

IMPLEMENTASI *HYBRID INTELLIGENCE SYSTEM* UNTUK KLASIFIKASI BIJI-BIJIAN DENGAN ALGORITMA PCA DAN KNN

Fajri Rinaldi Chan^{1*}; Agung Ramadhanu²

Magister Teknik Informatika, Pascasarjana^{1,2}
Universitas Putra Indonesia YPTK Padang, Padang, Indonesia^{1,2}
<https://upiypk.ac.id/>^{1,2}
fajririnaldichan@gmail.com^{1*}, agung_ramadhanu@upiypk.ac.id²
(*) Corresponding Author



Ciptaan disebarluaskan di bawah Lisensi Creative Commons Atribusi-NonKomersial 4.0 Internasional.

Abstract— Food security has become a pressing global issue with the increasing population and food consumption needs. Red kidney beans, peanuts, and sunflower seeds play a crucial role in meeting the nutritional needs of society and serving as raw materials for various industries. This study aims to develop a seed classification system based on the Principal Component Analysis (PCA) and K-Nearest Neighbor (KNN) algorithms. The system is designed to recognize three types of seeds—red kidney beans, peanuts, and sunflower seeds—to improve the efficiency and accuracy of the classification process compared to manual methods. The dataset consists of 58 seed image samples, divided into training data (48 samples) and test data (10 samples). The research stages include image preprocessing (cropping, background removal, and thresholding segmentation), feature extraction using PCA to reduce data dimensionality, and classification with KNN based on Euclidean distance. A value of $K=3$ is used in the KNN algorithm to determine the proximity between data points. The test results show a classification accuracy of 90%, with 9 out of 10 test data correctly classified. PCA successfully simplified high-dimensional data into two main components without significant information loss, while KNN demonstrated strong capability in distinguishing the three types of seeds. This research contributes to the development of an AI-based automatic classification system for the food industry, with broader potential applications in high-dimensional data processing across various fields.

Keywords: artificial intelligence, k-nearest neighbor, principal component analysis, seed classification.

Abstrak—Ketahanan pangan menjadi isu global yang mendesak dengan meningkatnya kebutuhan populasi dan konsumsi pangan. Kacang merah, kacang tanah, dan biji bunga matahari memiliki peran penting dalam memenuhi kebutuhan nutrisi masyarakat dan bahan baku industri. Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi biji-bijian berbasis algoritma *Principal Component Analysis* (PCA) dan *K-Nearest Neighbor* (KNN). Sistem ini dirancang untuk mengenali tiga jenis biji-bijian, yaitu kacang merah, kacang tanah, dan biji bunga matahari, dengan tujuan meningkatkan efisiensi dan akurasi proses klasifikasi dibandingkan metode manual. Dataset yang digunakan terdiri dari 58 sampel citra biji-bijian yang dibagi menjadi data latih (48 sampel) dan data uji (10 sampel). Tahapan penelitian mencakup *preprocessing* citra (*cropping*, penghapusan latar belakang, segmentasi *thresholding*), ekstraksi fitur menggunakan PCA untuk mereduksi dimensi data, dan klasifikasi dengan KNN berbasis jarak *Euclidean*. Nilai $K=3$ digunakan dalam algoritma KNN untuk menentukan kedekatan antar data. Hasil pengujian menunjukkan akurasi klasifikasi sebesar 90%, dengan 9 dari 10 data uji terklasifikasi dengan benar. PCA berhasil menyederhanakan data berdimensi tinggi menjadi dua komponen utama tanpa kehilangan informasi signifikan, sedangkan KNN menunjukkan kemampuan yang baik dalam membedakan tiga jenis biji-bijian. Penelitian ini memberikan kontribusi pada pengembangan sistem klasifikasi otomatis berbasis kecerdasan buatan untuk industri pangan, dengan potensi aplikasi yang lebih luas dalam pengolahan data berdimensi tinggi di berbagai bidang.

Kata kunci: kecerdasan buatan, k-nearest neighbor, principal component analysis, klasifikasi biji.

PENDAHULUAN

Ketahanan pangan merupakan salah satu isu strategis yang terus menjadi perhatian global, terutama dalam menghadapi tantangan peningkatan populasi dan kebutuhan pangan yang semakin meningkat. Biji-bijian seperti kacang merah, kacang tanah, dan biji bunga matahari memainkan peranan signifikan dalam memenuhi kebutuhan nutrisi masyarakat serta sebagai bahan baku industri makanan dan kosmetik (Simanullang et al., 2021). Namun, klasifikasi biji-bijian secara manual yang selama ini dilakukan masih menghadapi sejumlah kendala, termasuk waktu pemrosesan yang lama, akurasi rendah akibat subjektivitas manusia, serta biaya operasional yang tinggi. Kondisi ini menjadi tantangan besar, khususnya dalam memenuhi kebutuhan pasar yang terus meningkat. Oleh karena itu, diperlukan sistem klasifikasi berbasis teknologi otomatis untuk meningkatkan efisiensi, akurasi, dan konsistensi dalam proses pengolahan biji-bijian.

Kecerdasan buatan telah menunjukkan potensi besar dalam menyelesaikan berbagai permasalahan di berbagai bidang, termasuk dalam klasifikasi objek (Nurfritri et al., 2021). Algoritma pembelajaran mesin, seperti *Principal Component Analysis* (PCA) dan *K-Nearest Neighbor* (KNN), memberikan solusi yang menjanjikan dalam pengolahan data berdimensi tinggi dan klasifikasi berbasis pola (Krismawan & Rachmawanto, 2022). PCA berfungsi sebagai metode reduksi dimensi yang mampu mengidentifikasi fitur utama dalam *dataset*, sehingga mengurangi kompleksitas tanpa menghilangkan informasi penting (Dinanti & Purwadi, 2023) (Nuraini, 2022). Sementara itu, KNN merupakan algoritma klasifikasi berbasis jarak yang sederhana namun efektif dalam mengelompokkan data baru berdasarkan data yang sudah ada (Nurfritri et al., 2021) (Setiawan, 2022). Kombinasi kedua algoritma ini berpotensi menghasilkan sistem *hybrid* yang lebih efisien dan akurat dalam klasifikasi biji-bijian.

Beberapa penelitian sebelumnya telah mengaplikasikan PCA dan KNN dalam berbagai konteks pengenalan pola dan klasifikasi. Siti Raysyah, dkk. menggunakan PCA untuk reduksi dimensi dalam klasifikasi biji kopi berdasarkan fitur warna dan tekstur, menunjukkan bahwa reduksi hingga 40% tidak mengurangi akurasi secara signifikan (Raysyah et al., 2021). Penelitian ini membuktikan efektivitas PCA dalam mengelola data berdimensi tinggi, meskipun penggunaannya terbatas pada satu algoritma klasifikasi saja. Ulla Delfana Rosiani, dkk. memanfaatkan KNN untuk klasifikasi biji jagung berdasarkan warna dan bentuk, mencapai akurasi 88% (Rosiani et al.,

2019). Namun, penelitian ini masih menghadapi tantangan dalam pengolahan data berdimensi tinggi, yang menjadi kendala utama pada dataset yang lebih kompleks.

