

# Segmentasi Tomat Menggunakan Metode K-Means Clustering dan Pengolahan Citra Digital

Anindita Septiarini<sup>1</sup>, Hamdani Hamdani<sup>2</sup>, Yurni Khoemah<sup>3</sup>, Novianti Puspitasari<sup>4</sup>, Masna Wati<sup>5</sup>

Program Studi Informatika, Fakultas Teknik  
Universitas Mulawarman  
Samarinda, Indonesia

e-mail: <sup>1</sup>anindita@unmul.ac.id, <sup>2</sup>hamdani@unmul.ac.id, <sup>3</sup>yurninurul1419@gmail.com,  
<sup>4</sup>noviapuspitarsari @unmul.ac.id, <sup>5</sup>masnawati@fkti.unmul.ac.id

## Abstrak

Penerapan pengolahan citra digital pada bidang perkebunan telah banyak diteliti dan dikembangkan, misalnya pada sistem untuk mengidentifikasi kematangan buah dan mengendalikan robot pemanen buah. Pembuatan sistem tersebut membutuhkan proses utama yaitu segmentasi untuk membedakan antara area buah dan latar belakang (*background*) karena hanya area buah saja yang perlu diproses untuk menentukan kematangan buah. Penelitian ini bertujuan untuk menerapkan metode segmentasi tomat. Metode ini terdiri dari empat proses utama: akuisisi data, deteksi *region of interest* (ROI), *pre-processing*, dan segmentasi. Pengubahan ukuran dan *thresholding* dengan metode Otsu diterapkan pada deteksi ROI. Konversi ruang warna RGB menjadi HSV diterapkan pada *pre-processing*. Selanjutnya, K-Means clustering diterapkan pada segmentasi diikuti dengan implementasi operasi morfologi untuk membuang noise yang tersisa. Evaluasi kinerja metode segmentasi tomat yang dilakukan terhadap 160 citra menunjukkan rata-rata nilai *precision*, *recall*, dan *F-measure* yang diperoleh masing-masing mencapai 97%, 88%, dan 94%. Hasil tersebut menunjukkan bahwa metode segmentasi yang diterapkan mampu membedakan antara area buah dan *background*.

**Kata kunci:** *pre-processing*, segmentasi, *thresholding*, K-Means, operasi morfologi.

## Abstract

The implementation of image processing in the plantation sector has been widely researched and developed, for example, to identify fruit maturity and control fruit harvesting robots. These systems required the primary process, namely segmentation, to distinguish between the fruit area and the background because only the fruit area needs to be processed to determine fruit maturity. This study aims to apply the tomato segmentation method. This method consists of four main processes: data acquisition, region of interest (ROI) detection, *pre-processing*, and segmentation. Resizing and *thresholding* with the Otsu method were applied to ROI detection. RGB to HSV color space conversion was used in the *pre-processing*. Next, K-Means clustering was applied to the segmentation, followed by implementing morphological operations to remove the remaining noise. Evaluation of the performance of the tomato segmentation method on 160 images showed that the average *precision*, *recall*, and *F-measure* values obtained were 97%, 88%, and 94%, respectively. The results indicate the segmentation method applied can distinguish between the fruit area and the background.

**Keywords:** *pre-processing*, segmentation, *thresholding*, K-Means, morphology operation.

## 1. Pendahuluan

Sistem komputer, khususnya di bidang pertanian, telah banyak dikembangkan dalam beberapa tahun terakhir [1]. Berbagai aplikasi telah diterapkan untuk sistem berbasis citra dan video menggunakan teknik pengolahan citra sebagai data input utamanya. Misalnya, deteksi kematangan berbagai buah: apel [2], tomat [3], [4], dan kelapa sawit [5] yang dapat diterapkan pada sistem penyortiran kualitas buah atau mesin pemanenan otomatis serta pendeteksi penyakit daun [6]. Teknik pengolahan citra terkait pendeteksian kematangan buah yang dibutuhkan adalah proses segmentasi. Proses ini diperlukan untuk memisahkan area buah dan *background* karena hanya area buah saja yang perlu diproses untuk menentukan kematangan buah. Ketepatan dalam menentukan area buah dapat berpengaruh terhadap hasil klasifikasi

kematangan buah. Segmentasi beberapa jenis buah telah diterapkan pada apel [7], [8], mangga [9], delima [10], kelapa sawit [11], dan tomat [12], [13].

