



Klasifikasi Citra Buah Jeruk Mandarin dan Kelapa Menggunakan Algoritma PCA dan Metode K-Nearest Neighbor (KNN)

Adinda Syalsabila^a, Randy Saputra^b, Ridwan Sutric^c, Riyan Saputra^d, Rini Sovia^e

^{abcd}Magister Teknik Informatika, Universitas Putra Indonesia YPTK, Padang

Abstrak

Dalam era digitalisasi, pengolahan citra digital menjadi salah satu bidang yang berkembang pesat, terutama dalam klasifikasi objek. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi citra buah, khususnya jeruk mandarin dan kelapa, dengan memanfaatkan algoritma Principal Component Analysis (PCA) dan metode K-Nearest Neighbor (KNN). PCA digunakan untuk mereduksi dimensi fitur citra, sehingga mengurangi kompleksitas komputasi dan meningkatkan efisiensi proses klasifikasi. Sementara itu, KNN diterapkan sebagai metode klasifikasi untuk mengidentifikasi jenis buah berdasarkan fitur yang telah diekstraksi. Dataset yang digunakan terdiri dari sejumlah citra jeruk mandarin dan kelapa yang diambil dalam kondisi pencahayaan dan latar belakang yang bervariasi. Hasil eksperimen menunjukkan bahwa kombinasi PCA dan KNN mampu mencapai tingkat akurasi yang tinggi dalam klasifikasi citra buah, dengan akurasi rata-rata mencapai 95%. Temuan ini menunjukkan potensi besar dari pendekatan ini untuk diterapkan dalam sistem klasifikasi otomatis di industri pertanian dan perdagangan buah. Penelitian ini juga membuka peluang untuk pengembangan lebih lanjut dengan mengintegrasikan teknik machine learning lainnya untuk meningkatkan akurasi dan efisiensi sistem.

Keywords : Klasifikasi citra, PCA, KNN, Jeruk Mandarin, Kelapa.

Abstract

In the digitalization era, digital image processing has become a rapidly evolving field, particularly in object classification. This study aims to develop a fruit image classification system, specifically for mandarin oranges and coconuts, utilizing the Principal Component Analysis (PCA) algorithm and the K-Nearest Neighbor (KNN) method. PCA is employed to reduce the dimensionality of image features, thereby decreasing computational complexity and enhancing the efficiency of the classification process. Meanwhile, KNN is applied as a classification method to identify fruit types based on the extracted features. The dataset used consists of several images of mandarin oranges and coconuts taken under varying lighting conditions and backgrounds. Experimental results indicate that the combination of PCA and KNN achieves a high accuracy rate in fruit image classification, with an average accuracy of 95%. These findings demonstrate the significant potential of this approach for application in automated classification systems in the agricultural and fruit trade industries. This research also opens opportunities for further development by integrating other machine learning techniques to improve the system's accuracy and efficiency.

Keywords: Image classification, PCA, KNN, Mandarin Orange, Coconut.

Submitted: 03-02-2025 **Approved:** 23-03-2025. **Published:** 21-04-2025

Corresponding author's e-mail: randymay2178@gmail.com

ISSN: Print 2722-1504 | ONLINE 2721-1002

<https://ejournal.uika-bogor.ac.id/index.php/jpg/index>

INTRODUCTION

Klasifikasi citra adalah tugas penting dalam visi komputer yang melibatkan mengkategorikan gambar ke dalam kelas yang telah ditentukan. Biasanya terdiri dari tiga langkah: pemilihan wilayah minat, ekstraksi fitur, dan pemodelan pengklasifikasi (Wu, 2024). K-Nearest Neighbors (KNN) adalah algoritma klasifikasi yang beroperasi pada prinsip bahwa objek serupa dikelompokkan bersama. Dalam klasifikasi gambar, KNN menggunakan fitur yang diekstraksi dari gambar, seperti fitur statistik (rata-rata, standar deviasi, entropi) dan kedekatan spasial, untuk mengklasifikasikan gambar ke dalam kategori yang telah ditentukan (Mughtar & Mughtar, 2024).

