

PENERAPAN METODE HYBRID KNN DAN PCA DALAM KLASIFIKASI CABAI MERAH, CABAI HIJAU DAN JERUK MANDARIN

Nabilla Yasmin, Agung Ramadhanu

Magister Teknik Informatika, Universitas Putra Indonesia YPTK Padang
Jl. Raya Lubuk Begalung, Lubuk Begalung Nan XX, Padang, Indonesia
nbilayasmin2008@mail.com

ABSTRAK

Sistem klasifikasi otomatis berbasis citra digital semakin dibutuhkan dalam sektor pertanian untuk meningkatkan efisiensi seleksi dan distribusi produk. Pengklasifikasian cabai merah, cabai hijau, dan jeruk mandarin secara manual sering kali memakan waktu lama dan rentan terhadap kesalahan manusia. Oleh karena itu, penelitian ini mengusulkan penggunaan kombinasi Principal Component Analysis (PCA) dan K-Nearest Neighbors (KNN) untuk mengklasifikasikan ketiga jenis objek tersebut. Data yang digunakan terdiri dari 16 citra latih dan 10 citra uji, yang diproses melalui tahapan segmentasi, ekstraksi fitur, dan reduksi dimensi menggunakan PCA. Hasil klasifikasi menunjukkan tingkat akurasi mencapai 96,20%, dengan satu data jeruk mandarin yang gagal terdeteksi. Temuan ini membuktikan bahwa metode PCA dan KNN efektif dalam mengklasifikasikan citra dengan akurasi tinggi, memberikan kontribusi terhadap sistem klasifikasi otomatis yang dapat meningkatkan efisiensi dalam distribusi dan seleksi produk agrikultur serta mengurangi potensi kesalahan dalam penilaian kualitas produk secara manual.

Kata kunci : *Hybrid, K-Nearest Neighbors (KNN), Principal Component Analysis (PCA), Cabai Merah, Cabai Hijau, Jeruk Mandarin*

1. PENDAHULUAN

Indonesia dikenal sebagai negara agraris, di mana sektor pertanian memegang peranan strategis, khususnya dalam aktivitas berbasis tanaman pangan dan hortikultura [1].

Sektor ini tidak hanya melibatkan tenaga kerja dalam jumlah besar, tetapi juga menghasilkan produk yang menjadi bahan pangan utama bagi konsumsi nasional [2].

Salah satu komoditas penting dalam pertanian Indonesia adalah cabai (*Capsicum annum* L.), tanaman yang berasal dari Amerika Selatan dan termasuk ke dalam keluarga Solanaceae. Cabai memiliki nilai ekonomi yang tinggi baik untuk pasar ekspor maupun kebutuhan domestik [3].

Selain menjadi bahan utama dalam berbagai masakan tradisional, cabai juga kaya akan nutrisi seperti kalori, protein, lemak, kalsium, serta vitamin A, B1, dan C, yang menjadikannya sebagai komoditas bernutrisi dan bernilai tinggi [4].

Kajian literatur menunjukkan bahwa cabai mengandung berbagai senyawa bioaktif, dengan capsaicin sebagai komponen utama dari kelompok Capsaicinoid. Senyawa lainnya meliputi dihydrocapsaicin, homocapsaicin, nonivamide, nordihydrocapsaicin, homodihydrocapsaicin, dan nornordihydrocapsaicin. Senyawa-senyawa tersebut diketahui memiliki potensi untuk pengobatan berbagai kondisi, seperti nyeri otot, rematik, batuk, dan infeksi pada saluran pencernaan [5].

Selain manfaat di bidang medis, cabai juga menjadi bahan penting dalam industri makanan, baik sebagai sayuran maupun bumbu masakan [6].

Cabai keriting, misalnya, memiliki peluang ekspor yang besar, sehingga dapat berkontribusi pada

peningkatan pendapatan petani serta menciptakan lapangan kerja baru [7].