Tomat memiliki beberapa kandungan yang bermanfaat, seperti antioksidan serta vitamin C dan A untuk kesehatan mata. Selain itu, tomat dikonsumsi setiap hari sehingga menjadi salah satu tanaman yang memerlukan pemeriksaan tambahan. Sejumlah besar tomat juga diproduksi oleh produsen saus dan makanan olahan lainnya. Permintaan tomat berkualitas tinggi diperkirakan akan terus meningkat di masa depan. Tomat segar kualitasnya ditentukan oleh tingkat kematangan dan ukuran buahnya sehingga saat melakukan sortir hal tersebut harus dipertimbangkan.

Buah dapat disortir secara otomatis untuk tujuan kontrol kualitas, menghemat waktu dan biaya. Pada penerapan pengolahan citra digital diperlukan segmentasi khusus untuk membangun sistem tersebut. Tujuan dari proses segmentasi adalah untuk membedakan antara area tomat dan area sekitarnya atau *background*. Metode *K-Means clustering* merupakan salah satu metode yang telah banyak digunakan saat ini untuk melakukan segmentasi pada berbagai jenis objek [14], [15].

Banyak cara telah dilakukan untuk melakukan segmentasi berbagai buah pada penelitian sebelumnya. Segmentasi buah tomat dilakukan dengan menerapkan metode *K-Means clustering* dan deteksi tepi dengan operator Canny pada ruang warna HSV. Pada penelitian tersebut objek buah tomat berada pada *background* berwarna putih dan setiap citra hanya terdiri dari satu buah tomat. Kinerja dari metode segmentasi mencapai nilai akurasi 91,43% [12]. Segmentasi tomat diterapkan untuk klasifikasi tingkat kematangan buah. Metode segmentasi yang digunakan adalah Otsu pada berdasarkan ruang warna HSV. Hasil segmentasi menunjukkan terdapat *noise* berupa bagian area mahkota tomat yang terdeteksi sebagai area buah [16]. Selanjutnya, metode segmentasi tomat diterapkan menggunakan metode deteksi tepi dengan operator Canny dan teknik rekonstruksi pada *natural background*, di mana objek tomat bercampur dengan objek lain. Kinerja metode segmentasi tomat tersebut dievaluasi berdasarkan akurasi yang mencapai nilai 89,74% [3].

Selain itu, metode handal untuk segmentasi diuji pada tiga jenis buah yang berbeda: jeruk, anggur, dan lengkeng. Metode tersebut menggunakan transformasi wavelet untuk menyamakan iluminasi. Setelah itu, area buah disegmentasi menggunakan algoritma *K-Means clustering*. Dataset dikumpulkan pada dua kondisi hari yang berbeda yaitu pada kondisi cerah dan berawan. Nilai akurasi maksimum yang dicapai adalah 97,2% pada hari dalam kondisi cerah pada pengujian terhadap jeruk bali [15]. Selanjutnya, beberapa teknik pengolahan citra, seperti konversi model warna, thresholding, dan pemerataan histogram, digunakan untuk membuat metode segmentasi untuk penghitungan buah hijau pohon jeruk; kinerja metode menunjukkan hasil yang akurat terhadap buah yang berwarna hijau. Selanjutnya dilakukan penyaringan spasial dengan menggunakan operator Laplace dan Sobel serta Gaussian blur. Metode tersebut menimbulkan kesalahan dengan nilai toleransi sebesar 5% menggunakan pendekatan ini [17].

Banyak pekerjaan sebelumnya telah dilakukan untuk meningkatkan metode segmentasi buah yang saat ini digunakan [3], [7], [10], [13]. Namun, pencahayaan yang tidak merata masih menimbulkan *noise* yang dinilai belum memuaskan. Dataset pribadi digunakan untuk menerapkan metode segmentasi yang mampu secara otomatis mendeteksi area buah dan *background* pada citra tomat. Metode Otsu diperlukan untuk mendapatkan perkiraan kasar area tomat. Sementara itu, *K-Means clustering* digunakan untuk mendeteksi area tomat dan *background* pada proses segmentasi. Operator morfologi yang terdiri dari *closing* dan *opening* diperlukan untuk mengurangi area *noise* pada *background*. Ruang warna HSV dipilih untuk prosedur segmentasi karena itu yang paling sesuai untuk dataset yang digunakan pada penelitian ini.