Arif Lumute Unihehu, dkk. Klasifikasi Jenis Ikan Berbasis Jaringan Saraf Tiruan Menggunakan Algoritma *Principal Component Analysis* (Unihehu & Suharjo, 2021). PCA berhasil mereduksi data menjadi dua komponen utama, tetapi penelitian ini belum menggabungkan PCA dengan algoritma berbasis jarak seperti KNN, yang dapat meningkatkan akurasi klasifikasi. Sementara itu, Anwar Mujadin, dkk. Kemurnian Minyak Kayu Putih Berbasis *Electronic Nose* Menggunakan Metode PCA Dan *Neural Network*, menunjukkan akurasi hingga 95% (Mujadin et al., 2024). Meskipun demikian, pendekatan ini memerlukan waktu pelatihan yang cukup lama dibandingkan algoritma berbasis jarak seperti KNN.

Dari hasil tinjauan literatur tersebut, terlihat bahwa PCA dan KNN masing-masing memiliki keunggulan dalam berbagai aplikasi klasifikasi. Namun, integrasi kedua metode ini dalam konteks klasifikasi biji-bijian masih belum banyak dilakukan. Penelitian terdahulu cenderung fokus pada algoritma tunggal atau jenis biji tertentu, tanpa mengeksplorasi potensi sistem *hybrid* untuk meningkatkan performa klasifikasi (Yanti et al., 2024). Oleh karena itu, penelitian ini bertujuan untuk mengisi celah tersebut dengan mengembangkan sistem *hybrid* berbasis PCA dan KNN untuk klasifikasi tiga jenis biji-bijian, yaitu kacang merah, kacang tanah, dan biji bunga matahari. Sistem ini dirancang untuk meningkatkan akurasi dan efisiensi, sekaligus mengatasi keterbatasan proses manual yang ada saat ini.

Pengembangan sistem klasifikasi otomatis yang mampu memproses data berdimensi tinggi secara efisien menggunakan kombinasi PCA dan KNN diharapkan dapat memberikan solusi yang praktis dan ekonomis bagi industri pangan dalam meningkatkan efisiensi produksi serta menjaga kualitas produk (Putra et al., 2024). Evaluasi performa kombinasi PCA dan KNN dalam konteks klasifikasi biji-bijian, termasuk pengukuran akurasi, efisiensi waktu, dan konsistensi hasil klasifikasi, menjadi fokus utama dalam penelitian ini (Akbar et al., 2022). Dengan adanya sistem ini, proses klasifikasi dapat dilakukan secara otomatis tanpa membutuhkan intervensi manusia, sehingga meningkatkan efisiensi operasional.

Efisiensi dan akurasi dalam proses klasifikasi biji-bijian sangat penting untuk memastikan kualitas produk yang konsisten dan memenuhi standar yang ditetapkan. Penggunaan teknologi otomatis berbasis kecerdasan buatan dapat mengurangi ketergantungan pada proses manual

yang rentan terhadap kesalahan manusia dan variabilitas. Implementasi sistem klasifikasi otomatis berbasis PCA dan KNN dapat memberikan nilai tambah yang signifikan bagi industri pangan, baik dari segi efisiensi operasional maupun kualitas produk akhir.

Penelitian ini juga berpotensi memberikan kontribusi pada pengembangan metode klasifikasi yang lebih umum dan dapat diterapkan pada berbagai jenis biji-bijian lainnya. Dengan menggabungkan keunggulan PCA dalam reduksi dimensi dan KNN dalam klasifikasi berbasis jarak, sistem *hybrid* yang diusulkan diharapkan dapat mengatasi berbagai tantangan yang dihadapi dalam pengolahan data berdimensi tinggi dan klasifikasi objek (Hidayat et al., 2023). Hal ini membuka peluang untuk aplikasi yang lebih luas dalam berbagai bidang, termasuk pertanian, industri makanan, dan penelitian ilmiah.

Penelitian ini bertujuan untuk mengembangkan solusi praktis dan inovatif yang dapat meningkatkan efisiensi dan akurasi dalam klasifikasi biji-bijian, sekaligus memberikan kontribusi pada pengembangan ilmu pengetahuan di bidang kecerdasan buatan dan pembelajaran mesin. Dengan demikian, hasil penelitian ini diharapkan dapat memberikan manfaat yang signifikan bagi industri pangan dan masyarakat luas, serta mendorong kemajuan teknologi dalam pengolahan data dan klasifikasi objek.

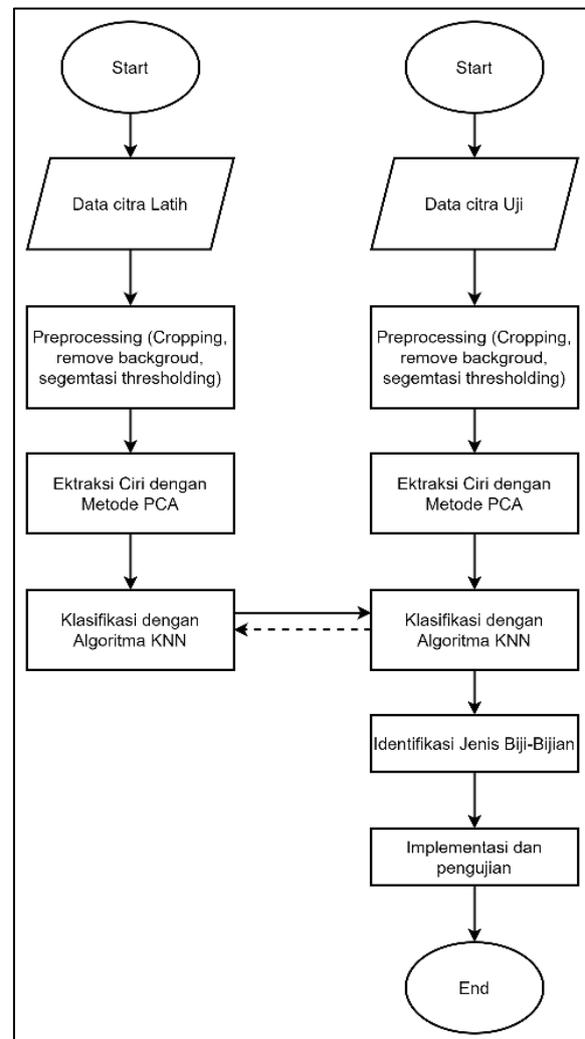
Penelitian ini diharapkan memberikan kontribusi langsung pada pengelolaan dan klasifikasi biji-bijian, khususnya dalam meningkatkan efisiensi produksi dan manajemen kualitas di industri pangan. Hasil penelitian ini juga dapat menjadi referensi bagi pengembangan sistem klasifikasi berbasis kecerdasan buatan di masa depan. Pendekatan *hybrid* ini tidak hanya relevan untuk klasifikasi biji-bijian tetapi juga berpotensi diterapkan pada berbagai aplikasi lain yang memerlukan pengolahan data berdimensi tinggi dan klasifikasi berbasis pola. Dengan demikian, penelitian ini tidak hanya memberikan solusi praktis tetapi juga berkontribusi pada pengembangan ilmu pengetahuan di bidang pembelajaran mesin dan kecerdasan buatan

BAHAN DAN METODE

Tahapan Penelitian

Penelitian ini dirancang sebagai penelitian eksperimental yang bertujuan untuk mengembangkan dan mengevaluasi sistem *hybrid* berbasis algoritma *Principal Component Analysis* (PCA) dan *K-Nearest Neighbor* (KNN) dalam klasifikasi biji-bijian. Sistem ini dirancang untuk

mengenali tiga jenis biji-bijian, yaitu kacang merah, kacang tanah, dan biji bunga matahari. Tahapan metodologi penelitian ini dijelaskan secara lengkap pada Gambar 1.