Selain cabai, jeruk juga merupakan komoditas hortikultura unggulan di Indonesia. Jeruk mandarin (*Citrus reticulata*) dikenal karena kandungan vitamin C yang tinggi serta produktivitasnya yang unggul. Berdasarkan data dari Badan Pusat Statistik Jawa Timur, produksi jeruk keprok (mandarin) pada tahun 2022 mencapai lebih dari 11 juta kuintal, menjadikannya salah satu jenis jeruk yang paling banyak dibudidayakan di Indonesia [8]. Jeruk mandarin memiliki karakteristik pertumbuhan yang cepat dan hasil panen yang melimpah dibandingkan dengan jenis jeruk lainnya [9].

Di era digital seperti saat ini, perkembangan teknologi telah memberikan kontribusi besar di berbagai bidang, termasuk pertanian [10].

Salah satu teknologi yang banyak digunakan adalah pengolahan citra digital, yang bertujuan mengelola piksel dalam gambar digital untuk memperoleh informasi tertentu sesuai kebutuhan [11].

Dalam konteks agrikultur dan industri makanan, klasifikasi berbasis citra memegang peranan penting, terutama dalam skala produksi besar. Klasifikasi manual seringkali membutuhkan waktu lama dan rentan terhadap kesalahan manusia, sehingga pengembangan sistem klasifikasi otomatis berbasis citra menjadi sangat diperlukan [12].

Penelitian ini menggunakan algoritma K-Nearest Neighbors (KNN) sebagai metode klasifikasi karena keunggulannya dalam menangani data non-linear. Namun, performa KNN sering kali terhambat oleh dimensi data yang tinggi, yang dapat menyebabkan overfitting dan peningkatan beban komputasi [13].

Oleh karena itu, Principal Component Analysis (PCA) digunakan sebagai teknik reduksi dimensi untuk menyederhanakan data tanpa kehilangan informasi penting [14].

Kombinasi metode PCA dan KNN diharapkan mampu memberikan solusi yang efisien dan akurat dalam klasifikasi citra cabai merah, cabai hijau, dan jeruk mandarin. Penelitian ini bertujuan untuk mengevaluasi performa metode PCA dan KNN dalam pengelompokan ketiga jenis objek tersebut. Dengan sistem klasifikasi otomatis ini, diharapkan proses seleksi dan pengelompokan produk dapat dilakukan secara lebih efisien, sehingga mendukung peningkatan produktivitas di sektor agrikultur dan distribusi komoditas.

2. TINJAUAN PUSTAKA

Penelitian ini didukung oleh berbagai referensi yang berkaitan dengan pengolahan citra digital, metode klasifikasi, serta penerapan algoritma PCA dan KNN dalam bidang agrikultur. Berikut adalah beberapa konsep utama yang menjadi dasar penelitian ini:

2.1. Metode PCA dan KNN

Principal Component Analysis (PCA) merupakan metode untuk mereduksi dimensi atribut dalam dataset, sehingga nilai yang terbentuk sangat berbeda dengan bentuk aslinya. Metode PCA digunakan untuk meringkas struktur dari dataset dengan dimensi yang banyak sehingga memiliki jumlah variabel yang lebih kecil. Kegunaan PCA selain untuk mereduksi dimensi yaitu dapat digunakan sebagai metode untuk menguji apakah setiap variabel dalam dataset saling terkait atau tidak terkait sama sekali [15].

Hasil reduksi dimensi kemudian menjadi input bagi algoritma K-Nearest Neighbors (KNN), yang digunakan untuk mengelompokkan citra berdasarkan kemiripan fitur. K-Nearest Neighbor merupakan algoritma klasifikasi objek berdasarkan data pembelajaran yang terdekat jarak atau mempunyai ciri-ciri benda yang paling umum. Ada juga beberapa langkah untuk menghitung algoritma K-NN, menentukan nilai K, menghitung jarak Euclidean (query instance) untuk setiap objek dari data pelatihan. Menyusun objek ke dalam kelompok-kelompok yang mempunyai jarak Euclidean terkecil. Kumpulkan label kelas Y (klasifikasi tetangga terdekat) [16].