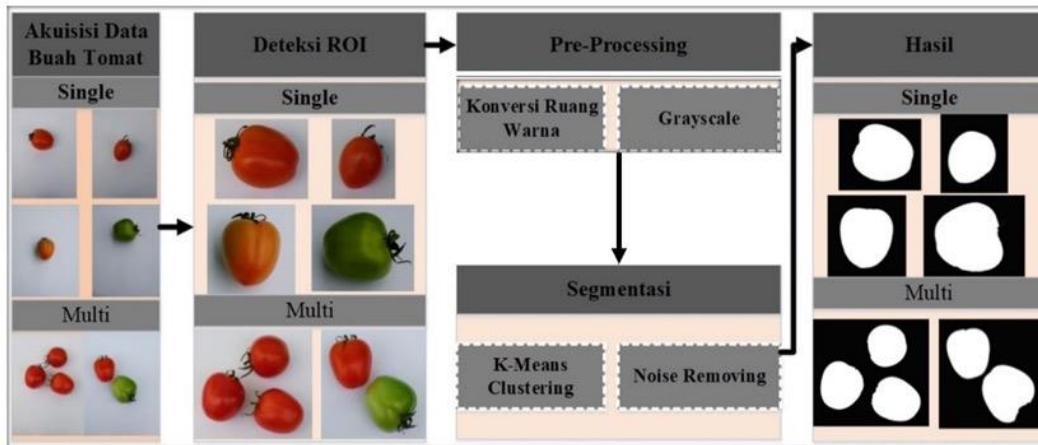
## 2. Metode Penelitian

Metode yang dibuat merupakan segmentasi buah tomat menggunakan metode *K-Means clustering*. Pada penelitian ini terdapat 4 proses utama yang dilakukan yaitu akuisisi data, deteksi ROI, *pre-processing*, dan segmentasi. Hasil dari penelitian ini berupa citra biner yang membedakan antara area buah tomat dan *background*. Gambaran umum proses utama metode ini dapat dilihat pada Gambar 1.

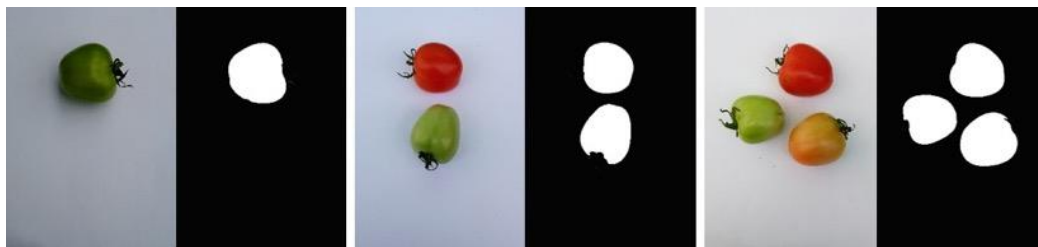
### 2.1. Akuisisi Data

Dataset yang digunakan terdiri dari 160 citra tomat yang memuat buah tomat tunggal maupun jamak (dalam satu citra terdiri dari beberapa buah tomat). Buah tomat terdiri dari tingkat kematangan dengan warna yang berbeda yaitu mentah (hijau), setengah matang (jingga), dan matang (merah). Citra dihasilkan menggunakan kamera digital yang disertakan dalam *smartphone* dan disimpan dalam format JPEG dengan *background* berwarna putih. Jarak antara kamera dan objek sekitar  $\pm 30$  cm. Akuisisi citra dilakukan dalam ruangan, dengan pencahayaan yang memadai dan merata. Selain itu, data *ground truth* dibuat terhadap semua citra untuk yang dibutuhkan pada tahap pengujian, seperti yang dilakukan dalam

[18]. Citra *ground truth* dibuat menggunakan aplikasi untuk melakukan *edit* gambar (*photoshop*) berupa citra biner. Contoh citra tomat dan *ground truth* disajikan pada Gambar 2.



