

1. PENDAHULUAN

1.1 Latar Belakang

Berkembang pesatnya teknologi saat ini telah membantu mengerjakan pekerjaan manusia dengan cepat pula, sehingga manusia dalam melakukan aktifitas sehari-hari tidak lepas dari teknologi. Dengan berkembang pesatnya teknologi maka banyak pula teknologi-teknologi baru yang juga bisa membantu manusia dalam beraktifitas. Namun teknologi yang ada saat ini, tidak membuat manusia meninggalkan tulisan tangan sebagai aktifitas kesehariannya. Dalam bidang teknologi, tulisan tangan merupakan disiplin ilmu yang mengklasifikasikan object dalam banyak kategori atau kelas. Tulisan tangan termasuk dalam bidang pengenalan pola (*pattern recognition*) dan sudah banyak diteliti.

Pengenalan angka tulisan tangan mudah dilakukan oleh manusia, namun tidak demikian dengan komputer. Hal ini dikarenakan banyaknya variasi bentuk dan ukuran setiap karakter angka yang dituliskan oleh manusia. Namun manusia juga membutuhkan waktu yang lama jika jumlah data atau angka yang dikenali terlalu banyak. Dengan banyaknya penelitian yang sudah dilakukan, dapat dibangun sebuah sistem yang dapat melakukan pengenalan dan interpretasi karakter angka tulisan tangan menggunakan bantuan komputer. Sistem ini dikenal dengan sebutan Optical Character Recognition (OCR).

OCR merupakan salah satu kemampuan komputer untuk mengenali dan menerjemahkan karakter pada sebuah dokumen fisik berupa gambar, foto, dan dokumen lainnya ke dalam teks komputer [1]. OCR dibagi menjadi 2 metode : *Metode online* menerapkan koordinat dua dimensi dari poin-poin penulisan direpresentasikan sebagai fungsi waktu dan urutan setiap garis yang dituliskan juga disimpan secara realtime untuk mengenali kata yang dituliskan, biasanya metode *online* menggunakan perangkat khusus seperti *Smartphone*, *tablet PC* atau *touch screen*. Sedangkan metode *offline* menerapkan konversi secara

otomatis tulisan pada sebuah citra menjadi karakter yang dapat diolah oleh komputer dan aplikasi perosesan teks dan tidak memerlukan perangkat khusus.

Berdasarkan latar belakang diatas, penulis membuat sebuah sistem yang dapat mengenali angka tulisan tangan (*offline system*) dengan beberapa tahap, *preprocessing* ditentukan titik pusat gravitasinya(*Center Of Gravity(COG)*). COG ini digunakan untuk membuat histogram, titik pertama yang menyentuh karakter mewakili lapisan pertama dari histogram, jika perpanjangan garis belum mencapai batas wilayah, titik berikutnya mewakili lapisan kedua dari histogram. Proses ini di ulang sampai garis menyentuh batas dari karakter. Dataset yang di gunakan untuk menentukan ciri yang terbentuk yaitu dataset Mixed National Institute of Standards and Technology (MNIST) dan dataset formulir model C1 dimana dataset formulir model C1 merupakan formulir yang digunakan dalam perhitungan suara KPU dan dalam sistem ini hanya mengambil angka tulisan tangan yang tertera di dalam formulir model C1.

1.2 Rumusan Masalah

Metode ekstraksi ciri yang digunakan dalam membangun sistem adalah *Star Layered Histogram Features* yang di adaptasi dari[9]. Rumusan masalah yang dibahas dalam tugas akhir ini adalah:

1. Bagaimana membangun sistem pengenalan angka tulisan tangan yang dapat mengenali angka tulisan tangan pada citra formulir model C1 dengan menggunakan fitur *Star Layered-Histogram*?
2. Bagaimana kinerja sistem pengenalan bila data testing ciri formulir model C1 diuji berdasarkan data training ciri MNIST?

Agar pengerjaan tugas akhir menjadi jelas, maka digunakan batasan-batasan sebagai berikut.

1. Karakter yang dapat dikenali sistem adalah karakter angka 0 sampai 9.
2. Data citra angka yang digunakan adalah:

- a. data set MNIST sebanyak 50000 sampel data training dan 10000 sampel data testing, dan
 - b. dataset formulir model C1 sebanyak 2100 sampel data training dan 550 sampel data testing.
3. Pembuatan dataset formulir C1 dilakukan dengan cara rotasi formulir dan cropping karakter angka tulisan tangan secara manual.
 4. Tidak menangani karakter yang melebihi batas kotak.
 5. Menggunakan variasi posisi natural (posisi *dataset* tidak diubah).
 6. Sistem pengenalan angka tulisan tangan yang dibangun bersifat offline.

