

1.PENDAHULUAN

1.1 Latar Belakang

Banyaknya artikel-artikel *text* menarik yang disajikan dalam format gambar seperti *.jpg dan *.bmp. Sangat disayangkan sekali, andaikan artikel penting yang berlembar-lembar tersebut, ingin dijadikan sebagai sumber dari sebuah karya tulis harus diketik ulang seluruhnya. Hal tersebut cukup memakan waktu dan tenaga. Oleh karena itu diperlukan suatu teknik untuk ‘mengkonversi’ teks yang berformat gambar menjadi format *.txt agar dapat di-copy dan di-edit. Teknik ini menggunakan sistem *Optical Character Recognition* (OCR) [1].

Optical Character Recognition (OCR) adalah sebuah sistem komputer yang dapat membaca huruf, baik yang berasal dari sebuah pencetak (printer atau mesin ketik) maupun yang berasal dari tulisan tangan. Optical Character Recognition (OCR) merupakan solusi untuk memudahkan usaha mendigitalisasikan informasi dan pengetahuan [1].

Salah satu metode dalam Optical Character Recognition (OCR) adalah metode struktur. Dalam metode yang berbasis struktur, setiap pola yang diproses dinyatakan sebagai gabungan beberapa struktur elementer. Proses pengenalan dilakukan dengan mencocokkan komposisi struktur elementer dengan struktur yang sudah disimpan memakai pendekatan karakteristik titik dan vektor. Sehingga memudahkan dan mempercepat dalam proses pengenalan dan pelatihan karakter.

Pendekatan karakteristik titik dan vektor merupakan metode untuk mengambil ciri dari karakter, sehingga dapat membedakan karakter yang satu dengan yang lain. Pendekatan ini dapat meningkatkan akurasi pengenalan karakter dari metode struktur. Sehingga dapat menghasilkan Optical Character Recognition (OCR) dengan akurasi lebih dari 90 persen .

Pada tugas akhir ini penulis mencoba membangun suatu aplikasi OCR dengan menggunakan pendekatan ciri titik dan vektor pada ekstraksi cirinya. Diharapkan metode yang digunakan dapat dijadikan referensi sebagai salah satu metode untuk mengidentifikasi karakter yang handal, dan aplikasi OCR yang dihasilkan memiliki tingkat akurasi lebih dari 90 persen.

1.2 Perumusan Masalah

Masalah yang akan dibahas dalam Tugas Akhir ini yaitu:

1. Bagaimana proses ekstraksi ciri menggunakan pendekatan struktur menggunakan karakteristik titik dan vektor
 2. Bagaimana proses pembelajaran terhadap input-input sampel karakter.
 3. Bagaimana proses pencocokan gambar input dengan sampel yang ada dalam database.

Pembatasan Masalah

Batasan-batasan masalah yang digunakan dalam tugas akhir ini adalah:

1. Format file masukan dalam format BMP dan JPG.
2. Text pada file gambar yang akan diinterpretasikan harus dalam posisi mendatar dan terpisah antar karakternya, serta bukan dalam format *italic*, *underline*, ataupun *strikethrough*.
3. Tidak menangani kemiringan text.
4. Menggunakan jenis font Times New Roman, MS Sans Serif dan Arial.
5. Resolusi minimal file gambar adalah 200 ppi.

1.3 Tujuan

Tugas Akhir ini bertujuan untuk :

1. Membangun suatu perangkat lunak untuk mengidentifikasi karakter pada suatu file gambar yang berasal dari hardcopy dokumen atau dari sumber lainnya, dengan menggunakan pendekatan berbasis struktur menggunakan karakteristik titik dan vektor pada ekstraksi cirinya.

2. Menganalisis performansi perangkat lunak OCR dengan parameter tingkat keakuratan identifikasi.

1.4 Metodologi Penyelesaian Masalah

Metodologi penyelesaian masalah yang akan dilakukan dalam menyelesaikan tugas akhir ini adalah:

1. Studi literatur

Studi literatur mengenai konsep-konsep pengenalan karakter dan pengolahan citra pada umumnya.

2. Analisa dan Perancangan perangkat lunak.

Menganalisis dan merancang perangkat lunak dengan pemrograman terstruktur dan membangun perangkat lunak dengan pendekatan berbasis struktur menggunakan karakteristik titik dan vektor.

3. Implementasi

Mengimplementasikan sistem menggunakan bahasa pemrograman Visual Basic .NET dengan pendekatan berbasis struktur menggunakan karakteristik titik dan vektor.

4. Pengujian sistem

Melakukan pengujian dengan mengambil sampel tulisan tangan 20 orang dan melakukan pengujian dengan perangkat lunak dengan pendekatan berbasis struktur menggunakan karakteristik titik dan vektor pada sistem.

5. Penyusunan laporan dan kesimpulan, menyusun laporan tugas akhir dan menarik kesimpulan akhir berdasarkan analisis yang dilakukan.