Sumber: (Hasil Penelitian, 2025)

Gambar 1. Tahapan Penelitian

Penelitian ini dirancang sebagai penelitian eksperimental yang bertujuan untuk mengembangkan dan mengevaluasi sistem *hybrid* berbasis algoritma *Principal Component Analysis* (PCA) dan *K-Nearest Neighbor* (KNN) dalam klasifikasi biji-bijian. Sistem ini dikembangkan untuk mengenali tiga jenis biji-bijian, yaitu kacang merah, kacang tanah, dan biji bunga matahari. Tahapan penelitian yang dijelaskan pada Gambar 1 mencakup proses pengolahan data, ekstraksi fitur, serta implementasi algoritma PCA dan KNN untuk meningkatkan akurasi klasifikasi.

Penelitian dimulai dengan pengumpulan data citra biji-bijian yang dibagi menjadi dua jenis, yaitu data citra latih dan data citra uji. Data citra

latih digunakan untuk melatih sistem dalam mengenali pola karakteristik dari masing-masing jenis biji-bijian, sementara data citra uji digunakan untuk menguji dan mengevaluasi kinerja sistem yang telah dibangun.

Tahap pertama dalam proses pengolahan data adalah preprocessing. Pada tahap ini, dilakukan serangkaian langkah untuk mempersiapkan data citra agar lebih mudah diolah (Nugraha et al., 2023). *Preprocessing* mencakup beberapa proses, yaitu cropping untuk memotong bagian citra sehingga hanya mencakup objek biji-bijian, penghapusan latar belakang (*remove background*) untuk menghilangkan elemen-elemen yang tidak relevan, serta segmentasi menggunakan teknik thresholding untuk memisahkan biji-bijian dari latar belakang secara jelas. Tahapan ini bertujuan untuk menghasilkan citra yang bersih dan siap untuk dianalisis lebih lanjut.

Setelah tahap *preprocessing*, dilakukan ekstraksi ciri menggunakan algoritma *Principal Component Analysis* (PCA). PCA digunakan untuk mereduksi dimensi data dengan memilih fitur-fitur utama yang paling signifikan dari citra biji-bijian. Dengan menggunakan PCA, data yang kompleks dapat disederhanakan tanpa kehilangan informasi penting yang diperlukan untuk proses klasifikasi.

Hasil dari ekstraksi ciri kemudian dimasukkan ke dalam proses klasifikasi menggunakan algoritma *K-Nearest Neighbor* (KNN). KNN adalah algoritma berbasis pembelajaran non-parametrik yang mengklasifikasikan data baru berdasarkan kedekatannya dengan data latih (Putra et al., 2024). Pada penelitian ini, algoritma KNN digunakan untuk membandingkan fitur-fitur hasil ekstraksi dari data uji dengan data latih, sehingga dapat menentukan jenis biji-bijian, apakah itu kacang merah, kacang tanah, atau biji bunga matahari.

Tahap berikutnya adalah identifikasi jenis biji-bijian berdasarkan hasil klasifikasi. Proses identifikasi ini bertujuan untuk memberikan hasil akhir berupa pengenalan jenis biji-bijian yang sesuai dengan ciri-ciri yang dimiliki.

Langkah terakhir dalam penelitian ini adalah implementasi dan pengujian sistem. Implementasi dilakukan dengan merancang sistem secara utuh, mulai dari integrasi *preprocessing*, ekstraksi ciri, hingga klasifikasi. Pengujian sistem bertujuan untuk mengevaluasi kinerja metode yang digunakan, terutama dalam hal akurasi, efisiensi, dan keandalannya dalam mengenali jenis biji-bijian. Pengujian dilakukan menggunakan data citra uji untuk memastikan sistem mampu memberikan hasil yang sesuai dengan kebutuhan.

Dengan tahapan-tahapan tersebut, penelitian ini berupaya menghasilkan sistem yang andal untuk

mengenali tiga jenis biji-bijian secara otomatis dan efisien, menggunakan pendekatan *hybrid* berbasis PCA dan KNN.

PCA (*Principal Component Analysis*)

Principal Component Analysis (PCA) adalah teknik yang digunakan untuk mengekstraksi fitur utama dari suatu kumpulan data dengan memisahkan atau mendekomposisi data sehingga koefisien yang tidak relevan dapat dihilangkan. Metode ini juga dikenal sebagai transformasi *Karhunen-Loève*, transformasi *Hotelling*, atau pendekatan *Eigenface* (Zanah & Sriani, 2024).

Analisis Komponen Utama (PCA) merupakan teknik yang digunakan untuk merestrukturisasi bentuk awal variabel menjadi sejumlah kecil variabel yang tidak berkorelasi, namun tetap mampu mewakili informasi dari kumpulan variabel tersebut (Zanah & Sriani, 2024) (Siregar et al., 2023). Analisis Komponen Utama (PCA) bertujuan untuk menggambarkan variabilitas dalam suatu kumpulan variabel (Tarigan et al., 2024). Variabel asli kemudian diubah atau disederhanakan menjadi sejumlah variabel baru. Berikut cara kerja beserta formulasi matematisnya:

1. Membentuk matriks parameter klasifikasi dari data yang akan dilatih, Misalkan *dataset* terdiri dari n sampel, masing-masing memiliki m fitur. *Dataset* tersebut direpresentasikan sebagai matriks x dengan ukuran $n \times m$, di mana setiap baris adalah sampel, dan setiap kolom adalah fitur:

$$X = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix} \quad (1)$$

2. Setiap fitur dalam *dataset* dinormalisasi untuk memastikan rata-rata setiap kolom menjadi nol. Normalisasi dilakukan dengan:

$$X_{normalized} = X - \mu \quad (2)$$

$$\mu = \frac{1}{n} \sum_{i=1}^n X_i \quad (3)$$

3. Setelah data dinormalisasi, matriks kovarian C dihitung untuk menentukan hubungan antara fitur-fitur:

$$C = \frac{1}{n-1} X_{normalized}^T X_{normalized} \quad (4)$$

C adalah matriks simetris berukuran $m \times m$, di mana setiap elemen C_{ij} merepresentasikan kovarian antara fitur i dan fitur j .