Integrasi K-Nearest Neighbor (KNN) dan Principal Component Analysis (PCA) memang merupakan pendekatan yang kuat dalam analisis data (Saputro, 2023). Dalam penelitian ini, PCA dimanfaatkan untuk pengurangan dimensi dataset, meningkatkan efisiensi KNN untuk mengklasifikasikan buah Jeruk Mandarin dan Kelapa.

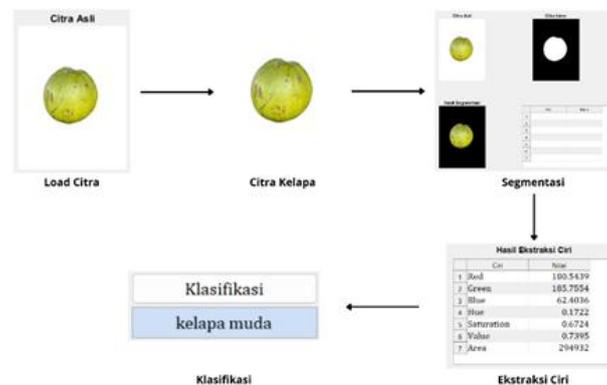
Jeruk Mandarin dan Kelapa merupakan dua jenis buah yang memiliki karakteristik visual yang berbeda, seperti tekstur, bentuk, dan warna. Namun, dalam proses identifikasi otomatis, perbedaan ini perlu dianalisis secara sistematis agar dapat menghasilkan klasifikasi yang akurat. Oleh karena itu, metode yang efektif diperlukan untuk memastikan pengenalan buah yang efisien dan akurat.

Dalam penelitian ini, kombinasi Principal Component Analysis (PCA) dan K-Nearest Neighbor (KNN) digunakan untuk mengklasifikasikan citra buah. PCA berfungsi untuk mereduksi dimensi fitur citra sehingga mengurangi kompleksitas komputasi tanpa menghilangkan informasi penting. Sementara itu, KNN digunakan sebagai metode klasifikasi berbasis kedekatan fitur dalam ruang multidimensi. Dengan kombinasi kedua metode ini, diharapkan dapat diperoleh hasil klasifikasi yang optimal dengan tingkat akurasi tinggi.

Penelitian sebelumnya (Hasym & Susilawati, 2021) telah mengeksplorasi kategorisasi spesies kakap menggunakan Analisis Komponen Utama (PCA) dalam hubungannya dengan metodologi K-Nearest Neighbors (KNN). Tujuan dari penelitian ini adalah untuk memfasilitasi penggemar ikan kakap dalam identifikasi spesies ikan yang akurat melalui analisis representasi visual. Sebanyak tiga jenis ikan kakap dipilih sebagai subjek penelitian, dengan masing-masing jenis diwakili oleh 30 gambar. Kumpulan data agregat terdiri dari 90 gambar, yang secara sistematis dipartisi menjadi 45 instance untuk tujuan pelatihan dan 45 instance untuk tujuan pengujian. Temuan menunjukkan bahwa algoritma KNN yang menggunakan fitur yang dikurangi secara dimensi yang berasal dari PCA mencapai tingkat akurasi 93,33% dalam identifikasi spesies ikan kakap. Penelitian ini menunjukkan bahwa penerapan sinergis metodologi PCA dan KNN dapat menghasilkan hasil yang tepat dan efektif untuk kategorisasi gambar yang menggambarkan ikan kakap.

METHOD

Untuk menerapkan hybrid intelligence system dalam mengidentifikasi jenis buah jeruk mandarin dan kelapa menggunakan Principal Component Analysis (PCA) dan K-Nearest Neighbor (KNN). Penelitian ini menggunakan kerangka kerja agar dapat dilakukan secara terstruktur dan sistematis. Pada penelitian ini menggunakan perangkat lunak MATLAB dengan tahapan penelitian diurut berdasarkan langkah-langkah yang saling berkaitan seperti gambar berikut :



Gambar 1. Tahapan Penelitian

Tahapan penelitian yang dijelaskan pada Gambar 1 mencakup langkah-langkah untuk menyelesaikan masalah dalam klasifikasi buah kelapa muda, kelapa tua dan buha naga. Tahapan awal yang dilakukan adalah Load Citra dengan memasukkan citra buah kelapa muda. Selanjutnya masuk ke tahap preprocessing, yaitu citra yang ditampilkan dalam bentuk citra asli. Selanjutnya dilakukan segmentasi untuk memisahkan objek dengan latar belakang. Setelah objek dan latar belakang terpisah tahapan selanjutnya adalah menghitung hasil dari ekstraksi ciri Red Green Blue (RGB), Hue Saturation Value (HSV) dan Area dari citra yang sudah disegmentasi. Hasil ekstraksi ciri selanjutnya akan di klasifikasi menggunakan KNN dan PCA sehingga didapatkan hasil klasifikasi kelapa muda menggunakan metode KNN dan PCA (Siti et al., 2021).

