

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Kemajuan teknologi telah banyak mempengaruhi berbagai aspek dalam kehidupan manusia, salah satunya adalah dalam pengelolaan informasi. Kebanyakan orang lebih memilih menyimpan informasi-informasinya dalam bentuk digital. Selain memiliki banyak kelebihan dibanding dengan informasi cetak, informasi digital juga akan memberi kemudahan dalam mengakses dan memanipulasi informasi tersebut.

Salah satu cara untuk mengkonversi informasi tersebut adalah dengan mengetik ulang informasi tersebut pada komputer. Namun ketika jumlah informasi pada media cetak berjumlah banyak, proses pengkonversian akan memakan banyak waktu dan tenaga. Dengan teknologi saat ini, proses pengkonversian informasi bisa dilakukan dengan menggunakan teknologi *Optical Character Recognition* (OCR). OCR adalah sebuah teknologi yang mengkonversi dokumen tangan pada sebuah citra hasil dari scan dan mengkonversinya menjadi karakter digital [1]. Salah satu dokumen yang biasa diperlukan untuk diambil informasinya adalah laporan skripsi mahasiswa. Informasi yang lengkap tentang laporan skripsi mahasiswa sebenarnya bisa didapatkan hanya dari halaman abstraknya saja. Tetapi terkadang untuk dokumen skripsi lama, abstrak skripsi yang tersedia hanya memiliki *hardcopy*-nya saja. Pustakawan harus memasukkan identitas pada suatu dokumen dengan cara mengisi data-data yang diperlukan kedalam sistem. Masalah yang timbul salah satunya adalah akan memakan waktu yang banyak jika jumlah dokumen banyak, masalah lain yang mungkin timbul adalah kesalahan pengetikan identitas dokumen [2]. Kelemahan tersebut dapat ditangani dengan cara mengubah informasi yang ada pada citra abstrak skripsi menjadi *softcopy* dalam

bentuk teks secara otomatis dan memperoleh informasi dalam dokumen tersebut secara otomatis. Untuk melakukan pengenalan tulisan pada citra abstrak skripsi bisa dilakukan dengan metode klasifikasi dengan SVM dan mengambil informasi yang ada pada abstrak skripsi tersebut dengan menggunakan *Rule Based System* .

Penelitian tentang pengenalan tulisan sudah banyak dilakukan, tetapi jarang ada penelitian yang mengenali tulisan pada dokumen penuh. Ada beberapa metode yang bisa digunakan untuk pengenalan karakter tulisan pada suatu citra, seperti dengan menggunakan metode *Artificial Neural Network* (ANN)[3], *K-Nearest Neighbour* (KNN) [4] dan *Support Vector Machine* (SVM) [5]. Penelitian sebelumnya mencoba membandingkan performa dari ANN dan SVM [6] dengan akurasi terbaik didapatkan dari metode SVM dengan hasil sebesar 94,43%, penelitian lain juga mencoba membandingkan metode KNN dan SVM [7] dengan SVM mendapatkan akurasi terbaik sebesar 93,13%. Jadi bisa disimpulkan bahwa metode *Support Vector Machine* mengklasifikasi karakter lebih baik sehingga menjadikan alasan kenapa penelitian ini akan menggunakan metode SVM untuk mengenali karakter tulisan pada citra abstrak skripsi.

Berdasarkan pemaparan diatas, penelitian ini dimaksudkan untuk menerapkan metode *Support Vector Machine* untuk mengenali karakter dan ekstraksi informasi menggunakan *Rules Based System* pada citra dokumen abstrak skripsi.

1.2 Rumusan Masalah

Berdasarkan latar belakang masalah diatas, maka rumusan masalahnya adalah bagaimana cara dalam mengenali karakter tulisan yang ada pada citra dokumen abstrak skripsi dengan menggunakan metode *Support Vector Machine* dan mendapatkan beberapa informasi yang terdapat pada abstrak skripsi.

1.3 Maksud dan Tujuan

Maksud penelitian ini adalah membuat sistem yang mengimplementasikan metode *Support Vector Machine* untuk mengenali karakter tulisan pada citra abstrak skripsi dan memperoleh informasi yang ada pada abstrak skripsi tersebut. Sedangkan tujuan yang ingin dicapai pada penelitian ini adalah mengetahui tingkat akurasi yang didapatkan dengan mengimplementasikan metode *Support Vector Machine* untuk mengenali karakter tulisan pada citra abstrak skripsi dan mengekstrasi informasi yang ada pada abstrak skripsi.

