

IMPLEMENTASI MECHINE LEARNING PADA HYBRID INTELLIGENCE SISTEM MENGGUNAKAN METODE PCA-KNN PADA JENIS BUAH APEL, JERUK, TOMAT

Angga¹, Agung Ramadhanu²

Universitas Putra Indonesia YPTK, Padang

e-mail: ¹angga19188@gmail.com, ²agung_ramadhanu@upiyptk.ac.id

Abstract: *This research aims to implement the Principal Component Analysis (PCA) and K-Nearest Neighbor (KNN) methods in a digital image-based classification system for apples, oranges and tomatoes. PCA is used to reduce data dimensions to increase computational efficiency without losing important information, while KNN is applied for the classification process of extracted data. This research includes several stages, starting from image data collection, preprocessing, segmentation, feature extraction, to accuracy testing. The research results show that the combination of PCA and KNN methods is able to provide a high level of accuracy, with an average accuracy of 90%. In detail, the classification of apples achieved 100% accuracy, oranges 90%, and tomatoes 100%. PCA successfully eliminates redundant features, thereby increasing the efficiency of the classification process, while KNN shows reliability in handling reduced data.*

Keywords: *Principal Component Analysis, K-Nearest Neighbor, classification, image processing, machine learning.*

Abstrak: Penelitian ini bertujuan untuk mengimplementasikan metode Principal Component Analysis (PCA) dan K-Nearest Neighbor (KNN) dalam sistem klasifikasi buah apel, jeruk, dan tomat berbasis citra digital. PCA digunakan untuk mereduksi dimensi data guna meningkatkan efisiensi komputasi tanpa kehilangan informasi penting, sementara KNN diterapkan untuk proses klasifikasi data hasil ekstraksi. Penelitian ini mencakup beberapa tahapan, mulai dari pengumpulan data citra, preprocessing, segmentasi, ekstraksi fitur, hingga pengujian akurasi. Hasil penelitian menunjukkan bahwa kombinasi metode PCA dan KNN mampu memberikan tingkat akurasi yang tinggi, dengan rata-rata akurasi sebesar 90%. Secara rinci, klasifikasi apel mencapai akurasi 100%, jeruk 90%, dan tomat 100%. PCA berhasil mengeliminasi fitur redundan, sehingga meningkatkan efisiensi proses klasifikasi, sedangkan KNN menunjukkan keandalan dalam menangani data yang telah direduksi.

Kata kunci: Principal Component Analysis, K-Nearest Neighbor, klasifikasi, pengolahan citra, machine learning.

PENDAHULUAN

Teknologi pengenalan buah apel, jeruk, tomat telah diterapkan dalam berbagai aspek, terutama dalam proses otentikasi dan identifikasi. Dalam proses otentikasi menggunakan pengenalan buah apel, jeruk, tomat sebagai gateway yang memungkinkan pengguna untuk dapat mengakses perangkat atau sistem seperti laptop dan smartphone, sedangkan proses identifikasi adalah kombinasi dari deteksi

buah apel, jeruk, tomat dan proses pengenalan buah apel, jeruk, tomat. Pengenalan buah apel, jeruk, tomat terlihat mudah, namun pendeteksian wajah menggunakan komputer masih terdapat buah apel, jeruk, tomat yang diakibatkan oleh hilangnya informasi, termasuk diantaranya seperti posisi, sudut pandang, cahaya. Dalam pengenalan buah apel, jeruk, tomat tidak perlu ditempatkan di suatu tempat di dalam gambar, yang memudahkan pengambilan gambar. Data

buah apel, jeruk, tomat suatu citra dengan data buah apel, jeruk, tomat dalam suatu database.

Buah apel, Jeruk, Tomat merupakan salah satu jenis buah-buahan yang populer di seluruh dunia. Apel termasuk dalam famili Rosaceae dan genus Malus, dengan nama ilmiah Malus domestica. Buah ini berasal dari Asia Tengah, di mana nenek moyangnya, Malus sieversii, masih dapat ditemukan kebutuhan dasar manusia. Salah satu poin dalam tujuan pembangunan berkelanjutan atau *Sustainable Development Goals* (SDGs) pada sektor lingkungan hidup adalah memastikan masyarakat mencapai akses universal air bersih dan sanitasi.