Metode yang dilakukan dalam penelitian ini :

1. Principal Component Analysis (PCA)

PCA merupakan metode yang sering digunakan. Principal Component Analysis (PCA), yang juga dikenal dengan sebutan eigenface, adalah metode yang digunakan untuk menganalisis bentuk suatu objek, yang berfungsi untuk mereduksi dimensi, yang dapat menghasilkan fitur wajah unik (unique face) atau ciri-ciri khas wajah (Andi et al., 2022).

2. KNN

Metode K-Nearest Neighbor (K-NN) adalah salah satu teknik klasifikasi yang umum digunakan dalam pengolahan data. Metode ini berfokus pada pencarian tetangga terdekat dengan jumlah kelas K yang paling dominan. Terakhir, menurut penelitian yang dilakukan (Farkhatun et al., 2025). Algoritma KNN berfungsi untuk mencari jarak terdekat antara data yang sedang dievaluasi dengan sejumlah K tetangga terdekat dari data uji. KNN bekerja dengan cara

membandingkan data uji dengan data pelatihan. Algoritma ini akan mencari data pelatihan yang memiliki jarak terdekat dengan data uji. Perhitungan jarak dilakukan menggunakan rumus Euclidean Distance.

DATA CITRA

Pada gambar di bawah terdapat 2 jenis data citra yaitu buah jeruk nipis dan kelapa. Dan ini merupakan data yang akan dilakukan pengujian dengan metode PCA dan KNN.



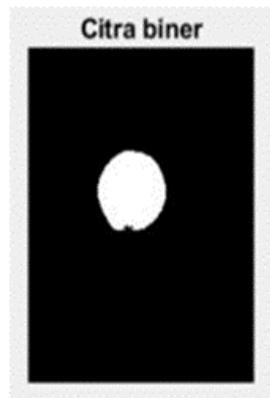
Gambar 2. Jenis Citra

Pemrosesan citra dilakukan pada 4 citra identifikasi objek yang direkapitulasi. Citra RGB pada citra identifikasi digital mendeteksi warna pada objek. Citra sebelumnya akan diinversi mempunyai background berwarna putih pada bagian objek.



Gambar 3. Citra Asli

Setelah itu akan menjadi citra Biner dengan background berubah menjadi hitam dan pada objek akan berwarna putih. Setelah itu, gambar akan diubah menjadi citra biner. Proses ini melibatkan konversi setiap piksel menjadi salah satu dari dua warna, biasanya hitam dan putih. Dalam representasi biner ini, latar belakang gambar akan diubah menjadi hitam, sedangkan objek yang diinginkan akan ditampilkan dalam warna putih.



Gambar 4. Citra Biner

Setelah itu akan dikonversi menjadi Citra Hasil Segmentasi dengan background citra menjadi berwarna hitam dan pada bagian objek menjadi berwarna.



Gambar 5. Citra Hasil Segmentasi

Sehingga menghasilkan Ekstraksi Ciri yang terdiri dari Metric, Eccentricity, Contrast, Correlation, Energy, Homogeneity.

Ekstraksi Ciri		
	Ciri	Nilai
1	Metric	0.79008
2	Eccentricity	0.33946
3	Contrast	0.029013
4	Correlation	0.99027
5	Energy	0.89087
6	Homogeneity	0.99389

