

# Implementasi *K-Means Clustering* pada Citra Digital Tomat untuk Identifikasi Kondisi Segar dan Busuk

Nasrul Mahruf Aznawi <sup>1,\*</sup>, Muhammad Irham Setiadi <sup>1</sup>, Zarifah Aina <sup>1</sup>, Setti Manullang <sup>1</sup>,  
Shafira Nur Rahmadiyah <sup>1</sup>

\* Korespondensi: e-mail: [nasrumahrufaznawi@gmail.com](mailto:nasrumahrufaznawi@gmail.com)

<sup>1</sup> Ilmu Komputer; Fakultas Sains dan Teknologi; Universitas Islam Negeri Sumatera Utara; Jl. Lap. Golf No.120, Kp. Tengah, Kec. Pancur Batu, Kabupaten Deli Serdang, Sumatera Utara 20353, Telp/Fax : +62-61-6615683; e-mail: [nasrumahrufaznawi@gmail.com](mailto:nasrumahrufaznawi@gmail.com), [muhhammadirhamsetiadi92@gmail.com](mailto:muhhammadirhamsetiadi92@gmail.com), [zarifahaina1@gmail.com](mailto:zarifahaina1@gmail.com), [setimanulang40@gmail.com](mailto:setimanulang40@gmail.com), [shafirarahmadiyah@gmail.com](mailto:shafirarahmadiyah@gmail.com)

Submitted : 17 Maret 2025  
Revised : 7 April 2025  
Accepted : 25 April 2025  
Published : 30 Mei 2025

## Abstract

*The manual identification of fresh and rotten tomatoes has still relied on human visual observation, which tends to be inconsistent and time-consuming. This study aimed to develop a tomato image classification system using the K-Means Clustering method based on color, shape, and texture features to automatically identify fresh and rotten conditions. The Dataset consisted of 500 tomato images for training and 60 tomato images for testing, equally representing fresh and rotten conditions. The process involved converting the images into L\*a\*b and grayscale formats, performing segmentation using K-Means, and extracting shape and texture features for the classification process. The testing results showed that the system successfully classified fresh and rotten tomatoes with an accuracy rate of 95%, with both precision and recall exceeding 93% for each class. These findings indicated that the K-Means method could be effectively applied in tomato image processing to support the sorting process of agricultural products. This research contributed to the development of a digital image-based classification system that could be integrated into smart agriculture systems.*

**Keywords:** Digital Image Processing, K-Means Clustering, Shape, Texture, Tomato Classification

## Abstrak

Proses identifikasi kondisi tomat segar dan busuk secara manual masih bergantung pada pengamatan visual manusia yang tidak konsisten dan memerlukan waktu. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi citra tomat menggunakan metode *K-Means Clustering* berdasarkan fitur warna, bentuk dan tekstur guna mengidentifikasi kondisi segar dan busuk secara otomatis. *Dataset* yang digunakan terdiri dari 500 citra tomat untuk pelatihan dan 60 citra tomat untuk pengujian, yang masing-masing mewakili kondisi segar dan busuk. Proses melibatkan konversi citra ke format L\*a\*b dan *grayscale*, segmentasi menggunakan *K-Means*, serta ekstraksi fitur bentuk dan tekstur untuk proses klasifikasi. Hasil pengujian menunjukkan bahwa sistem mampu mengklasifikasikan tomat segar dan busuk dengan tingkat akurasi sebesar 95%, dengan presisi dan *recall* di atas 93% untuk kedua kelas. Temuan ini menunjukkan bahwa metode *K-Means* dapat digunakan secara efektif dalam pengolahan citra tomat untuk mendukung proses sortasi produk pertanian. Penelitian ini memberikan kontribusi pada pengembangan sistem klasifikasi berbasis citra digital yang dapat diintegrasikan dalam sistem pertanian cerdas.

