**Bab I**

**Pendahuluan**

1. **Latar Belakang**

*Optical Character Recognition* (OCR) berfungsi untuk memindai citra digital menjadi teks yang ada didalam citra tersebut. Dengan adanya OCR, citra yang berisi tulisan tangan, tulisan mesin ketik atau teks computer dapat dimanipulasi. Teks yang discan dengan OCR dapat dicari kata per-kata atau per-kalimat. Kinerja pengenalan huruf pada OCR tergantung pada kualitas input teks yang dimasukkan, dan teknik yang berbeda yang digunakan untuk meningkatkan pengenalan dari karakter. Bahasa arab digunakan oleh lebih dari 1 miliyar pengguna di dunia. Jika sistem OCR tersedia untuk karakter arab, akan memiliki manfaat yang banyak.

Telah banyak penelitian tentang pengenalan huruf latin, jepang dan cina, tetapi sedikit penelitian tentang huruf Arab, ini dikarenakan terdapat kesulitan dalam pengenalan huruf Arab, yakni pada bagian segmentasi, karakteristik tulisan Arab yang bersambung membuat tahap segmentasi huruf menjadi sulit. Oleh karena itu perlu dirancang sebuah teknik segmentasi yang canggih, agar bisa memisahkan tulisan Arab yang bersambung.

Penelitian sebelumnya tentang pengenalan huruf arab telah berkembang, seperti yang dilakukan Albadr (2013) dengan mengekstraksi 24 fitur, diantaranya adalah *boundary chain code,* dan klasifikasi dengan menggunakan *decision tree* yang dihasilkan oleh algoritma C4.5 menghasilkan akurasi pengenalan sebesar 48%. Kemudian beberapa penelitian yang menggunakan *Hidden Markov Model* (HMM)dan *Support Vector Machine* (SVM) sebagai pengklasifikasi. Dalam penelitiam yang dilakukan Anugrah Redja Kusuma (2006) menggunakan HMM mencapai tingkat akurasi 50% hingga 60%. Penelitian menggunakan SVM dengan akurasi 99% dilakukan oleh Mehmmood Abdulla Abd. Berdasarkan hal tersebut, maka penulis ingin membangun sistem pengenalan huruf arab dalam rangkaian kata, dengan menggunakan klasifikasi *Hidden Markov Model* (HMM)dan *Support Vector Machine* (SVM) untuk dibandingkan hasil pengenalan dari keduanya.

1. **Tujuan**

Berdasarkan latar belakang dan rumusan masalah, berikut tujuan yang ingin dicapai dalam penelitian ini.

1. Merancang sistem yang dapat mengenali huruf arab dalam rangkaian kata. Jika diberikan masukan berupa kata atau kalimat dalam bahasa Arab, maka sistem akan mampu mengenali huruf tersebut.
2. Melakukan sistem pengenalan huruf Arab melalui *Hidden Markov Model* (HMM) dan *Support Vector Machine* (SVM)
3. **Batasan Masalah**

Hal-hal yang menjadi batasan pada penelitian ini adalah sebagai berikut.

1. Bahasa pemograman yang digunakan adalah Bahasa Java ™ SE Runtime Environtment (build 1.8.0.72-b15).
2. Tulisan huruf arab yang digunakan adalah dari tulisan cetak dengan font Times New Roman dengan besar font 72.
3. Laptop dengan processor core i3 dengan RAM 8GB.
4. **Metodologi Penelitian**

Tahapan-tahapan kegiatan yang dilakukan dalam penelitian ini dijelaskan sebagai berikut.

1. Studi literatur

Mempelajari dan menganalisis berbagai sumber informasi seperti buku – buku dan literatur referensi yang serupa dengan penelitian yang akan dilakukan

1. Analisis desain

Berdasarkan hasil studi literatur akan dibuat analisis desain dalam menyelesaikan proses perbandingan algoritma untuk setiap metode.

1. Pembuatan sistem

Hasil perancangan diimplementasikan dengan menggunakan Java sebagai pemrosesan citra.

1. Pengujian dan evaluasi

Berdasarkan aplikasi yang telah dibuat kemudian dilakukan uji coba sistem dan  
mengevaluasi sistem sesuai dengan tujuan penelitian

1. Pengambilan data

Setelah pengujian dan didapatkan hasil yang sesuai dengan tujuan penelitian  
kemudian dilakukan pengambilan data

1. Penulisan laporan

Penulisan laporan hasil pengujian dan pengambilan data dibukukan dalam  
laporan penelitian

1. **Sistematika Penelitian**

Dalam penulisan laporan tesis ini terdiri dari lima bab dengan masing–masing bab menguraikan beberapa hal yang terkait dengan perancangan yang dilakukan.

Bab I Pendahuluan. Bab ini mengemukakan latar belakang, tujuan, batasan masalah, metode penelitian, dan sistematika penulisan.

Bab II Tinjauan Pustaka. Bab ini menjelaskan teori dasar mengenai *Optical Character Recognition* (OCR), *Freeman Chain Code, Thinning dan Hidden Markov Model* (HMM) dan *Support Vector Machine* (SVM).

Bab III Perancangan Sistem dan Eksperimen. Bab ini membahas perancangan awal sistem dan eksperimen pengenalan huruf Arab, perancangan sistem dan eksperimen yakni ekstraksi fitur, dan perancangan sistem dan eksperimen untuk pengklasifikasi, serta perangkat lunak untuk pengenalan huruf Arab.

Bab IV Pengujian dan Pembahasan. Bab ini membahas pengujian sistem pengenalan huruf arab yang menghasilkan data yang dibutuhkan sebagai analisis pembahasan perbandingan metode.

Bab V Penutup. Bab ini mengemukakan bagian penutup dari pelaporan penelitian yang meliputi kesimpulan dan saran agar dapat dikembangkan dengan metode lainnya untuk sistem kerja yang sama.

Dalam bab selanjutnya akan disampaikan teori-teori dasar hasil dari peninjauan beberapa literatur yang digunakan. Literatur tersebut dapat berupa buku-buku. makalah penelitian yang telah dipublikasikan, dan sebagainya yang dijadikan sebagai referensi dalam melakukan penelitian ini.