Gambar 4. Gambaran umum proses utama pada metode segmentasi buah tomat



Gambar 2. Contoh citra tomat dan *ground truth*

**2.2. Deteksi ROI**

Proses deteksi ROI bertujuan untuk membentuk sub-citra yang diawali dengan pengubahan ukuran citra tomat dengan ukuran asli  $3456 \times 4608$  piksel menjadi  $1728 \times 2304$  piksel (diperkecil 0,5 dari ukuran asli) untuk mempercepat waktu komputasi proses selanjutnya. Metode Otsu diterapkan untuk memperoleh perkiraan area tomat diikuti dengan penentuan batas area untuk pembentukan citra ROI. Gambar 3 menunjukkan citra hasil setiap tahapan pada proses deteksi ROI. Metode Otsu bertujuan untuk memperoleh nilai threshold yang optimal dengan terlebih dahulu mengubah citra ROI menjadi citra grayscale. Adapun langkah-langkah algoritme metode Otsu telah dijelaskan pada [19], [20]. Citra ROI dibentuk berdasarkan citra biner yang dihasilkan dengan batasan kotak berwarna merah yang ditunjukkan pada Gambar 3(c).



Gambar 3. Contoh citra hasil tahapan pada deteksi ROI: (a) citra asli, (b) citra *grayscale*, (c) metode Otsu serta perkiraan area tomat, dan (d) citra ROI

**2.3. Pre-Processing**

*Pre-processing* diterapkan untuk meningkatkan kualitas citra agar dapat mencapai hasil segmentasi yang optimal. Metode yang dilakukan adalah dengan mengubah ruang warna RGB menjadi ruang warna HSV pada citra hasil deteksi ROI. Selanjutnya, citra HSV tersebut diubah menjadi citra grayscale. Citra 4 menampilkan contoh citra hasil konversi ruang warna RGB menjadi HSV. Konversi ruang warna RGB ke ruang warna HSV didefinisikan sebagai berikut menggunakan Persamaan. (1) – (4) [7], [21]:

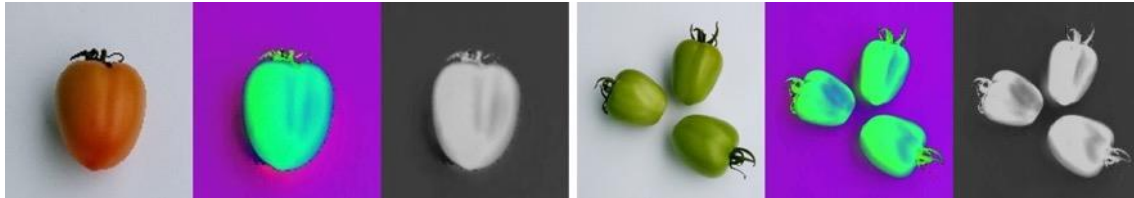
$$H = \begin{cases} \theta, & B \leq G \\ 360 - \theta, & B > G \end{cases} \quad (1)$$

di mana,

$$\theta = \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{[x(R-G)^2+(R-B)(G-B)]^{1/2}} \right\} \quad (2)$$

$$S = \begin{cases} 0, & \max(R, G, B) = 0 \\ 1 - \frac{\min(R, G, B)}{\max(R, G, B)}, & \text{otherwise} \end{cases} \quad (3)$$

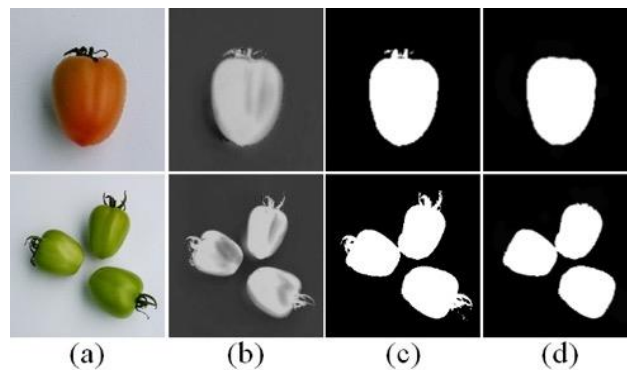
$$V = \max(R, G, B) \quad (4)$$



Gambar 4. Contoh citra hasil *pre-processing* pada ruang warna RGB menjadi HSV dan grayscale.