**Kata kunci:** Bentuk, Klasifikasi Tomat, *K-Means Clustering*, Pengolahan Citra Digital, Tekstur

## 1. Pendahuluan

Tomat atau dalam nama latin yaitu *Solanum lycopersicum* merupakan salah satu komoditas pertanian yang memiliki nilai ekonomi tinggi serta tingkat konsumsi yang cukup besar di kalangan masyarakat. Tomat merupakan salah satu komoditas hortikultura yang merupakan komoditi pertanian yang mempunyai potensi serta peluang untuk menjadi tanaman yang mampu meningkatkan kesejahteraan petani (Pitaloka, 2020). Namun demikian, kualitas fisik tomat sangat rentan mengalami penurunan, terutama akibat proses fisiologis alami serta pengaruh lingkungan selama tahap pascapanen. Buah tomat memiliki sifat *perishable* yaitu mudah rusak (Baladraf, 2024). Oleh sebab itu, penanganan pascapanen buah tomat adalah kebutuhan akan metode yang andal dan efisien dalam mengidentifikasi kondisi buah, khususnya dalam membedakan antara tomat yang masih segar dan yang telah mengalami pembusukan.

Perkembangan teknologi informasi dan pengolahan citra digital telah membuka peluang besar dalam penerapan metode otomatis untuk menilai kualitas buah berbasis fitur visual. Citra digital memungkinkan dilakukannya ekstraksi terhadap fitur-fitur penting seperti warna, bentuk, serta tekstur permukaan buah, yang dapat dijadikan indikator representatif terhadap kondisi fisiologis tomat. Misalnya, penelitian oleh (Denih et al., 2023) menunjukkan bahwa penggunaan ruang warna CIELAB dan metode *K-Means Clustering* mampu mengklasifikasikan tingkat kematangan tomat dengan akurasi rata-rata sebesar 92,70%. Meski demikian, tantangan utama yang masih dihadapi dalam penelitian ini adalah bagaimana melakukan pengelompokan data citra secara efektif guna membedakan tomat segar dari tomat busuk secara otomatis dan akurat. Seperti penelitian yang dilakukan oleh (Saleh et al., 2023) dalam mengimplementasikan *Machine Learning* dan *Fuzzy C-Means Clustering* untuk Identifikasi Penyakit Tomat, namun hasil akurasi pada penelitian tersebut berada pada angka 88,87%, yang menunjukkan perlu adanya peningkatan dalam metode pengelompokan.

Salah satu pendekatan yang telah banyak diterapkan dalam proses pengelompokan data adalah algoritma *K-Means Clustering*. Algoritma ini termasuk ke dalam metode *unsupervised learning*, yang bekerja dengan mengclusterkan data berdasarkan kemiripan fitur tanpa memerlukan data berlabel. Algoritma ini mempartisi data ke dalam *cluster* sehingga data yang memiliki karakteristik yang sama dikelompokkan ke dalam satu *cluster* yang sama dan data yang mempunyai karakteristik yang berbeda di kelompokkan ke dalam *cluster* yang lain (Utomo et al., 2024). Dalam konteks citra digital tomat, algoritma *K-Means* dapat dimanfaatkan untuk mengelompokkan piksel atau objek berdasarkan karakteristik warna, bentuk dan tekstur, sehingga memungkinkan identifikasi area yang menunjukkan indikasi pembusukan secara otomatis. Penelitian lain yang dilakukan (Rizki et al., 2024) mengaplikasikan *K-Means Clustering* untuk mengidentifikasi buah stroberi yang matang berdasarkan bentuk dan warna, menghasilkan akurasi sebesar 100%.

Penelitian ini bertujuan untuk menerapkan algoritma *K-Means Clustering* pada citra digital tomat guna mengidentifikasi kondisi kesegaran buah dengan memanfaatkan fitur warna,

bentuk dan tekstur sebagai dasar pengelompokan. Permasalahan ini penting untuk diselesaikan agar proses identifikasi kualitas tomat dapat dilakukan secara otomatis dan akurat, terutama dalam membedakan tomat segar dan busuk. Melalui tahapan utama yang meliputi *preprocessing* citra, ekstraksi fitur visual, serta pengelompokan dengan algoritma *K-Means*, diharapkan sistem yang dikembangkan mampu mengklasifikasikan tomat segar dan busuk dengan akurasi yang reliabel dan dapat diterapkan sebagai alat bantu dalam proses sortasi buah tomat. Hal ini tidak hanya mendukung peningkatan akurasi klasifikasi buah visual berbasis teknologi, tetapi juga menjadi solusi praktis dalam mendukung penanganan pasca panen komoditas hortikultura seperti tomat.

