9	7276	318,1492783	0,336796288
10	7153	316,8355698	0,299497806

d K-Means Clustering

Hasil ekstraksi ciri digunakan acuan dalam proses *cluster* menggunakan metode K-Means. Pada proses clutering terjadi 3 kali perulangan pada proses perhitungan manual. Hasil akhir proses *clustering* terlihat pada Tabel 2.

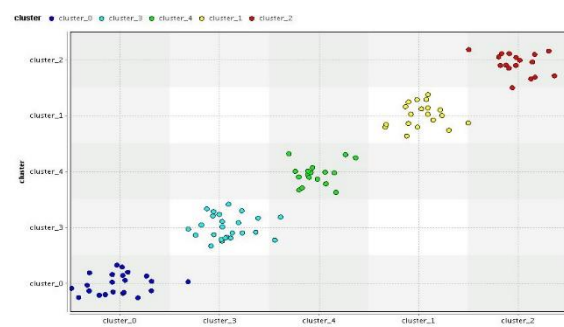
Tabel 2 Hasil Clustering

No	dc1			dc5	Jarak	C1	...	C4	C5
1	7432,41			5174165	7432,41	V			
2	3747,8			5060879	3747,8	V			
3	30357,7			5581031	30357,7	V			
4	7463,97			4419946	7463,97	V			

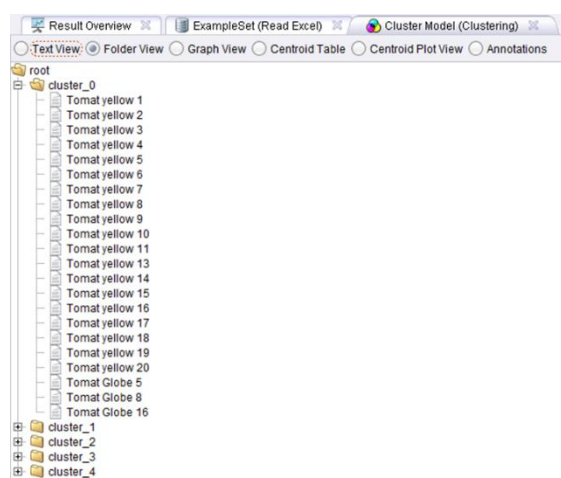
5	24,7078			4798479	24,7077	V	
6	7432,41			5174165	7432,41	V	
7	10042,5			5238006	10042,5	V	
8	40337,1			3953431	40337,1	V	
9	1539,91			4963084	1539,91	V	
10	7028,96			4431195	7028,96	V	
11	3604,73			4533265	3604,73	V	
12	292567			2720781	25960,5	V	
13	8609,96			4392743	8609,96	V	
14	24,7078			4798479	24,7077	V	
15	30357,7			5581031	30357,7	V	

Pada perulangan ke 3, tidak terdapat perubahan anggota *clustering* dan tidak terdapat perubahan nilai pusat cluster. Hal ini menunjukkan bahwa proses *clustering* telah selesai.

Proses *clustering* menggunakan aplikasi Rapid Miner, menghasilkan *cluster* seperti terlihat pada Gambar 6. Gambar 7 memberikan informasi keanggotaan tomat berdasarkan jenisnya.

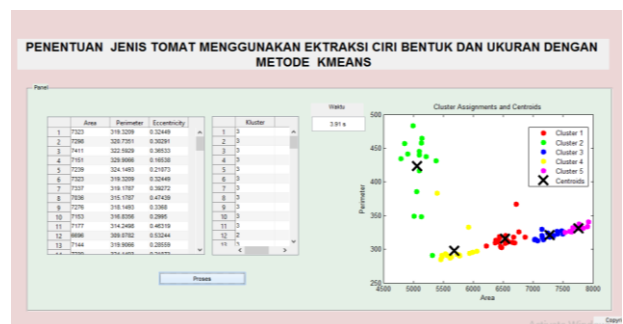


Gambar 6 Grafik Hasil *Clustering*



Gambar 7 Keanggotaan Tomat Berdasarkan *Cluster*

Proses *cluster* menggunakan GUI Matlab, memberikan hasil *cluster*, waktu pemrosesan dan grafik *clustering* jenis tomat. Tampilan GUI Matlab terlihat pada Gambar 8 sebagai berikut:



Gambar 8 Tampilan GUI Matlab

4.2 Pembahasan

Hasil dari penelitian yang sudah di uji bahwa ciri bentuk dan ukuran mampu menjadi acuan untuk membedakan tomat dengan baik dengan metode K-means serta penggunaan metode ini memiliki waktu *clustering* yang cepat seperti pernyataan Yohanes (Yohannes, 2016). Hasil penelitian waktu proses pengelompokan dengan K-means mendapatkan rata – rata 3,5s dari 5 kali percobaan. Pada perbandingan perhitungan manual metode K-means, aplikasi Rapid Miner dan Matlab didapatkan akurasi sebanyak 89%, dengan 11 titik kesalahan pada penentuan jenis Tomat.

5. Simpulan

Berdasarkan penelitian yang sudah dilakukan bahwa penentuan dengan menggunakan ekstraksi ciri bentuk dan ukuran mampu membedakan antara jenis tomat. Metode K-means dapat menentukan jenis tomat berdasarkan ciri bentuk dan ukuran dengan baik, hasil *clustering* dari matlab dan Rapid Miner menunjukan hasil yang sama yaitu 11 titik kesalahan dengan akurasi 89%. Saran penelitian selanjutnya menambahkan jumlah dataset dan jumlah jenis tomat, menggunakan ekstraksi ciri baru seperti ciri warna dan tekstur serta membandingkan hasil *clustering* menggunakan algoritma lainnya.

Daftar Pustaka

- Acharya, T., & Ray, A. K. (2005). *Image Processing Principles and Applications*. John Wiley & Sons.
- Aulia, A. (2019). *Tomat Adalah Buah Paling Populer Di Dunia*. <https://mediakom.kemkes.go.id>
- Darmi, Y., & Setiawan, A. (2016). Penerapan Metode Clustering K-Means Dalam Pengelompokan

- Penjualan Produk. *Media Infotama*, 12(2), 148–157.
- Deshpande, A., & Mandi, I. I. T. (2020). Robust k - means ++. *Proceedings of the 36th Conference on Uncertainty in Artificial Intelligence (UAI)*, 124.
- Hartiningtyas, S., Ruslianto, I., & Hidayati, R. (2018). Klasifikasi Jenis Mangga Berdasarkan Fitur Bentuk Dan Warna Daun Dengan Menggunakan Metode K-Nearest Neighbor Berbasis Android. *Jurnal Coding, Sistem Komputer Untan*, 6(1), 12–23.
- Huang, S., Kang, Z., Xu, Z., & Liu, Q. (2021). Robust deep k-means: An effective and simple method for data clustering. *Pattern Recognition*, 117, 6.
- Keshavarzpour, F., & Achakzai, A. K. K. (2013). Fruit Shape Classification in Cantaloupe Using the Analysis of Geometrical Attributes. *World Engineering & Applied Sciences Journal*, 4(1). <https://doi.org/10.5829/idosi.weasj.2013.4.1.1109>
- Raja Sekar L, Ambika N, D. V. and, & T, K. (2018). Fruit Classification System Using Computer Vision : A Review. *International Journal of Trend in Research and Development*, 5(1), 22–26.
- Shaumi, R. F. (2018). *Rancang Bangun Alat Penyortir Buah Tomat Berdasarkan Ukuran Berbasis Raspberry PI 3. 3.*
- Sivi Anisa, N., & Herdian Andika, T. (2020). Sistem Identifikasi Citra Daun Berbasis Segmentasi Dengan Menggunakan Metode K-Means Clustering. *Aisyah Journal Of Informatics and Electrical Engineering (A.J.I.E.E)*, 2(1), 9–17. <https://doi.org/10.30604/jti.v2i1.22>
- Yohannes. (2016). Analisis Perbandingan Algoritma Fuzzy C-Means. *Annual Research Seminar 2016*, 2(1), 151–155.