Benda jauh atau dekat dapat dihitung dengan Euclidean jarak, dimana dua jarak vektor berukuran n , misalnya $x = (x_1, x_2, x_3, \dots, x_n)$ dan $y = (y_1, y_2, y_3, \dots, y_n)$ kita peroleh Persamaan (1) sebagai berikut:

$$Dist(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

2.2. Objek Penelitian

Objek penelitian ini adalah cabai merah, cabai hijau, dan jeruk mandarin dipilih sebagai objek karena

merupakan komoditas penting dalam sektor agrikultur Indonesia. Klasifikasi otomatis dapat meningkatkan efisiensi dalam seleksi produk dan mengurangi kesalahan dalam identifikasi visual.

2.3. Flowchart Metode

Tahapan penelitian disusun dalam bentuk flowchart untuk memperjelas proses dari pengumpulan data hingga klasifikasi akhir yang terlihat pada gambar 1.



Gambar 1. Flowchart Penelitian

3. METODE PENELITIAN

Metode yang digunakan untuk mengembangkan sistem klasifikasi otomatis berbasis pengolahan citra digital adalah dengan metode PCA dan KNN yang mampu membedakan citra cabai merah, cabai hijau, dan jeruk mandarin.

3.1. Tahapan Metode Penelitian

Tahapan penelitian meliputi pengumpulan data citra menggunakan perangkat digital dengan standar resolusi tinggi, preprocessing untuk meningkatkan kualitas citra menggunakan teknik seperti median filter dan transformasi ruang warna, serta segmentasi citra untuk memisahkan objek dari latar belakang. Fitur warna dan tekstur diekstraksi untuk merepresentasikan karakteristik unik setiap objek yang akan digunakan dalam klasifikasi.

Implementasi metode dilakukan dengan perangkat lunak MATLAB, melalui tahapan berikut:

3.2. Pengumpulan Data

Data citra yang digunakan dalam penelitian ini dikumpulkan menggunakan perangkat digital dengan resolusi tinggi untuk memastikan kualitas gambar yang optimal. Pengambilan citra dilakukan dalam kondisi pencahayaan seragam guna meminimalkan noise dan variasi warna yang dapat mempengaruhi proses klasifikasi. Dataset terdiri dari citra cabai merah, cabai hijau, dan jeruk mandarin yang diambil dari berbagai sudut dan jarak tertentu untuk

meningkatkan representasi data yang lebih akurat dalam sistem klasifikasi.

3.3. Preprocessing (Load Citra)

Citra digital cabai merah, cabai hijau, dan jeruk mandarin dimuat ke dalam sistem sebagai dataset penelitian. Citra diambil dalam kondisi pencahayaan seragam untuk meminimalkan noise. Setiap citra disimpan dalam format RGB dengan ukuran resolusi yang seragam.

3.4. Segmentasi

Segmentasi dilakukan untuk memisahkan objek utama dari latar belakang citra. Proses ini bertujuan untuk fokus pada area yang relevan, sehingga fitur yang diekstraksi benar-benar merepresentasikan karakteristik objek. Teknik segmentasi berbasis thresholding digunakan untuk mempermudah pemisahan berdasarkan perbedaan warna dan intensitas piksel.

3.5. Ekstraksi Ciri

Setelah segmentasi, ciri-ciri utama dari objek citra diekstraksi. Fitur yang diekstraksi meliputi karakteristik warna (mean, standar deviasi warna) dan tekstur (GLCM - Gray Level Co-occurrence Matrix). Data fitur yang dihasilkan digunakan sebagai representasi numerik dari masing-masing objek.

