

## BAB II TINJAUAN PUSTAKA

### 2.1 Penelitian Terdahulu

Penelitian terdahulu dijelaskan dengan berbentuk tabel seperti berikut:

Tabel 2. 1 Tabel Penelitian Terdahulu

No	Judul	Masalah	Kelebihan	kekurangan	Solusi
1.	Digitalisasi Arsip Untuk Efisiensi Penyimpanan dan <i>Akseibilitas</i> (Bakhtiar dkk., 2019)	Bagaimana cara mengubah catatan atau rekaman ke dalam bentuk digital agar meningkatkan efisiensi dalam penyimpanan dan memudahkan akses saat dibutuhkan	Menyimpan dokumen untuk keperluan jangka pendek dan panjang, arsip dapat diakses dengan cepat dan mudah saat diperlukan, membatasi akses hanya kepada pihak yang berwenang, menyimpan salinan untuk mengatasi bencana atau kerusakan, serta meningkatkan efisiensi penggunaan peralatan penyimpanan arsip	Sistem yang digunakan untuk mengelola objek-objek tersebut harus mampu melestarikannya dan merekonstruksinya sebagai replika yang identik dengan rekaman aslinya.	Pengelolaan arsip digital dalam organisasi memerlukan kebijakan dan prosedur pengendalian sistematis yang efektif untuk menghindari berbagai masalah dan risiko.
2.	Perancangan Sistem <i>Text Extraction</i> menggunakan <i>library Tesseract</i> OCR untuk mengambil Nomor Induk Kependudukan pada foto Kartu Tanda Penduduk(Sulistiwo & Nisraha Saian, 2019)	Proses pemasukan Nomor Induk Kependudukan (NIK) masih dilakukan secara manual, di mana nasabah harus menyetikkan NIK mereka sendiri. Hal ini berpotensi menyebabkan kesalahan saat menginput NIK, atau mungkin juga disebabkan oleh kurang teliti petugas saat membaca NIK dari Kartu	Sebanyak 81% responden mengungkapkan bahwa aplikasi pemindaian KTP yang menggunakan <i>library Tesseract</i> OCR dapat memudahkan nasabah dalam proses pembukaan rekening secara online	Sebanyak 19% responden menyatakan bahwa aplikasi ini masih memerlukan pengembangan dan perbaikan lebih lanjut	Solusi untuk mengatasi masalah tersebut adalah dengan mengembangkan aplikasi yang menggunakan teknologi <i>Tesseract</i> untuk mengekstrak teks dari gambar.

		Tanda Penduduk (KTP) nasabah			
3.	Efektivitas Penerapan Aplikasi Pelaporan Kedatangan dan Keberangkatan Kapal(SIKAPAL) di Kantor Imigrasi Kelas I TPI Bengkulu(Rosmaya dkk., 2021)	Observasi awal menunjukkan bahwa masih terdapat kendala dalam penerapan aplikasi pelaporan kedatangan dan keberangkatan kapal (SiKapal) di lapangan	Penelitian ini menunjukkan bahwa penerapan Aplikasi SiKapal telah berjalan efektif berdasarkan lima indikator efektivitas.	Kendala yang ditemui dalam penerapannya meliputi belum mendapatkan izin untuk memindai MRZ (Machine Readable Zone), ketidakstabilan kondisi jaringan, dan belum tersedianya aplikasi versi untuk Website dan iOS.	Menyajikan dan merangkum hasil indikator efektivitas dari implementasi aplikasi SiKapal.
1.	Sistem Pengenalan Karakter Dokumen Secara Otomatis Menggunakan Metode <i>Optical Character Recognition</i> (Nurhaliza & ETP, 2022)	Saat ini, proses pemeriksaan berbagai dokumen dilakukan secara manual, yang mengakibatkan sering terjadi kesalahan dalam membaca isi dokumen, mengidentifikasi nama penerima barang, dan memerlukan waktu yang cukup lama.	Hasil pengujian menggunakan 5 citra dokumen menunjukkan tingkat akurasi pengenalan karakter sebesar 98.78% dan waktu proses rata-rata 1.29 detik. Dengan demikian, dapat disimpulkan bahwa metode OCR dapat diterapkan untuk mengenali karakter pada dokumen izin distribusi alat kesehatan.	Diperlukan analisis lebih lanjut untuk mengevaluasi sejauh mana metode OCR dapat mengenali berbagai jenis huruf dengan akurasi yang memadai.	Mengimplementasikan Optical Character Recognition (OCR) untuk mengidentifikasi karakter pada dokumen izin distribusi alat kesehatan.
2.	<i>Citizen Id Card Detection using Image Processing and Optical Character Recognition</i> (Satyawan dkk., 2019).	Input data pelanggan dilakukan secara manual, yang tidak efisien karena memakan waktu yang cukup lama untuk menginput	Percobaan dengan data pelatihan pada <i>tesseract</i> 4.0 menunjukkan bahwa akurasi deteksinya berkisar antara 90 hingga 100% saat menggunakan teknik yang disarankan	Kualitas dokumen yang dimasukkan mempengaruhi performa OCR, sehingga apabila dokumen yang dimasukkan kurang berkualitas	Kami mengajukan suatu pendekatan pemrosesan citra guna mengidentifikasi kartu tanda penduduk, terutama yang berlaku untuk KTP

data satu per satu	akan dapat mempengaruhi performa dari OCR	Indonesia. Dengan menguji berbagai model pengenalan yang kami terapkan menggunakan kerangka kerja <i>Tesseract</i> , kami menyusun sebuah antarmuka situs web yang mampu mendeteksi dan mengenali KTP, yang akan menjadi berguna dalam hasil pemindaian.
--------------------	---	--