## **2. Metode Penelitian**

Penelitian ini berfokus pada pengolahan citra digital untuk mengelompokkan kondisi buah tomat berdasarkan warna dan tekstur. Algoritma *K-Means Clustering* diimplementasikan sebagai metode utama dalam melakukan klasifikasi terhadap citra buah tomat kedalam dua *cluster*, yakni segar dan busuk.

### **2.1. Alat dan Bahan Penelitian**

#### **MATLAB**

MATLAB merupakan sebuah *platform* pemrograman dan komputasi numerik yang berguna untuk analisis data, pemodelan, simulasi, visualisasi dan pengembangan algoritma. MATLAB digunakan dalam penelitian ini karena pada MATLAB tersedia fungsi-fungsi atau perintah/*coding* yang berkaitan dengan pengolahan citra. Selain perintah/*coding* MATLAB juga menyediakan *form* yang lebih representatif untuk membuat tampilan citra (Ridhoi, 2023).

#### **Dataset**

*Dataset* adalah kumpulan data terstruktur yang digunakan untuk melatih, menguji, dan memvalidasi model dalam berbagai bidang, termasuk pengolahan citra (Hermawan, 2024). *Dataset* yang digunakan dalam penelitian ini merupakan kumpulan citra digital tomat yang diperoleh dari website Roboflow Universe dengan nama *Dataset Tomato Detection Fresh or Rotten*. *Dataset* terdiri dari dua kelas yakni tomat segar dan tomat busuk. Setiap citra memiliki ukuran atau resolusi yang seragam dan tersimpan dalam format RGB. *Dataset* terdiri dari 500 citra buah tomat untuk data *training* yang terbagi menjadi 250 tomat segar dan 250 tomat busuk, dan 60 citra buah tomat berbeda terbagi menjadi 30 tomat segar dan 30 tomat busuk untuk data *testing*.

### **2.2. Tahapan Penelitian**

Tahapan penelitian ini dilakukan secara sistematis sesuai alur kerja sistem yang telah dirancang. Berikut adalah tahapan-tahapan utama dalam proses identifikasi kondisi buah tomat menggunakan metode *K-Means Clustering*.



Sumber: Hasil Penelitian (2025)

Gambar 1. Tahapan Penelitian

Pada tahap awal dalam penelitian ini adalah menginput citra RGB tomat, pada tahap ini menyediakan data visual yang akan diproses lebih lanjut. Kemudian citra RGB dikonversi ke ruang warna  $L^*a^*b$ , ruang warna  $L^*a^*b$  dipilih karena baik dalam memisahkan informasi warna dan pencahayaan dari citra. Selanjutnya, dilakukan segmentasi warna pada komponen  $a^*$  dan  $b^*$  menggunakan algoritma *K-Means Clustering*. Tujuannya adalah untuk memisahkan objek tomat dari latar belakang berdasarkan kedekatan warna ( $a^*$  dan  $b^*$ ). Hasil dari proses ini adalah citra tersegmentasi hanya berisi area objek dari buah tomat.

Tahap selanjutnya, citra hasil segmentasi dikonversi ke format *grayscale* untuk menyederhanakan informasi warna, lalu menjadi citra biner (hitam-putih) untuk keperluan ekstraksi ciri bentuk. Proses konversi ini bertujuan agar mengurangi kompleksitas dari visual dan meningkatkan akurasi ketika perhitungan fitur bentuk. Kemudian dilakukan ekstraksi ciri menjadi dua kategori yaitu fitur bentuk dan fitur tekstur. Pada fitur bentuk terdapat *Metric* (keliling/luas) dan *Eccentricity* (kesimetrian bentuk), yang digunakan untuk membedakan tomat busuk yang cenderung berbentuk tidak simetris. Lalu penggunaan fitur tekstur dihitung dengan menggunakan metode *Gray Level Co-occurrence Matrix* (GLCM) dengan parameter *Contrast*, *Correlation*, *Energy* dan *Homogeneity*, yang merepresentasikan Tingkat keseragaman dan ketidakteraturan permukaan dari objek. GLCM juga cukup populer digunakan dalam fitur tekstur, seperti (Karim, M & Karim, A 2021) dalam penelitiannya mereka menggunakan fitur tekstur GLCM dalam mengklasifikasikan buah.