3.6. Klasifikasi

Data hasil ekstraksi ciri direduksi dimensinya menggunakan Principal Component Analysis (PCA) untuk menyederhanakan struktur data dan menghilangkan redundansi. Hasil dari PCA kemudian menjadi input bagi algoritma K-Nearest Neighbors (KNN), yang digunakan untuk mengelompokkan citra berdasarkan kesamaan fitur. KNN bekerja dengan mengukur jarak antara data uji dan data latih, menggunakan parameter jumlah tetangga terdekat (k).

3.7. Pengujian

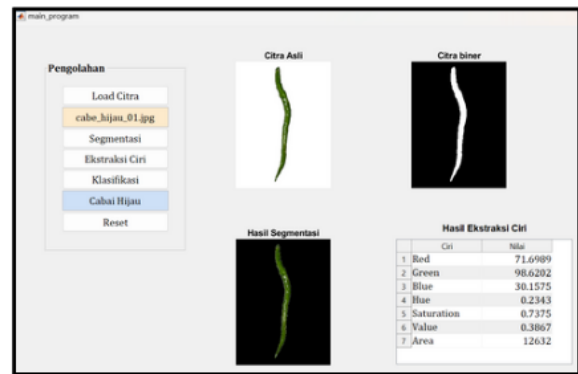
Pengujian dilakukan untuk mengetahui tingkat akurasi dari implementasi algoritma PCA dan KNN. Akurasi merupakan tingkat kedekatan nilai prediksi terhadap nilai perhitungan yang sebenarnya. Perhitungan akurasi untuk mengetahui tingkat kebenaran prediksi didapatkan menggunakan persamaan sebagai berikut :

$$\text{Akurasi} = \frac{\text{Jumlah Data yang Terklasifikasi Benar}}{\text{Total Data Uji}} \times 100\%$$

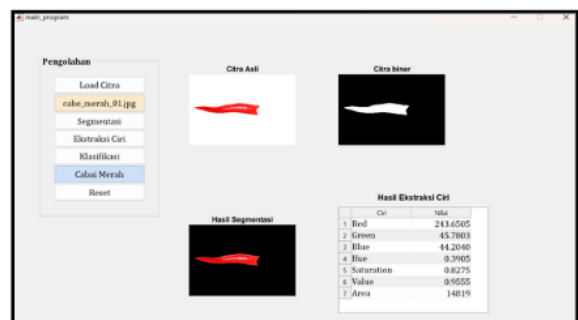
(2)

4. HASIL DAN PEMBAHASAN

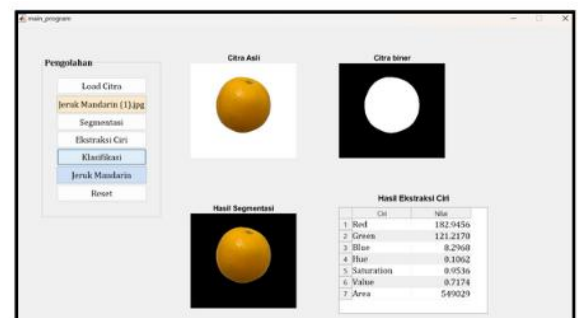
Pada tahap pengujian, hasil klasifikasi citra yang dihasilkan oleh sistem dapat dilihat dalam bentuk output yang ditampilkan melalui MATLAB pada Gambar berikut.



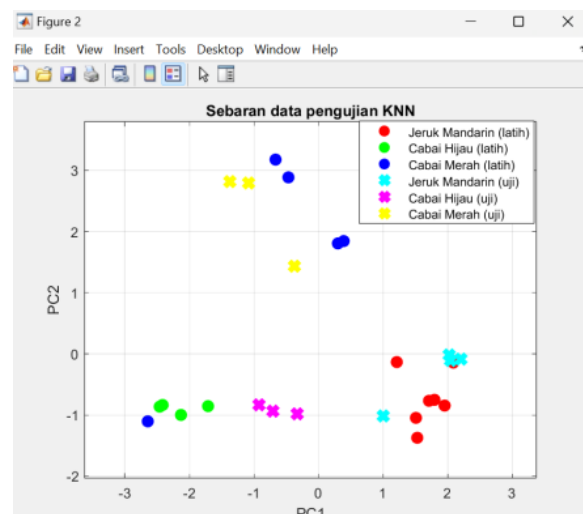
Gambar 2. Hasil Klasifikasi Citra Cabai Hijau