4. *Eigenvektor* dan *eigenvalue* dari matriks kovarian CCC dihitung untuk menentukan komponen utama:

$$Cv = \lambda v \quad (5)$$

v adalah *eigenvektor* (komponen utama) dan λ adalah *eigenvalue* (variansi yang dijelaskan oleh *eigenvektor* tersebut). *Eigenvektor* disusun berdasarkan *eigenvalue* dalam urutan menurun, dengan *eigenvektor* pertama memiliki *eigenvalue* terbesar.

5. Untuk mereduksi dimensi, hanya k komponen utama pertama yang dipilih (dengan $k < m$). Data asli diproyeksikan ke ruang baru dengan menggunakan komponen utama yang dipilih:

$$X_{reduce} = X_{normalized} V_k \quad (6)$$

di mana V_k adalah matriks berisi k *eigenvektor* terpilih (ukuran $m \times k$), X_{reduce} adalah matriks hasil reduksi (ukuran $m \times k$) (Sriani et al., 2024).

K-Nearest Neighbors (KNN)

K-Nearest Neighbors (KNN) merupakan algoritma pembelajaran terawasi yang digunakan untuk melakukan klasifikasi dengan mempertimbangkan k tetangga terdekat (Rulaningtyas & Winarno, 2023) (Sugiharto et al., 2022) Rumus yang digunakan untuk perhitungan KNN dapat dilihat pada Persamaan. Dengan $d(i, j)$ adalah nilai jarak, x_i adalah nilai pada fitur 1, dan x_j nilai pada fitur (Rasyid et al., 2022).

$$d(x_{baru}, x_i) = \sqrt{\sum_{j=1}^m (x_{baru_j} - x_{ij})^2} \quad (7)$$

Dimana d adalah Jarak *Euclidean*, m adalah Jumlah fitur dalam dataset, x_{baru_j} adalah Fitur ke- j dari sampel data baru dan x_{ij} adalah Fitur ke- j dari data latih ke- i (Sabry, 2023).

HASIL DAN PEMBAHASAN

Dataset

Dalam penelitian ini, peneliti mengembangkan sebuah sistem untuk mengklasifikasikan jenis biji-bijian. Algoritma *K-Nearest Neighbor* digunakan dalam proses klasifikasi data, sementara *Principal Component Analysis* diterapkan untuk menganalisis sebaran *dataset*. Untuk mendukung penelitian ini, peneliti memanfaatkan pemrograman MATLAB R2019b dalam pembuatan sistem klasifikasi biji-bijian. Jenis

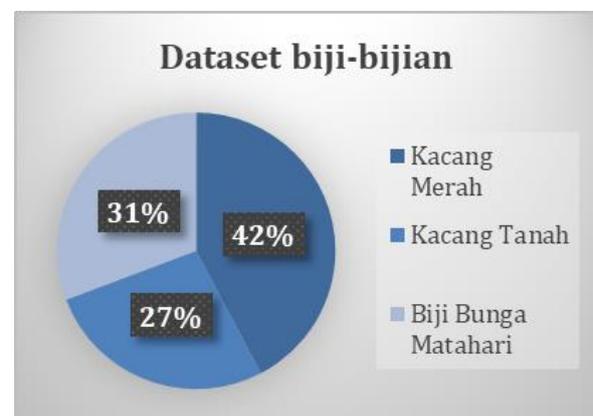
biji-bijian yang digunakan adalah kacang merah, kacang tanah, dan biji bunga matahari. Tabel 1 adalah contoh jenis biji-bijian yang digunakan dalam penelitian ini.

Tabel 1. Tabel Kelas Biji-bijian

Kelas	Definisi	Gambar
Kacang Merah	Citra berupa biji-bijian berjenis kacang merah	
Kacang Tanah	Citra berupa biji-bijian berjenis kacang tanah	
Biji Matahari	Citra berupa biji-bijian berjenis biji bunga matahari	

Sumber: (Hasil Penelitian, 2025)

Dataset biji-bijian merupakan kumpulan yang terdiri dari 58 sampel yang dibedakan menjadi tiga kelas yaitu kacang merah, kacang tanah, biji bunga matahari, yang terdiri dari *sample* data pada Gambar 2.



Sumber: (Hasil Penelitian, 2025)

Gambar 2. Persebaran *Sample Dataset*

Pada Gambar 2 di atas memperlihatkan persebaran *dataset* dari 58 sampel 42% diantaranya adalah sampel kacang merah dan 27% adalah kacang tanah, dan 31% adalah *sample* biji bunga matahari.

Pre-Processing

Tahapan *pre-processing* dilakukan untuk mempersiapkan data citra biji-bijian agar optimal untuk analisis. Proses dimulai dengan *cropping*

untuk memfokuskan area objek, diikuti dengan penghapusan latar belakang untuk menghilangkan elemen yang tidak relevan. Selanjutnya, dilakukan segmentasi *thresholding* untuk memisahkan objek biji-bijian dari latar belakang, menghasilkan citra biner yang jelas (Chan et al., 2024). Terakhir, citra dinormalisasi untuk menyamakan skala dan ukuran setiap sampel.

Setelah *pre-processing*, dataset yang terdiri dari 58 sampel dibagi menjadi data *training* sebanyak 48 sampel, dan data *testing* sebanyak 10 sampel. Pembagian dilakukan secara stratifikasi untuk memastikan distribusi proporsional setiap kelas (kacang merah, kacang tanah, dan biji bunga matahari). Data *training* digunakan untuk melatih model, sementara data *testing* digunakan untuk mengevaluasi kinerja model secara objektif. Seperti Tabel 2 dibawah ini:

Tabel 2. Jumlah data latih dan data uji

Sample	Data Latih	Data Uji
Kacang Merah	21	4
Kacang Tanah	12	3
Biji Bunga Matahari	15	3

Sumber: (Hasil Penelitian, 2025)

Ekstraksi Fitur dan Klasifikasi

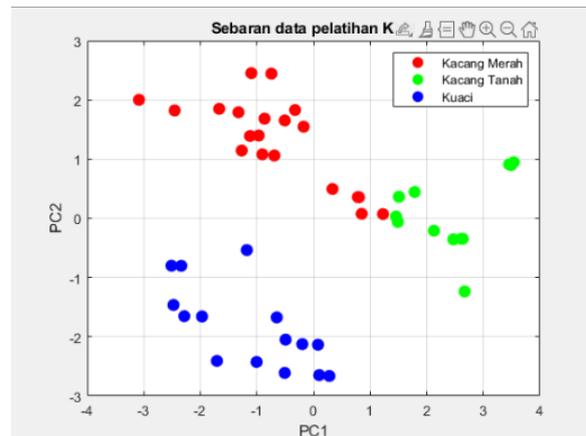
Melakukan ekstraksi fitur menggunakan metode PCA dengan tujuan mereduksi ukuran menjadi dua, yakni dengan memilih dua eigen yang paling dominan dari vektor fitur input berdimensi empat. Setelah proses ekstraksi fitur PCA dilakukan, ukuran data menjadi lebih ringkas (Hasan, 2023). Perhitungan KNN menggunakan *Euclidean Distance* untuk menghitung jarak terdekat dari dua PC dan kelas latihnya (Fitri et al., 2023). Nilai K yang digunakan dalam penelitian ini adalah K=3.