Kemudian dari proses ekstraksi ciri, fitur-fitur yang diperoleh kemudian digunakan sebagai dasar untuk mengklasifikasikan kondisi buah tomat ke dalam dua *cluster* segar atau busuk. Proses klasifikasi dilakukan dengan membandingkan nilai dari fitur terhadap ambang batas yang telah ditentukan berdasarkan hasil pengujian *Dataset*. Lalu fitur reset disediakan agar dapat mengosongkan tampilan citra yang telah diproses dan mengulangi proses dengan citra yang berbeda.

### ***K-Means Clustering***

*K-Means Clustering* merupakan algoritma *unsupervised learning* yang sangat populer dalam proses pengelompokan data (*clustering*) berdasarkan kemiripan fitur tertentu, seperti warna, intensitas, atau tekstur. Proses kerja algoritma ini terdiri dari beberapa tahap, dimulai

dengan inisialisasi sejumlah *centroid* (pusat kluster) secara acak, kemudian dilanjutkan dengan pengelompokan data ke kluster terdekat menggunakan jarak *Euclidean*. Setelah itu, *centroid* diperbarui berdasarkan rata-rata dari seluruh anggota kluster, dan proses ini diulang hingga posisi *centroid* stabil atau jumlah iterasi maksimum tercapai. Dalam pengolahan citra digital, seperti pada penelitian ini, *K-Means* banyak digunakan untuk segmentasi objek dalam gambar agar dapat dipisahkan dari latar belakang dengan akurat (Septiarini et al., 2022).

### **Model K-Means**

Perancangan *Model K-Means* dilakukan dengan menggunakan fungsi *kmeans* yang telah tersedia di MATLAB. Dalam prosesnya dimulai dengan mengubah citra buah tomat dari warna asli (RGB) ke model warna  $L^*a^*b$ , model ini dipilih karena lebih baik dalam membedakan warna berdasarkan cara manusia dalam melihat warna. Setelah citra diubah, ambil bagian  $a^*$  dan  $b^*$  dari citra yaitu bagian yang mewakili perbedaan warna. Tentukan *clustering*  $k=2$  yang mana akan membagi citra menjadi 2 kelompok berdasarkan kemiripan warna. Setelah kelompok terbentuk, kita bisa mengevaluasi seberapa baik pengelompokan tersebut dengan cara menghitung seberapa rapat data di dalam setiap kelompok dan seberapa jelas pemisahan antar kelompok. Semua perhitungan dan visualisasi hasilnya dapat dilakukan dengan mudah di MATLAB (Yolanda et al., 2024).

### **3. Hasil dan Pembahasan**

Dalam prosesnya dilakukan pengujian *Dataset* training untuk buah tomat segar dan busuk mendapatkan hasil akurasi pelatihan yang disajikan dalam Tabel 1.

Tabel 1. Hasil *Training* Tomat Segar

Ciri	Nilai
<i>Metric</i>	0.62063
<i>Eccentricity</i>	0.54483
<i>Contrast</i>	0.032077
<i>Correlation</i>	0.98589
<i>Energy</i>	0.59577
<i>Homogeneity</i>	0.99531

Sumber: Hasil Penelitian (2025)

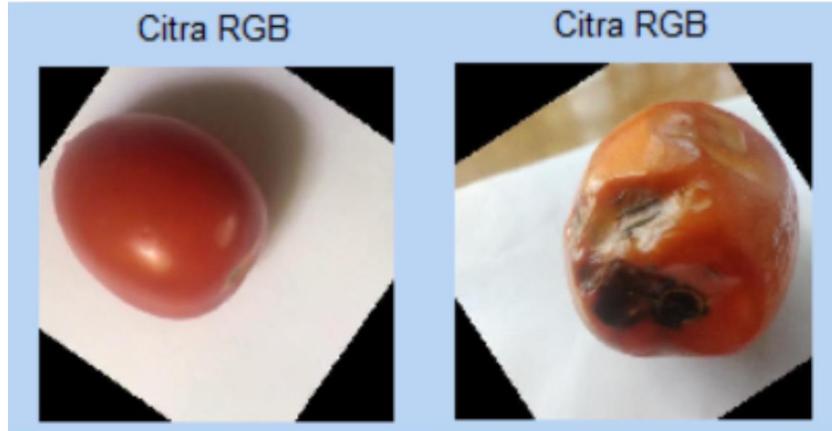
Tabel 2. Hasil *Training* Tomat Busuk

Ciri	Nilai
<i>Metric</i>	0.40868
<i>Eccentricity</i>	0.93563
<i>Contrast</i>	0.10661
<i>Correlation</i>	0.98048
<i>Energy</i>	0.37548
<i>Homogeneity</i>	0.98165