Gambar 3. Hasil Klasifikasi Citra Cabai Hijau



Gambar 4. Hasil Klasifikasi Citra Cabai Hijau



Gambar 5. Hasil Sebaran Data Pengujian KNN

Gambar-gambar tersebut menunjukkan hasil segmentasi dan klasifikasi berdasarkan fitur warna dan tekstur yang telah diekstraksi. Penelitian ini menggunakan total 26 citra, yaitu 16 data latih, yang terdiri dari 7 citra jeruk mandarin, 4 citra cabai merah, dan 5 citra cabai hijau, serta 10 data uji, yang terdiri dari 4 citra jeruk mandarin, 3 citra cabai merah, dan 3 citra cabai hijau.

Pada gambar tersebut, masing-masing citra diberi label yang sesuai dengan kelas yang diprediksi oleh sistem. Hasil pengujian data citra dapat dilihat pada Tabel 1 dan Tabel 2.

Tabel 1. Tabel Data Latih

Jenis Citra	Jumlah Data Citra	Jumlah Data Citra Berhasil	Persentase (%)
Jeruk Mandarin	7 data	7 data	100%
Cabai Merah Keriting	4 data	4 data	100%
Cabai Hijau Keriting	5 data	5 data	100%

Tabel 2. Tabel Data Uji

Jenis Citra	Jumlah Data Citra	Jumlah Data Citra Berhasil	Persentase (%)
Jeruk Mandarin	4 data	3 data	75%
Cabai Merah Keriting	3 data	3 data	100%
Cabai Hijau Keriting	3 data	3 data	100%

Berdasarkan hasil klasifikasi menggunakan metode kombinasi Principal Component Analysis (PCA) dan K-Nearest Neighbors (KNN), sistem berhasil mengidentifikasi sebagian besar citra dengan akurasi tinggi dengan data citra yang terklasifikasi sebanyak 25 data citra dari total 26 data citra.

Perhitungan Akurasi citra sebagai berikut

$$Akurasi = \frac{25}{26} \times 100\% = 96,2\%$$

(3)

Hasil evaluasi sistem klasifikasi menunjukkan tingkat akurasi yang sangat baik, yaitu sebesar 96,20%. Angka ini menggambarkan kemampuan sistem dalam mengklasifikasikan citra dengan benar, baik pada data latih maupun data uji. Dengan akurasi yang tinggi, sistem berhasil mengidentifikasi cabai merah, cabai hijau, dan jeruk mandarin dengan tingkat kesalahan yang sangat rendah. Hasil ini menunjukkan bahwa metode Principal Component Analysis (PCA) yang dipadukan dengan K-Nearest Neighbors (KNN) dapat memberikan performa yang efisien dan akurat dalam klasifikasi citra digital untuk objek agrikultur. Meskipun ada sedikit kesalahan klasifikasi pada citra jeruk mandarin, secara keseluruhan, sistem ini dapat

diandalkan untuk aplikasi praktis dalam seleksi dan pengelompokan produk agrikultur.

Analisis lebih lanjut menunjukkan bahwa metode PCA berhasil mereduksi dimensi data tanpa kehilangan informasi penting, memungkinkan KNN untuk melakukan klasifikasi dengan tingkat presisi yang tinggi. Namun, hasil juga mengindikasikan bahwa pencahayaan yang tidak seragam dapat memengaruhi ekstraksi fitur warna, yang pada akhirnya berdampak pada akurasi klasifikasi.

