

Klasifikasi Mutu Tomat dan Potensi Umur Simpan Berdasarkan Fitur Warna-Tekstur Menggunakan Random Forest

Intan Noviyanti*, Esti Wijayanti, Evanita

Fakultas Teknik, Teknik Informatika, Universitas Muria Kudus, Kudus

Jl. Lingkar Utara, Kayuapu Kulon, Gondangmanis, Kec. Bae, Kabupaten Kudus, Jawa Tengah, Indonesia

Email: ¹*202251103@std.umk.ac.id, ²esti.wijayanti@umk.ac.id, ³evanita@umk.ac.id

Email Penulis Korespondensi: 202251103@std.umk.ac.id

Submitted: 13/06/2026; Accepted: 04/07/2026; Published: 05/07/2026

Abstrak—Kerusakan tomat pascapanen masih menjadi tantangan akibat proses penilaian mutu yang dilakukan secara manual dan bersifat subjektif sehingga berpotensi menyebabkan ketidakkonsistenan hasil sortasi dan kesalahan dalam memperkirakan umur simpan. Penelitian ini bertujuan mengembangkan sistem klasifikasi mutu tomat dan prediksi potensi umur simpan berbasis pengolahan citra digital menggunakan algoritma Random Forest. Penelitian ini menggunakan 936 citra tomat dan 450 citra non-tomat yang dikumpulkan secara mandiri. Fitur yang digunakan terdiri atas fitur warna Red Green Blue (RGB), Hue Saturation Value (HSV), serta fitur tekstur Gray Level Co-occurrence Matrix (GLCM). Proses klasifikasi dilakukan menggunakan Random Forest Classifier untuk tiga kategori mutu, yaitu Buruk, Sedang, dan Baik, sedangkan prediksi umur simpan dilakukan menggunakan Random Forest Regressor. Hasil pengujian menunjukkan bahwa model klasifikasi memperoleh akurasi 96,81%, presisi 96,82%, recall 96,81%, dan F1-score 96,81%. Model regresi menghasilkan nilai Mean Absolute Error (MAE) sebesar 0,0621, Root Mean Square Error (RMSE) sebesar 0,1152, dan R^2 sebesar 0,8752, sedangkan cross-validation menghasilkan akurasi rata-rata $95,83\% \pm 1,24\%$ yang menunjukkan performa model yang stabil. Analisis feature importance menunjukkan bahwa fitur warna memberikan kontribusi terbesar pada kedua model, dengan g_mean sebagai fitur paling dominan dalam klasifikasi mutu maupun prediksi umur simpan. Penelitian ini memberikan kontribusi berupa pengembangan sistem penilaian mutu tomat yang mampu mengklasifikasikan mutu sekaligus memprediksi potensi umur simpan berbasis pengolahan citra digital menggunakan Random Forest, serta analisis feature importance untuk mengidentifikasi karakteristik visual yang paling berpengaruh terhadap kinerja model. Hasil penelitian menunjukkan bahwa pendekatan tersebut berpotensi mendukung proses sortasi dan pengelolaan pascapanen tomat secara lebih objektif dan efisien.

Kata Kunci: Random Forest; Klasifikasi Mutu Tomat; Prediksi Umur Simpan; Ekstraksi Fitur; Pengolahan Citra

Abstract—Postharvest tomato deterioration remains a major challenge due to manual and subjective quality assessment, which may lead to inconsistent sorting results and inaccurate shelf-life estimation. This study aims to develop a tomato quality classification system and predict potential shelf life based on digital image processing using the Random Forest algorithm. The study employed 936 tomato images and 450 non-tomato images collected independently. The extracted features consisted of Red Green Blue (RGB) and Hue Saturation Value (HSV) color features, as well as Gray Level Co-occurrence Matrix (GLCM) texture features. Tomato quality was classified into three categories, namely Poor, Medium, and Good, using a Random Forest Classifier, while shelf-life prediction was performed using a Random Forest Regressor. The classification model achieved an accuracy of 96.81%, precision of 96.82%, recall of 96.81%, and an F1-score of 96.81%. The regression model produced a Mean Absolute Error (MAE) of 0.0621, a Root Mean Square Error (RMSE) of 0.1152, and an R^2 value of 0.8752, while cross-validation yielded an average accuracy of $95.83\% \pm 1.24\%$, indicating stable model performance. Feature importance analysis revealed that color features contributed the most to both models, with g_mean identified as the most influential feature for tomato quality classification and shelf-life prediction. This study contributes to the development of a tomato quality assessment system capable of simultaneously classifying tomato quality and predicting shelf-life potential based on digital image processing using the Random Forest algorithm. In addition, feature importance analysis is employed to identify the visual characteristics that have the greatest influence on model performance. The results demonstrate that the proposed approach has the potential to support tomato sorting and postharvest management processes in a more objective and efficient manner.

Keywords: Random Forest; Tomato Quality Classification; Shelf-life Prediction; Feature Extraction; Digital Image Processing

1. PENDAHULUAN

Tomat (*Solanum lycopersicum* L.) merupakan salah satu komoditas hortikultura yang memiliki nilai ekonomi tinggi dan berperan penting dalam memenuhi kebutuhan pangan masyarakat Indonesia [1]. Berdasarkan data Badan Pusat Statistik tahun 2024, produksi tomat nasional mencapai 1,02 juta ton dengan kontribusi terbesar berasal dari Provinsi Jawa Timur dan Sumatera Utara [2]. Peningkatan permintaan pasar domestik maupun ekspor mendorong perlunya pengelolaan mutu tomat yang lebih baik, terutama pada tahap pascapanen. Namun, kerusakan dan penurunan kualitas tomat masih menjadi permasalahan utama dalam rantai distribusi hortikultura. Kementerian Pertanian Republik Indonesia melaporkan bahwa kehilangan hasil pascapanen tomat dapat mencapai 20–25% dari total produksi [3]. Kondisi tersebut dipengaruhi oleh proses sortasi yang masih dilakukan secara manual melalui pengamatan visual sehingga bersifat subjektif, menghasilkan penilaian mutu yang tidak konsisten, serta menyulitkan dalam memperkirakan potensi umur simpan buah.

Mutu tomat dipengaruhi oleh karakteristik visual yang berubah selama proses pematangan, seperti warna kulit, tekstur permukaan, dan tingkat kekerasan buah. Perubahan warna dari hijau menuju merah menunjukkan peningkatan kandungan likopen dan karotenoid, sedangkan perubahan tekstur berkaitan dengan degradasi dinding sel selama proses respirasi dan pematangan [4]. Karakteristik tersebut dapat dimanfaatkan sebagai indikator untuk menentukan kualitas konsumsi maupun potensi umur simpan tomat [5].

Perkembangan teknologi computer vision dan machine learning memberikan peluang untuk mengotomatisasi proses penilaian mutu buah melalui analisis citra digital [6]. Pendekatan berbasis citra digital dinilai mampu memberikan proses penilaian mutu yang lebih cepat, objektif, dan konsisten dibandingkan metode manual yang masih bergantung pada pengamatan visual manusia [7]. Berbagai penelitian menunjukkan bahwa fitur warna maupun tekstur mampu merepresentasikan karakteristik visual buah dengan baik. Ruang warna Red Green Blue (RGB) banyak digunakan untuk menggambarkan intensitas warna dasar, sedangkan Hue Saturation Value (HSV) mampu merepresentasikan tingkat kematangan secara lebih stabil terhadap perubahan intensitas cahaya [8]. Di sisi lain, fitur tekstur Gray Level Co-occurrence Matrix (GLCM) dapat menggambarkan pola permukaan buah sehingga informasi warna dan tekstur yang dikombinasikan berpotensi menghasilkan representasi citra yang lebih komprehensif dibandingkan penggunaan satu jenis fitur saja [9]. Selain itu, algoritma Random Forest dikenal memiliki kemampuan klasifikasi yang baik [10], tahan terhadap overfitting, serta mampu memberikan interpretasi model melalui analisis feature importance.

