

BAB I

PENDAHULUAN

1.1 Latar Belakang Masalah

Teks merupakan jenis karangan atau tulisan yang berisi paparan kejadian sesuai dengan konteks dan tujuan yang ingin dibahas di dalam tulisan tersebut. Teks atau tulisan sering di temukan diberbagai tempat seperti nama jalan, nama toko, spanduk, petunjuk arah, peringatan dan lain sebagainya. Teks berperan penting dalam kehidupan sehari-hari karena teks memberikan informasi. Hal ini akan menjadi masalah pada seorang yang memiliki masalah pada penglihatan atau biasa disebut tunanetra dan pada seorang wisatawan mancanegara yang tidak mengerti bahasa sekitar [1]. Untuk itu saat ini banyak dibangun suatu perangkat atau sebuah sistem yang dapat membantu seorang dengan masalah pada penglihatannya agar melakukan aktivitas sehari-hari.

Deteksi teks adalah proses mendeteksi teks yang ada dalam gambar, diikuti dengan mengitarinya dengan kotak pembatas persegi panjang. Deteksi teks dapat dilakukan menggunakan teknik berbasis gambar atau teknik berbasis frekuensi. Dalam teknik berbasis gambar, gambar disegmentasi menjadi beberapa segmen. Setiap segmen adalah komponen *pixel* yang terhubung dengan karakteristik serupa. Fitur statistik dari komponen yang terhubung digunakan untuk mengelompokkannya dan membentuk teks. Pendekatan pembelajaran mesin seperti mesin vektor pendukung dan jaringan saraf convolutional digunakan untuk mengklasifikasikan komponen ke dalam teks dan non-teks.

Ketersediaan perangkat seperti smartphone, kamera dan lainnya yang dapat menghasilkan gambar dan video, menjadikan deteksi teks pada gambar menjadi luas. Sebagian besar informasi tekstual yang ada pada gambar diharuskan secara otomatis diproses untuk pengenalan obyek. Ini membutuhkan otomatis deteksi, segmentasi dan pengakuan entitas teks visual dalam gambar. Teks dalam gambar berisi informasi berguna yang dapat digunakan untuk memahami gambar.

Teks dalam gambar diklasifikasikan sebagai *document image*, *scene text* dan *caption image*. *Document image* diperoleh dari scan sampul buku, dokumen cetak dan lain-lain. *Scene text*

juga disebut sebagai teks grafis yang dapat ditemukan di gambar yang berisi iklan seperti nama jalan, rambu lalu lintas dan lain-lain. *Caption image* berisi teks yang dimasukkan ke dalam gambar atau biasa disebut teks buatan [1]. Mendeteksi dan melokalisasi teks pada gambar menjadi tantangan dalam komputer seperti pencahayaan, *low resolution*, blur. Selain itu, tantangan lain efek kamera kabur, latar belakang yang kompleks.

Penelitian ini menggunakan metode *directional discrete cosine transform* (DDCT) untuk mendeteksi teks pada *scene text image*. Umumnya, 2D Discrete Cosine Transform (DCT) diterapkan pada blok gambar yang tidak tumpang tindih dengan ukuran persegi ($N \times N$). DCT 2D (DCT konvensional) diimplementasikan secara terpisah dengan dua titik N DCT, satu disepanjang arah horizontal dan satu lagi sepanjang arah vertikal [2]. Menurut sistem visual manusia (HVS), kedua arah horizontal dan vertikal sangat penting. DCT 2D memberikan hasil terbaik untuk blok gambar di mana tepi horizontal dan/atau vertikal mendominasi. Tepian selain arah horizontal dan vertikal juga sangat penting dan orientasi tepi ini sangat bervariasi dari satu blok gambar ke yang lain. Dalam situasi ini, *Directional Discrete Cosine Transform* (DDCT) adalah pilihan terbaik daripada DCT konvensional untuk menghindari banyak komponen AC non-nol dalam domain frekuensi. DDCT yang mana DCT pertama diterapkan sepanjang arah yang mendominasi dalam blok gambar, sedangkan DCT titik-N kedua dapat diterapkan sesuai dengan pengaturan koefisien setelah N-point DCT pertama. Akhir-akhir ini, DDCT secara luas digunakan untuk pengkodean gambar dan video [3].

Sementara untuk *dataset* yang akan digunakan adalah *dataset Focused Scene Text Robust Reading Competition ICDAR2013*. *Dataset* ini dipilih karena skenario khas untuk aplikasi membaca teks dan menerjemahkan teks, di mana pengguna secara eksplisit mengarahkan fokus kamera pada konten teks yang menarik. *Dataset ICDAR2013* hampir sama dengan *dataset* di *ICDAR2011 Robust Reading Competition*. Perbedaan dari *dataset ICDAR2001* adalah revisi teks *groundtruth* di beberapa gambar. Sebagai tambahan sejumlah kecil gambar digandakan selama pelatihan dan *test set* dikecualikan. Jadi,

dataset ICDAR2013 adalah himpunan bagian dari dataset ICDAR2011 [4].

