

# BAB I

## Pendahuluan

### 1.1 Latar Belakang Penelitian

*Optical Character Recognition (OCR)* adalah salah satu teknologi identifikasi otomatis hasil *machine learning* yang dapat merubah kumpulan piksel yang membentuk teks pada citra menjadi karakter atau teks digital dan bisa diubah pada komputer. *OCR* terdiri dari beberapa komponen yaitu *scanning*, *preprocessing*, ekstraksi fitur adalah untuk menangkap karakteristik penting dari simbol, dan pengenalan karakter berdasarkan hasil deteksi [1]. Dalam pengembangannya, banyak penelitian yang berfokus dalam lokalisasi segmentasi karakter deteksi untuk mempercepat kinerja *OCR* agar dalam pengenalan sebuah karakter atau kata, sistem tidak perlu mendeteksi keseluruhan citra dan hanya “membaca” daerah-daerah yang sudah tersegmentasi untuk mempercepat proses identifikasi karakter. Lokalisasi atau pendeteksian teks sendiri memiliki beragam permasalahan pada subjek citra yang berbeda

Secara strukturnya teks terbagi menjadi beberapa klasifikasi, pada klasifikasi berdasarkan teksturnya teks citra terbagi menjadi *Natural Scene Image* dan *Document Image*. Dalam pendekatannya, lokalisasi teks citra pada kedua tipe tekstur juga memiliki cara yang berbeda untuk mengekstrak teks dari citra. Pada *Natural Scene Image* digunakannya pendekatan *Top-down*, *Bottom-up*, dan *Hybrid*. Sedangkan pada *Document Image* digunakan pendekatan berdasarkan baris, kata, dan karakter [2]. Sehingga *Natural Scene Image* maupun *Document Image* tidak hanya memiliki fokus permasalahan yang berbeda, cara pemecahan masalahnya pun berbeda-beda. Meskipun memiliki focus masalah maupun cara penyelesaian yang berbeda, dalam mengatasi permasalahan lokalisasi teks, keduanya mengalami proses evolusi yang mana dalam proses deteksi teks terciptanya dua metode yaitu *Traditional Method* dan *Deep Learning Method* [3]. *Traditional Method* adalah metode yang secara garis besar menggunakan dua pendekatan yaitu *Sliding Window Approache* dan *Connected-Component Analytic Approache*. Sedangkan pada metode *Deep Learning* beberapa cara yang digunakan yaitu pendekatan *Bounding-Box Regression*, *Segmentation*, dan *Hybrid* [4].

Meski terklasifikasi sebagai permasalahan yang berbeda, *Natural Scene Image* maupun *Document Image*, bukan berarti cara penyelesaian *Natural Scene Image* tidak dapat mendeteksi teks pada *Document Image*. Hal ini dikarenakan adanya permasalahan *Natural Scene Image* pada *Document Image* sehingga memerlukan pemecahan solusi yang digunakan pada *Natural Scene Image* begitu juga sebaliknya. Pada penelitian Shuang Liu dkk (2022), mereka memaparkan alasan tidak menggunakan solusi *Natural Scene Image*, sistem *Text Flow*, dalam mendeteksi baris teks pada *Document Image Handwriting* atau tulis tangan dikarenakan tata letak karakter, kata, maupun kalimat pada dokumen tertulis memiliki spasi maupun ukuran karakter yang beragam bahkan ketinggian sebuah karakter ada yang bisa sejajar dengan karakter diatas maupun dibawahnya. Akan tetapi Liu juga berpendapat bahwa sistem *Text Flow* sendiri bisa atau bahkan efektif dalam mendeteksi teks pada *Document Image* [5].