Beberapa penelitian terdahulu telah menerapkan pengolahan citra digital untuk klasifikasi mutu maupun tingkat kematangan buah. Fatah et al menggunakan fitur warna RGB dan metode K-Nearest Neighbor untuk klasifikasi kematangan tomat dengan tingkat keberhasilan sebesar 66,66%, namun penelitian tersebut masih menggunakan jumlah data yang terbatas serta hanya memanfaatkan fitur warna [5]. Faradita dan Harahap menerapkan Jaringan Saraf Tiruan Perceptron berbasis histogram warna untuk klasifikasi tomat dengan tingkat keberhasilan sebesar 43,33%, sehingga menunjukkan bahwa penggunaan fitur warna sederhana belum mampu menghasilkan performa yang optimal [11]. Iskandar et al memanfaatkan fitur tekstur GLCM dengan metode Support Vector Machine pada klasifikasi tingkat kematangan buah nanas dan memperoleh akurasi sebesar 86%, namun penelitian tersebut belum mengombinasikan informasi warna sebagai representasi visual buah [12]. Khultsum dan Subekti berhasil menerapkan Random Forest dengan kombinasi fitur warna dan tekstur haralick untuk klasifikasi penyakit daun tomat dengan akurasi mencapai 96%, akan tetapi objek penelitian masih terbatas pada daun dan belum membahas mutu buah [13]. Sementara itu, Argadianata et al menggunakan Random Forest untuk klasifikasi kualitas buah apel berdasarkan parameter fisik berupa ukuran, berat, kemanisan, dan tingkat kematangan dengan akurasi sebesar 88,5%, tetapi masih bergantung pada pengukuran manual sehingga belum memanfaatkan citra digital sebagai sumber informasi utama [14].

Berdasarkan penelitian-penelitian tersebut, masih terdapat beberapa kesenjangan penelitian (research gap). Pertama, sebagian besar penelitian hanya berfokus pada klasifikasi mutu atau tingkat kematangan tanpa mengintegrasikan prediksi potensi umur simpan dalam satu sistem. Kedua, penelitian yang secara khusus mengombinasikan fitur warna RGB, HSV, dan fitur tekstur GLCM pada objek tomat masih sangat terbatas, padahal kombinasi ketiga fitur tersebut berpotensi memberikan representasi visual yang lebih lengkap. Ketiga, sebagian besar penelitian hanya berorientasi pada peningkatan akurasi model tanpa melakukan analisis feature importance sehingga kontribusi masing-masing fitur terhadap proses klasifikasi maupun prediksi belum dapat dijelaskan secara interpretatif. Padahal, informasi mengenai fitur yang paling berpengaruh dapat menjadi dasar dalam pengembangan sistem yang lebih efisien sekaligus meningkatkan pemahaman mengenai karakteristik visual yang menentukan mutu tomat.

Berdasarkan kesenjangan tersebut, penelitian ini bertujuan mengembangkan sistem klasifikasi mutu tomat dan prediksi potensi umur simpan berbasis citra digital menggunakan algoritma Random Forest dengan kombinasi fitur warna RGB, HSV, dan tekstur GLCM. Selain mengevaluasi performa model klasifikasi dan regresi, penelitian ini juga menganalisis feature importance untuk mengidentifikasi fitur visual yang memberikan kontribusi terbesar terhadap proses klasifikasi mutu dan prediksi potensi umur simpan tomat. Kontribusi penelitian ini terletak pada penyediaan pendekatan penilaian mutu tomat yang mengintegrasikan klasifikasi mutu dan prediksi potensi umur simpan dalam satu kerangka berbasis citra digital, disertai interpretasi karakteristik visual yang paling berpengaruh terhadap kinerja model melalui analisis feature importance.

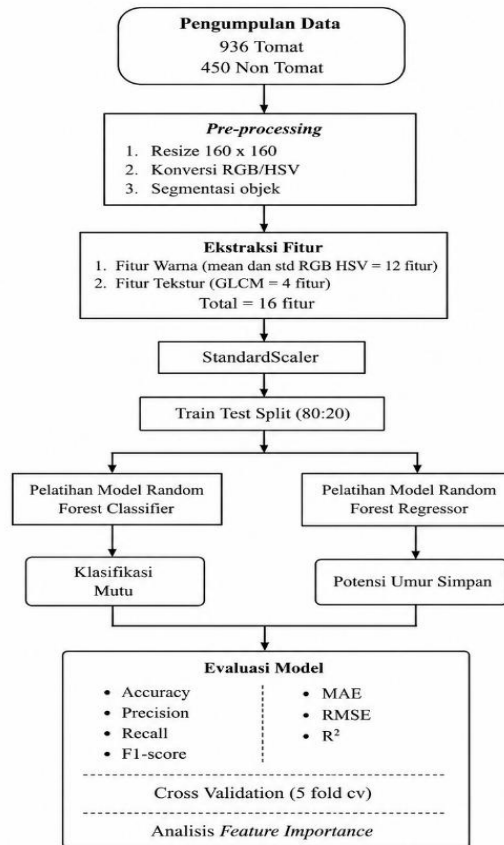
2. METODOLOGI PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode eksperimental untuk mengembangkan model klasifikasi mutu tomat dan prediksi potensi umur simpan berbasis citra digital berdasarkan fitur warna dan tekstur menggunakan algoritma Random Forest. Implementasi dilakukan menggunakan Python 3.12 dengan pustaka OpenCV, scikit-image, scikit-learn, NumPy, Pandas, dan Matplotlib, sedangkan antarmuka sistem dibangun menggunakan framework Gradio. Seluruh proses pelatihan dan pengujian model dijalankan pada platform Google Colaboratory karena menyediakan lingkungan komputasi yang stabil dan mendukung eksperimen machine learning berbasis citra digital [15].

2.1 Tahapan Penelitian

Tahapan penelitian terdiri atas proses pengumpulan data, pre-processing citra, ekstraksi fitur warna dan tekstur, normalisasi fitur menggunakan StandardScaler, pembagian data menggunakan train-test split dengan rasio 80:20, pelatihan model klasifikasi mutu, pelatihan model prediksi umur simpan, serta evaluasi model. Alur penelitian

dirancang secara sistematis [16] untuk menghasilkan model klasifikasi mutu dan estimasi umur simpan tomat berbasis citra digital yang dapat dilihat pada Gambar 1.