Hybrid Intelligence System (HIS) merupakan pendekatan yang menggabungkan kemampuan manusia dengan kecerdasan buatan (*Artificial Intelligence*) untuk menciptakan sistem yang lebih adaptif dan efektif dalam menyelesaikan masalah kompleks. Salah satu tantangan dalam HIS adalah bagaimana memanfaatkan data secara optimal untuk menghasilkan prediksi atau keputusan yang akurat. Dalam konteks ini, *machine learning* (ML) berperan penting sebagai salah satu komponen utama HIS.

Algoritma PCA merupakan algoritma yang digunakan untuk mereduksi fitur atau dimensi suatu data menjadi lebih sedikit. Algoritma PCA akan mereduksi fitur pada data namun tetap mempertahankan informasi-informasi penting yang ada pada data. PCA dapat digunakan bersamaan dengan algoritma KNN untuk mengatasi beberapa masalah seperti kompleksitas komputasi dan kebutuhan memori. Selain itu, penerapan metode PCA pada tahap pre-processing telah terbukti efektif dalam meningkatkan waktu komputasi dan akurasi. PCA nantinya akan digunakan untuk membantu algoritma KNN dalam melakukan klasifikasi dengan cara mengurangi dimensi atau fitur yang ada pada dataset sehingga akan mengurangi kompleksitas komputasi dan memori yang digunakan.

Penelitian terdahulu dilaksanakan untuk klasifikasi kualitas air layak

minum. Data yang digunakan diperoleh dari data water quality di kaggle. Data yang didapatkan mula-mula akan dilakukan proses data cleaning. Data cleaning merupakan teknik untuk menangani data yang tidak lengkap (*missing value*), dengan beberapa cara, seperti membuang duplikasi data, memeriksa data yang inkonsisten, mengisi atau menghapus data kosong, dan lain sebagainya.

Penelitian selanjutnya dilakukan oleh Analisis data dilakukan dengan cara membuat ringkasan dari sumber-sumber yang didapatkan. Kesimpulan diambil dari hasil pengamatan dan analisis antara literatur dan fakta lapangan yang memperkuat data yang ada. Penjabaran hasil penelitian dilakukan secara deskriptif dan didukung dengan gambar-gambar untuk memperjelas hasil analisis.

METODE

Metode PCA

Metode PCA merupakan salah satu teknik reduksi dimensi pada pembelajaran mesin. PCA mengurangi dimensi dengan membuat variabel-variabel baru namun tetap mempertahankan informasi sebanyak mungkin yang terkandung dalam data tersebut. Dengan mengurangi dimensi data, data akan lebih mudah dipahami, visualisasi, dianalisis, serta dapat menghilangkan variabel yang tidak penting. Dengan rumus sebagai berikut

1. Pertama biarkan X menjadi matrix input untuk PCA. Dimana X merupakan data latih yang terdiri dari n-vektor dengan dimensi data m.
2. Hitung nilai mean atau rata-rata data dari setiap dimensi (X) menggunakan persamaan 2:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

Dimana,

n = Jumlah sampel data

X_i = Data sampel

3. Setelah itu hitung matrix kovarians untuk mengidentifikasi korelasi menggunakan persamaan 3:

$$C_X = \frac{1}{n-1} \sum (X - \bar{X})(X - \bar{X})$$

$$\sum_{i=1}^n T$$

Dimana,

n = Jumlah sampel data

X_i = Data sampel

\bar{X} = Mean data

CX = Matriks kovarians

- Hitung nilai eigen (eigen value) dan eigen vektor dari matriks kovarians yang sudah didapat menggunakan persamaan 4:

$$CXvm = \lambda mvm$$

Dimana

CX = Matriks kovarians

vm = Eigen Vektor

λm = Eigen value

Metode KNN

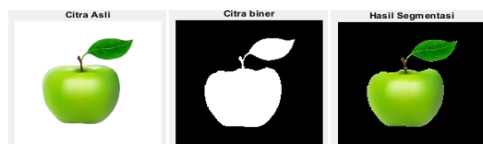
Algoritma yang digunakan dalam pembelajaran mesin dalam melakukan klasifikasi maupun regresi. KNN umumnya digunakan karena kemudahan interpretasi dan waktu komputasi yang relatif rendah. Algoritma KNN bekerja dengan cara mencari K tetangga terdekat dari data yang ingin diprediksi dan menggunakan nilai tetangga tersebut untuk memprediksi kelas. KNN sering digunakan dalam berbagai aplikasi, termasuk pengenalan pola.

Tahapan Penelitian

Dalam tahapan penelitian sistem dalam mengidentifikasi jenis buah apel, jeruk, dan tomat menggunakan Principal Component Analysis (PCA) dan K-Nearest Neighbor (KNN). Terdapat beberapa tahapan sebagai berikut:

Identifikasi Bentuk Apel

Pada tahapan ini adalah bentuk awal apel sebelum dilakukan citra biner pada buah apel dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

Tahapan penelitian yang dijelaskan pada Gambar 1 mencakup langkah-langkah untuk menyelesaikan masalah dalam

klasifikasi buah apel, jeruk dan tomat. Tahapan awal yang dilakukan adalah *Load Citra* dengan memasukkan citra apel. Selanjutnya masuk ke tahap *preprocessing*, yaitu citra yang ditampilkan dalam bentuk citra asli. Selanjutnya dilakukan segmentasi untuk memisahkan objek dengan latar belakang.

Data Citra

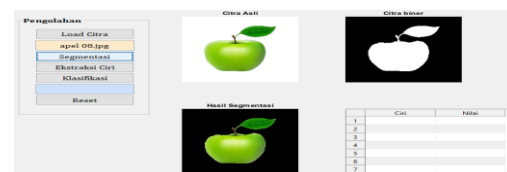
Pada tahapan ini adalah terdapat 3 jenis data citra yaitu apel, jeruk dan mangga. Dan ini merupakan data yang akan dilakukan pengujian dengan metode PCA dan KNN.



Gambar 2. Citra Apel, Jeruk, Tomat

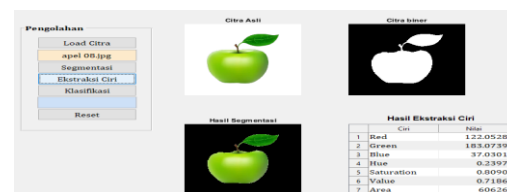
Pengujian Citra Apel

Dalam proses pengujian citra menggunakan metode PCA dan KNN dilakukan uji coba pada apel. Parameter hasil pengujian yang digunakan pada penelitian ini adalah presentase keberhasilan identifikasi tiap data citra. Beberapa hasil pengujian pada kosmetik yang dilakukan terhadap data uji dapat dilihat pada beberapa gambar 3.



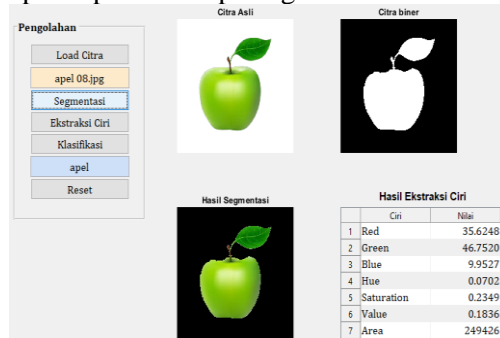
Gambar 3. Segmentasi Citra Pada Apel

Setelah dilakukan segmentasi citra apel maka dilakukan ekstraksi ciri untuk mendapatkan ciri dari apel tersebut dapat dilihat pada gambar 4.