*Text Flow* sendiri merupakan system yang diciptakan Tian dk (2016) untuk mengatasi permasalahan deteksi baris teks mandarin dalam *Natural Scene Image*, yang mana pada saat itu pendeteksian teks mandarin, yang bisa memiliki banyak garis dalam satu karakter, masih jarang dikembangkan. *Text Flow* menggunakan metode tradisional *Sliding Window* tipe *Bottom-up*, yang mana dalam sistemnya hanya digunakan dua langkah yaitu penentuan karakter kandidat menggunakan *Cascade Boosting* dan pengelompokan sebagai kelompok baris teks menggunakan *Min-Cost Flow Network*, sehingga didapatkan nilai *F-Score* sebesar 80,89% pada kompetisi *ICDAR2011* [6]. Selain *Text Flow* ada juga penelitian tradisional *Bottom-up Pipeline* yang pengelompokan baris teksnya tidak terlalu berbeda dengan penggunaan *Min-Cost Flow* dan metode penentuan karakter kandidat yang berbeda [7], [8] yang menggunakan metode pendekatan *Connected-Component Analytic* algoritma *Maximally Stable Extremal Region* dengan hasil akurasi yang sedikit lebih tinggi daripada *Text Flow* pada dataset yang sama.

Penerapan dari alur tradisional *Bottom-up Pipeline* sendiri lebih banyak berfokus kepada permasalahan *Natural Scene Image*. Adapun digunakan dalam permasalahan *Document Image* menghasilkan nilai yang lebih rendah dibandingkan penerapan pada *Natural Scene Image* yaitu sebesar *F-Score* 62% dengan menggunakan *Tesseract* sebagai penentu karakter kandidat dan *Min-Cost Flow*

*Network* sebagai pengelompokan baris teks [9]. Objek yang digunakan pada penelitian Indra dkk (2021) sendiri menggunakan sertifikat yang mana memiliki karakteristik kesulitan dari *Natural Scene Image* dengan tercampurnya objek *non-text* dan latar belakang yang rumit juga karakteristik *Document Image* yang memiliki cukup banyak teks dan *layout* tersendiri. Indra sendiri menyatakan bahwa hasil akurasi yang didapat dipengaruhi oleh proses segmentasi yang tidak memuaskan dan menyarankan untuk mengganti metode segmentasinya. Meski dikatakan metode *Bottom-up* yang biasa digunakan pada *Natural Scene Image* bisa dengan efektif mendeteksi teks pada *Document Image*, apakah metode tersebut bisa memberikan kinerja yang lebih baik bila diselesaikan permasalahan segmentasinya atau metode tersebut memang kesulitan dalam mendeteksi teks yang beragam dan memiliki kuantitas yang banyak. Oleh karena itu penelitian ini akan menggunakan judul **“Perbandingan Akurasi Deteksi Baris Min-Cost Flow Network dan Tesseract Pada Citra Digital Berdasarkan Perbedaan Metode Segmentasi”**.

## 1.2 Perumusan Masalah Penelitian

Permasalahan umum yang dikaji berdasarkan latar belakang di atas adalah sebagai berikut:

1. Bagaimana penerapan deteksi baris teks pada *Min-Cost Flow Network* dan *Tesseract* terhadap keragaman kondisi teks pada teks citra dengan integrasi metode segmentasi yang berbeda?
2. Bagaimana perbandingan akurasi deteksi baris teks pada *Min-Cost Flow Network* dan *Tesseract* terhadap keragaman kondisi teks pada teks citra dengan integrasi metode segmentasi yang berbeda?
3. Apakah integrasi dengan setiap metode segmentasi yang dibandingkan dapat meningkatkan akurasi *Min-Cost Flow Network* terhadap keragaman kondisi teks pada teks citra?
4. Apakah integrasi pada setiap metode segmentasi yang dibandingkan dapat membatasi terdeteksinya objek *non-character*?

### 1.3 Tujuan Penelitian

Adapun tujuan dari penelitian ini adalah sebagai berikut:

1. Mengetahui bagaimana penerapan deteksi baris teks pada *Min-Cost Flow Network* dan *Tesseract* terhadap keragaman kondisi teks pada teks citra dengan integrasi metode segmentasi yang berbeda
2. Mengetahui bagaimana perbandingan akurasi deteksi baris teks pada *Min-Cost Flow Network* dan *Tesseract* terhadap keragaman kondisi teks pada teks citra dengan integrasi metode segmentasi yang berbeda
3. Mengetahui apakah integrasi dengan setiap metode segmentasi yang diusulkan dapat meningkatkan akurasi *Min-Cost Flow Network* terhadap keragaman kondisi teks pada teks citra
4. Mengetahui apakah integrasi pada setiap metode segmentasi yang dibandingkan dapat membatasi terdeteksinya objek *non-character*

