

Implementasi Pendataan Warga RW 007 Penggilingan Jakarta Timur Menggunakan Metode *OCR Tesseract*

Wahyu Saputro¹, Seli Amelia¹

¹ Program Studi Teknik Informatika, STIKOM CKI
Jl. Raden Inten II No.8, RT.5/RW.14, Duren Sawit, Kec. Duren Sawit, Kota Jakarta Timur, DKI Jakarta
Email: wahyudahsyat@gmail.com

Naskah Masuk: 29 Juli 2022; Diterima: 09 Agustus 2022; Terbit: 17 Maret 2023

ABSTRAK

Abstrak – Pada wilayah Penggilingan RW 007 Jakarta Timur di zaman modern seperti sekarang ini sudah banyak penemuan dari para ilmuwan, warga, dan mahasiswa untuk memudahkan kegiatan warga. Penemuan ini bertujuan agar warga RW 007 dapat dengan mudah menjalankan kegiatan disetiap harinya. Namun tidak banyak penemuan yang bisa mempermudah pekerjaan manusia, salah satunya dalam bidang pendataan. Data yang telah diolah secara konvensional terkadang menemui kendala akibat kesalahan dari manusia. Banyak peluang untuk memecahkan yang ditimbulkan oleh kesalahan manusia (*Human Error*) tersebut. Kemungkinan terjadinya kesalahan berasal dari beberapa situasi. Pendataan warga wilayah khususnya di RW 007 masih belum menggunakan sistem yang terbaru. Tidak hanya itu, ketika terjadi bencana alam seperti banjir, gempa, dan tanah longsor, berkas-berkas yang masih berupa konvensional kemungkinan besar akan rusak dan membutuhkan waktu untuk diperbaiki. Dari penjelasan diatas, penulis mengusulkan sebuah sistem pendataan penduduk yang berbasis teknologi. Pembuatan aplikasi pendataan penduduk berbasis website ini menggunakan fitur *Optical Character Recognition* untuk menunjang pembuatan aplikasi pendataan, yaitu suatu jenis *scan* gambar dua dimensi dengan fungsionalitas utama yaitu dapat dengan mudah dibaca.

Kata Kunci: *Optical Character Recognition, TesseractJs, Pendataan*

ABSTRACT

Abstract- – *In the Milling area of RW 007, East Jakarta. In modern times like now, there have been many discoveries from scientists, residents and students to facilitate residents' activities. This discovery is so that the residents of RW 007 can easily carry out their daily activities. but there are not many discoveries that can facilitate human work, one of which is in the field of data collection. Data that has been processed conventionally sometimes encounters obstacles as a result of human error. There are many opportunities to solve these problems caused by human error (Human Error). The possibility of errors occurring stems from several situations. Data collection for residents of the area, especially in RW 007, is still using a system that is not yet up-to-date. not only that when natural disasters such as floods, earthquakes, landslides occur. Files that are still conventional are likely to be corrupted and take time to repair. From the explanation above, the author proposes a technology-based population data collection system. In making a website-based population data collection application. It uses the Optical Character Recognition feature to support the creation of data collection applications, which is a type of two-dimensional image scan with the main functionality being easy to read.*

Keywords: *Optical Character Recognition, TesseractJs, Citizen data collection*

Copyright © 2023 Jurnal Teknik Elektro dan Komputasi (ELKOM)

1. PENDAHULUAN

Perkembangan teknologi pada era moden sudah banyak penemuan berasal dari kalangan ilmuwan, warga dan pelajar untuk menginovasikan sebuah temuan yang bertujuan mempermudah aktifitas masyarakat. yang akan terjadi diharapkan dapat membantu aktivitas masyarakat sehari-hari. Namun, tidak banyak penemuan yang dapat memudahkan pekerjaan manusia. Rangkaian fakta akurat diharapkan dapat memberikan hal yang luar biasa bagi warga jika diolah dengan baik. Data yang selama ini diolah secara konvensional, secara teori, kini sering menemui kendala karena *human error*. RW 007 masih menggunakan sistem manual. Tak hanya itu, ketika data harus dimasukkan kedalam komputer, membutuhkan tambahan waktu dan ketika terjadinya bencana alam seperti banjir, gempa, tanah longsor. Berkas yang masih berupa konvensional

kemungkinan besar akan rusak dan membutuhkan waktu untuk diperbaiki. Dari pemikiran di atas, penulis mengusulkan sistem kependudukan berbasis teknologi. Perangkat lunak berbasis web sepenuhnya digunakan untuk pengambilan foto dan komunitas internet untuk menambahkan informasi ke server.

