BAB II
LANDASAN TEORI

2.1 Penelitian Terkait

Pada penulisan ini penulis juga menggali informasi dari beberapa peneliti sebelumnya sebagai bahan referensi dalam membuat penelitian sebagai berikut :

Pada penelitian sebelumya sudah menggunakan metode Optical Character Recognition (OCR) salah satunya Penelitian fokus pada identifikasi plat nomor kendaraan bermotor dengan standar Republik Indonesia dengan studi kasus sistem parkir Itenas Bandung pengenalan karakter menggunakan metode Optical Character Recognition (OCR) dan setelahnya akan ada proses character error handling untuk meningkatkan akurasi identifikasi karakter pada plat nomor (Sugeng, Utoro, & Prabowo, 2020).

Menurut peneliti sebelumnya dengan judul “Penerapan Optical Character Recognition untuk Pengenalan Variasi Teks pada Media Presentasi Pembelajaran” menjelaskan bahwa teks, karakter yang umumnya terdiri dari alfabet dan angka (Nugraha, 2024).

Menurut peneliti sebelumnya Gottschalk menyatakan bahwa dokumen (dokumentasi) dalam pengertiannya yang lebih luas berupa setiap proses pembuktian yang didasarkan atas jenis sumber apapun, baik itu yang bersifat tulisan, lisan, gambaran, atau arkeologis (Mohamad Ali Murtadho, 2016).

Peneliti sebelumnya mejelaskan bahwa, SIM (Surat Ijin Mengemudi) adalah bukti registrasi dan identifikasi yang diberikan olehPolri kepada seseoraang yang telah memenuhi persyaratan administrasi, sehat jasmani dan rohani, memahami peraturan lalu lintas dan trampil mengemudikan kendaraan bermotor (Syach, 2019).

Kesimpulan dari Penelitian-penelitian sebelumnya menunjukkan bahwa teknologi OCR memberikan kontribusi besar dalam meningkatkan efisiensi pengelolaan data di berbagai bidang. Dari pengolahan citra digital hingga pengelolaan dokumen resmi seperti SIM, teknologi ini menunjukkan potensi yang besar untuk dioptimalkan melalui peningkatan akurasi, pengolahan citra yang lebih baik, dan aplikasi pada berbagai konteks studi kasus.

2.2 Optical Character Recognition (OCR)

Optical Character Recognition (OCR) Optical Character Recognition (OCR) adalah metode yang memungkinkan komputer untuk mengenali dan mengekstrak teks dari gambar atau dokumen digital. Menurut Peneliti terdahulu, (Sonita & Khairunnisyah, 2018), Optical Character Recognition (OCR) atau sering disebut dengan Pengenalan Pola karakter yang merupakan salah satu cabang dari pengenalan pola. Salah satu metode pengenalan pola angka atau huruf banyak digunakan algoritma Neural Network Backpropagation, yang sistem kerjanya seperti otak manusia yang dapat mengenali pola setelah mendapatkan pelatihan atau pengecekan terhadap beberapa contoh terlebih dahulu. metode ini memanfaatkan algoritma pemrosesan citra untuk mendeteksi karakter dan mengonversinya menjadi teks yang dapat diedit atau diproses lebih lanjut secara otomatis (Nurhaliza et al., 2022). OCR telah banyak diterapkan dalam berbagai bidang, termasuk pengelolaan dokumen, sistem arsip digital, serta aplikasi perbankan dan keamanan. Proses kerja metode ocr pada umumnya terdiri dari beberapa tahap utama, yaitu:

a. **Cropping**

Pemotongan bagian tertentu dari Citra yang dilakukan secara interaktif oleh Penulis. Dalam proses ini, Penulis secara manual memilih area yang ingin dipotong seperti Identitas yang terdapat pada SIM.

b. **Processing Citra**

Pada tahap ini, citra yang telah di cropping dikonversi ke grayscale, lalu di segmentasi menggunakan metode thresholding outsu. untuk memisahkan teks dari latar belakang serta memperbaiki struktur objek dalam citra biner, menghilangkan noise, serta memperbaiki bentuk karakter pada identitas teks.

c. **Bounding Box**

Hasil citra yang sudah disegmentasikan tahapan selanjutkan bounding box, yang bekerja mendeteksi tiap karakter teks yang terdapat dari citra sim yang sudah di processing.

d. **Optical Character Recognition (OCR)**

Pada metode ini hasil karakter teks yang telah dideteksi oleh bounding box, di ektrak oleh metode ocr.

2.3 Identifikasi

Identifikasi adalah proses mengenali, menentukan, dan membedakan suatu objek, individu, atau informasi berdasarkan karakteristik tertentu (Kedokteran & Kesehatan, 2015). Proses ini digunakan dalam berbagai bidang, seperti pengolahan data, dan sistem informasi. Identifikasi bertujuan untuk memastikan bahwa suatu entitas dapat dikenali dengan benar. Identifikasi dapat dilakukan dengan berbagai metode, tergantung pada tujuan dan lingkup penggunaannya, di antaranya. Identifikasi Dokumen, mengenali teks dan informasi dalam dokumen resmi seperti SIM, dan paspor dengan menggunakan teknologi seperti Optical Character Recognition (OCR). Dalam konteks penelitian ini, identifikasi dilakukan dengan menggunakan teknologi OCR untuk mengenali teks pada dokumen Surat Izin Mengemudi (SIM), sehingga dapat meningkatkan efisiensi dan akurasi dalam pengelolaan data.

2.4 Surat Izin Mengemudi (SIM)

SIM dalam konteks penelitian ini, OCR digunakan untuk mengenali teks pada dokumen Surat Izin Mengemudi (SIM). Implementasi teknologi ini diharapkan dapat mengurangi kesalahan input data akibat proses manual serta meningkatkan efisiensi waktu dalam perpanjangan SIM (Saptahadi & Setiawati, 2022). Faktor utama yang mempengaruhi keberhasilan OCR dalam mengenali teks pada SIM meliputi, Kualitas Gambar, Resolusi gambar yang rendah atau pencahayaan yang buruk dapat menurunkan akurasi pengenalan