Pemilihan nilai K = 3 dalam algoritma *K-Nearest Neighbor* (KNN) didasarkan pada pertimbangan untuk mengurangi kemungkinan *overfitting* dan memastikan kestabilan klasifikasi. Jika K terlalu kecil (misalnya, K = 1), model menjadi terlalu sensitif terhadap *outlier*, sehingga dapat mengakibatkan hasil klasifikasi yang tidak stabil. Sebaliknya, jika K terlalu besar, model dapat kehilangan kepekaan terhadap pola lokal dan cenderung melakukan klasifikasi yang lebih global, yang dapat mengurangi akurasi untuk *dataset* dengan distribusi kelas yang kompleks.

Dalam penelitian ini, K = 3 dipilih karena memberikan keseimbangan antara akurasi dan kestabilan klasifikasi. Dengan memilih tiga tetangga terdekat, model dapat melakukan voting mayoritas untuk menentukan kelas dari suatu data uji, sehingga lebih tahan terhadap *noise* dibandingkan dengan K = 1 atau K = 2.

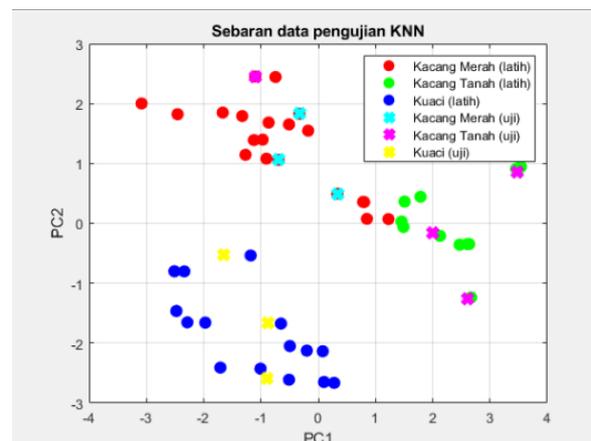
Sedangkan untuk *Principal Component Analysis* (PCA), metode ini digunakan untuk mereduksi dimensi fitur dari empat menjadi dua dengan memilih dua eigen dominan yang memiliki varians terbesar. Dengan demikian, PCA dapat menghilangkan redundansi dalam data, mempercepat perhitungan KNN, serta meningkatkan efisiensi komputasi tanpa kehilangan informasi yang signifikan.

Pada Gambar 3, persebaran data pelatihan berdasarkan perhitungan KNN dengan *Euclidean Distance* menunjukkan bagaimana data diklasifikasikan berdasarkan jarak terdekat dari dua *Principal Components* (PCs). Penggunaan *Euclidean Distance* sebagai metrik jarak dalam KNN didasarkan pada kesederhanaan dan efektivitasnya dalam mengukur kedekatan antara titik data dalam ruang berdimensi rendah yang telah direduksi oleh PCA.



Sumber: (Hasil Penelitian, 2025)

Gambar 3. Persebaran Data Latih pada KNN



Sumber: (Hasil Penelitian, 2025)

Gambar 4. Persebaran Data Uji pada KNN

Pada sebaran data pelatihan terlihat bahwa data pelatihan telah direduksi dimensinya menjadi dua komponen utama, yaitu *Principal Component 1*

(PC1) dan *Principal Component 2* (PC2). Data ini terdiri dari tiga kelas biji-bijian, yaitu kacang merah (ditandai dengan titik merah), kacang tanah (titik hijau), dan kuaci atau biji bunga matahari (titik biru). Sebaran data ini menunjukkan bahwa setiap kelas memiliki distribusi tertentu yang memungkinkan KNN untuk mengklasifikasikan data baru berdasarkan jaraknya terhadap data pelatihan. Sebaran yang terpisah antar kelas menunjukkan bahwa fitur-fitur yang dihasilkan dari PCA cukup representatif untuk membedakan jenis biji-bijian.

Sementara itu, sebaran data pengujian pada Gambar 4 memperlihatkan kombinasi data pelatihan dan data pengujian. Data pelatihan ditampilkan dengan simbol lingkaran, sedangkan data pengujian ditampilkan dengan simbol bintang. Data kacang merah ditunjukkan dengan warna merah, kacang tanah dengan warna hijau atau cyan (untuk uji), dan kuaci dengan warna biru atau kuning (untuk uji). Dalam ruang dua dimensi ini, data pengujian tampak tersebar di sekitar data pelatihan, menunjukkan posisi relatif data baru yang akan diklasifikasikan oleh algoritma KNN.

Sebaran data pengujian menunjukkan bahwa sebagian besar data uji berada dekat dengan kelas data latih yang relevan. Hal ini mengindikasikan bahwa algoritma KNN diharapkan mampu memberikan klasifikasi yang akurat karena pemisahan antar kelas yang cukup jelas. Dengan memanfaatkan kedekatan antar data berdasarkan jarak, KNN dapat memprediksi kelas data uji secara efektif.

Secara keseluruhan hasil dari proses reduksi dimensi menggunakan PCA dan klasifikasi menggunakan KNN. PCA berhasil menyederhanakan data berdimensi tinggi menjadi dua dimensi yang tetap mempertahankan informasi penting untuk proses klasifikasi. Dengan distribusi data yang terlihat terorganisir, sistem berbasis PCA dan KNN dapat diharapkan memberikan performa yang baik dalam mengklasifikasikan biji-bijian seperti kacang merah, kacang tanah, dan kuaci.

Implementasi dan Pengujian

Penelitian ini menggunakan metode *K-Nearest Neighbor* (KNN) untuk melakukan proses klasifikasi data. Implementasi KNN dilakukan dengan menghitung jarak terdekat antara data uji dan data latih. Dalam penelitian ini, terdapat label yang menunjukkan jenis biji-bijian, serta atribut yang diperoleh melalui proses ekstraksi fitur citra. Atribut-atribut tersebut mencakup komponen warna seperti R (*red*), G (*green*), B (*blue*), H (*hue*), S (*saturation*), V (*value*), serta ukuran Area.

Pada tahap ini, nilai-nilai ekstraksi fitur dari data latih dihitung secara menyeluruh. Sebagai ilustrasi, Tabel 2 menampilkan 10 sampel data uji digunakan untuk membangun model klasifikasi yang dirancang untuk mengenali tiga jenis biji-bijian, yaitu kacang merah, kacang tanah, dan biji bunga matahari.