1.3 Tujuan

Berdasarkan permasalahan yang telah dirumuskan, maka tujuan tugas akhir ini adalah sebagai berikut:

1. Membangun sistem pengenalan angka tulisan tangan yang dapat mengenali angka tulisan tangan pada citra formulir model C1 dengan menggunakan Star Layered-Histogram.
2. Menguji dan menganalisis tingkat akurasi pengenalan data set ciri formulir model C1 yang diuji berdasarkan dataset ciri MNIST.

1.4 Metodologi Penyelesaian Masalah

Metodologi yang digunakan dalam menyelesaikan tugas akhir ini adalah sebagai berikut :

1. Studi Literatur

Mencari referensi dan mempelajari metode yang digunakan, yaitu *Star Layered Histogram Features* dari buku, paper, maupun artikel di internet.

2. Pengumpulan Data

Pada tahap ini, dataset MNIST format CSV diunduh dari situs web <http://www.pjreddie.com/projects/mnist-in-csv/>. Citra *scan* formulir model C1 juga diunduh dari situs web resmi KPU RI di alamat <http://pilpres2014.kpu.go.id/c1.php>. Citra *scan* formulir model C1 yang telah diunduh dipotong-potong pada bagian tulisan tangan angka di dalam kotak formulir, kemudian dibagi ke dalam dataset latihan dan dataset uji.

3. Perancangan Sistem

Membuat rancangan desain sistem mulai dari desain tahap *preprocessing* citra, proses ekstraksi ciri, dan pengenalan ciri.

4. Implementasi Sistem

Melakukan implementasi sistem menggunakan perangkat lunak MATLAB R2015b sesuai rancangan yang telah dibuat.

5. Pengujian dan Analisis Hasil Implementasi

Sistem yang telah dibuat diuji berdasarkan enam skenario pengujian, diantaranya sebagai berikut :

- a. Skenario 1, menguji pengaruh ekstraksi ciri fitur *Star Layered Histogram* menggunakan data latihan MNIST kemudian di ujikan kepada data uji C1 menggunakan parameter terbaik dengan menghitung tingkat akurasi.
- b. Skenario 2, menguji dan menganalisis pengenalan data *testing* ciri fitur *Star Layered Histogram* pada dataset formulir model C1.
- c. Skenario 3, menguji dan menganalisis pengenalan data *testing* ciri fitur *Star Layered Histogram* pada dataset gabungan antara MNIST formulir model C1.
- d. Skenario 4, menguji dan menganalisis pengenalan data *testing* ciri fitur *Star Layered Histogram* menggunakan dataset C1.

- e. Skenario 5, menguji dan menganalisis pengenalan data *testing* ciri fitur *Star Layered Histogram* dengan 16 ciri tambahan menggunakan dataset C1.
- f. Skenario 6, menguji dan menganalisis pengenalan data *testing* ciri fitur *Star Layered Histogram* menggunakan klasifikasi yang berbeda.

Pengujian skenario 1 sampai dengan skenario 5 dilakukan dengan menggunakan metode klasifikasi K-NN, sedangkan skenario 6 menggunakan metode klasifikasi SVM untuk mengukur akurasi dari hasil ekstraksi ciri yang menjadi fokus dalam tugas akhir ini. Kemudian analisis dilakukan berdasarkan hasil yang didapat.

6. Penyusunan Laporan Akhir

Menarik kesimpulan berdasarkan hasil implementasi dan pengujian, dan membuat buku laporan tugas akhir beserta dokumentasi selama proses kerja.

1.5 Sistematika Penulisan Tugas Akhir

Sistematika yang digunakan dalam penulisan laporan akhir tugas akhir ini terdiri dari 5 bab, yaitu sebagai berikut.

BAB I PENDAHULUAN

Bab ini membahas latar belakang pemilihan metode, rumusan masalah dan batasan masalah, tujuan penelitian, metodologi yang digunakan dalam penelitian tugas akhir, dan sistematika penulisan laporan tugas akhir.

BAB II DASAR TEORI

Bab ini membahas teori-teori dasar yang mendukung pembangunan sistem pengenalan angka tulisan tangan. Teori dasar yang dibahas antara lain mengenai representasi Citra Digital, Pengenalan Karakter Optis (OCR), *Image Preprocessing*, Ekstraksi Ciri dengan fitur *Star Layered Histogram*, dan Klasifikasi secara umum.

BAB III PERANCANGAN SISTEM

Kebutuhan perangkat untuk implementasi sistem dan rancangan desain *preprocessing* citra, metode pengambilan ciri menggunakan fitur *Star Layered Histogram*, dan penggunaan metode klasifikasi diuraikan dalam bab ini.

BAB IV PENGUJIAN DAN ANALISIS

Bab ini membahas pengujian terhadap sistem pengenalan angka tulisan tangan yang telah dibangun berdasarkan skenario pengujian, dan melakukan analisis pengaruh metode dan hasilnya terhadap akurasi kinerja sistem.

BAB V PENUTUP

Berisi kesimpulan dari implementasi, pengujian, dan analisis yang telah dilakukan dalam tugas akhir ini, dan saran-saran untuk pengembangan selanjutnya.