Sebagai upaya perbaikan, optimasi segmentasi citra dapat diterapkan dengan teknik yang lebih adaptif agar objek dapat dipisahkan dengan lebih presisi dari latar belakangnya. Selain itu, penggunaan parameter K yang lebih optimal dalam KNN dapat diuji lebih lanjut untuk meningkatkan akurasi klasifikasi, terutama dalam menangani objek dengan karakteristik visual yang mirip.

5. KESIMPULAN DAN SARAN

Berdasarkan penelitian yang telah dilakukan, dapat disimpulkan bahwa kombinasi Principal Component Analysis (PCA) dan K-Nearest Neighbors (KNN) efektif dalam klasifikasi citra cabai merah, cabai hijau, dan jeruk mandarin. Sistem yang dikembangkan menunjukkan tingkat akurasi yang sangat baik, yaitu 96,2%, dengan data citra yang berhasil diklasifikasikan sebanyak 25 dari total 26 data citra. Hasil klasifikasi pada data latih menunjukkan tingkat akurasi sebesar 100% untuk semua kategori objek, sementara hasil klasifikasi pada data uji menunjukkan bahwa 3 dari 4 citra jeruk mandarin diklasifikasikan dengan benar, menghasilkan akurasi 75%, sedangkan cabai merah dan cabai hijau masing-masing memiliki akurasi 100%. Meskipun demikian, terdapat satu kasus di mana citra jeruk mandarin tidak terdeteksi dengan baik, yang mengindikasikan perlunya perbaikan pada tahap segmentasi dan ekstraksi fitur. Evaluasi hasil menunjukkan bahwa metode PCA membantu dalam reduksi dimensi data tanpa kehilangan informasi penting, sementara KNN memberikan klasifikasi yang akurat berdasarkan fitur warna dan tekstur.

DAFTAR PUSTAKA

- [1] I. Irma, M. Muchtar, R. Adawiyah, and S. Sarimuddin, "Klasifikasi Tingkat Kematangan Cabai Merah Keriting Menggunakan Svm Multiclass Berdasarkan Ekstraksi Fitur Warna," *J. Inform. dan Tek. Elektro Terap.*, vol. 12, no. 3, pp. 1747–1755, 2024, doi: 10.23960/jitet.v12i3.4430.
- [2] E. Asrianti, "Marketing Analysis of Red Chili Farming in Maku Village Sub District Dolo," *J. Agroekbis*, vol. 2, no. 6, pp. 660–666, 2021.
- [3] T. Widowati, N. Nuriyana, L. Nurjanah, S. J. R. Lekatompessy, and R. Simarmata, "Pengaruh Bahan Baku Kompos terhadap Pertumbuhan dan Produksi Cabai Merah Keriting (*Capsicum annuum* L.)," *J. Ilmu Lingkung.*, vol. 20, no. 3,