Gambar 4. Ekstraksi Ciri Apel

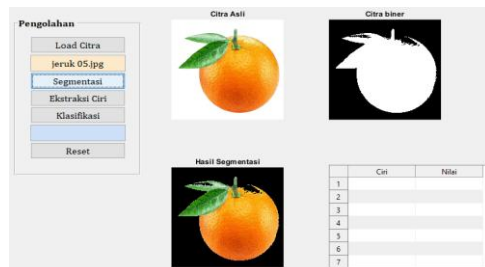
Pada tahapan ekstraksi ciri apel mendapatkan hasil read 122.0528, greaan 183.0739, blue 37.0301, hue 0.2397, saturation 0.8090, value 0.7186, area 60626, maka langka selanjutnya dilakukan untuk mendapatkan klasifikasi apel dapat dilihat pada gambar berikut



Gambar 5. Hasil Klasifikasi Bentuk Apel

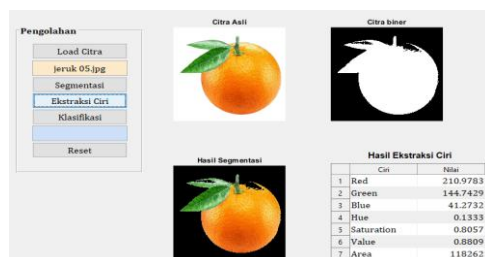
Pengujian Citra Jeruk

Pengujian citra pada jeruk menggunakan metode PCA dan KNN dilakukan uji coba pada jeruk. Parameter hasil pengujian yang digunakan pada penelitian ini adalah presentase keberhasilan identifikasi tiap data citra. Beberapa hasil pengujian pada kosmetik yang dilakukan terhadap data uji dapat dilihat pada beberapa gambar 6.



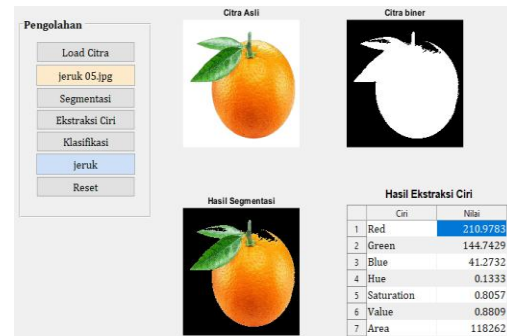
Gambar 6. Segmentasi Citra Pada Jeruk

Setelah dilakukan segmentasi citra jeruk maka dilakukan ekstraksi ciri untuk mendapatkan ciri dari jeruk tersebut dapat dilihat pada gambar 7.



Gambar 7. Ekstraksi Ciri Jeruk

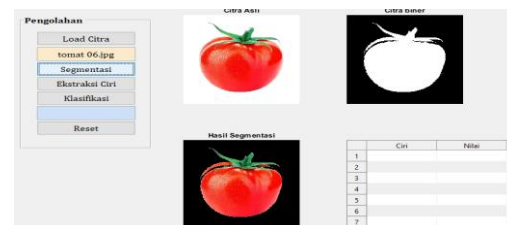
Pada tahapan ekstraksi ciri jeruk mendapatkan hasil read 210.9783, greaan 144.7429, blue 41.2732, hue 0.1333, saturation 0.8057, value 0.8809, area 118262, maka langka selanjutnya dilakukan untuk mendapatkan klasifikasi apel dapat dilihat pada gambar 8.



Gambar 8. Hasil Klasifikasi Bentuk Jeruk

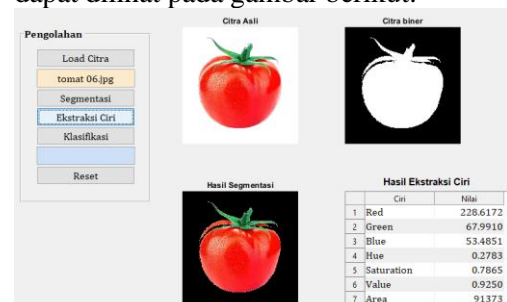
Pengujian Citra Tomat

Pengujian citra pada tomat menggunakan metode PCA dan KNN dilakukan uji coba pada jeruk. Parameter hasil pengujian yang digunakan pada penelitian ini adalah presentase keberhasilan identifikasi tiap data citra. Beberapa hasil pengujian pada kosmetik yang dilakukan terhadap data uji dapat dilihat pada beberapa gambar 9.