**2.4. Segmentasi**

Proses segmentasi tomat bertujuan untuk membedakan antara area tomat dan *background* pada citra. Metode yang diterapkan pada proses ini ada *K-Means clustering* pada citra dengan ruang warna HSV. *K-Means clustering* dipilih karena telah berhasil digunakan dalam beberapa penelitian sebelumnya yang berkaitan dengan segmentasi objek buah [10], [22]. Ruang warna HSV dipilih karena secara visual, hasil *clustering* memiliki lebih sedikit *noise*. Selanjutnya, dilakukan *noise removing* dengan menerapkan operasi morfologi yaitu *closing* dan *opening*. karena dimungkinkan hasil *clustering* masih memuat *noise*. Gambar 5 menunjukkan hasil segmentasi menggunakan metode *K-Means clustering* pada ruang warna HSV. Langkah-langkah dari metode *K-Means clustering* adalah sebagai berikut [14], [22]: sejumlah sampel data digunakan untuk menentukan pusat cluster ( $C_i$ ) secara acak (piksel dalam citra). Selanjutnya, setiap titik diklasifikasikan berdasarkan kedekatannya dengan pusat cluster terdekat menggunakan fungsi jarak. Lakukan perhitungan ulang rata-rata aritmatika dari setiap cluster untuk memperbarui pusat cluster baru. Pada tahap akhir lakukan perulangan terhadap klasifikasi setiap titik berdasarkan kedekatan dan perhitungan ulang rata-rata aritmatika dari setiap cluster



Gambar 5. Citra hasil pada setiap langkah proses segmentasi: (a) RGB, (b) grayscale, (c) *K-Means clustering*, dan (d) operasi morfologi

**3. Hasil dan Pembahasan**

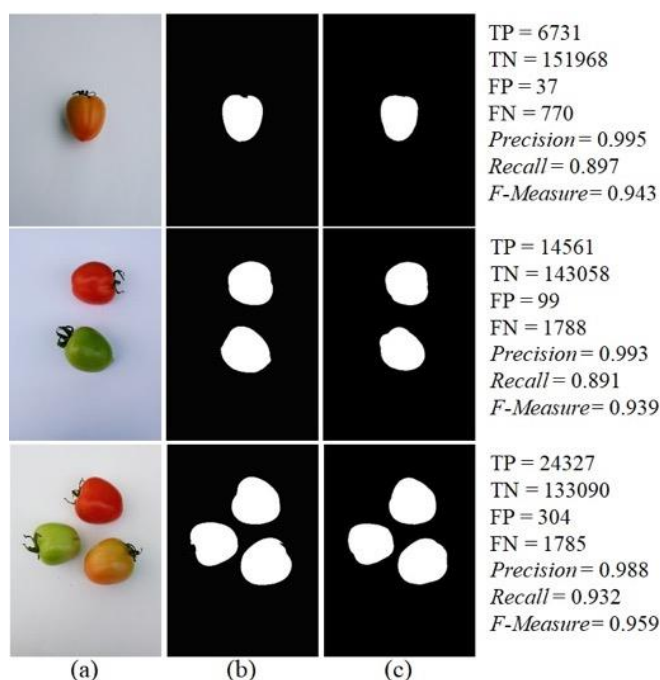
Evaluasi metode segmentasi tomat perlu diterapkan untuk mengetahui kinerja dari metode yang diterapkan. Hal ini dilakukan dengan membandingkan citra hasil segmentasi dengan citra *ground truth*. Pada penelitian ini kinerja metode dinilai berdasarkan tiga parameter evaluasi yaitu: *precision*, *recall*, dan *F-measure*. Nilai parameter berada diantara 0 hingga 100. Jika nilai parameter mendekati 100 menunjukkan metode yang diterapkan bekerja dengan baik jika nilai mendekati 0 artinya kinerja metode buruk. Tiga parameter evaluasi dihitung menggunakan Persamaan (5) - (7) sebagai berikut [22]:

$$Precision = \frac{T_p}{T_p + F_p} \quad (5)$$

$$Recall = \frac{T_p}{T_p + F_n} \quad (6)$$

$$F\text{-Measure} = 2 \left( \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \quad (7)$$

*True positive* (TP) adalah jumlah piksel berwarna putih yang merupakan area tomat dari citra hasil metode segmentasi dan citra *ground truth*. Sedangkan *false positive* (FP) adalah jumlah piksel berwarna putih pada citra hasil metode segmentasi yang merupakan area tomat, tetapi pada citra *ground truth* merupakan area *background*. Sebaliknya, *false negative* (FN) adalah jumlah piksel berwarna putih pada citra *ground truth*, yang merupakan area tomat, tetapi pada citra hasil metode segmentasi merupakan area *background*. Hasil perhitungan kinerja metode segmentasi area tomat ditunjukkan pada Gambar 6.