Tabel 3. Table Ekstraksi ciri data uji

Hasil Segmentasi	R	G	B	H	S	V	Area	Label
	129.4165	37.4550	48.6367	0.9426	0.7315	0.5076	119485	Kacang Merah
	145.3512	51.8004	58.9866	0.8490	0.6632	0.5700	95802	Kacang Merah
	157.6456	80.4854	87.9702	0.8998	0.5357	0.6184	86380	Kacang Merah
	131.6979	62.5739	76.7300	0.9480	0.5995	0.5553	93843	Kacang Merah
	177.0045	112.5532	75.0983	0.0671	0.6068	0.6941	93770	Kacang Tanah
	189.7366	128.5141	99.9558	0.0546	0.4882	0.7441	69754	Kacang Tanah
	202.4033	140.837	46.7038	0.0995	0.7801	0.7937	28028	Kacang Tanah
	106.5248	100.9323	84.2053	0.1536	0.2390	0.4186	140774	Biji Bunga Matahari
	112.8341	111.0940	104.6199	0.2014	0.1068	0.442	197226	Biji Bunga Matahari
	99.7622	82.7471	63.4518	0.0877	0.4192	0.3912	285189	Biji Bunga Matahari

Sumber: (Hasil Penelitian, 2025)

Berdasarkan hasil perhitungan ekstraksi fitur yang ditampilkan pada Tabel 3, proses klasifikasi dapat dilakukan menggunakan metode *K-Nearest Neighbor* (KNN). Klasifikasi ini

melibatkan penghitungan jarak terdekat antara data menggunakan rumus *Euclidean Distance*.

Berdasarkan perhitungan tersebut, klasifikasi dilakukan dengan menggunakan nilai $K=3$. Hasil klasifikasi untuk data uji. Proses klasifikasi ini kemudian diulangi untuk seluruh data uji, yang berjumlah data. Setelah seluruh data uji diproses, hasil klasifikasi akhir diperoleh dan dirangkum pada Tabel 4.

Tabel 4. Hasil Pengujian

Data Uji	Kelas Asli	Hasil KNN	Keterangan
1	Kacang merah	Kacang merah	Akurat
2	Kacang Merah	Kacang merah	Akurat
3	Kacang Merah	Kacang merah	Akurat
4	Kacang Merah	Kacang Tanah	Tidak Akurat
5	Kacang Tanah	Kacang Tanah	Akurat
6	Kacang Tanah	Kacang Tanah	Akurat
7	Kacang Tanah	Kacang Tanah	Akurat
8	Biji Bunga Matahari	Biji Bunga Matahari	Akurat
9	Biji Bunga Matahari	Biji Bunga Matahari	Akurat
10	Biji Bunga Matahari	Biji Bunga Matahari	Akurat

Sumber: (Hasil Penelitian, 2025)

Dari 10 data uji yang dianalisis menggunakan metode KNN, sebanyak 9 citra berhasil diklasifikasikan dengan akurat, sementara 1 citra mengalami kesalahan dalam klasifikasi.

Kesalahan klasifikasi pada satu data uji kemungkinan disebabkan oleh kedekatan fitur dengan kelas lain, distribusi data pelatihan yang tidak merata, pemilihan nilai $K = 3$, atau hilangnya informasi penting akibat reduksi dimensi menggunakan PCA. Data yang salah diklasifikasikan mungkin memiliki atribut warna atau ukuran yang lebih mirip dengan kelas lain berdasarkan perhitungan *Euclidean Distance*, sehingga KNN memberikan hasil yang salah. Untuk mengurangi kesalahan ini, sistem dapat diperbaiki dengan menambah jumlah data latih agar model lebih memahami variasi tiap kelas, mengoptimalkan nilai K dengan eksperimen berbagai nilai seperti $K = 5$ atau $K = 7$, serta memilih fitur yang lebih relevan seperti tekstur atau bentuk menggunakan metode ekstraksi fitur tambahan seperti GLCM atau LBP. Selain itu, normalisasi data dapat dilakukan agar skala setiap fitur seragam dalam perhitungan jarak, dan pendekatan *hybrid* dengan metode lain seperti *Support Vector Machine (SVM)* atau *Random Forest* dapat dipertimbangkan untuk meningkatkan akurasi klasifikasi. dapat diketahui dari hasil pengujian yaitu dirumuskan sebagai berikut (Marisa et al., 2022):

$$Akurasi = \frac{\text{Jumlah data benar}}{\text{jumlah seluruh data}} \times 100\% \quad (8)$$

Maka mendapatkan tingkat akurasi sebagai berikut :

$$Akurasi = \frac{9}{10} \times 100\% = 90\%$$

Tingkat Akurasi yang diperoleh pada sistem klasifikasi tingkat kematangan buah kopi dengan metode *K-Nearest Neighbor (KNN)* adalah 90%.

Tingkat akurasi sebesar 90% menunjukkan bahwa sistem klasifikasi menggunakan metode *K-Nearest Neighbor (KNN)* dengan $K = 3$ memiliki performa yang cukup baik dalam mengklasifikasikan biji-bijian berdasarkan fitur warna dan ukuran. Akurasi ini mengindikasikan bahwa dari 10 data uji, sebanyak 9 data diklasifikasikan dengan benar, sementara 1 data mengalami kesalahan klasifikasi, yang kemungkinan disebabkan oleh kedekatan fitur dengan kelas lain, distribusi data latih yang kurang merata, atau pemilihan parameter yang belum optimal.

Diskusi

Tingkat akurasi 90% yang diperoleh dalam penelitian ini menunjukkan bahwa kombinasi *Principal Component Analysis (PCA)* dan *K-Nearest Neighbor (KNN)* cukup efektif dalam klasifikasi biji-bijian. Jika dibandingkan dengan penelitian sebelumnya, hasil ini menunjukkan keunggulan dalam penggunaan PCA untuk reduksi dimensi tanpa mengorbankan akurasi secara signifikan. Misalnya, penelitian oleh Siti Raysyah yang menerapkan PCA dalam klasifikasi biji kopi berdasarkan fitur warna dan tekstur menemukan bahwa reduksi hingga 40% tidak menyebabkan penurunan akurasi yang signifikan, yang mendukung efektivitas PCA dalam mengelola data berdimensi tinggi (Raysyah et al., 2021).

Selain itu, penelitian oleh Ulla Delfana yang menggunakan KNN untuk klasifikasi biji jagung berdasarkan warna dan bentuk mencapai akurasi 88%, yang sedikit lebih rendah dibandingkan penelitian ini (Rosiani et al., 2019). Hal ini menunjukkan bahwa integrasi PCA dalam penelitian ini membantu meningkatkan akurasi dengan mereduksi dimensi dan menghilangkan fitur yang kurang relevan. Sebaliknya, penelitian oleh Arif Lumute menggunakan PCA dalam klasifikasi jenis ikan berbasis jaringan saraf tiruan (ANN), tetapi belum mengombinasikan PCA dengan metode berbasis jarak seperti KNN, yang berpotensi meningkatkan akurasi lebih lanjut (Unihehu & Suharjo, 2021).

Sementara itu, penelitian oleh Anwar dalam klasifikasi kemurnian minyak kayu putih berbasis *Electronic Nose* dengan PCA dan *Neural Network* mencapai akurasi hingga 95%, lebih tinggi

dibandingkan penelitian ini (Mujadin et al., 2024). Namun, pendekatan tersebut memiliki kelemahan pada waktu pelatihan yang lebih lama dibandingkan metode berbasis jarak seperti KNN yang digunakan dalam penelitian ini. Dengan demikian, hasil penelitian ini menunjukkan bahwa kombinasi PCA dan KNN merupakan solusi yang baik dalam klasifikasi biji-bijian, dengan akurasi yang kompetitif dibandingkan penelitian sebelumnya, serta tetap mempertahankan efisiensi komputasi yang lebih baik dibandingkan metode berbasis jaringan saraf tiruan.