Pada pembuatan aplikasi pendataan penduduk berbasis web ini memakai fitur *Optical Character Recognition Tesseract* untuk pembuatan aplikasi pendataan, yaitu suatu jenis *scan* gambar dua dimensi dengan fungsionalitas utama yaitu dapat dengan mudah dibaca oleh pemindai. Nomor KTP digunakan sebagai identitas warga, dari proses pemindaian nomor KTP akan didapat nomor KTP.

Secara umum terdapat dua hal utama yang mempengaruhi proses OCR yaitu: mekanisme ekstraksi ciri dan mekanisme pengenalan. Mekanisme ekstraksi ciri dilakukan untuk mendapatkan ciri atau identitas dari suatu karakter atau huruf. Proses pengenalan bertujuan untuk mencocokkan pola huruf yang berasal dari inputan dengan pola yang ada dalam basis pengetahuan. Dari latar belakang yang peneliti utarakan di atas maka peneliti berkeinginan untuk mengimplementasikan *Optical Character Recognition tesseract* melalui penelitian dengan judul “Implementasi Pendataan Warga RW 007 Penggilingan Jakarta Timur Menggunakan Metode OCR Tesseract”.

2. KAJIAN PUSTAKA

2.1 Penelitian Terkait

Penelitian yang terkait dengan mengenai implementasi pendataan warga RW 007 Penggilingan Jakarta Timur menggunakan OCR *Tesseract*. Pada penelitian yang berjudul “Analisis Sistem *Optical Character Recognition (OCR)* Pada Dokumen Digital Menggunakan Metode *Tesseract*” menjelaskan bahwa suatu alat yang dapat digunakan untuk mengkonversi cetakan karakter menjadi teks digital, tanpa harus menyetik ulang. Dengan menggunakan OCR ini, gambar yang bertulis tangan, mesin ketik, dapat di cari per kata atau kalimat yang dapat diganti atau dimanipulasi dan diberikan barcode. Dalam mengubah gambar menjadi text, langkah-langkah yang dilakukan oleh algoritma OCR, yaitu: akuisisi gambar, pra-pemrosesan, segmentasi, ekstraksi fitur, klasifikasi, dan pasca-pemrosesan. Sistem OCR dapat digunakan dalam berbagai aplikasi praktis seperti pengenalan plat nomor, perpustakaan pintar, pengenalan karakter berbagai macam bahasa. OCR merupakan sub bidang dari *Pattern Recognition (PR)* yang berkaitan dengan pengenalan karakter [1].

Pada penelitian lain yang berjudul “Ekstraksi Karakter Citra Menggunakan *Optical Character Recognition* Untuk Pencetakan Nomor Kendaraan Pada Struk Parkir” diperoleh hasil bahwa karakter hasil ekstraksi tersebut dapat ikut dicetak pada struk parkir sebagai salah satu faktor validasi kepemilikan kendaraan yang diharapkan dapat meningkatkan keamanan dan kenyamanan pengguna sarana parkir yang menitipkan kendaraannya. Salah satu bidang ilmu dalam pengolahan citra digital adalah pengenalan pola (*pattern recognition*). Pengenalan otomatis, deskripsi, dan pengelompokan objek atau data ke dalam jumlah kelas bermakna disebut sebagai pengenalan pola. Pengenalan pola dalam pengolahan citra digital umum digunakan untuk mendeteksi bentuk objek dan mencocokkan pola dari objek tersebut terhadap objek lain. Salah satu cabang dari bidang ilmu pengenalan pola dalam pengolahan citra digital adalah *Optical Character Recognition (OCR)* [2].