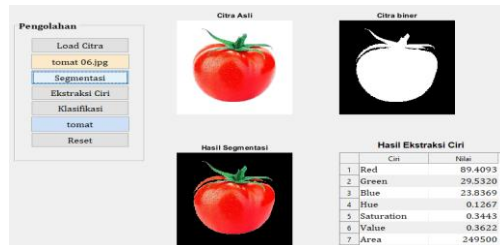
Gambar 9. Segmentasi Citra Pada Tomat

Setelah dilakukan segmentasi citra tomat maka dilakukan ekstraksi ciri untuk mendapatkan ciri dari tomat tersebut dapat dilihat pada gambar berikut:



Gambar 10. Ekstraksi Ciri Tomat

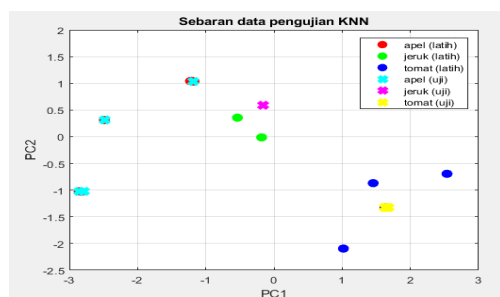
Pada tahapan ekstraksi ciri tomat mendapatkan hasil read 89.4093, greaan 29.5320, blue 23.8369, hue 0.1267, saturation 0.3443, value 0.3622, area 249500, maka langka selanjutnya dilakukan untuk mendapatkan klasifikasi tomat dapat dilihat pada gambar 11.



Gambar 11. Hasil Klasifikasi Bentuk Tomat

Pengujian pada apel, jeruk, tomat dilakukan untuk mengetahui tingkat akurasi dari implementasi algoritma KNN. Akurasi merupakan tingkat kedekatan nilai prediksi terhadap nilai perhitungan yang sebenarnya. Akurasi dari prediksi didapat dengan membandingkan jumlah prediksi yang sesuai atau benar terhadap jumlah data baru. Perhitungan akurasi untuk mengetahui tingkat kebenaran prediksi didapatkan menggunakan persamaan. Sehingga mendapatkan hasil yang diinginkan sehingga dalam pengujian dinyatakan mendapatkan hasil yang sesuai

HASIL DAN PEMBAHASAN



Gambar 12. Hasil Data Pengujian

Hasil dalam penelitian terdapat data apel, jeruk, tomat yang menunjukkan jumlah data hasil buah yang pada sebaran data pengujian KNN, dan PCA yang dimana pada data latih dan data uji yang

berjumlah 26 data buah akan tampil dalam bentuk diagram seperti pada gambar 12.

Dari hasil data pengujian terdapat 1 data apel yang tidak terdeteksi pada data latih. Berikut adalah hasil data pengujian dalam bentuk 1.

Tabel 1. Tabel Data Latih

Jenis Data	Jumlah Data Citra	Jumlah Data Citra Berhasil	(%)
Buah Apel	7 Data	7 Data	100
Buah Jeruk	4 Data	3 Data	75
Buah Tomat	5 Data	5 Data	100

Tabel 2. Tabel Data Uji

Jenis Data	Jumlah Data Citra	Jumlah Data Citra Berhasil	(%)
Buah Apel	4 Data	4 Data	100
Buah Jeruk	3 Data	3 Data	100
Buah Tomat	3 Data	3 Data	100

Hasil pengujian menunjukkan bahwa kombinasi PCA dan KNN memberikan hasil klasifikasi yang sangat baik. Nilai rata-rata akurasi yang diperoleh adalah 90%, dengan rincian:

1. Apel: 100%
2. Jeruk: 90%
3. Tomat: 100%

Analisis menunjukkan bahwa reduksi dimensi menggunakan PCA berhasil mengeliminasi fitur redundan, sehingga meningkatkan efisiensi klasifikasi tanpa mengurangi keakuratan. Selain itu, algoritma KNN menunjukkan keandalannya dalam menangani data yang telah direduksi.