KESIMPULAN

Penelitian ini berhasil mengembangkan sistem klasifikasi otomatis untuk tiga jenis biji-bijian, yaitu kacang merah, kacang tanah, dan biji bunga matahari, menggunakan kombinasi algoritma *Principal Component Analysis* (PCA) dan *K-Nearest Neighbor* (KNN). PCA digunakan untuk mereduksi dimensi data menjadi dua komponen utama, yang mampu mempertahankan informasi penting tanpa mengurangi akurasi. Sementara itu, KNN diterapkan untuk mengklasifikasikan data berdasarkan jarak terdekat antara data uji dan data latih. Dengan metode ini, penelitian menunjukkan bahwa sistem *hybrid* berbasis PCA dan KNN efektif dalam menangani data berdimensi tinggi dan menghasilkan klasifikasi yang akurat.

Evaluasi kinerja sistem menunjukkan bahwa pendekatan *hybrid* ini mampu memberikan akurasi yang tinggi dalam klasifikasi biji-bijian. Proses ekstraksi fitur menggunakan PCA berhasil menyederhanakan *dataset* yang kompleks, sedangkan algoritma KNN memberikan hasil klasifikasi yang konsisten berdasarkan kedekatan data. Sebaran data latih dan uji yang terorganisir menunjukkan bahwa fitur yang dihasilkan PCA cukup representatif untuk membedakan jenis biji-bijian. Kombinasi metode ini terbukti dapat meningkatkan efisiensi proses klasifikasi sekaligus mengurangi keterlibatan manual yang rawan terhadap kesalahan.

Implementasi algoritma KNN dilakukan dengan menghitung jarak terdekat menggunakan *Euclidean Distance* untuk menentukan kelas data uji berdasarkan data latih. Hasil pengujian pada 10 data uji menunjukkan bahwa sembilan sampel berhasil diklasifikasikan dengan akurat, sementara satu sampel mengalami kesalahan klasifikasi. Tingkat akurasi yang diperoleh adalah 90%, yang menunjukkan bahwa sistem mampu memberikan hasil yang baik dalam mengenali jenis biji-bijian. Distribusi data yang dihasilkan oleh PCA menunjukkan pemisahan yang cukup jelas antara

kelas, memungkinkan algoritma KNN untuk bekerja secara optimal dalam mengelompokkan data.

Keseluruhan hasil penelitian menunjukkan bahwa kombinasi metode PCA dan KNN dapat digunakan secara efektif untuk klasifikasi citra biji-bijian. PCA berhasil mempertahankan informasi penting dari data berdimensi tinggi, sementara KNN menunjukkan keandalan dalam memprediksi kelas data berdasarkan jarak terdekat. Dengan tingkat akurasi 90%, sistem ini dapat menjadi alternatif yang efisien dan sederhana untuk pengelompokan jenis biji-bijian, khususnya dalam aplikasi berbasis citra digital. Penelitian ini memberikan kontribusi signifikan terhadap pengembangan metode klasifikasi berbasis citra dengan menggunakan pendekatan kombinasi algoritma sederhana namun efektif.

Untuk penelitian lanjutan, terdapat beberapa rekomendasi yang dapat dilakukan guna meningkatkan performa sistem klasifikasi ini. Pertama, eksplorasi metode ekstraksi fitur tambahan seperti *Gray-Level Co-occurrence Matrix* (GLCM) atau *Local Binary Pattern* (LBP) dapat dilakukan untuk menambah informasi tekstur pada biji-bijian, yang dapat meningkatkan akurasi klasifikasi. Kedua, pengujian dengan jumlah dataset yang lebih besar dan variasi pencahayaan yang berbeda dapat membantu meningkatkan generalisasi model dalam berbagai kondisi lingkungan.

Ketiga, eksperimen dengan nilai K yang lebih optimal dalam algoritma KNN atau penggunaan teknik *weighted* KNN dapat diteliti lebih lanjut untuk meningkatkan akurasi dan mengurangi kesalahan klasifikasi. Keempat, kombinasi metode lain seperti *Support Vector Machine* (SVM) atau *Random Forest* dengan PCA dapat diuji untuk melihat apakah pendekatan ini dapat meningkatkan performa sistem dibandingkan dengan KNN.

Selain itu, penggunaan *deep learning* dengan *Convolutional Neural Network* (CNN) dapat menjadi pendekatan yang lebih canggih untuk klasifikasi citra biji-bijian, terutama jika *dataset* yang digunakan lebih kompleks dan dalam jumlah besar. Terakhir, pengembangan sistem otomatis berbasis aplikasi atau perangkat IoT untuk klasifikasi biji-bijian secara *real-time* dapat menjadi inovasi yang berguna untuk industri pertanian dan pangan.

REFERENSI

Akbar, J. M., Sabirin, M., Nugraha, G. S., & Alamsyah, N. (2022). Penerapan Metode Principal Component Analysis (Pca) Dan K-Nearest Neighbors (Knn) Untuk Klasifikasi Data Kanker Paru-Paru. *Jurnal Teknologi Informasi*,