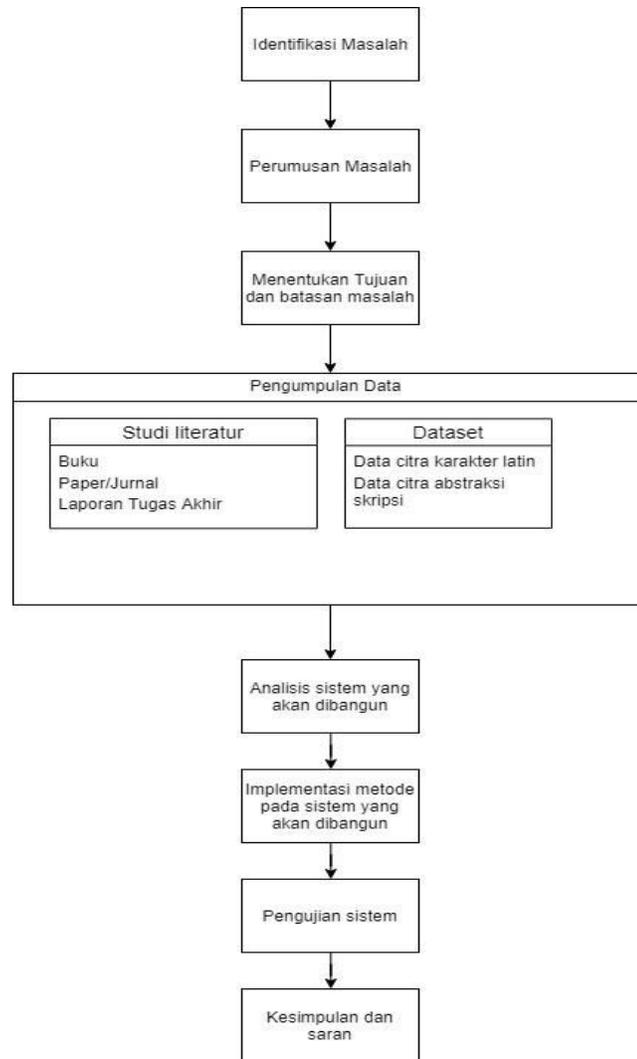
1.4 Batasan Masalah

Agar penelitian yang dilakukan lebih terarah sesuai dengan tujuan yang diinginkan, maka berikut adalah batasan masalah yang digunakan:

1. Citra abstrak dari skripsi hasil *scan* atau foto .
2. *Font* tulisan yang dilatih dan diuji adalah *Times New Roman*
3. Karakter unik yang digunakan pada pelatihan adalah `.,/:[]%*()'#"`
4. Data uji yang digunakan diambil dari abstrak skripsi mahasiswa Teknik Informatika Universitas Komputer Indonesia.

1.5 Metode Penelitian

Metode yang akan dipakai dalam penelitian ini adalah metode kuantitatif, dimana penelitian didasarkan terhadap pengukuran-pengukuran/kuantitas dan penelitian dilakukan dengan hipotesis dan tujuan yang jelas. Hasil akhir akan berupa suatu angka persentase keakurasian sistem pengklasifikasian. Alur penelitian yang akan dilakukan sebagai berikut.



Gambar 1.1 Alur Penelitian

Penelitian akan dimulai dengan pengidentifikasian masalah, kemudian dilanjutkan dengan merumuskan masalah, menentukan tujuan dan batasan masalah, pengumpulan data, analisis sistem yang akan dibangun, implementasi, pengujian dan membuat kesimpulan dan saran.

