BAB 1
PENDAHULUAN

1.1 Latar Belakang Masalah

Optical Character Recognition (OCR) adalah suatu alat yang digunakan untuk menterjemahkan karakter pada citra digital menjadi format teks [1]. OCR sudah banyak digunakan dalam penerapannya diantaranya pengenalan plat nomor, tulisan tangan huruf Jawa, huruf Jepang, dan aksara Bali, bahkan masih banyak lagi penerapan lainnya dari penggunaan OCR. Pada pengolahan citra, untuk mendapatkan hasil yang baik dalam pengenalan, maka digunakan dua metode yakni metode untuk ekstraksi ciri dan metode untuk pengenalan.

Metode ekstraksi ciri pada penelitian ini akan menggunakan *Principal Components Analysis* (PCA). Tujuan dari PCA adalah untuk mereduksi dimensi yang besar dari ruang data (*observed variables*) menjadi dimensi intrinsik yang lebih kecil dari ruang fitur (*independent variables*), yang dibutuhkan untuk mendeskripsikan data lebih sederhana [2]. PCA dapat digunakan sebagai ekstraksi ciri pada pengenalan angka tulisan tangan [3]. Penggunaan PCA dalam mengenali pola bergantung pada dimensi dan jumlah zona [3]. Sedangkan metode pembelajaran yang digunakan dalam penelitian ini adalah metode *K-Nearest Neighbor* (K-NN). K-NN adalah metode klasifikasi yang didasarkan pada kelas dari tetangga terdekat. K-NN dapat digunakan untuk tulisan tangan angka *arabic* (indian) [4]. Nilai akurasi terbaik dari penggunaan metode K-NN adalah 86% dengan k=1 [4].

Pengenalan tulisan tangan merupakan salah satu bentuk kemampuan komputer dalam mengenali pola. Penelitian dalam bidang pengenalan tulisan tangan telah berkembang dalam kurun waktu yang cukup lama. Penelitian tersebut dilakukan karena semakin banyak dipergunakan model tulisan tangan dalam kehidupan sehari-hari, seperti identifikasi dokumen-dokumen penting, bukti pengesahan dalam dunia perbankan, dan hal lainnya yang melibatkan tulisan tangan di dalamnya.

Berdasarkan penelitian mengenai pengenalan tulisan tangan yang telah dilakukan ditemukan kendala pada proses pengenalannya, hal ini dikarenakan beberapa huruf alfabet memiliki kemiripan bentuk dengan huruf lainnya [5], serta pada penerapannya pengenalan tulisan tangan baru diimplementasikan pada kasus yang secara umum. Pada penelitian ini pengenalan tulisan tangan akan diimplementasikan pada kasus penilaian esai. Penelitian mengenai penilaian esai pun telah banyak dilakukan akan tetapi terbatas hanya pada teks digital saja [6]. Sehingga dalam penelitian ini peneliti akan membuat prototipe penerapan algoritma *K-Nearest Neighbor* dan *Principal Components Analysis* pada pengenalan tulisan tangan dalam penilaian jawaban esai. Kemudian untuk akurasi, akan dilakukan perhitungan dengan menghitung banyaknya karakter tulisan tangan yang dapat dikenali.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan maka identifikasi masalah dalam penelitian ini adalah bagaimana mengetahui tingkat akurasi dalam pengenalan tulisan tangan pada penilaian jawaban esai menggunakan algoritma *K-Nearest Neighbor* dan *Principal Component Analysis*.

1.3 Maksud dan Tujuan

Maksud dari penelitian yang dilakukan dan penulisan tugas akhir ini adalah untuk mengukur tingkat akurasi agoritma *K-Nearest Neighbor* dengan ekstraksi ciri *Principal Components Analysis* pada pengenalan tulisan tangan dalam kasus penilaian jawaban esai. Adapun tujuan dari penelitian ini adalah untuk mengetahui akurasi pada pengenalan tulisan tangan menggunakan algoritma *K-Nearest Neighbor* dan *Principal Components Analysis* dan dapat memberikan penilaian jawaban esai menggunakan algoritma *Rabin Karp*.