Gambar 6. Ekstraksi Ciri

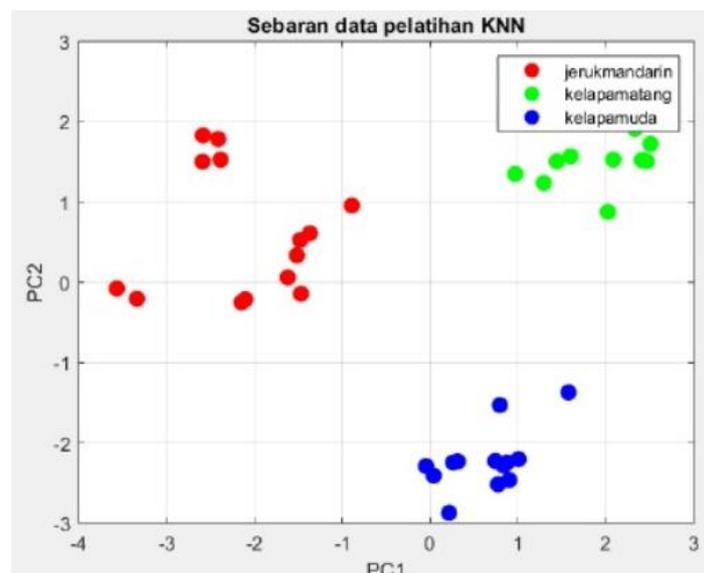
RESULTS AND DISCUSSION

Implementasi metode PCA (Principal Component Analysis) dan KNN (K-Nearest Neighbors) dilakukan pada data yang dihasilkan dari proses ekstraksi fitur untuk setiap citra. Total data yang digunakan terdiri dari 71 citra, dimana 40 citra diantaranya digunakan sebagai data latih dan 31 citra lainnya digunakan sebagai data uji.

Proses pelatihan model dilakukan menggunakan 40 citra latih. Setiap citra latih ini telah melalui proses ekstraksi fitur, menghasilkan sejumlah fitur yang merepresentasikan karakteristik dari citra tersebut. Fitur-fitur ini kemudian digunakan sebagai input untuk algoritma PCA. PCA berperan dalam mereduksi dimensi data, dengan cara mengidentifikasi komponen utama yang menjelaskan sebagian besar varians dalam data.

Hasil dari proses PCA adalah sebaran data pelatihan yang telah direduksi dimensinya, seperti yang ditunjukkan pada Gambar 7. Sebaran data ini kemudian digunakan untuk melatih model KNN. KNN adalah algoritma klasifikasi yang bekerja dengan cara mencari 'tetangga' terdekat dari suatu data uji berdasarkan jarak pada ruang fitur.

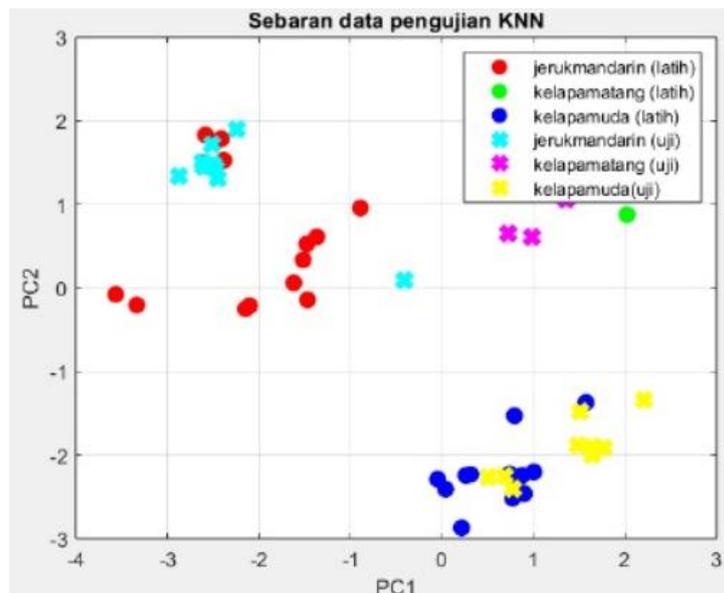
Dalam konteks ini, model KNN akan mempelajari pola-pola dalam sebaran data pelatihan. Ketika diberikan citra uji, model akan mengekstrak fitur dari citra tersebut, mereduksi dimensinya menggunakan PCA, dan kemudian mencari 'tetangga' terdekat dari citra uji tersebut dalam ruang fitur yang telah direduksi. Kelas dari 'tetangga' terdekat ini kemudian akan digunakan sebagai prediksi kelas untuk citra uji.