- pp. 665–671, 2022, doi: 10.14710/jil.20.3.665-671.
- [4] S. Chairunnissa, D. Akbar, N. Yasmin, and A. Ramadhanu, “Pengelolaan Citra Cabai Keriting: Kombinasi Median Filtering dan Algoritma K-Means untuk Pengelompokan Berbasis Fitur,” vol. 0738, no. 4, pp. 6031–6040.
- [5] N. I. Ahmad, Y. N. Bunga, and Y. Bare, “Etnobotani Tanaman Cabai Merah Keriting (*Capsicum Annum* L.) Di Desa Waiwuring, Kecamatan Witihamu Kabupaten Flores Timur,” *Spizaetus J. Biol. dan Pendidik. Biol.*, vol. 2, no. 2, p. 8, 2021, doi: 10.55241/spibio.v2i2.46.
- [6] A. Jalan, S. Hasanuddin, K. Konawe, and S. Tenggara, “Pengaruh Penggunaan Pupuk pada Produksi Cabai Keriting di Desa Toriki Kecamatan Anggaberu Kabupaten Konawe Sarty Syarbiah perkembangannya . Nutrisi yang diberikan pada tanaman berupa pupuk yang mengandung,” vol. 1, no. 2, 2024.
- [7] Ahmad Izuddin Holis, Hery Haryanto, and Mulat Isnaini, “Populasi dan Intensitas Serangan Lalat Buah (*Bactrocera* spp.) pada Pertanaman Cabai Keriting (*Capsicum annum* L.) Di Desa Darmasari Kabupaten Lombok Timur,” *J. Ilm. Mhs. Agrokomplek*, vol. 2, no. 1, pp. 161–170, 2023, doi: 10.29303/jima.v2i1.2332.
- [8] Syifa, U. Salsabilla, D. Kusumawati Hari, and Fitriana, “Karakteristik Carbon Dots (C-dots) Dari Buah Jeruk Mandarin (*Citrus Reticulata*) yang Disintesis Menggunakan Metode Hidrotermal,” *J. Inov. Fis. Indones.*, vol. 13, pp. 60–73, 2024.
- [9] R. Amalia, H. Haris, and R. S. Nurlaela, “PENGARUH KONSENTRASI GULA DAN WAKTU PEMASAKAN TERHADAP KARAKTERISTIK KIMIA, SENSORI, DAN AKTIVITAS ANTIOKSIDAN SELAI JERUK MANDARIN (The Effect of Sugar Concentration and Cooking Time to Chemical, Sensory, and Antioxidant Activity of Mandarin Orange),” vol. xx, no. September, pp. 79–92, 2024.
- [10] A. R. Muslikh, I. Akbar, D. R. I. M. Setiadi, and H. M. M. Islam, “Multi-label Classification of Indonesian Al-Quran Translation based CNN, BiLSTM, and FastText,” *Techno.Com*, vol. 23, no. 1, pp. 37–50, 2024, doi: 10.62411/tc.v23i1.9925.
- [11] L. Hakim, H. R. Rahmanto, S. P. Kristanto, and D. Yusuf, “Klasifikasi Citra Motif Batik Banyuwangi Menggunakan Convolutional Neural Network,” *J. Teknoinfo*, vol. 17, no. 1, p. 203, 2023, doi: 10.33365/jti.v17i1.2342.
- [12] R. I. Borman, I. Ahmad, and Y. Rahmanto, “Klasifikasi Citra Tanaman Perdu Liar Berkhasiat Obat Menggunakan Jaringan Syaraf Tiruan Radial Basis Function,” *Bull. Informatics Data Sci.*, vol. 1, no. 1, p. 6, 2022, doi: 10.61944/bids.v1i1.3.
- [13] Aminatus Syarifah, Aditya Akbar Riadi, and Arief Susanto, “Klasifikasi Tingkat Kematangan Jambu Bol Berbasis Pengolahan,” vol. 7, no. 1, pp. 27–35, 2022.
- [14] D. Novianto and T. Sugihartono, “Sistem Deteksi Kualitas Buah Jambu Air Berdasarkan Warna Kulit Menggunakan Algoritma Principal Component Analysis (Pca) dan K-Nearest Neighbor (K-NN),” *J. Ilm. Inform. Glob.*, vol. 11, no. 2, pp. 42–47, 2020, doi: 10.36982/jiig.v11i2.1223.
- [15] Baiq Nurul Azmi, Arief Hermawan, and Donny Avianto, “Analisis Pengaruh Komposisi Data Training dan Data Testing pada Penggunaan PCA dan Algoritma Decision Tree untuk Klasifikasi Penderita Penyakit Liver,” *JTIM J. Teknol. Inf. dan Multimed.*, vol. 4, no. 4, pp. 281–290, 2023, doi: 10.35746/jtim.v4i4.298.
- [16] I. Istiadi, A. Y. Rahman, and A. D. R. Wisnu, “Identification of Tempe Fermentation Maturity Using Principal Component Analysis and K-Nearest Neighbor,” *Sinkron*, vol. 8, no. 1, pp. 286–294, 2023, doi: 10.33395/sinkron.v8i1.12006.