Berdasarkan tabel penelitian di atas ditemukan research gap:

1. Perlu memanfaatkan OCR (*Optical Character Recognition*) pada penelitian yang akan dilakukan, berdasarkan penelitian dari (Satyawan dkk., 2019), data pelatihan pada tesseract 4.0 menunjukkan bahwa akurasi deteksinya berkisar antara 90 hingga 100%. Serta melihat hasil dari penelitian (Nurhaliza & ETP, 2022) bahwa uji coba 5 citra dokumen menunjukkan tingkat akurasi pengenalan karakter sebesar 98.78% dan waktu proses rata-rata 1.29 detik. Dengan itu, penggunaan OCR menjadi solusi terbaik dalam pendeteksian karakter yang ada pada data paspor.
2. Perlu proses deteksi menggunakan MRZ pada pemecahan masalah yang akan dilakukan, karena menurut saran dari (Rosmaya dkk., 2021) yang telah menerapkan sistem informasi dalam mengelola data kedatangan dan

keberangkatan kapal asing, deteksi MRZ sangat dibutuhkan untuk mempermudah agen perusahaan dalam mengelola data kapal asing untuk mendapatkan data paspor lebih mudah, serta penggunaan media *mobile* platform untuk implementasi aplikasi.

## **2.2 Teori Terkait**

### **2.2.1. Digitalisasi**

Digitalisasi adalah proses mengubah materi analog konvensional, seperti buku, peta, dan dokumen kertas lainnya, menjadi bentuk digital dan elektronik. Menurut Business Dictionary, digitalisasi merujuk pada konversi informasi analog dalam berbagai bentuk (seperti teks, foto, suara, dll.) menjadi format digital menggunakan perangkat elektronik yang sesuai seperti pemindai atau *chip* komputer khusus. Hal ini memungkinkan informasi tersebut untuk diolah, disimpan, dan ditransmisikan melalui rangkaian, perangkat, dan jaringan digital (Bakhtiar dkk., 2019). Menurut Kamus Besar Bahasa Indonesia (KBBI), digitalisasi adalah proses menerapkan atau menggunakan sistem digital. Secara singkat Digitalisasi adalah proses mengubah entitas fisik yang berbasis analog menjadi bentuk virtual dan digital.

### **2.2.2. OCR**

OCR (Optical Character Recognition) adalah teknologi di mana mesin dapat secara otomatis mengidentifikasi karakter melalui koneksi optik. Teknologi OCR sering digunakan untuk mengkonversi buku dan surat ke dalam format berkas lunak, seperti PDF dan lainnya. Selain itu, OCR juga dapat digunakan untuk

mengotomatisasi pencatatan, mempublikasikan teks di situs web, serta dalam bidang penelitian pengenalan pola, kecerdasan buatan, dan *computer vision*. Selain itu, teknologi OCR memiliki kemampuan untuk melakukan pencarian kata atau frasa, melakukan penyuntingan teks, dan mengimplementasikan teknik seperti mesin penerjemah, teks-ke-suara, dan penambangan teks(Sulistiyo & Nugraha Saian, 2019).

### **2.2.3. MRZ**

MRZ(*Machine Readable Zone*) adalah zona pada dokumen perjalanan, termasuk paspor, yang dirancang untuk dapat dibaca oleh mesin secara otomatis sesuai standar yang ditentukan seperti ICAO(*International Civil Aviation Organization*). Selain ICAO, terdapat beberapa standar untuk dokumen yang dapat dibaca mesin, dan di setiap standar tersebut, terdapat variasi MRZ yang berbeda. Setiap variasi memiliki ciri-ciri yang tetap, termasuk karakteristik geometris seperti rasio aspek zona. Beberapa negara telah mengadopsi standar MRZ khusus mereka sendiri, yang sejalan dengan karakteristik geometris MRZ tersebut(Savelyev dkk., 2022). ICAO(*International Civil Aviation Organization*) merupakan standar global pembuatan sebuah paspor yang telah banyak digunakan banyak negara saat ini untuk memastikan konsistensi, keamanan, dan kemudahan identifikasi di seluruh dunia. Indonesia bergabung dengan resmi menjadi anggota ICAO pada 1 Mei 2004(Candra Susanto & Keke, 2019) maka dasar proses parsing MRZ mengikuti standar ICAO.



10. 1 : Jumlah digit tanggal lahir

11. F : Jenis kelamin

12. 940623 : Tanggal Kadaluwarsa format YYMMDD

13. ZE184226 : Data opsional negara penerbit

Penggunaan dan tata cara penulisan baris MRZ di atas mengikuti standar ICAO sesuai pada dokumen mesin yang dapat dibaca(International Civil Aiation Organization, 2021).