Gambar 1. Workflow Pipeline Diagram

Tahap pertama adalah pengumpulan data. Data yang digunakan dalam penelitian ini merupakan dataset mandiri yang dikumpulkan secara langsung melalui pengambilan foto menggunakan kamera ponsel Samsung Galaxy A04s. Pengambilan foto dilakukan dengan latar belakang kertas HVS putih untuk meminimalkan gangguan dari objek di luar tomat. Total citra tomat yang terkumpul sebanyak 936 citra yang terbagi menjadi tiga kelas mutu dengan masing-masing kelas berjumlah 312 citra. Selain citra tomat, penelitian ini juga mengumpulkan 450 citra non-tomat yang terdiri dari berbagai objek yang memiliki bentuk menyerupai tomat. Penentuan kelas mutu tomat dalam penelitian ini didasarkan pada kombinasi parameter warna dan tekstur yang mengacu pada standar yang telah ditetapkan. Parameter warna digunakan untuk merepresentasikan tingkat kematangan tomat berdasarkan tahapan kematangan USDA/UC Davis, yaitu mature green, breaker, turning, pink, dan red. Sementara itu, parameter tekstur dan kondisi fisik digunakan untuk merepresentasikan mutu tomat berdasarkan tingkat kekerasan, cacat fisik, dan kerusakan buah yang mengacu pada Standar Nasional Indonesia mutu tomat (SNI 01-3162-1992). Berdasarkan kombinasi kedua parameter tersebut, penelitian ini menetapkan tiga kelas mutu tomat yaitu mutu baik, mutu sedang, dan mutu buruk. Tomat dengan warna merah segar, tekstur masih keras, dan tanpa kerusakan fisik dikategorikan sebagai mutu baik. Tomat dengan warna hijau hingga oranye serta kondisi fisik yang masih layak dikategorikan sebagai mutu sedang karena masih berpotensi melanjutkan proses pematangan selama penyimpanan. Adapun tomat yang menunjukkan gejala pelunakan berlebih, memar, atau pembusukan dikategorikan sebagai mutu buruk. Sedangkan untuk penetapan estimasi umur simpan untuk setiap kelas mutu mengacu pada penelitian Badebo dan Eysa [17] yang menyatakan bahwa pada penyimpanan suhu ruang (25-28°C) menggunakan wooden box, tomat dengan tingkat kematangan full ripe (merah) memiliki umur simpan 4,3 hari, half ripe (oranye) 6 hari, dan mature green (hijau) 7,3 hari. Pola menurunnya umur simpan seiring meningkatnya tingkat kematangan ini juga dikonfirmasi oleh Deglas [18] dalam konteks penyimpanan di Indonesia serta Fertiasari et al. [19] yang menyatakan bahwa tomat hanya bertahan maksimal 7 hari pada suhu kamar. Untuk mendukung proses pembelajaran model regresi secara kontinu, nilai target umur simpan setiap citra diberikan variasi kecil dalam rentang yang wajar sesuai karakteristik visual tomat dalam setiap kelas mutu. Pendekatan ini bertujuan agar model dapat mempelajari hubungan antara fitur citra dengan potensi umur simpan secara lebih akurat.

Tahap kedua adalah pre-processing citra. Pre-processing citra dilakukan untuk menyeragamkan dimensi data agar dapat digunakan pada proses pelatihan model [20]. Seluruh citra diubah ukurannya menjadi 160×160 piksel menggunakan fungsi `cv2.resize()` dari pustaka OpenCV. Selanjutnya, citra dikonversi dari ruang warna

BGR ke RGB dan HSV. Ruang warna RGB digunakan untuk merepresentasikan intensitas warna dasar citra, sedangkan HSV digunakan karena lebih mampu memisahkan informasi warna dan pencahayaan sehingga efektif dalam proses segmentasi objek [1]. Segmentasi area buah dilakukan menggunakan metode color thresholding berbasis HSV untuk memisahkan objek tomat dari latar belakang [21]. Masker yang dihasilkan kemudian diproses dengan operasi morfologi closing untuk menghilangkan noise.

Tahap ketiga adalah ekstraksi fitur menggunakan kombinasi fitur warna dan tekstur. Fitur warna diekstraksi menggunakan nilai rata-rata (mean) dan standar deviasi (std) pada masing-masing kanal RGB dan HSV untuk merepresentasikan karakteristik warna tomat selama proses pematangan, sehingga menghasilkan 12 fitur warna. Sementara itu, fitur tekstur diekstraksi menggunakan metode GLCM [22] dengan parameter jarak 1 piksel dan sudut 0° , menghasilkan 4 fitur tekstur yaitu contrast, homogeneity, energy, dan correlation karena mampu merepresentasikan pola permukaan dan kondisi fisik objek citra [23]. Total fitur yang digunakan dalam penelitian ini adalah 16 fitur.

Sebelum proses pelatihan model, seluruh fitur dinormalisasi menggunakan StandardScaler untuk menyamakan skala setiap fitur sehingga proses pembelajaran model menjadi lebih stabil. Dataset kemudian dibagi menggunakan metode train-test split dengan rasio 80:20 menjadi data pelatihan dan data pengujian.

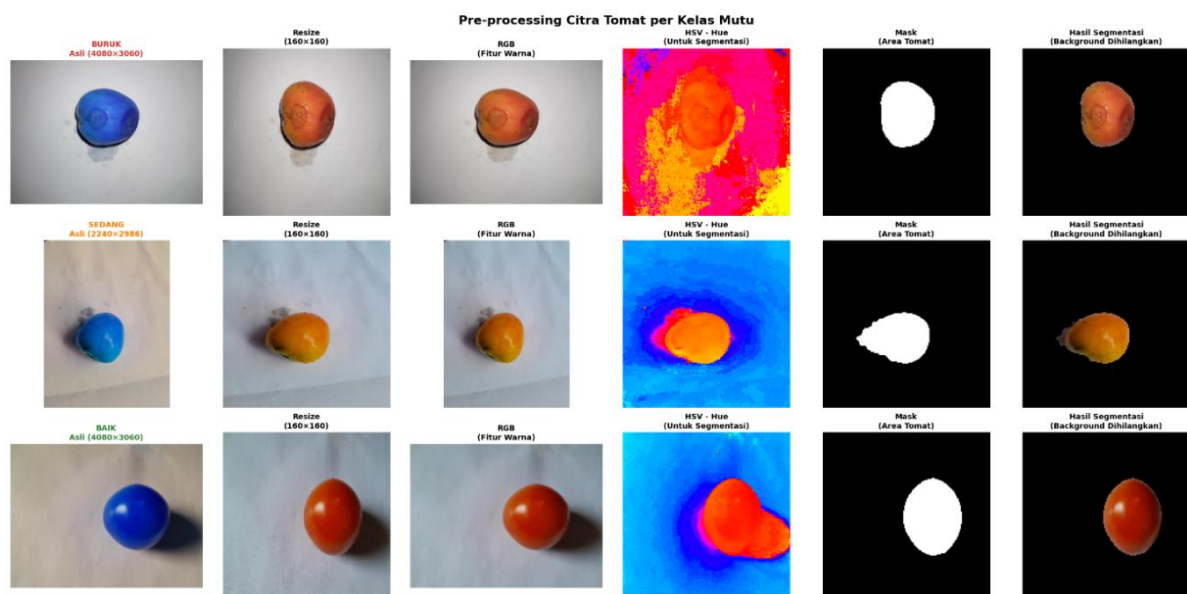
Tahap selanjutnya adalah pelatihan model. Data hasil ekstraksi fitur digunakan pada dua model yang berbeda, yaitu Random Forest Classifier untuk mengklasifikasikan mutu tomat ke dalam kategori buruk, sedang, dan baik, serta Random Regressor untuk memprediksi potensi umur simpan tomat berdasarkan karakteristik visual citra. Algoritma Random Forest dipilih karena mampu menangani data multidimensi, mengurangi risiko overfitting, serta memiliki performa yang baik pada proses klasifikasi citra berbasis fitur warna dan tekstur [13].