Sumber: Hasil Penelitian (2025)

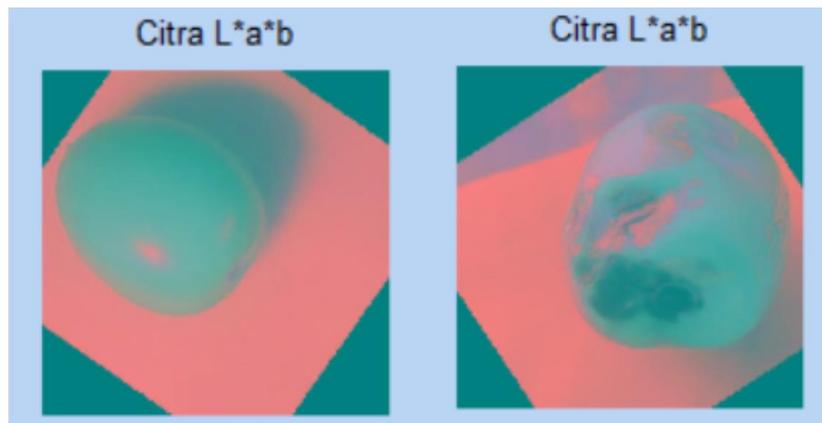
### 3.1. Hasil Segmentasi Citra

Setelah dilakukan konversi citra dari ruang warna RGB ke  $L^*a^*b$ , proses segmentasi dengan algoritma *K-Means Clustering* berhasil memisahkan objek buah tomat dari latar belakang pada seluruh *Dataset*. Gambar 2, 3, dan 4 merupakan hasil proses segmentasi.



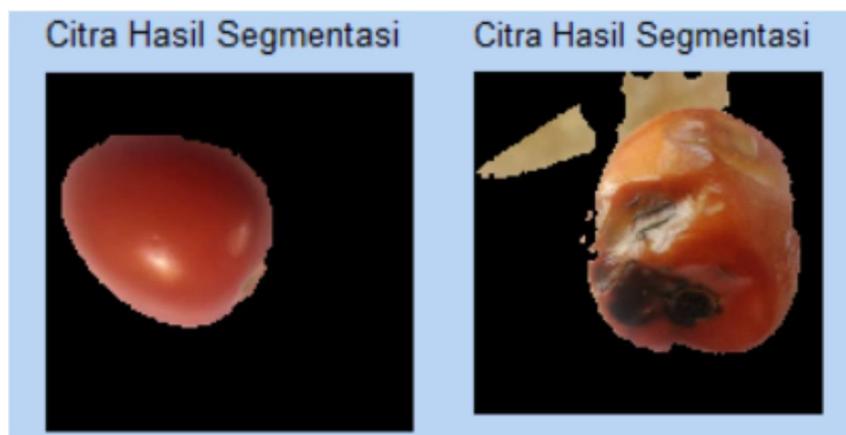
Sumber: Hasil Penelitian (2025)

Gambar 2. Citra RGB Tomat Segar dan Busuk



Sumber: Hasil Penelitian (2025)