Menurut Penelitian yang dilakukan oleh Efendi, dkk yang berjudul “Penerapan Metode Ekstraksi Image Ke Text Dengan *Optical Character Recognition* Untuk Otomatis Data Kependudukan”, pada saat ini pemanfaatan e-KTP masih belum dilaksanakan secara optimal, baik di pemerintahan maupun di instansi atau perusahaan swasta. Akibatnya, jika dibutuhkan maka e-KTP tersebut harus di fotocopy terlebih dahulu kemudian data e-KTP tersebut di input kembali ke database organisasi yang memerlukan baik pemerintah maupun swasta. Hal ini memerlukan kerja dua kali dan penyimpanan fisik e-KTP memiliki kelemahan dalam asumsi yang sulit dipenuhi, yaitu independensi feature area dan penyimpanan yang tidak baik dapat merusak fotocopy e-KTP. Fotocopy e-KTP juga bisa rusak karena lembab dan pencariannya juga agak sulit. Dengan bantuan mesin scanner maka pengambilan e-KTP dapat dilakukan dan disimpan pada komputer. Pengambilan data e-KTP agar dapat diinput dalam database diperlukan *Tesseract Optical Character Recognition* dan bantuan *image magic* mampu mengekstrak citra ataupun *image* dan keluarannya berupa text yang langsung dimasukkan kolom yang sesuai dengan fungsi data tersebut dengan sistem berbasis website sehingga mempermudah layanan administrasi [3].

Penelitian lainnya yaitu penelitian yang berjudul “Aplikasi Kalkulator Tangan Sederhana Menggunakan *Optical Character Recognition (OCR)*. Penelitian ini bertujuan untuk membuat aplikasi kalkulator tulisan tangan sederhana menggunakan *Optical Character Recognition (OCR)* dengan menambahkan metode segmentasi proyeksi histogram. Dengan adanya penelitian ini, diharapkan dapat menerapkan metode *Image Processing* dan *Optical Character Recognition* untuk menerjemahkan dan menghitung citra atau gambar tulisan tangan yang mengandung operasi bilangan matematika dengan akurat [4].

Menurut Penelitian yang berjudul “Navigasi Robot Beroda Berdasarkan Pengenalan Teks Untuk Melakukan Pergerakan Menggunakan Metode *Optical Character Recognition*”, sistem navigasi menjadi bagian yang sangat penting pada robot beroda agar mampu bergerak secara mandiri. Selain itu, navigasi juga dapat diartikan sebagai suatu kemampuan untuk memandu pergerakan dari satu posisi ke posisi lainnya melalui penentuan posisi dan arah gerakannya. Selain itu juga harus mampu memahami, menganalisis dan mengetahui posisi serta situasi dirinya dalam lingkungan. Berdasarkan permasalahan tersebut maka dilakukan penelitian mengenai sistem navigasi robot beroda untuk melakukan pergerakan berdasarkan pengenalan teks menggunakan metode *Optical Character Recognition (OCR)* [5].

2.2 Implementasi

Implementasi adalah pelaksanaan dari sebuah rencana yang disusun secara matang dan terperinci, implementasi biasanya dilakukan setelah perencanaan dianggap sudah selesai. Jadi dapat disimpulkan implementasi merupakan suatu kegiatan yang dilakukan untuk proses menguji data dan menerapkan suatu sistem yang sudah dirancang dengan konsep yang sudah ditentukan [6].

2.3 Pendataan

Pendataan merupakan proses, cara, perbuatan mendata. Arti lainnya dari pendataan artinya pengumpulan data. Pendataan berasal berasal kata dasar data. Pendataan artinya sebuah homonim karena artinya adalah mempunyai ejaan dan pelafalan yang sama tetapi maknanya tidak sama. Pendataan mempunyai arti pada kelas nomina atau istilah benda menjadi akibatnya pendataan mampu menyatakan nama dari seorang, daerah, atau seluruh benda dan segala yang dibendakan.

2.4 *Optical Character Recognition*

Optical Character Recognition artinya aplikasi computer yang di gunakan untuk mengidentifikasi citra huruf maupun angka agar di konversi dalam bentuk file [7].