1.5.1 Metode Pengumpulan Data

Metode yang dilakukan dalam pengumpulan data penelitian adalah sebagai berikut:

a. Studi Literatur

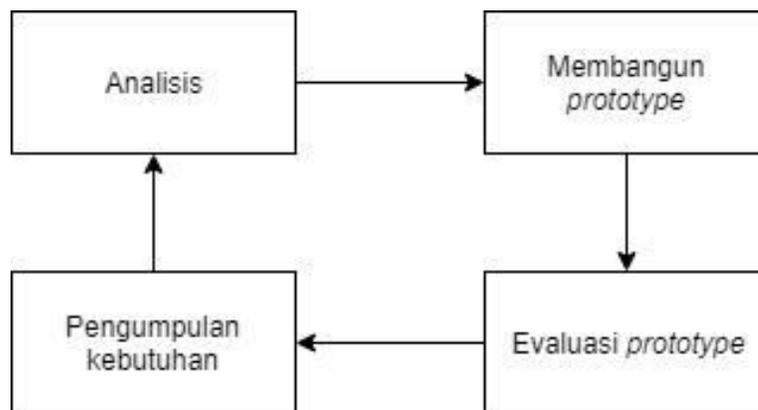
Pengumpulan data dengan cara mengumpulkan jurnal, literatur, paper, dan bahan bacaan yang ada kaitannya dengan judul penelitian.

b. Pengumpulan data latih dan data uji

Data yang akan dipakai untuk pelatihan adalah data citra dari tiap karakter latin dan untuk data uji akan digunakan citra hasil foto dan *scan*.

1.5.2 Metode Pembangunan Perangkat Lunak

Metode pembangunan perangkat lunak yang akan digunakan dalam penelitian ini adalah metode *prototype*. Karena pada pembangunan sistem pengenalan karakter tulisan didasarkan atas analisis, pengembangan, pengujian dan menyempurnakan sistem kembali ke tahap pengembangan. *Prototyping* adalah metode pengembangan perangkat lunak yang cepat dan pengujian terhadap model kerja dari aplikasi baru melalui proses interaksi dan berulang-ulang. Berikut adalah tahapan-tahapan dalam metode prototyping:



Gambar 1.2 Metode Pembangunan Perangkat Lunak

1. Tahapan-tahapan yang terdapat pada metode *prototype* adalah sebagai berikut: Pengumpulan kebutuhan

Kegiatan yang dilakukan pada tahap ini adalah mengumpulkan dan menganalisis kebutuhan data dengan cara studi literatur hal-hal yang

berkaitan dengan metode-metode yang akan digunakan pada penelitian ini. Selain itu juga untuk mengumpulkan sampel data yang akan digunakan.

2. Analisis

Analisis dilakukan mulai dari data yang akan digunakan hingga proses-proses yang akan dilakukan pada sistem pengenalan karakter tulisan.

3. Pembangunan *prototype*

Pada tahap ini proses-proses dan kebutuhan sistem yang telah dianalisa pada tahap sebelumnya akan diimplementasikan. Sehingga akan menciptakan sebuah *prototype* baru yang siap dievaluasi.

4. Evaluasi *prototype*

Pada tahap ini *prototype* yang telah dibuat pada tahapan sebelumnya akan diuji untuk melihat apakah *prototype* sudah berjalan dengan baik. Jika masih ada kekurangan, maka *prototype* akan diperbaiki dengan cara kembali ke tahapan-tahapan sebelumnya. Proses akan terus berulang sampai *prototype* dirasa berhasil mendekati harapan.

1.6 Sistematika Penulisan

Sistematika penulisan laporan penelitian ini disusun untuk memberikan gambaran tentang penelitian yang dijalankan. Sistematika penulisan sebagai berikut.

BAB I PENDAHULUAN

Dalam bab ini akan membahas tentang latar belakang permasalahan, mencoba mengidentifikasi permasalahan yang dihadapi, menentukan tujuan dan kegunaan penelitian, yang kemudian diikuti dengan pembatasan masalah, serta sistematika penulisan.

BAB II TINJAUAN PUSTAKA

Dalam bab ini akan membahas berbagai konsep dasar dan teori-teori pendukung yang berkaitan dengan penelitian yang akan dilakukan.

BAB III ANALISIS DAN PERENCANAAN SISTEM

Bab ini berisi analisis pembangunan sistem seperti gambaran umum sistem, analisis kebutuhan fungsional dan analisis kebutuhan non fungsional.

BAB IV IMPLEMENTASI DAN PENGUJIAN SISTEM

Bab ini berisi pembahasan mengenai implementasi dalam bahasa pemrograman yaitu implementasi kebutuhan perangkat keras dan perangkat lunak dan tahap-tahap dalam melakukan pengujian perangkat lunak.

BAB V KESIMPULAN DAN SARAN

Bab ini berisi hal-hal yang bisa disimpulkan dari hasil keseluruhan penelitian yang dibuat, serta saran untuk pengembangan sistem selanjutnya.