Gambar 3. Citra  $L^*a^*b$  Tomat Segar dan Busuk

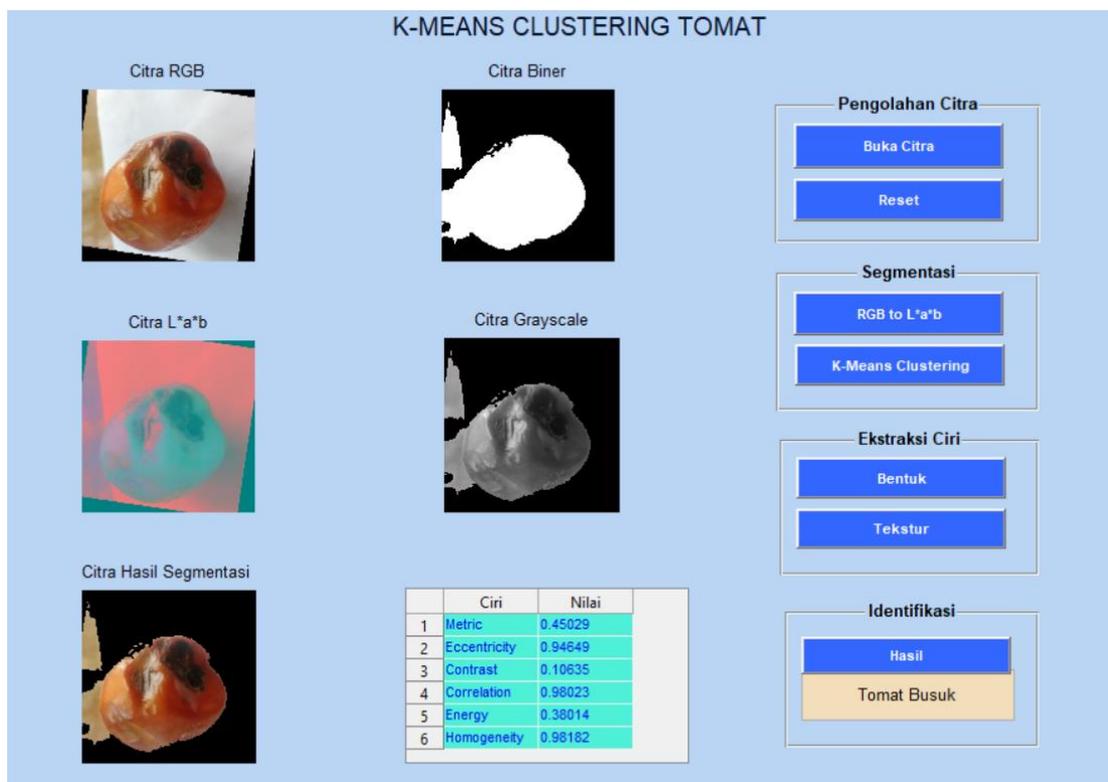


Sumber: Hasil Penelitian (2025)

Gambar 4. Citra Hasil Segmentasi Tomat Segar dan Busuk

### 3.2. Analisis Citra Tomat Busuk

Pada Gambar 5 tekstur dari citra tomat busuk memiliki *Contrast* 0.10635 yang mana nilai tersebut cukup tinggi, menunjukkan adanya variasi intensitas piksel akibat area membusuk. Kemudian *Homogeneity* 0.98182 cukup tinggi, namun lebih rendah dibandingkan dengan tomat segar, hal tersebut menunjukkan ketidakteraturan pada tekstur permukaan dari tomat. *Energy* dan *Correlation* juga memiliki nilai yang rendah dari tomat segar, mendukung indikasi permukaan tomat yang rusak akibat pembusukan. Bentuk dari buah tomat yang busuk memiliki *Metric* 0.45029 cukup rendah yang mengindikasikan bentuk buah tidak bulat seperti tomat pada umumnya. *Eccentricity* 0.94649 sangat tinggi, nilai tersebut menunjukkan bentuk yang tidak simetris, umum pada buah tomat yang membusuk.