Dengan demikian, dengan adanya penelitian ini diharapkan mampu menghasilkan kompleksitas yang lebih baik untuk melokalisasi teks.

1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, dapat dirumuskan permasalahan yang dihadapi yaitu bagaimana akurasi lokalisasi teks pada *scene text image* dengan menggunakan metode *directional discrete cosine transform*.

1.3 Tujuan dan Manfaat

1.3.1 Tujuan Penelitian

Tujuan yang ingin dicapai dalam penelitian ini adalah menentukan akurasi metode *Directional Discrete Cosine Transform* (DDCT) untuk melokalisasi letak teks pada *scene text image*.

1.3.2 Manfaat Penelitian

Manfaat yang ingin dicapai dalam penelitian ini adalah diharapkan metode yang diusulkan bisa menjadi rekomendasi teknik lokalisasi teks pada gambar yang lebih efisien.

1.4 Batasan Masalah

Pada penelitian ini terdapat beberapa batasan masalah sebagai berikut:

1. Metode yang digunakan pada penelitian ini adalah *Directional Discrete Cosine Transform* (DDCT) menggunakan mode 3 dan ukuran blok 8.
2. *Dataset* yang digunakan pada penelitian ini adalah *Focused Scene Text* ICDAR2013 bagian tes set yang berjumlah 233 gambar
3. Ukuran gambar yang digunakan untuk lokalisasi teks adalah 640x480 *pixel*
4. Format gambar yang digunakan untuk lokalisasi teks adalah berekstensi .jpg
5. *Tools* yang digunakan adalah *Matlab R2015a*

1.5 Metodologi Penelitian

Secara umum, penelitian dilakukan dengan beberapa tahapan, yaitu:

a. Studi Literatur

Pada tahap ini dilakukan pengumpulan beberapa informasi yang dibutuhkan serta mempelajari informasi-informasi tersebut sehingga dapat digunakan untuk membantu dalam menyelesaikan pengerjaan penelitian ini. Informasi didapatkan dari membaca buku, mempelajari literatur, makalah, jurnal ilmiah dan juga informasi yang berasal dari beberapa situs web yang berkaitan dengan penelitian, yaitu meliputi:

1. Metode *Directional Discrete Cosine Transform* (DDCT)
2. Proyeksi horizontal dan vertikal
3. *Connected Component*
4. Operasi Morfologi

b. Pengumpulan Data

Pada tahap ini dilakukan untuk memperoleh data. Data yang digunakan adalah *dataset Focused Scene Text Robust Reading Competition ICDAR2013*

c. Analisa dan Perancangan Sistem

Pada tahap ini dilakukan analisa terhadap metode yang akan digunakan seperti *Directional Discrete Cosine Transform* (DDCT), proyeksi, *connected component* dan operasi morfologi. Selanjutnya akan dibangun perancangan sistem yang dibangun dengan menentukan alur berjalannya sistem yang akan digunakan dalam penelitian ini.

d. Implementasi Sistem

Pada tahap ini, dilakukan pengimplementasian metode kedalam program sesuai dengan analisa dan perancangan yang telah dilakukan seperti membuat fungsi-fungsi seperti *Directional Discrete Cosine Transform* (DDCT), proyeksi, *connected component* dan operasi morfologi yang akan mempermudah membuat dalam pembuatan program selanjutnya.

e. Uji Coba

Sistem yang telah dibuat akan diuji coba terhadap fitur yang telah dibuat sebelumnya. Uji coba juga dilakukan untuk mengetahui akurasi dari metode pembelajaran.

f. Analisa dan Evaluasi

Pada tahap ini, hasil uji coba sistem akan di analisa untuk mengetahui hasil akurasi dari metode yang digunakan dan untuk mengetahui kelebihan dan kekurangan metode yang digunakan. Jika hasil yang di dapat kurang dari apa yang diharapkan maka akan dilakukan evaluasi untuk sistem yang dibuat.

g. Dokumentasi

Pada tahap ini dilakukan penyusunan laporan yang berisi dokumentasi penelitian dari tahap awal hingga akhir.

1.6 Sistematika Penulisan

Secara garis besar materi laporan skripsi ini terbagi dalam beberapa bab yang tersusun sebagai berikut:

BAB I PENDAHULUAN

Bab ini berisikan latar belakang dari penelitian yang akan dilakukan, rumusan masalah, tujuan dan manfaat, batasan masalah, metodologi penelitian dan sistematika penelitian.

BAB II STUDI PUSTAKA

Pada bab studi pustaka dijelaskan mengenai materi-materi yang menjadi dasar pedoman penyusunan penelitian ini.

BAB III PERANCANGAN SISTEM

Bab ini menjelaskan mengenai alur dari sistem yang dirancang setelah melakukan analisa permasalahan.

BAB IV IMPLEMENTASI SISTEM

Bab ini menjelaskan implementasi sistem, dimana sistem diuji coba dengan berbagai parameter-parameter dan menganalisis tingkat efektifitas dari sistem.

BAB V PENUTUP

Bab penutup berisi kesimpulan dari hasil uji coba sistem dan saran untuk pengembangan selanjutnya.