### 1.4 Batasan Masalah Penelitian

Untuk memfokuskan penelitian dan memperjelas penyelesaian sehingga mudah dipahami dan penyusunannya lebih terarah, maka dilakukan pembatasan masalah sebagai berikut:

1. Teks citra yang digunakan hanya menggunakan format berkas jpg
2. Kondisi teks citra yang akan dideteksi hanya citra dengan teks horizontal - teks dari kiri ke kanan
3. Teks citra yang akan dideteksi berupa dokumen teks citra infografis pada dataset *Infographics VQA* [10] sebanyak 358 gambar dengan batas ukuran  $3500px \times 3500px$
4. Metode segmentasi yang digunakan untuk perbandingan hanya akan menggunakan kombinasi segmentasi metode *Projection Profile* dan *CCL*; dan segmentasi menggunakan *Maximally Stable Extremal Region (MSER)* dan Algoritma *Canny*, dengan *CCL* sebagai pembanding utama

## 1.5 Manfaat Penelitian

Adapun manfaat yang diperoleh dari penelitian ini adalah sebagai berikut:

1. Mengetahui metode segmentasi yang tepat untuk mendapatkan kualitas tertinggi deteksi baris teks pada *Min-Cost Flow Network* dan *Tesseract*
2. Mengetahui kinerja metode-metode segmentasi dalam kombinasi dengan *Min-Cost Flow Network* dan *Tesseract* dalam mendeteksi baris teks

## 1.6 Kerangka Pemikiran Penelitian

Kerangka pemikiran dari penelitian tugas akhir ini dapat dilihat pada Gambar 1:



Gambar 1 Kerangka Pemikiran

Penelitian ini didasari oleh keadaan sistem deteksi baris pada citra digital yang menggunakan *Min-Cost Flow Network*. Pada penelitian Tian (2014) didapati nilai *F-Score* 80.89% menyatakan bahwa *Min-Cost Flow Network* dengan penggunaan *Cascade Boosting* sebagai pendeteksi kandidat karakter dapat digunakan untuk mendeteksi baris teks pada citra digital *Natural Scene Image*. Akan tetapi pada penelitian Indra (2021) didapati nilai *F-Score* sebesar 62% dengan *Tesseract* sebagai pendeteksi kandidat karakter pada citra digital *Document Image* dan mengalami *error* seperti kesulitan dalam mendeteksi karakter yang berdempet, terdeteksinya *noise*, dan terdeteksinya huruf “i” atau “j” sebagai dua buah objek terpisah meski “i” adalah satu kesatuan karakter. Indra (2021) berpendapat bahwa nilai akurasi tersebut dapat ditingkatkan dengan peningkatan kualitas metode segmentasinya sehingga masalah-masalah tersebut dapat terselesaikan. Oleh karena itu penelitian ini akan membandingkan perbedaan hasil akurasi deteksi baris berdasarkan perbedaan metode segmentasi yang digunakan.

### 1.7 Metodologi Penelitian

Metode penelitian yang akan digunakan pada penelitian ini adalah metode kuantitatif dengan *Experimental Research*. Riset eksperimen merupakan *Research that allows for the causes of behavior to be determined*- penelitian yang memungkinkan untuk menentukan penyebab perilaku [11]–[13]. *Experimental Research* adalah penelitian yang diberikan perlakuan (treatment) dengan minimal dilakukannya memanipulasi/mengontrol salah satu variabel (*independent variable*) sedangkan variabel atau faktor lainnya tetap konstan (*dependent variable*) dengan *independent variable* yang akan digunakan adalah hasil deteksi teks metode segmentasi yang berbeda-beda pada penerapan deteksi baris teks menggunakan *Tesseract* dan *Min Cost Flow Network* dengan proses yang digunakan yaitu, *preprocessing (Grayscale, Smoothing, Thresholding)*, *Segmentation* menggunakan metode *CCL*, deteksi kandidat karakter menggunakan *Tesseract* dan deteksi baris teks menggunakan *Min Cost Flow Network*, proses segmentasi menggunakan metode *CCL* inilah yang merupakan *dependent variable* yang digunakan dalam penelitian ini. Alur eksperimen ini diawali dengan dilakukannya studi literatur untuk mendapatkan penelitian-penelitian serupa tentang deteksi teks dan peningkatan akurasi dari sebuah metode deteksi teks, dan diakhiri dengan evaluasi