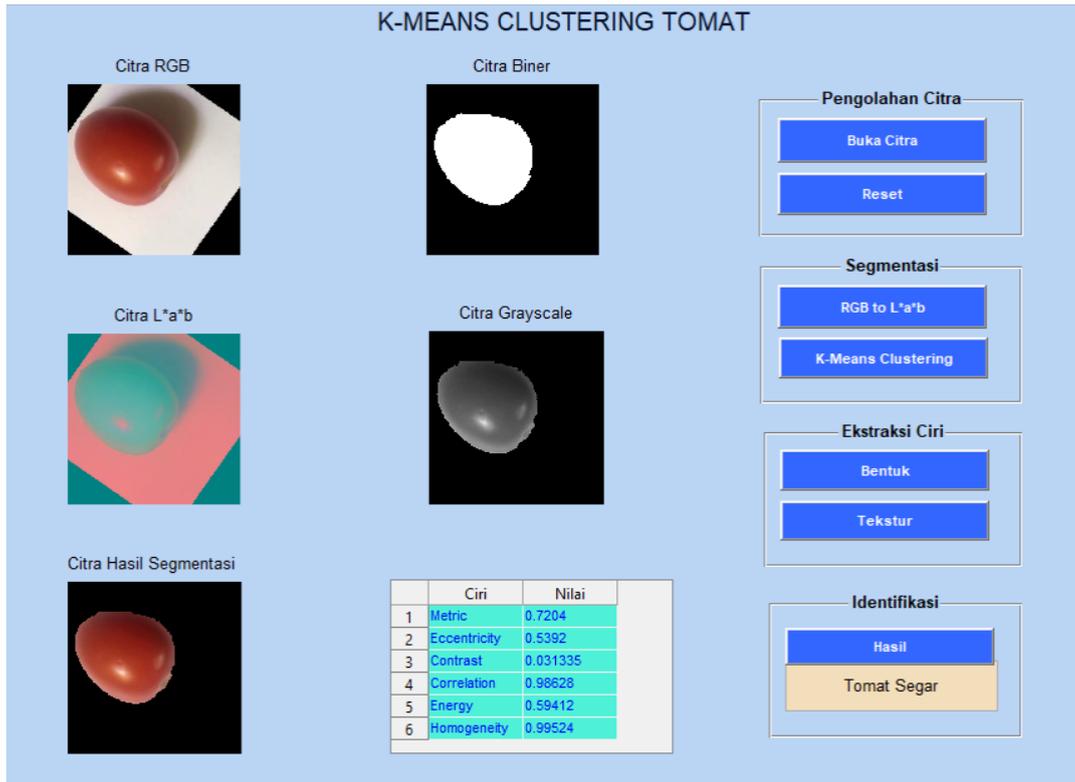


Sumber: Hasil Penelitian (2025)

Gambar 5. Hasil *K-Means Clustering* Tomat Busuk

### 3.3 Analisis Citra Tomat Segar

Citra pada Gambar 6 menunjukkan nilai tekstur yakni *Contrast* 0.031335, cukup rendah yang menandakan permukaan lebih rata dan seragam, menunjukkan kondisi buah tomat segar. Lalu *Homogeneity* 0.99524, sangat tinggi menunjukkan permukaan buah tomat yang halus dan tidak rusak akibat pembusukan. *Correlation* dan *Energy* juga memiliki nilai yang tinggi, Nilai tersebut mendukung bahwa citra tidak memiliki tekstur yang kasar akibat busuk. Citra buah tomat segar pada Gambar 6 memiliki bentuk dengan *Metric* 0.7204, nilai yang lebih tinggi dibanding dengan *Metric* pada citra tomat busuk, menandakan bentuk buah tomat yang mendekati bentuk bulat ideal. *Eccentricity* 0.5392, cukup rendah menunjukkan bentuk yang simetri dan juga bentuk bulat khas dari buah tomat yang segar.



Sumber: Hasil Penelitian (2025)

Gambar 6. Hasil *K-Means Clustering* Tomat Segar

### Confusion Matrix

Untuk menilai performa *K-Means* pada klasifikasi di penelitian ini, digunakan *confusion matrix* yang membandingkan hasil prediksi sistem dengan label sebenarnya (*ground truth*).

Tabel 3. *Confusion Matrix* Hasil Citra Tomat

	Prediksi Segar	Prediksi Busuk
Sebenarnya Segar (TP, FN)	28	2
Sebenarnya Busuk (FP, TN)	1	29

Sumber: Hasil Penelitian (2025)

*True Positive (TP)* : 28

*False Negative (FN)* : 2

*False Positive (FP)* : 1

*True Negative (TN)* : 29

### Akurasi (*Accuracy*)

Persentase total prediksi yang benar terhadap seluruh data.

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} = \frac{28 + 29}{28 + 29 + 1 + 2} = \frac{57}{60} = 0.95$$

### Presisi (*Precision*)

Seberapa tepat *K-Means* dalam memprediksi kelas tertentu.

$$\text{Presisi Segar} = \frac{TP}{TP + FP} = \frac{28}{28 + 1} = \frac{28}{29} = 0.9655$$

$$\text{Presisi Busuk} = \frac{TN}{TN + FN} = \frac{29}{29 + 2} = \frac{29}{31} = 0.9355$$

### **Recall (Sensitivity)**

Seberapa baik K-Means dalam menangkap semua data dari kelas sebenarnya.

$$\text{Recall Segar} = TP / TP + FN = 28 / 28 + 2 = 28 / 30 = 0.9333$$

$$\text{Recall Busuk} = TN / TN + FP = 29 / 29 + 1 = 29 / 30 = 0.9667$$

Klasifikasi citra buah tomat menggunakan metode *K-Means Clustering* menunjukkan performa yang baik dengan akurasi nilai sebesar 95%, presisi dan *recall* dengan nilai diatas 93% untuk tomat segar dan busuk. Nilai tersebut menunjukkan model mampu mengklasifikasikan tomat segar dan busuk secara seimbang dan akurat.