Gambar 6. Contoh hasil segmentasi: (a) citra asli, (b) citra *ground truth*, dan (c) citra hasil metode *K-Means clustering*

Gambar 6 menunjukkan terjadinya *under-segmentation* pada baris satu dan dua contoh akibat area tomat yang salah diklasifikasikan sebagai *background*. Kesalahan tersebut dapat dilihat dari jumlah nilai FN lebih besar dari FP. Pada penelitian ini *under-segmentation* lebih sering terjadi. Berdasarkan hasil eksperimen, kinerja metode segmentasi yang ditunjukkan berdasarkan parameter precision, recall, dan F-Measure menghasilkan nilai rata-rata masing-masing sebesar 99%, 88%, dan 94%.

#### 4. Kesimpulan

Penelitian ini mengembangkan metode segmentasi tomat dengan data masukan berupa citra yang memuat objek tomat tunggal maupun jamak. Proses segmentasi diperlukan untuk membedakan antara area tomat dan *background* untuk beberapa tujuan tertentu. Metode Otsu, *K-Means clustering*, dan operasi morfologi digunakan untuk melakukan segmentasi tomat. Metode tersebut diterapkan dalam proses deteksi ROI dan segmentasi pada ruang warna HSV. Terdapat 160 citra tomat yang memuat tomat tunggal maupun multi dengan tiga jenis tingkat kematangan (mentah, setengah matang, dan matang) untuk mengevaluasi kinerja dari metode segmentasi. Kinerja metode segmentasi diuji dengan menggunakan tiga parameter yaitu precision, recall, dan F-Measure yang masing-masing mencapai nilai rata-rata 97%, 88%, dan 94%. Berdasarkan hasil eksperimen ditunjukkan bahwa *under-segmentation* lebih sering terjadi terutama pada bagian dekat kelopak buah. Metode yang diterapkan terbatas hanya diuji dengan dataset yang digunakan pada penelitian ini. Dengan demikian, ruang lingkup pekerjaan ini dapat masih dapat dikembangkan untuk mengatasi kesalahan yang terjadi serta dapat diterapkan untuk dataset jenis tanaman selain tomat.

#### Daftar Pustaka

- [1] J. Zheng, H. Fu, W. Li, W. Wu, L. Yu, S. Yuan, W.Y.W. Tao, T.K. Pang, and K.D. Kanniah, Growing status observation for oil palm trees using unmanned aerial vehicle (uav) images, *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 95–121, 2021.