Tahap terakhir adalah evaluasi model menggunakan metrik accuracy, precision, recall, dan F1-score pada model klasifikasi, sedangkan model regresi dievaluasi menggunakan Mean Absolute Error (MAE), Root Mean Square Error (RMSE), dan koefisien determinasi (R^2). Untuk memastikan konsistensi dan ketahanan model, dilakukan 5-fold cross validation pada data latih dengan menghitung rata-rata akurasi serta standar deviasi hasil pengujian. Selain itu juga dilakukan analisis feature importance untuk mengetahui kontribusi masing-masing fitur terhadap proses klasifikasi mutu dan prediksi potensi umur simpan.

3. HASIL DAN PEMBAHASAN

3.1 Hasil Pre-processing Citra

Tahap pre-processing dilakukan untuk menyeragamkan data citra sebelum proses ekstraksi fitur dan pelatihan model. Tahapan ini meliputi resize citra, konversi ruang warna, segmentasi objek tomat, serta pemisahan area buah dari latar belakang. Hasil visualisasi tahapan pre-processing dapat dilihat pada Gambar 2.



Gambar 2. Hasil Pre-processing Citra

Berdasarkan Gambar 2, tahapan pre-processing yang meliputi proses resize, konversi ruang warna RGB–HSV, dan segmentasi berhasil mempersiapkan citra untuk proses ekstraksi fitur. Proses resize menghasilkan ukuran citra yang seragam sehingga setiap sampel memiliki representasi yang konsisten, sedangkan konversi ke ruang warna HSV membantu memisahkan informasi warna dari intensitas pencahayaan sehingga karakteristik visual tomat lebih mudah dianalisis. Hasil segmentasi menunjukkan bahwa area utama objek tomat pada seluruh kelas mutu berhasil dipisahkan dari latar belakang putih dengan baik meskipun masih terdapat sedikit noise pada

beberapa citra. Namun, noise tersebut tidak mendominasi area objek sehingga tidak mengganggu proses ekstraksi fitur warna RGB, HSV, maupun tekstur Gray Level Co-occurrence Matrix (GLCM). Hasil ini menunjukkan bahwa tahapan pre-processing yang diterapkan mampu menghasilkan citra yang representatif untuk mendukung proses klasifikasi mutu dan prediksi potensi umur simpan, sejalan dengan penelitian sebelumnya yang menyatakan bahwa segmentasi berbasis HSV efektif dalam mempertahankan informasi objek dan meningkatkan kualitas analisis citra digital [21].

3.2 Hasil Ekstraksi Fitur

Ekstraksi fitur berhasil dilakukan pada seluruh citra, menghasilkan 16 fitur yang terdiri atas fitur warna (mean dan std dari RGB dan HSV) serta fitur tekstur GLCM (contrast, homogeneity, energy, correlation). Tabel 1 menyajikan contoh nilai fitur dari tiga sampel yang mewakili masing-masing kelas mutu.

Tabel 1. Nilai Fitur dari Tiga Kelas Mutu Tomat

Fitur	Baik	Sedang	Buruk
r_mean	183.0562	163.3784	166.7144
g_mean	146.5335	111.2722	60.7111
b_mean	171.6596	33.7039	60.0999
r_std	22.8049	39.4622	36.0469
g_std	48.3464	37.5281	28.2070
b_std	61.1914	53.8628	24.5523
h_mean	124.6377	25.3369	84.5324
s_mean	65.6384	200.6024	164.3655
v_mean	192.5491	165.0369	166.7144
h_std	45.5192	29.3348	88.2332
s_std	59.6394	79.9372	40.3714
v_std	25.5986	39.9372	36.0469
glcm_contrast	0.0701	0.0515	0.1328
glcm_homogeneity	0.9670	0.9762	0.9560
glcm_energy	0.5265	0.6372	0.4602
glcm_correlation	0.9782	0.9731	0.9782

Berdasarkan Tabel 1, setiap kelas mutu menunjukkan karakteristik visual yang berbeda. Kelas baik cenderung memiliki nilai intensitas warna hijau (g_mean) yang lebih tinggi dan tekstur yang lebih homogen, sedangkan kelas buruk menunjukkan nilai contrast yang lebih tinggi yang mengindikasikan permukaan buah lebih tidak merata akibat proses pematangan lanjut atau kerusakan. Perbedaan karakteristik tersebut menunjukkan bahwa fitur warna dan tekstur mampu merepresentasikan perubahan fisiologis tomat selama proses pematangan. Hasil ini sejalan dengan penelitian sebelumnya yang menyatakan bahwa fitur warna merupakan indikator utama dalam identifikasi tingkat kematangan buah, sedangkan fitur tekstur berperan sebagai informasi pelengkap untuk membedakan kondisi permukaan objek [23]. Penggunaan kombinasi fitur RGB, HSV, dan GLCM pada penelitian ini menghasilkan representasi citra yang lebih komprehensif dibandingkan penelitian Fatah et al. [5] dan Faradita et al. [11] yang hanya memanfaatkan fitur warna, sehingga diharapkan dapat meningkatkan kemampuan model dalam membedakan mutu tomat serta mendukung prediksi potensi umur simpan secara lebih akurat.

3.3 Hasil Pelatihan Model

Model klasifikasi mutu tomat dikembangkan menggunakan algoritma Random Forest Classifier dengan memanfaatkan 16 fitur hasil ekstraksi yang terdiri atas fitur warna RGB, HSV, dan tekstur GLCM. Sebelum proses pelatihan, seluruh fitur dinormalisasi menggunakan StandardScaler untuk menyamakan skala data, kemudian dataset dibagi menggunakan metode train-test split dengan rasio 80:20 sehingga diperoleh 748 citra data latih dan 188 citra data uji. Model yang terbentuk mampu mempelajari karakteristik visual tomat untuk mengklasifikasikan mutu ke dalam tiga kategori, yaitu Baik, Sedang, dan Buruk.

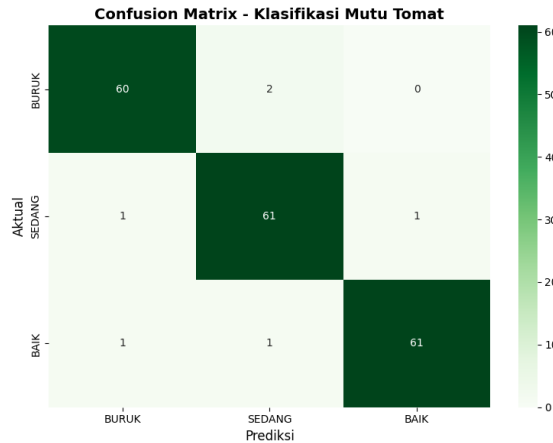
Selain klasifikasi mutu, penelitian ini juga mengembangkan model Random Forest Regressor untuk memprediksi potensi umur simpan berdasarkan fitur yang sama. Target regresi berupa nilai umur simpan kontinu yang disusun berdasarkan karakteristik visual setiap kelas mutu sehingga model dapat mempelajari hubungan antara perubahan warna, tekstur, dan estimasi umur simpan. Data dibagi dengan rasio yang sama, yaitu 80% untuk pelatihan dan 20% untuk pengujian sebelum dilakukan proses evaluasi model.