## **4. Kesimpulan**

Penelitian ini berhasil menerapkan metode *K-Means Clustering* pada citra digital tomat untuk mengidentifikasi kondisi segar dan busuk. *K-Means* mampu menentukan kondisi tomat dengan tingkat akurasi sebesar 95% pada data uji yang terdiri dari 60 citra tomat (30 segar dan 30 busuk), angka tersebut menunjukkan bahwa metode ini cukup efektif dalam pengelompokan berdasarkan karakteristik visual. Kedepannya dapat dilakukan pengembangan lebih lanjut dapat mencakup penerapan metode klasifikasi berbasis *Deep Learning* serta pengujian dengan varietas tomat yang lebih beragam.

## **Daftar Pustaka**

- Baladraf, T. T. (2024). Analisis Kerugian Pasca Panen melalui Commodity System Assesment Method pada Komoditas Tomat di Indonesia. *Prosiding Seminar Nasional Pembangunan Dan Pendidikan Vokasi Pertanian*, September, 387–396. <https://doi.org/https://doi.org/10.47687/snppvp.v5i1.1121>
- Denih, A., Negara, T. P., & Marzuki, I. (2023). Analysis of Tomato Ripeness by Color and Texture Using Cielab and *K-Means Clustering*. *Komputasi: Jurnal Ilmiah Ilmu Komputer Dan Matematika*, 20(2), 148–159. <https://doi.org/10.33751/komputasi.v20i2.8311>
- Karim, M. H. A., & Karim, A. A. (2021). Using Texture Feature in Fruit Classification. *Engineering and Technology Journal*, 39(1B), 67–79. <https://doi.org/10.30684/etj.v39i1b.1741>
- Hermawan, G. (2024). Memahami Peran *Dataset* dalam Penelitian Kecerdasan Buatan : Kualitas , Aksesibilitas , dan Tantangan. *Preprint*, October. <https://doi.org/10.13140/RG.2.2.34468.49288>
- Pitaloka, D. (2020). Hortikultura: Potensi, Pengembangan Dan Tantangan. *Jurnal Teknologi Terapan: G-Tech*, 1(1), 1–4. <https://doi.org/10.33379/gtech.v1i1.260>
- Ridhoi, A. (2023). Penerapan Pengolahan Citra Untuk Perbaikan Gambar 2 Dimensi Dengan Menggunakan Matlab. *Joutica*, 8(1), 64–69. <https://doi.org/10.30736/informatika.v8i1.1063>
- Rizki, M., Furqan, M., & Sriani. (2024). Application of The *K-Means Clustering* Algorithm to Identify Strawberry Fruit Ripe. *International Journal of Information System & Technology*, 8(158), 130–138.

- Saleh, A., Ridwan, A., & Gibran, M. K. (2023). Machine Learning and Fuzzy C-Means Clustering for the Identification of Tomato Diseases. *Indonesian Journal of Computer Science*, 12(5), 2401–2413. <https://doi.org/10.33022/ijcs.v12i5.3379>
- Septiarini, A., Hamdani, H., Khoemah, Y., Puspitasari, N., & Wati, M. (2022). Segmentasi Tomat Menggunakan Metode *K-Means Clustering* dan Pengolahan Citra Digital. *Seminar Nasional CORIS 2022*, 113–118. <https://corisindo.stikom-bali.ac.id/penelitian/index.php/semnas/article/view/22>
- Utomo, I., Zufria, I., & Hasibuan, M. S. (2024). PENERAPAN ALGORITMA *K-MEANS CLUSTERING* UNTUK PENCATATAN PENCAPAIAN KINERJA KARYAWAN. *Journal of Science and Social Research*, VII(4), 1960–1969.
- Yolanda, Rosa, I., & Ramadhanu, A. (2024). Identifikasi Cerdas Apel Fuji dan Apel Hijau : Pendekatan *K-Means Clustering* untuk Segmentasi Buah. *Journal of Education Research*, 5(3), 4232–4239. <https://doi.org/https://doi.org/10.37985/jer.v5i3.1703>