dari akurasi *F-Score* pada deteksi baris teks menggunakan *Tesseract* dan *Min Cost Flow Network* maupun deteksi baris teks menggunakan *Tesseract* dan *Min Cost Flow Network* yang sudah diintegrasikan dengan solusi metode segmentasi berdasarkan pencarian melalui studi literatur. Desain yang digunakan pada *experimental research* ini adalah metode eksperimen murni (*true experimental*) jenis *The Posttest Only Control Group Design*, yaitu menganalisis hasil penelitian berdasarkan hasil uji coba akhir berdasarkan kelompok uji coba [14]. Kelompok uji coba yang digunakan yaitu kelompok metode segmentasi *CCL* dan dua kelompok metode segmentasi yang diajukan.

Tabel 1 Desain Penelitian: *The Posttest Only Control Group Design*

Kelompok Kontrol (K.K)	→ X0 → O2
Kelompok Eksperimen (K.E)	→ XE → O2

Keterangan: X0 = Segmentasi *CCL*  
 XE = Segmentasi pengajuan  
 O2 = Posttest

Dapat dilihat pada Tabel 1, penelitian ini akan menggunakan dua kelompok variable yaitu kelompok kontrol (K.K) dan kelompok eksperimen (K.E). Masing-masing kelompok tersebut akan mengalami beberapa perlakuan yang sama hingga proses integrasi segmentasi akan terjadinya perbedaan hasil antara K.K dengan K.E. Dengan adanya variabel pengaruh X0 untuk K.K, dan variable XE untuk K.E akan memberikan hasil yang berbeda pada masing-masing kelompok. Berdasarkan perbedaan tersebut proses penelitian tetap berlanjut dengan perilaku yang sama pada K.K maupun K.E, kemudian akan dilihat pada hasil akhir atau *Posttest* (O2) apakah dengan diberikannya perbedaan perlakuan X0 dan X1 akan memberi hasil yang berbeda atau tidak.

### 1.8 Sistematika Penulisan

Sistematika penulisan dokumen tugas akhir ini terdiri dari lima bab yaitu BAB I Pendahuluan; BAB II Kajian Literatur; BAB III Metodologi Penelitian; BAB IV Hasil dan Pembahasan; BAB V Kesimpulan dan Saran.

BAB I terdiri dari tujuh sub bab yaitu Latar Belakang Penelitian; Perumusan Masalah Penelitian; Tujuan Penelitian; Batasan Masalah Penelitian; Manfaat Penelitian; Kerangka Pemikiran Penelitian; Metodologi Penelitian; dan Sistematika Penulisan.

BAB II terdiri tiga sub bab *The State of The Art* yang berisi tentang penelitian-penelitian terdahulu yang serupa dengan permasalahan pada penelitian ini; Tinjauan Pustaka memaparkan sumber pustaka yang menjadi permasalahan penelitian; dan Landasan Teori yang membahas permasalahan teori, metode, konsep maupun model dalam penelitian ini.

BAB III menjelaskan tentang langkah-langkah penelitian yang terdiri dari enam sub bab yaitu Studi Literatur; Analisis deteksi baris teks menggunakan Min-Cost Flow Network; Pengumpulan Data; Analisis Data; Integrasi Metode Segmentasi; Evaluasi.

BAB IV terdiri dari Hasil dan Pembahasan. Bab ini membahas bagaimana pemaparan dari hasil penelitian yang menjawab rumusan permasalahan pada BAB I.

BAB V merupakan simpulan dari hasil penelitian serta saran penulis terhadap penelitian untuk mengevaluasi hasil penelitian.