3.4 Evaluasi Model

3.4.1 Evaluasi Klasifikasi Mutu

Setelah proses pelatihan, performa model dievaluasi menggunakan confusion matrix dan metrik accuracy, precision, recall, serta F1-score. Hasil confusion matrix ditunjukkan pada Gambar 3. Berdasarkan Gambar 3, model berhasil mengklasifikasikan 182 dari 188 citra uji dengan benar atau mencapai akurasi 96,81%. Sebagian

besar prediksi berada pada diagonal utama matriks, yaitu 60 citra kelas buruk, 61 citra kelas sedang, dan 61 citra kelas baik. Kesalahan klasifikasi hanya terjadi pada enam citra yang umumnya berada pada kelas dengan karakteristik visual yang saling berdekatan, sedangkan kesalahan antara kelas mutu ekstrem relatif sangat sedikit. Hasil evaluasi kuantitatif disajikan pada Tabel 2.



Gambar 3. Confusion Matrix Klasifikasi Mutu Tomat

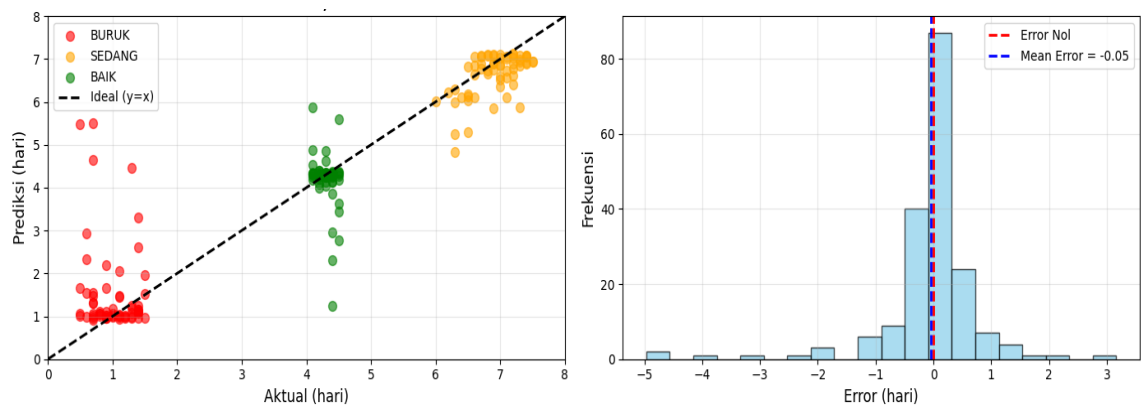
Tabel 2. Metrik Evaluasi Klasifikasi Mutu Tomat

Kelas	Precision	Recall	F-1 Score	Support
Baik	98.39%	96.83%	97.60%	63
Sedang	95.31%	96.83%	96.06%	63
Buruk	96.77%	96.77%	96.77%	62
Akurasi			96.81%	188

Berdasarkan Tabel 2, model Random Forest memperoleh akurasi 96,81%, dengan nilai precision, recall, dan F1-score masing-masing sebesar 96,82%, 96,81%, dan 96,81%. Nilai tersebut menunjukkan bahwa kombinasi fitur warna RGB, HSV, dan tekstur GLCM mampu merepresentasikan karakteristik mutu tomat secara efektif sehingga menghasilkan performa klasifikasi yang tinggi.

3.4.2 Evaluasi Prediksi Umur Simpan

Visualisasi hubungan antara nilai aktual dan nilai prediksi umur simpan ditunjukkan pada Gambar 4.



Gambar 4. Hasil Prediksi Umur Simpan

Berdasarkan Gambar 4, sebagian besar titik prediksi berada di sekitar garis ideal ($y = x$), yang menunjukkan bahwa model mampu menghasilkan estimasi umur simpan yang mendekati nilai aktual pada ketiga kelompok mutu tomat. Selain itu, histogram distribusi error memperlihatkan bahwa sebagian besar kesalahan prediksi terkonsentrasi di sekitar nilai nol, sehingga model tidak menunjukkan kecenderungan yang dominan untuk menghasilkan overestimation maupun underestimation. Hasil evaluasi regresi disajikan pada Tabel 3.

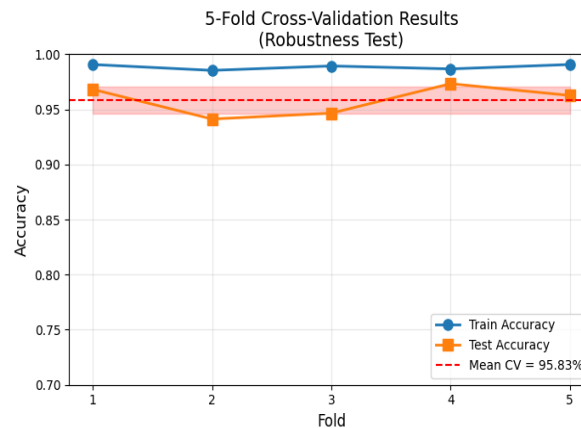
Tabel 3. Metrik Evaluasi Regresi Umur Simpan

Metrik	Nilai	Interpretasi
MAE	0,0621	Rata-rata error prediksi
RMSE	0,1152	Besarnya deviasi kesalahan prediksi
R	0,8752	87,52% variasi dapat dijelaskan

Berdasarkan Tabel 3, model Random Forest menghasilkan nilai MAE sebesar 0,0621 dan RMSE sebesar 0,1152 yang menunjukkan tingkat kesalahan prediksi yang relatif rendah. Selain itu, nilai koefisien determinasi (R^2) sebesar 0,8752 menunjukkan bahwa model mampu menjelaskan 87,52% variasi data umur simpan. Hasil tersebut mengindikasikan bahwa hubungan antara fitur warna RGB, HSV, serta tekstur GLCM dengan umur simpan tomat dapat dipelajari dengan baik oleh model Random Forest Regressor.

3.5 Hasil Cross Validation

Untuk menguji konsistensi dan kemampuan generalisasi model, dilakukan pengujian menggunakan 5-Fold Cross-Validation pada model Random Forest. Hasil pengujian pada setiap fold divisualisasikan pada Gambar 5.