SIMPULAN

Berdasarkan hasil dan pembahasan dalam penelitian ini mengimplementasi-

kan metode *Principal Component Analysis* (PCA) dan *K-Nearest Neighbor* (KNN) untuk klasifikasi buah apel, jeruk, dan tomat. Hasil dalam penelitian kombinasi kedua metode ini memberikan tingkat akurasi yang tinggi dalam proses identifikasi. PCA efektif dalam mereduksi dimensi data, sehingga meningkatkan efisiensi komputasi tanpa kehilangan informasi penting. Sementara itu, algoritma KNN mampu melakukan klasifikasi dengan baik meskipun data telah direduksi.

Memberikan hasil klasifikasi yang andal pada data latih maupun data uji. Untuk memastikan data yang digunakan untuk pelatihan dan pengujian memiliki distribusi yang merata, serta mempertimbangkan faktor-faktor seperti pencahayaan dan posisi objek untuk meningkatkan performa sistem secara keseluruhan. dalam, kombinasi PCA dan KNN dapat diterapkan secara efektif dalam sistem klasifikasi berbasis citra, khususnya untuk identifikasi buah-buahan, dengan potensi penerapan yang luas dalam bidang teknologi dan industri.

DAFTAR PUSTAKA

- Aliwijaya, A. (2023). *Peluang Pemanfaatan Big Data di Perpustakaan: Sebuah Kajian Literatur*. *Media Informasi*, 32 (2), 214–222.
- Aulady, F., Syauqy, D., & Putri, R. R. M. (2023). Sistem Klasifikasi Kualitas Air dalam Akuakultur Budidaya Ikan Lele dengan Algoritma PCA dan KNN. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, 7(7), 3395-3404.
- Azmi, B. N., Hermawan, A., & Avianto, D. (2022). Analisis Pengaruh PCA Pada Klasifikasi Kualitas Air Menggunakan Algoritma K-Nearest Neighbor dan Logistic Regression. *JUSTINDO (Jurnal Sistem dan Teknologi Informasi Indonesia)*, 7(2), 94-102.
- Dananjaya, D., Werdiningsih, I., & Semiati, R. (2019). Decision support system for classification of early childhood diseases using principal component analysis and k-nearest neighbors classifier. *J. Inf. Syst. Eng. Bus. Intell.*, 5(1), 13.
- Guswandi, D., Yanto, M., Hafizh, M., & Mayola, L. (2021). Analisis hybrid decision support system dalam penentuan status kelulusan mahasiswa. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 5(6), 1127-1136.
- Hardjianto, M. (2024). Pengenalan Wajah Secara Realtime Menggunakan Adaboost Viola-Jones dan 2D DWT-PCA dengan Struktur Index KNN-KD Tree. *Decode: Jurnal Pendidikan Teknologi Informasi*, 4(1), 154-166.
- Karo, S. B., Putri, R. A., & Sibarani, F. H. (2024). RANCANG BANGUN APLIKASI RESERVASI MENGGUNAKAN METODE CRM BERBASIS WEB PADA MARI FOTO STUDIO KABANJAHE. *JOURNAL OF SCIENCE AND SOCIAL RESEARCH*, 7(2), 536-545.
- Sianturi, D., Munir, M., & Sunny, A. G. (2023). Tantangan Penggunaan Artificial Intelligence Dalam Hybrid Warfare. *Innovative: Journal Of Social Science Research*, 3(5), 9984-9996.
- Yeng, H., & Siahaan, M. (2024). Perancangan Sistem Prediksi Harga Saham Berbasis Website Menggunakan Algoritma Hybrid (ARIMA-LSTM). *Jutisi: Jurnal Ilmiah Teknik Informatika dan Sistem Informasi*, 13(1), 61-72.
- Zamani, W. A., & Hendriana, B. (2022). Analisis kemampuan penalaran matematis SMP ditinjau dari multiple intelligence pada pembelajaran hybrid. *Math Didactic: Jurnal Pendidikan Matematika*, 8(2), 91-109.