

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Pengenalan wajah didefinisikan sebagai bagian dari identifikasi biometrik yang mengekstraksi fitur-fitur wajah untuk mengenali seseorang secara unik, menyimpannya untuk keperluan identifikasi [1]. Pengenalan wajah telah mendapatkan popularitas karena aplikasinya dalam pengawasan di bandara, jalan, sekolah, penegakan hukum, dan keamanan informasi. Berbeda dengan identifikasi retina dan sidik jari, yang tidak praktis selama pandemi COVID-19 karena persyaratan jarak sosial, pengenalan wajah menawarkan alternatif yang aman dan efektif. Teknologi ini memanfaatkan fitur unik manusia dan memberikan keseimbangan antara keamanan dan privasi, sehingga sangat berharga dalam memastikan keamanan tanpa invasi privasi yang signifikan [2]. Pada pengenalan wajah, peneliti menghadapi tantangan seperti variasi kondisi pencahayaan, *noise* gambar, skala, oklusi, penyamaran, orientasi sembarang, ekspresi wajah, dan perbedaan usia, yang menyebabkan tingkat akurasi pengenalan wajah rendah, dan menjadikan pengenalan wajah sebagai masalah yang kompleks [3]. Teknologi pengenalan wajah yang efektif memerlukan ekstraksi fitur dan dataset besar [4], tetapi seiring waktu, peningkatan data dapat menyebabkan kebutuhan ruang penyimpanan lebih besar dan waktu pemrosesan lebih lama, yang dapat mempengaruhi kinerja secara negatif [5].

Solusi yang diberikan pada penelitian ini adalah menggunakan klustering K-Means untuk mengatasi permasalahan ruang penyimpanan dan waktu pemrosesan yang lama, maka digunakan metode pengelompokan kluster menggunakan K-Means dengan cara membagi dataset untuk mengurangi waktu komputasi tanpa mengorbankan akurasi secara signifikan. Pada penelitian sebelumnya yang dilakukan oleh Pengcheng Wei dkk., saat menggunakan metode K-Means hasil yang didapatkan adalah tingkat pengenalan untuk kondisi normal setelah klustering menggunakan metode K-Means yaitu 91% [6]. Pada penelitian serupa pada tahun 2020 yang dilakukan oleh Wang Changyuan dkk., saat menggunakan metode klasifikasi PCA+KNN, tingkat akurasi yang didapatkan dari algoritma KNN dengan menggunakan dataset ORL adalah 45% [7]. Maka pada penelitian ini, untuk menurunkan beban komputasi digunakan metode K-Means dan KNN. K-Means digunakan sebagai metode pembagi data, dan KNN digunakan sebagai metode klasifikasi data.

Harapannya, hasil dari penelitian ini dapat untuk menggunakan algoritma dan KNN untuk menurunkan beban komputasi tanpa menurunkan tingkat akurasi secara signifikan.

## 1.2 Maksud dan Tujuan

Maksud penelitian ini adalah untuk mengurangi beban komputasi. Berikut adalah beberapa poin penting tujuan dari penelitian ini:

1. Penelitian ini menerapkan algoritma K-Means dan KNN untuk mengurangi beban komputasi dengan menerapkan strategi pembagian dataset atau klastering menggunakan algoritma K-Means dan KNN.
2. Melihat performa dari algoritma K-Means dan KNN dari pengenalan wajah berdasarkan waktu komputasi dan akurasi.

## 1.3 Batasan Masalah

Adapun pada penelitian ini terdapat batasan masalah antara lain :

1. Pada metode *K-Means Clustering* jumlah kluster dibatasi menjadi 2 kluster.
2. Dataset yang digunakan untuk penelitian ini adalah dataset wajah yang diperoleh dari website Kaggle yaitu AT&T, Yale Face, dan Georgia Tech Face.
3. Perhitungan jarak menggunakan *Euclidean Distance*

## 1.4 Metode Penelitian

Berikut adalah metode penelitian yang digunakan pada penelitian ini:

1. Studi Literatur  
Studi literatur bertujuan untuk mempelajari teori dasar mengenai Pengolahan Citra Digital, teori pendukung mengenai metode *K-Means Clustering* dan *K-Nearest Neighbor* serta mempelajari bahasa pemrograman dan *library* yang digunakan.
2. Pengambilan Data  
Data citra wajah diambil pada *website* Kaggle.com, dengan menggunakan dataset AT&T, Yale Face, dan Georgia Tech Face yang tersedia di Internet.

### 3. Pengujian

Pengujian merupakan lanjutan untuk memastikan bahwa source code telah bekerja dengan baik. Pada tahap pengujian, pengenalan wajah terdiri dari dua tahap yaitu tahap pelatihan (*training*) dan tahap pengujian (*testing*).

### 4. Analisis dan Kesimpulan

Pada tahap analisis dan kesimpulan dilakukan analisis terhadap pengujian sistem yang sudah dibuat dan mengambil kesimpulan dari penelitian ini.

## 1.5 Sistematika Penulisan

Sistematika penyusunan laporan skripsi ini dilakukan dengan sebagai berikut:

### **BAB I PENDAHULUAN**

Materi yang dibahas mengenai latar belakang pengambilan topik penelitian, maksud dan tujuan dilakukannya penelitian, batasan masalah, metode penelitian, dan sistematika penulisan skripsi.

### **BAB II DASAR TEORI**

Bab ini membahas dan menjelaskan teori-teori yang mendukung dan mendasari penulisan skripsi ini. Mulai dari *state of the art*, pengolahan citra digital, K-Nearest neighbor, K-Means Clustering, dataset yang digunakan.

### **BAB III PERANCANGAN SISTEM**

Bab ini fokus pada implementasi konkret dari metode-metode yang dipilih dalam penelitian, memberikan penjelasan rinci tentang kebutuhan perangkat untuk menjalankan program, perhitungan jarak menggunakan Euclidean Distance, dan tahapan perancangan sistem.

### **BAB IV IMPLEMENTASI DAN PENGUJIAN**

Bab ini membahas analisis mendalam terhadap setiap metode yang telah diimplementasikan, meliputi perbandingan antara kelebihan dan kekurangan masing-masing metode yang digunakan dalam penelitian.

### **BAB V KESIMPULAN DAN SARAN**

Penjelasan terkait kesimpulan yang didapatkan dari masalah yang dibahas berdasarkan hasil penelitian yang telah dilakukan pada bab sebelumnya. Serta, memberikan kritik dan saran yang dapat menunjang pengembangan penelitian selanjutnya.