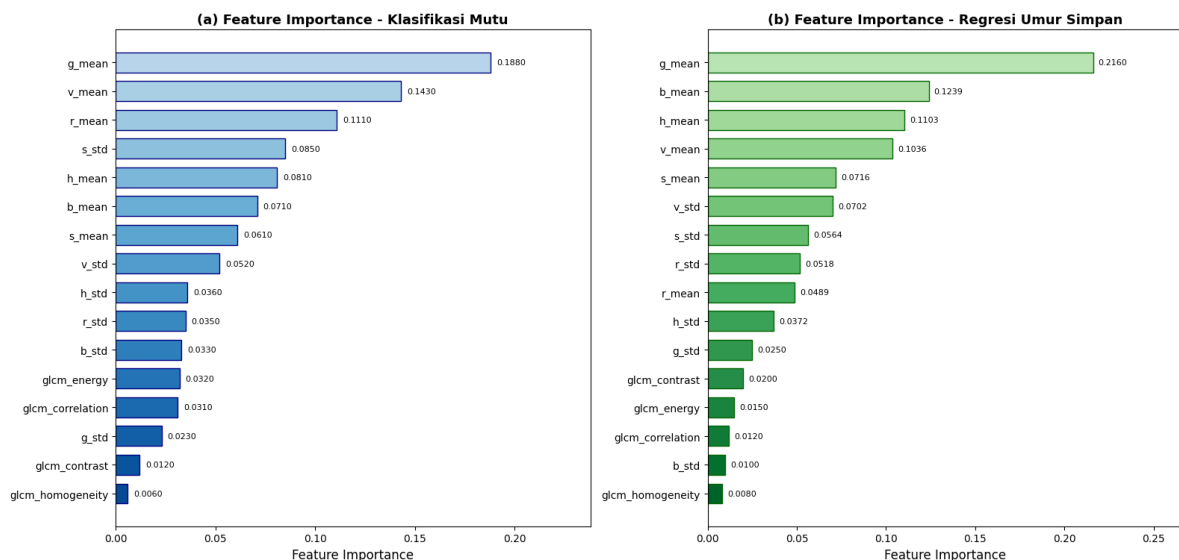


Gambar 5. Hasil 5-Fold Cross-Validation

Berdasarkan Gambar 5, nilai train accuracy berada pada rentang 98,53%–99,07%, sedangkan test accuracy berada pada rentang 94,12%–97,33%. Rata-rata test accuracy yang diperoleh sebesar 95,83% $\pm$ 1,24%, dengan nilai F1-score berkisar antara 0,9414–0,9733. Perbedaan yang relatif kecil antara performa data latih dan data uji menunjukkan bahwa model mampu mempelajari pola data dengan baik tanpa mengalami overfitting yang signifikan. Selain itu, standar deviasi yang rendah mengindikasikan bahwa variasi performa antar fold relatif kecil sehingga model memiliki tingkat stabilitas yang baik. Hasil tersebut juga konsisten dengan pengujian utama yang menghasilkan akurasi 96,81%, sehingga menunjukkan bahwa performa model tidak bergantung pada satu pembagian data tertentu. Temuan ini sejalan dengan penelitian Khultsum dan Subekti bahwa algoritma Random Forest memiliki kemampuan generalisasi yang baik pada klasifikasi berbasis fitur warna dan tekstur [13].

3.6 Analisis Feature Importance

Untuk memahami kontribusi masing-masing fitur terhadap proses klasifikasi mutu tomat dan prediksi umur simpan, dilakukan analisis feature importance pada model Random Forest. Analisis ini bertujuan untuk mengidentifikasi fitur yang paling berpengaruh dalam pengambilan keputusan model sehingga dapat diketahui karakteristik visual tomat yang memiliki hubungan paling kuat dengan mutu dan potensi umur simpan. Hasil analisis feature importance ditunjukkan pada Gambar 6.



Gambar 6. Feature Importance pada Random Forest

Berdasarkan Gambar 6, nilai feature importance yang lebih tinggi menunjukkan bahwa fitur tersebut memiliki peran yang lebih besar dalam proses pembentukan keputusan pada setiap pohon keputusan (decision tree). Berdasarkan hasil klasifikasi mutu, fitur g_mean memiliki nilai feature importance tertinggi sebesar 0,188, diikuti oleh v_mean (0,143), r_mean (0,111), s_std (0,085), dan h_mean (0,081). Sementara itu, fitur tekstur GLCM seperti energy (0,032), correlation (0,031), contrast (0,012), dan homogeneity (0,006) memberikan kontribusi yang relatif lebih kecil dibandingkan fitur warna. Hasil tersebut menunjukkan bahwa proses penentuan mutu tomat lebih banyak dipengaruhi oleh karakteristik warna, terutama komponen rata-rata kanal hijau (g_mean), yang merepresentasikan perubahan warna tomat selama proses pematangan dari hijau menuju oranye hingga merah. Temuan ini sejalan dengan penelitian Putra et al. [24] yang menyatakan bahwa kombinasi fitur RGB dan HSV mampu membedakan tingkat kematangan buah sawit berdasarkan perubahan karakteristik warna secara efektif, sehingga informasi warna menjadi faktor utama dalam proses identifikasi tingkat kematangan. Selain itu, penelitian Hadi et al. [23] juga menunjukkan bahwa fitur warna memberikan kontribusi yang lebih dominan dibandingkan fitur tekstur pada klasifikasi kematangan buah menggunakan ekstraksi RGB dan GLCM.

Pada model regresi umur simpan, fitur yang memiliki kontribusi terbesar adalah g_mean (0,2160), diikuti oleh b_mean (0,1239), h_mean (0,1103), v_mean (0,1036), dan s_mean (0,0716). Hasil tersebut menunjukkan bahwa estimasi umur simpan tomat juga lebih dipengaruhi oleh informasi warna dibandingkan tekstur. Dominasi fitur g_mean mengindikasikan bahwa proporsi warna hijau masih menjadi indikator penting dalam memprediksi potensi umur simpan, karena tomat dengan intensitas hijau yang lebih tinggi umumnya berada pada tingkat kematangan yang lebih rendah sehingga memiliki masa simpan yang lebih panjang. Sebaliknya, semakin dominan warna merah dan semakin tinggi tingkat kecerahan buah, umur simpan cenderung semakin pendek akibat proses pematangan yang telah berlangsung lebih lanjut. Temuan ini didukung oleh penelitian Mirfan et al. [25] yang menyatakan bahwa kombinasi fitur warna HSV dan tekstur GLCM mampu merepresentasikan karakteristik kematangan buah secara efektif, namun informasi warna memiliki peran utama dalam membedakan tingkat kematangan dibandingkan tekstur permukaan. Dengan demikian, hasil analisis feature importance pada penelitian ini menunjukkan bahwa fitur warna, khususnya g_mean , merupakan fitur yang paling berpengaruh baik pada proses klasifikasi mutu maupun prediksi umur simpan tomat, sedangkan fitur tekstur GLCM berperan sebagai fitur pendukung untuk meningkatkan kemampuan model dalam mengenali karakteristik visual tomat secara lebih komprehensif.

3.7 Implementasi Sistem

Model yang telah dikembangkan kemudian diimplementasikan dalam bentuk aplikasi web sederhana menggunakan framework Gradio. Antarmuka sistem terdiri atas tiga menu utama, yaitu Panduan, Sistem Utama, dan History. Menu Panduan berisi petunjuk penggunaan aplikasi, menu Sistem Utama digunakan untuk melakukan proses analisis citra tomat, sedangkan menu History menyimpan riwayat hasil prediksi yang telah dilakukan. Implementasi ini bertujuan untuk mempermudah pengguna dalam mengakses sistem klasifikasi mutu dan prediksi potensi umur simpan melalui antarmuka yang sederhana dan mudah digunakan. Implementasi sistem ditunjukkan pada Gambar 7.



Gambar 7. Implementasi Sistem Klasifikasi Mutu Tomat dan Potensi Umur Simpan Berbasis Gradio

Pada menu Sistem Utama, pengguna cukup mengunggah citra tomat, kemudian sistem secara otomatis melakukan proses pre-processing, ekstraksi fitur warna RGB, HSV, dan tekstur GLCM, dilanjutkan dengan prediksi menggunakan model Random Forest Classifier dan Random Forest Regressor yang telah dilatih

sebelumnya. Hasil analisis ditampilkan dalam bentuk kategori mutu tomat, estimasi potensi umur simpan, tingkat keyakinan (confidence score), probabilitas setiap kelas mutu, serta rekomendasi penanganan berdasarkan hasil prediksi. Hasil tersebut menunjukkan bahwa model yang dikembangkan tidak hanya memiliki performa yang baik pada tahap evaluasi, tetapi juga dapat diimplementasikan dalam bentuk aplikasi yang mudah dioperasikan dan berpotensi mendukung proses penilaian mutu tomat secara lebih objektif, cepat, dan efisien.