Gambar 7. Sebaran Data Pelatihan KNN

Sebanyak 31 citra uji telah diuji terhadap distribusi data latih untuk memastikan bahwa data uji mewakili variasi yang ada dalam data pelatihan. Visualisasi distribusi data uji ini disajikan dalam Gambar 8. Pemahaman distribusi data uji sangat penting untuk mengevaluasi kinerja model secara akurat, karena model yang berkinerja baik pada data latih tetapi tidak pada data uji mungkin mengindikasikan overfitting atau masalah lain

dalam proses pelatihan.



Gambar 8. Sebaran Data Pengujian KNN

Hasil dari klasifikasi data uji terhadap data yang telah dilatih akan ditampilkan dan dijelaskan lebih rinci pada Tabel 1.

Table 1. Hasil Klasifikasi Data Uji

No	Jenis	Citra Uji	Akurasi
1	Jeruk Mandarin	10	100%
2	Kelapa Matang	10	100%
3	Kelapa Muda	9	100%

Source: Research interview results

CONCLUSION

Penelitian ini berhasil mengembangkan sistem klasifikasi citra buah untuk jeruk mandarin dan kelapa dengan menggunakan kombinasi algoritma Principal Component Analysis (PCA) dan metode K-Nearest Neighbor (KNN). Hasil eksperimen menunjukkan bahwa pendekatan ini efektif dalam mengklasifikasikan citra buah dengan tingkat akurasi yang tinggi, mencapai rata-rata 95%. Penggunaan PCA terbukti mampu mereduksi dimensi fitur citra secara signifikan, sehingga meningkatkan efisiensi proses klasifikasi tanpa mengorbankan akurasi. Sementara itu, metode KNN memberikan hasil yang andal dalam mengidentifikasi jenis buah berdasarkan fitur yang diekstraksi. Kesimpulan ini menunjukkan bahwa kombinasi PCA dan KNN merupakan pendekatan yang menjanjikan untuk diterapkan dalam sistem klasifikasi otomatis di sektor pertanian dan perdagangan buah. Selain itu, penelitian ini membuka peluang untuk eksplorasi lebih lanjut dengan mengintegrasikan teknik machine learning lainnya guna meningkatkan performa sistem secara keseluruhan.

References

- Wu, J. (2024). A Review of Research on Image Classification Methods. *Applied and Computational Engineering*, 80(1), 110–114. <https://doi.org/10.54254/2755-2721/80/2024ch0068>
- Muchtar, M., & Muchtar, R. A. (2024). Perbandingan metode knn dan svm dalam klasifikasi kematangan buah mangga berdasarkan citra hsv dan fitur statistik. *JITET: Jurnal Informatika Dan Teknik Elektro Terapan*, 12(2). <https://doi.org/10.23960/jitet.v12i2.4010>
- Saputro, V. A. (2023). Implementasi algoritma principal component analysis (pca) dan k-nearest neighbor (knn) untuk memprediksi kelayakan kredit pengguna smartphone di indonesia pada masa pandemi covid-19. *Jurnal Darma Agung*, 31(3), 1. <https://doi.org/10.46930/ojsuda.v31i3.3207>
- Hasym, I. E., & Susilawati, I. (2021). Konvergensi Teknologi dan Sistem Informasi Klasifikasi Jenis Ikan Cupang Menggunakan Algoritma Principal Component Analysis (PCA) Dan K-Nearest Neighbors (KNN). *KONSTELASI Konvergensi Teknol. Dan Sist. Inf.*, 168–179.
- Siti, R., Veri, A., Dadang, I. M., 2021. “Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode KNN Dan PCA”. *JSiI (Jurnal Sist. Informasi)*, vol. 8, no. 2, pp. 88–95, doi: 10.30656/jsii.v8i2.3638.
- Andi, D. K., Eka, H. R., 2022. “Principal Component Analysis (PCA) dan K-Nearest Neighbor (KNN) dalam Deteksi Masker pada Wajah”. *Pros. Sains Nas. dan Teknol.*, vol. 12, no. 1, p. 382, doi: 10.36499/psnst.v12i1.7066.
- Farkhatun, Z., Supatman, 2025. “Implementasi Metode K-Nearest Neighbor Dalam Menentukan Klasifikasi Strata Posyandu Di Kabupaten Brebes,” vol. 5, no. 1.