1.4 Batasan Masalah

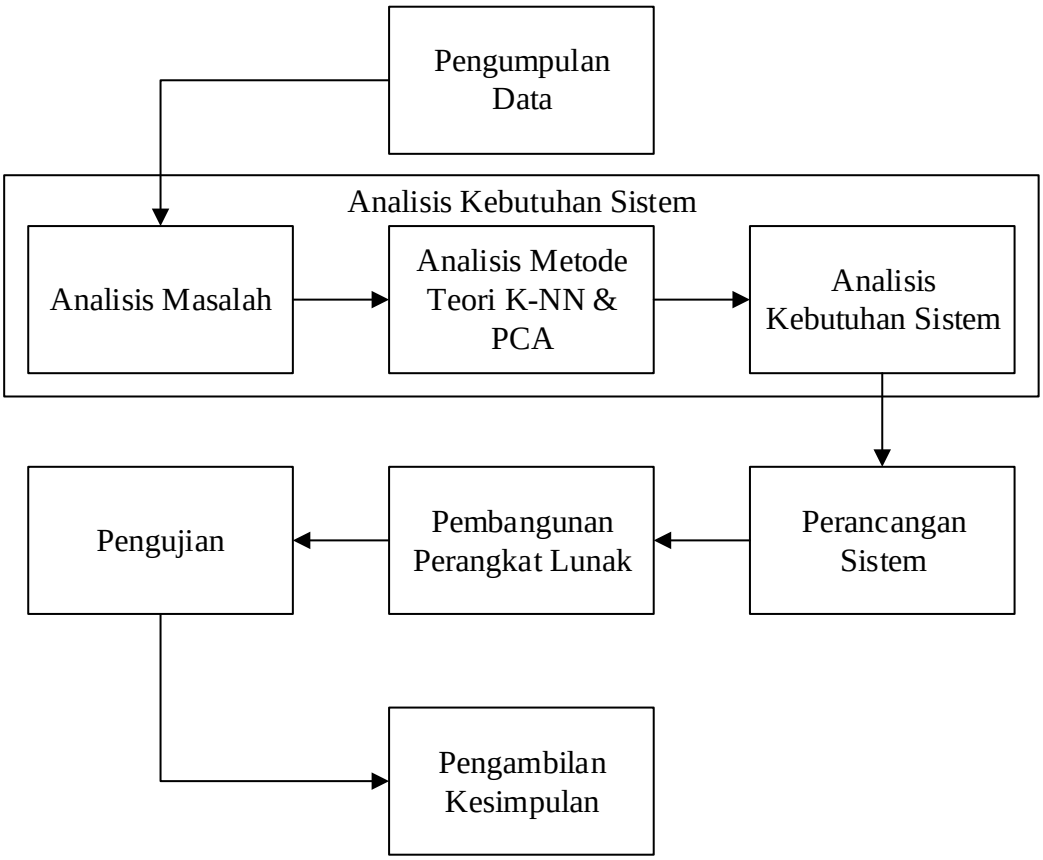
Batasan masalah dalam penelitian ini adalah sebagai berikut :

1. Data masukan merupakan file citra tulisan tangan hasil scan
2. Pengenalan hanya berupa karakter A-Z, a-z, 0-9
3. Tulisan tangan yang dapat dikenali tidak mencakup tulisan bersambung, melainkan hanya tulisan cetak

4. Pertanyaan esai yang akan digunakan bersifat *definitive*
5. Kertas yang digunakan adalah kertas HVS dengan *background* putih dan ditulis menggunakan pena hitam
6. Evaluasi jawaban ujian ditentukan oleh kesamaan teks jawaban dengan kunci jawaban
7. Metode yang digunakan untuk ekstraksi ciri adalah *Principal Components Analysis* (PCA)
8. Metode yang digunakan untuk klasifikasi adalah *K-Nearest Neighbor* (KNN)
9. Metode yang digunakan untuk evaluasi jawaban adalah *Rabin Karp*

1.5 Metodologi Penelitian

Metodologi penelitian yang digunakan dalam penelitian ini adalah metode penelitian kuantitatif dengan metode eksperimen [7]. Hasil akhir yang didapatkan berupa persentase keakuratan pengenalan tulisan tangan dan nilai yang didapatkan dari hasil evaluasi jawaban ujian. Berikut adalah gambar alur dari penelitians.