- Komputer Dan Aplikasinya (JTika)*, 4(2), 144–151.
<http://jtika.if.unram.ac.id/index.php/JTIKA/>
- Chan, F. R., Annas, F., Yuspita, Y. E., & Darmawati, G. (2024). Implementation of Convolutional Neural Networks (CNN) in An Emotion Detection System for Measuring Learning Concentration Levels. *Knowbase: International Journal of Knowledge in Database*, 04(01), 49–61.
- Dinanti, A., & Purwadi, J. (2023). Analisis Performa Algoritma K-Nearest Neighbor dan Reduksi Dimensi Menggunakan Principal Component Analysis. *Jambura Journal of Mathematics*, 5(1), 155–165.
<https://doi.org/10.34312/jjom.v5i1.17098>
- Fitri, E. N., Ariansyah, M. H., Winarno, S., Budiman, F., (2023). Performance of K-Means Clustering and Knn Classifier in Fish Feed Seller Determination Models. *J. Tek. Inform.*
<https://www.academia.edu/download/119655849/308.pdf>
- Hasan, M. A. (2023). Classification of Apple Types Using Principal Component Analysis and K-Nearest Neighbor. *International Journal of Information System Technology and Data Science*, 1(1), 15–22.
<http://ejournal.lotusaruna.id/index.php/ijistdas/article/view/11>
- Hidayat, S., Ramadhan, H. M. T., & ... (2023). Comparison of K-Nearest Neighbor and Decision Tree Methods using Principal Component Analysis Technique in Heart Disease Classification. *Indonesian Journal*
<https://www.jurnal.yoctobrain.org/index.php/ijodas/article/view/70>
- Krismawan, A. D., & Rachmawanto, E. H. (2022). Principal Component Analysis (PCA) dan K-Nearest Neighbor (KNN) dalam Deteksi Masker pada Wajah. *Prosiding Sains Nasional Dan Teknologi*, 12(1), 382.
<https://doi.org/10.36499/psnst.v12i1.7066>
- Marisa, F., Maukar, A. L., Farhan, A., Widodo, E. A., Sa, I., & Dasilva, R. T. L. (2022). Pengukuran Tingkat Kematangan Kopi Arabika Menggunakan Algoritma K-Nearest Neighbour. *JIMP : Jurnal Informatika Merdeka Pasuruan*, 6(3), 4–8.
- Mujadin, A., Rifaldi, S. P., Samijayani, O. N., & Suyono, H. (2024). Pengujian Kemurnian Minyak Kayu Putih Berbasis Electronic Nose Menggunakan Metode PCA Dan Neural Network. *BULLET : Jurnal Multidisiplin Ilmu*, 3(01), 40–47.
- Nugraha, R. A., Hidayat, E. W., & Shofa, R. N. (2023). Klasifikasi Jenis Buah Jambu Biji Menggunakan Algoritma Principal Component Analysis dan K-Nearest Neighbor. *Gener.*
<https://www.academia.edu/download/110957018/3143.pdf>
- Nuraini, R. (2022). Klasifikasi Citra Jenis Kapasitor Menggunakan Kombinasi Algoritma K-Nearest Neighbor dan Principal Component Analysis. *Journal of Computer System and Informatics (JoSYC)*, 3(3), 133–140.
<http://ejournal.seminar-id.com/index.php/josyc/article/view/1694>
- Nurfitri, K., Pradana, A. D., & Widaningrum, I. (2021). PENERAPAN ALGORITMA PRINCIPAL COMPONENT ANALYSIS (PCA) DAN K-NEAREST NEIGHBORS (KNN) PADA KLASIFIKASI. *Jurnal Rekayasa Teknologi Dan Komputasi*, 1(1), 1–19.
- Putra, A., Abdullah, D., & Daud, M. (2024). Klasterisasi Kualitas Biji Kopi Berdasarkan Taraf Penyusutan Menggunakan Metode K-Harmonic Means dengan Validasi Silhouette Index dan C-index. *Jurnal Janitra Informatika Dan Sistem*
<https://janitra.org/index.php/home/article/view/203>
- Rasyid, N., Novella, T. J., & Nuraminah, A. (2022). IMPLEMENTASI METODE K-NEAREST NEIGHBOUR DALAM MEMREDIKSI CURAH HUJAN DI KOTA BOGOR. *JURNAL RISET RUMPUN ILMU TEKNIK*, 1(1).
<https://prin.or.id/index.php/JURRITEK/article/view/592>
- Raysyah, S. R., Veri Arinal, & Dadang Iskandar Mulyana. (2021). Klasifikasi Tingkat Kematangan Buah Kopi Berdasarkan Deteksi Warna Menggunakan Metode Knn Dan Pca. *JSil (Jurnal Sistem Informasi)*, 8(2), 88–95.
<https://doi.org/10.30656/jsii.v8i2.3638>
- Rosiani, U. D., Mentari, M., & Prastya, A. N. (2019). Klasifikasi Kualitas Biji Jagung Berdasarkan Deteksi Warna dan Bentuk Menggunakan Metode K-Nearest Neighbor. *Seminar Informatika Aplikatif 2019*.
<http://jurnalti.polinema.ac.id/index.php/SIAP/article/view/641/222>
- Rulaningtyas, R., & Winarno, W. (2023). Classification of cervical cancer cells using the K-nearest neighbor (KNN) method based on geometric feature extraction. *AIP Conference Proceedings*, 1.
<https://pubs.aip.org/aip/acp/article-abstract/2858/1/030003/2906771>
- Sabry, F. (2023). *K Nearest Neighbor Algorithm: Fundamentals and Applications*. One Billion Knowledgeable.
- Setiawan, A. (2022). Perbandingan Penggunaan Jarak Manhattan, Jarak Euclid, dan Jarak Minkowski dalam Klasifikasi Menggunakan

- Metode KNN pada Data Iris. *Jurnal Sains Dan Edukasi Sains*.
<https://ejournal.uksw.edu/juses/article/view/6097>
- Simanullang, G., Ngadeni, A., & Hartiyana, T. (2021). FORMULASI SEDIAAN SABUN PELEMBAB TRANSPARAN YANG MENGANDUNG MINYAK BIJI BUNGA MATAHARI (Sunflowerseed Oil). *Pharmacoscript*, 4(1), 10–31.
<https://doi.org/10.36423/pharmacoscript.v4i1.604>
- Siregar, S. D., Ginting, Y. U. R., Sintami, N., Butar-Butar, H. S., (2023). Implementation of KNN algorithm in classifying diabetic ulcers in patients with diabetes mellitus. *Jurnal Mantik*.
<http://www.iocscience.org/ejournal/index.php/mantik/article/view/3928>
- Sriani, S., Lubis, A. H., & Nasution, R. R. (2024). PENERAPAN DATA MINING UNTUK PREDIKSI PENJUALAN PRODUK ELEKTRONIK TERLARIS MENGGUNAKAN METODE K-NEAREST NEIGHBOR. *JOURNAL OF SCIENCE*
<https://jurnal.goretanpena.com/index.php/JSSR/article/view/2153>
- Sugiharto, A., Wirawan, P. W., Nugroho, F. A., & ... (2022). Comparison of SVM, Random Forest and KNN Classification By Using HOG on Traffic Sign Detection. *2022 6th International Conference on Informatics and Computational Sciences (ICICoS)*.
<https://ieeexplore.ieee.org/abstract/document/9930588/>
- Tarigan, E., Hutauruk, H., Siahaan, I. S., & ... (2024). Analisis Reduksi Dimensi Produktivitas Jagung di Provinsi-Provinsi Indonesia Menggunakan PCA. *Interdisciplinary Explorations in Research Journal*, 2(3), 1967–1976.
<https://shariajournal.com/index.php/IERJ/article/download/778/410>
- Unihehu, A. L., & Suharjo, I. (2021). The Klasifikasi Jenis Ikan Berbasis Jaringan Saraf Tiruan Menggunakan Algoritma Principal Component Analysis (PCA). *Jurnal Ilmiah Ilmu Komputer*, 7(2), 27–32.
<https://doi.org/10.35329/jiik.v7i2.200>
- Yanti, R., Chan, F. R., & Ramadhanu, A. (2024). Penerapan Image Processing untuk Identifikasi Jenis Pisang Emas dan Pisang Kapas Menggunakan Metode K-Means Clustering. *Journal of Education Research*, 0738(4), 4938–4943.
- Zanah, N., & Sriani, S. (2024). Classification of Types of Dental Disease Using Principal Component Analysis (PCA) and K-Nearest Neighbor (K-NN) Methods. *JINAV: Journal of Information and Visualization*.
<https://jinav.org/index.php/jinav/article/view/2778>