3.8 Pembahasan

Hasil penelitian menunjukkan bahwa kombinasi fitur warna RGB, HSV, dan tekstur GLCM yang diproses menggunakan algoritma Random Forest mampu menghasilkan performa klasifikasi mutu tomat yang tinggi dengan akurasi sebesar 96,81%. Nilai tersebut lebih tinggi dibandingkan penelitian Fatah et al. [5] yang menggunakan fitur warna RGB dan metode K-Nearest Neighbor dengan tingkat keberhasilan sebesar 66,66%, serta penelitian Faradita dan Harahap [11] yang menerapkan Jaringan Saraf Tiruan berbasis histogram warna dengan tingkat keberhasilan sebesar 43,33%. Peningkatan performa ini menunjukkan bahwa kombinasi fitur warna dan tekstur mampu memberikan representasi visual yang lebih lengkap dibandingkan penggunaan fitur warna saja.

Dibandingkan penelitian Iskandar et al. [12] yang menggunakan fitur tekstur GLCM dan Support Vector Machine untuk klasifikasi kematangan buah nanas dengan akurasi sebesar 86%, penelitian ini menghasilkan akurasi yang lebih tinggi. Hal tersebut mengindikasikan bahwa informasi warna memiliki peran penting dalam membedakan tingkat mutu tomat karena perubahan warna merupakan indikator utama proses pematangan buah. Temuan ini diperkuat oleh hasil feature importance yang menunjukkan bahwa fitur warna, khususnya g_mean , memiliki kontribusi terbesar dalam proses klasifikasi maupun prediksi umur simpan.

Hasil penelitian ini juga sejalan dengan penelitian Khultsum dan Subekti [13] yang memperoleh akurasi sekitar 96% menggunakan Random Forest pada klasifikasi penyakit daun tomat berbasis kombinasi fitur warna dan tekstur. Kesamaan hasil tersebut menunjukkan bahwa algoritma Random Forest memiliki kemampuan yang baik dalam menangani data citra pertanian yang memiliki karakteristik visual kompleks. Namun demikian, penelitian ini memiliki perbedaan karena berfokus pada mutu buah tomat sekaligus mengintegrasikan prediksi potensi umur simpan dalam satu sistem.

Selain menghasilkan performa klasifikasi yang tinggi, penelitian ini juga mampu memprediksi potensi umur simpan dengan nilai R^2 sebesar 0,8752. Hasil tersebut menunjukkan bahwa karakteristik visual tomat memiliki hubungan yang kuat dengan umur simpan buah. Integrasi klasifikasi mutu dan prediksi umur simpan menjadi kontribusi utama penelitian ini karena sebagian besar penelitian terdahulu hanya berfokus pada klasifikasi mutu atau tingkat kematangan tanpa mengestimasi umur simpan secara simultan. Dengan demikian, sistem yang dikembangkan tidak hanya mampu melakukan identifikasi mutu secara otomatis, tetapi juga memberikan informasi tambahan yang dapat dimanfaatkan dalam proses sortasi, penyimpanan, dan distribusi tomat secara lebih efektif.

4. KESIMPULAN

Penelitian ini berhasil mengembangkan sistem klasifikasi mutu tomat dan prediksi potensi umur simpan berbasis pengolahan citra digital menggunakan algoritma Random Forest dengan memanfaatkan kombinasi fitur warna RGB, HSV, dan tekstur Gray Level Co-occurrence Matrix (GLCM). Hasil penelitian menunjukkan bahwa kombinasi fitur tersebut mampu merepresentasikan karakteristik visual tomat secara efektif sehingga menghasilkan performa klasifikasi yang tinggi dengan akurasi sebesar 96,81%, precision 96,82%, recall 96,81%, dan F1-score 96,81%. Pada prediksi potensi umur simpan, model regresi menghasilkan nilai Mean Absolute Error (MAE) sebesar 0,0621, Root Mean Square Error (RMSE) sebesar 0,1152, dan koefisien determinasi (R^2) sebesar 0,8752, yang menunjukkan bahwa model mampu mempelajari hubungan antara karakteristik visual tomat dan estimasi umur simpan dengan baik. Hasil analisis feature importance memperlihatkan bahwa fitur warna, khususnya g_mean , memberikan kontribusi paling besar baik pada proses klasifikasi mutu maupun prediksi potensi umur simpan, sehingga mengindikasikan bahwa perubahan warna merupakan indikator utama dalam penentuan kualitas tomat. Selain itu, pengujian menggunakan 5-Fold Cross-Validation menghasilkan rata-rata akurasi $95,83\% \pm 1,24\%$, yang menunjukkan bahwa model memiliki performa yang stabil dan kemampuan generalisasi yang baik. Model yang dikembangkan juga telah berhasil diimplementasikan dalam bentuk aplikasi web berbasis Gradio sehingga dapat digunakan untuk melakukan klasifikasi mutu dan estimasi potensi umur simpan secara otomatis melalui antarmuka yang sederhana. Meskipun demikian, penelitian ini masih memiliki keterbatasan karena dataset diperoleh dari proses pengambilan citra dengan kondisi pencahayaan dan latar belakang yang relatif terkontrol serta estimasi umur simpan didasarkan pada rentang acuan dari penelitian sebelumnya. Oleh karena itu, penelitian selanjutnya dapat memperluas variasi dataset pada berbagai kondisi lingkungan, menggunakan data umur simpan hasil pengamatan langsung, serta mengeksplorasi metode deep learning atau pendekatan hibrida untuk meningkatkan akurasi dan kemampuan adaptasi sistem pada kondisi nyata.

REFERENCES

- [1] S. W. Agusta and W. Kaswidjanti, "The Implementation of Color Feature Extraction and Gray Level Co- occurrence