Gambar 1. 1 Alur Penelitian

1.5.1 Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini yaitu :

1. Studi Literatur

   Studi literatur dilakukan dengan membaca jurnal, paper, buku online terkait topik pengenalan tulisan tangan, metode dalam pengolahan citra, evaluasi jawaban esai, dan klasifikasi tulisan tangan menggunakan metode K-NN dan PCA.

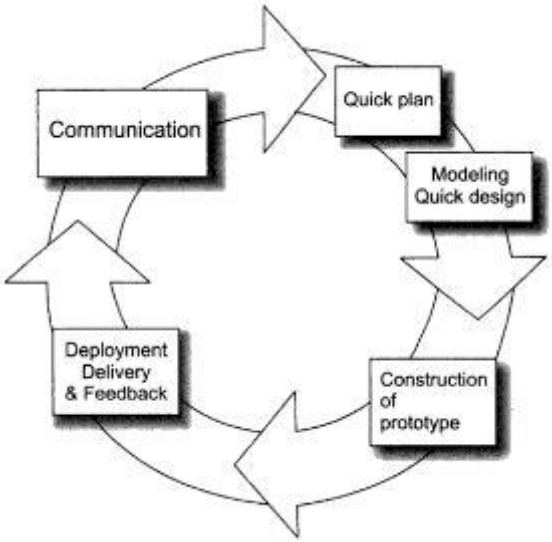
2. Pengumpulan sampel tulisan tangan

   Pada pengumpulan sampel karakter tulisan tangan dari 10 orang. Data sampel akan digunakan sebagai data training saat melakukan pelatihan dengan metode K-NN.

1.5.2 Metode Pembangunan Perangkat Lunak

Dalam penelitian ini metode pembangunan perangkat lunak yang digunakan adalah model prototipe. Berikut adalah gambaran dari metode pembangunan

perangkat lunak. Metode pembangunan perangkat lunak dapat dilihat pada gambar 1.2.



Gambar 1. 2 Model Pembangunan Perangkat Lunak [8]

Berikut adalah langkah-langkah yang dilakukan dalam proses Model Prototipe.

1. *Communication*

   Tahap ini dilakukan untuk menganalisa fungsional dari kebutuhan sistem.

2. *Quick plan*

   Tahap ini dilakukan untuk merancang fungsional kebutuhan sistem yang telah dianalisa pada tahap *Communication.*

3. *Modeling Quick design*

   Tahap ini dilakukan untuk membuat model basis data dan model antarmuka berdasarkan hasil perancangan dari tahap *Quik plan*.

4. *Contruction of prototype*

   Tahap ini dilakukan untuk pembangunan perangkat lunak yang akan diimplementasikan kedalam *sourcecode* dari hasil *Modeling Quick design.*

5. *Deployment Delivery & Feedback*

   Tahap ini dilakukan untuk pengembangan dan evaluasi terhadap sistem yang telah dibangun berdasarkan *Contruction of prototype*.

1.6 Sistematika Penulisan

Sistematika penulisan bab-bab yang digunakan dalam laporan tugas akhir ini adalah sebagai berikut.

BAB 1 PENDAHULUAN

Bab ini berisi pembahasan masalah umum yang berhubungan dengan penyusunan laporan tugas akhir, yang meliputi latar belakang, identifikasi masalah, maksud dan tujuan, batasan masalah, metodologi penelitian, dan sistematika.

BAB 2 LANDASAN TEORI

Bab ini berisi pembahasan mengenai teori-teori yang berhubungan dengan sistem yang akan dibangun seperti Kecerdasan Buatan, Pengolahan Citra, *Principal Component Analysis* (PCA), *K-Nearest Neighbor* (K-NN), Unified Modeling Language (UML), dan juga metode dalam pengujian.

BAB 3 ANALISIS DAN PERANCANGAN

Bab ini berisi tentang pembahasan analisa dan perancangan dari sistem yang akan dibangun diantaranya analisis sistem, analisis data masukan, analisis pada processing dan klasifikasi, analisis data keluaran, analisis non fungsional, perancangan database, perancangan antar muka perancangan jaringan semantik.

BAB 4 IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi hasil implementasi dari hasil analisa dan perancangan yang telah dibuat disertai juga hasil pengujian dari sistem yang dibangun ini yang dan pengujian dilakukan menggunakan pengujian *black* box.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan tentang hasil penelitian dan saran untuk penelitian lebih lanjut dengan kasus yang sama.