- 
- [2] J. Llorens, R. Sanz-cortiella, A. Escol, J. Morros, J. Ruiz- hidalgo, and J. R. Rosell-polo, "Fruit detection in an apple orchard using a mobile terrestrial laser scanner Jordi Gen," *Biosyst. Eng.*, vol. 7, 2019.
- [3] A. Septiarini, D.N. Siswoyo, H. Hamdani, M. Wati, J.A. Widians, and N. Puspitasari, "Tomato Segmentation on Natural Background Using Multi Operation of Edge Detection And Reconstruction," in 2021 IEEE URUCON, 2021, pp. 571–575.
- [4] M. H. Malik, T. Zhang, H. Li, M. Zhang, S. Shabbir, and A. Saeed, "Mature Tomato Fruit Detection Algorithm Based on improved HSV and Watershed Algorithm," *IFAC-PapersOnLine*, vol. 51, no. 17, pp. 431–436, 2018.
- [5] A. Septiarini, A. Sunyoto, H. Hamdani, A. Ahmad, F. Utaminigrum, and H. R. Hatta, "Machine vision for the maturity classification of oil palm fresh fruit bunches based on color and texture features," *Sci. Hortic.*, vol. 286, pp. 110245, 2021.
- [6] H. Hamdani, A. Septiarini, A. Sunyoto, S. Suyanto, and F. Utaminigrum, "Detection of oil palm leaf disease based on color histogram and supervised classifier," *Optik*, vol. 245, pp. 167753, 2021.
- [7] J. Lv, H. Ni, Q. Wang, B. Yang, and L. Xu, "A segmentation method of red apple image," *Sci. Hortic.*, vol. 256, no. June, p. 108615, 2019.
- [8] J. Lv, F. Wang, L. Xu, Z. Ma, and B. Yang, "A segmentation method of bagged green apple image," *Sci. Hortic.*, vol. 246, no. November 2018, pp. 411–417, 2019.
- [9] C. Zheng, P. Chen, J. Pang, X. Yang, C. Chen, S. Tu, and Y. Xu, "A mango picking vision algorithm on instance segmentation and key point detection from RGB images in an open orchard," *Biosyst. Eng.*, vol. 206, pp. 32–54, 2021.
- [10] M. Fashi, L. Naderloo, and H. Javadikia, "The relationship between the appearance of pomegranate fruit and color and size of arils based on image processing," *Postharvest Biol. Technol.*, vol. 154, no. April, pp. 52–57, 2019.
- [11] A. Septiarini, H. Hamdani, H. R. Hatta, and K. Anwar, "Automatic image segmentation of oil palm fruits by applying the contour-based approach," *Sci. Hortic*, vol. 261, pp. 108939, 2020.
- [12] A. Septiarini, H. Hamdani, S. U. Sari, H. Rahmania Hatta, N. Puspitasari, and W. Hadikurniawati, "Image Processing Techniques For Tomato Segmentation Applying K-Means Clustering and Edge Detection Approach," 2021 International Seminar on Machine Learning, Optimization, and Data Science (ISMODE), 2022, pp. 92-96.
- [13] R. Xiang, "Image segmentation for whole tomato plant recognition at night," *Comput. Electron. Agric.*, vol. 154, no. June, pp. 434–442, 2018.
- [14] J. Qiao, X. Cai, Q. Xiao, Z. Chen, P. Kulkarni, C. Ferris, S. Kamarthi, and S. Sridhar, "Data on mri brain lesion segmentation using k-means and gaussian mixture model-expectation maximization," *Data in Brief*, vol. 27, p. 104628, 2019.
- [15] C. Wang, Y. Tang, X. Zou, W. SiTu, and W. Feng, "A robust fruit image segmentation algorithm against varying illumination for vision system of fruit harvesting robot," *Optik*, vol. 131, pp. 626–631, 2017.
- [16] A. Septiarini, H. Hamdani, M.S. Sauri, and J.A. Widians, "Image processing for maturity classification of tomato using otsu and manhattan distance methods," *Jurnal Informatika*, vol. 16, no. 1, 2022.
- [17] W. Maldonado and J. C. Barbosa, "Automatic green fruit counting in orange trees using digital images," *Comput. Electron. Agric.*, vol. 127, pp. 572–581, 2016.
- [18] S. Sun, H. Song, D. He, and Y. Long, "An adaptive segmentation method combining MSRRCR and mean shift algorithm with K-means correction of green apples in natural environment," *Inf. Process. Agric.*, vol. 6, no. 2, pp. 200–215, 2019.
- [19] I. A. Muwakhid and D. Nurdiyah, "Otsu Method for Image Finish Segmentation With Components of Hue Saturation Value," *Transformatika*, vol. 15, no. 2, pp. 67–73, 2018.
- [20] A. B. Kaswar, A. Akram, And N. Risal, "Klasifikasi Tingkat Kematangan Buah Markisa Menggunakan Jaringan Syaraf Tiruan Berbasis Pengolahan Citra Digital," *J. Embed. Syst. Secur. Intell. Syst.*, Vol. 01, No. May, Pp. 1–8, 2020.
- [21] F. Garcia-lamont, J. Cervantes, A. López, and L. Rodriguez, "Segmentation of images by color features: A survey," *Neurocomputing*, vol. 292, pp. 1–27, 2018.
- [22] A. Septiarini, H. Hamdani, E. Setyaningsih, E. Arisandy, S. Suyanto, and E. Winarno, "Automatic Segmentation of Optic Nerve Head by Median Filtering and Clustering Approach," in 2021 13th International Conference on Information & Communication Technology and System (ICTS), 2021, pp. 118–122.
-