- Matrix Combination in K-Nearest Neighbor Classification Method for Tomato Leaf Disease Identification,” *Telemat. J. Inform. dan Teknol. Inf.*, vol. 20, no. 2, pp. 250–262, 2023, doi: 10.31515/telematika.v20i2.10009.
- [2] B. P. Statistik, “Statistik Hortikultura 2023,” Jakarta, 2023. [Online]. Available: <https://www.bps.go.id/id/statistics-table/2/NjEjMg==/production-of-vegetables.html> 1.143.788 ton (1,02 jt ton)
- [3] K. Pertanian, Statistik Pertanian (Agricultural Statistics) 2024. Jakarta: Pusat Data dan Sistem Informasi Pertanian Kementerian Pertanian Republik Indonesia, 2024.
- [4] M. B. Sulthan and I. Wahyudi, “Klasifikasi Kematangan Buah Tomat Menggunakan Fitur RGB dan HSI Berbasis Backpropagation,” *J. Apl. Teknol. Inf. dan Manaj.*, vol. 4, no. 1, pp. 84–95, 2023, doi: <https://doi.org/10.31102/jatim.v4i1.2177>.
- [5] I. R. M. Fatah, A. H. Ginting, and W. T. Ina, “Klasifikasi Tingkat Kematangan Buah Tomat Berdasarkan Warna,” *J. Tek. Elektro*, vol. 1, no. 1, pp. 20–25, 2024, [Online]. Available: <https://elektro.ejournal.web.id/index.php/elektro/article/view/112>
- [6] J. Rusman, B. Z. Haryati, and A. Michael, “Optimisasi Hiperparameter Tuning Pada Metode Support Vector Untuk Klasifikasi Tingkat Kematangan Buah Kopi,” *J-Icon J. Inform. dan Komput.*, vol. 11, no. 2, pp. 195–202, 2023, doi: 10.35508/jicon.v11i2.12571.
- [7] D. A. Salim, Y. B. Roza, A. Ramadhanu, and Putra, “Evaluasi Kualitas dan Kematangan Mangga Menggunakan Analisis Citra Digital dengan Euclidean Distance Fokus pada Buah Hijau dan Kuning,” *Indones. J. Comput. Sci.*, vol. 3, no. 2, pp. 57–64, 2024, doi: <https://doi.org/10.31294/jics.v3i2>.
- [8] F. Santoso and E. Hartati, “Penggunaan Algoritma Random Forest dalam Klasifikasi Buah Segar dan Busuk,” *Algoritma. J. Mhs. Tek. Inform.*, vol. 3, no. 1, pp. 133–140, 2022, doi: <https://doi.org/10.35957/algoritme.v3i1.3404>.
- [9] C. Nyasulu, A. Diattara, A. Traore, and C. Ba, “A comparative Study of Machine Learning-based Classification of Tomato Fungal Diseases : Application of GLCM Texture Features,” *Heliyon*, vol. 9, no. 11, p. e21697, 2023, doi: 10.1016/j.heliyon.2023.e21697.
- [10] A. Setiyawan, D. Syauqy, and S. R. Akbar, “Rancang Bangun Sistem Klasifikasi Padi Siap Dipanen dengan Parameter Warna Padi dan Warna Daun Menggunakan Metode Random Forest,” *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 6, no. 7, pp. 3228–3235, 2022.
- [11] N. A. Faradita and L. S. Harahap, “Identifikasi Tingkat Kematangan Buah Tomat Melalui Warna dengan Penerapan Jaringan Saraf Tiruan (JST),” *Polyg. J. Ilmu Komput. dan Ilmu Pengetah. Alam*, vol. 2, no. 6, pp. 71–78, 2024, doi: <https://doi.org/10.62383/polygon.v2i6.292>.
- [12] Y. Akbar, D. Iskandar, and F. C. W. S, “Klasifikasi Tingkat Kematangan Buah Nanas Berdasarkan Tekstur Gray Level CO- Occurencematrix dengan Metode Support Vector Machine,” *Remik Ris. dan E-Jurnal Manaj. Inform. Komput.*, vol. 8, no. 1, pp. 371–382, 2024, doi: <http://doi.org/10.33395/remik.v8i1.13403>.
- [13] U. Khultsum and A. Subekti, “Penerapan Algoritma Random Forest dengan Kombinasi Ekstraksi Fitur Untuk Klasifikasi Penyakit Daun Tomat,” *J. Media Inform. Budidarma*, vol. 5, no. 1, pp. 186–193, 2021, doi: 10.30865/mib.v5i1.2624.
- [14] A. P. Argadianata, D. A. Fatah, and H. Sukri, “Klasifikasi Kualitas Buah Apel Menggunakan Metode Random Forest,” *JATI (Jurnal Mhs. Tek. Inform.*, vol. 9, no. 2, pp. 2016–2022, 2025, doi: 10.36040/jati.v9i2.12854.
- [15] S. Lina, M. Sitio, S. Kom, and M. Kom, *Machine Learning Berbasis Pohon Keputusan: Implementasi Decision Tree dan Random Forest di Google Colab*. Cilacap: Media Pustaka Indo, 2026.
- [16] D. Rofianto, K. Amaliah, T. K. Khoerunissa, and M. Fitri, “Metode Ekstraksi Fitur dan Klasifikasi Visual Untuk Identifikasi Kualitas Pangan Lokal Berbasis Citra Digital,” *J. Ilmu Komput. Agri-Informatika*, vol. 12, no. 2, pp. 178–188, 2025, doi: 10.29244/jika.12.2.178-188.
- [17] E. Dadebo and W. W. Eyses, “Effect of Harvesting Stage and Types of Storage on the Quality and Shelf- Life of Tomato Fruit,” *J. Agric. Sci. Food Res.*, vol. 13, no. 4, pp. 1–8, 2022, doi: 10.35248/2593-9173.22.13.501.
- [18] W. Deglas, “Pengaruh Suhu Penyimpanan dan Tingkat Kematangan terhadap Umur Simpan Buah Tomat,” *J. Ilmu Pangan dan Has. Pertan.*, vol. 7, no. 1, pp. 49–60, 2023, doi: 10.26877/jiphp.v7i1i.15460.
- [19] R. Fertiasari, S. Arditian, S. Yuliani, N. Nurhafiza, and P. Aryasari, “Perubahan Fisiologi Buah Tomat (*Solanum Lycopersicum*) terhadap Suhu Kamar dan Umur Simpan yang Memengaruhi Mutu,” *J. Food Secur. Agroindustry*, vol. 1, no. 3, pp. 97–104, 2023, doi: 10.58184/jfsa.v1i3.125.
- [20] R. Ramli, Evanita, and A. A. Riadi, “Classification of Rice Leaf Diseases Using Support Vector Machine with HSV and GLCM-Based Feature Extraction,” *J. Appl. Informatics Comput.*, vol. 9, no. 5, pp. 2329–2337, 2025, doi: <https://doi.org/10.30871/jaic.v9i5.10403>.
- [21] A. Septiarini, H. Hamdani, Y. Khoemah, N. Puspitasari, and M. Wati, “Segmentasi Tomat Menggunakan Metode K-Means Clustering dan Pengolahan Citra Digital,” in *Seminar Nasional Corisindo*, 2022, pp. 113–118.
- [22] R. Dijaya, *Buku Ajar Pengolahan Citra Digital*, vol. 11, no. 1. Sidoarjo, 2023. [Online]. Available: <https://doi.org/10.21070/2023/978-623-464-075-5>
- [23] H. P. Hadi and E. H. Rachmawanto, “Ekstraksi Fitur Warna dan GLCM pada Algoritma KNN untuk Klasifikasi Kematangan Rambutan,” *JIP (Jurnal Inform. Polinema)*, vol. 8, no. 3, pp. 63–68, 2022, doi: <https://doi.org/10.33795/jip.v8i3.949>.
- [24] R. P. Putra, J. Jumadi, and D. Lianda, “Pengolahan Citra Digital Untuk Mengidentifikasi Tingkat Kematangan Buah Kelapa Sawit Berdasarkan Warna RGB Dan HSV dengan Menggunakan Metode Self Organizing Map (SOM),” *J. Media Infotama*, vol. 20, no. 1, pp. 98–105, 2024, doi: 10.34128/jsi.v6i2.242.
- [25] B. Eden, W. Asrul, and M. Jumarlis, “Automated Ripeness Detection of Oil Palm Fruit Using a Hybrid GLCM-HSV-KNN Model,” *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 9, no. 5, pp. 1066–1077, 2025, doi: <https://doi.org/10.29207/resti.v9i5.6683>.