2.5 *Tesseract*

Tesseract-OCR adalah *engine open source optical character recognition* yang dikembangkan oleh Ray Smith. *Tesseract-OCR* dipilih untuk digunakan sebagai *engine optical character recognition* pada kali ini karena akurasi dan kecepatan scanning pada gambar yang baik. Aplikasi ini akan dikembangkan menggunakan *Tesseract-OCR* [8].

2.6 Visual Studio Code

Visual Studio Code adalah editor source code yang dikembangkan Microsoft buat Windows, Linux serta MacOS. Ini termasuk dukungan buat debugging, GIT Control yang disematkan, penyorotan sintaks, penyelesaian kodecerdas, cuplikan, dan kode refactoring.

2.7 *JavaScript*

JavaScript adalah bahasa scripting untuk membuat konten dinamis di halaman atau aplikasi web, game, serta membangun web server. agar bisa membuat elemen dinamis untuk interaksi user, seperti menu drop-down atau grafik animasi. Sejarah *JavaScript* dimulai pada tahun 1995, diciptakan oleh *Brendan Eich* di *Netscape Communications*. Awalnya, *JavaScript* hanya digunakan untuk keperluan internal dengan web browser perusahaan tersebut, *Netscape Navigator*. Namanya dulunya adalah *LiveScript*, lalu diubah menjadi *JavaScript* agar bisa menjadi “teman” bagi bahasa pemrograman Java milik mitra mereka, *Sun Microsystems*.

2.8 *Express JS*

Dengan *Express JS*, programmer tidak perlu lagi membuat aplikasi web menggunakan module http bawaan *Node.JS*. Framework dengan serangkaian fitur menarik, seperti routing, rendering view, dan didukung oleh middleware ini bisa menghemat waktu dalam mengembangkan aplikasi *Node.js*.

2.9 *Nodejs*

Nodejs adalah sebuah runtime environment untuk *JavaScript* yang bersifat *open-source* dan *cross-platform*. Dengan *Node.js*, kamu dapat menjalankan kode *JavaScript* di mana saja, tidak hanya terbatas pada lingkungan browser.

2.10 Storage

Storage ialah penyimpanan, kawasan penyimpanan, serta media yang digunakan buat menyimpan data yang diolah oleh komputer. Fungsi storage adalah menyimpan data atau file dalam jangka panjang dan pendek. Jadi storage adalah tempat menyimpan instruksi acara serta data buat keperluan penggunaan menggunakan komputer.

2.11 Website

Website merupakan gugusan page web yang saling terhubung dan seluruh file saling terkait. Web terdiri dari laman atau halaman serta kumpulan page yang dinamakan homepage. Homepage berada pada posisi teratas dengan page-laman terkait berada di bawahnya. biasanya, setiap laman pada bawah homepage (*child page*) berisi hyperlink ke laman lain dalam web.

3. METODE PENELITIAN

Dalam penelitian ini yang akan di gunakan berupa KTP RW 007 Penggilingan Jakarta Timur yang mau di teliti sebanyak 15 KTP dan setiap KTP memiliki 19 kategori diantaranya sebagai berikut:

K1 = NIK

K2 = Nama

K3 = Tempat Lahir

K4 = Tanggal Lahir

K5 = Jenis Kelamin

K6 = Golongan Darah

K7 = Alamat

K8 = Kelurahan / Desa

K9 = Kecamatan

K10 = Kabupaten

K11 = Provinsi

K12 = RT

K13 = RW

K14 = Agama

K15 = Status Perkawinan

K16 = Pekerjaan

K17 = Kewarganegaraan

K18 = Berlaku Hingga

K19 = Permanen

3.1.1 Sumber Data

Sumber data sangat krusial pada penelitian dan lebih lanjut digunakan peneliti buat memperoleh data-data penelitian sehingga meminimalkan saat serta porto. asal data dalam penelitian ini menggunakan asal data primer serta sumber data sekunder.

3.1.2 Dokumentasi

Metode dokumentasi digunakan untuk mengumpulkan data warga RW 007 Penggilingan Jakarta Timur.

3.1.3 Observasi

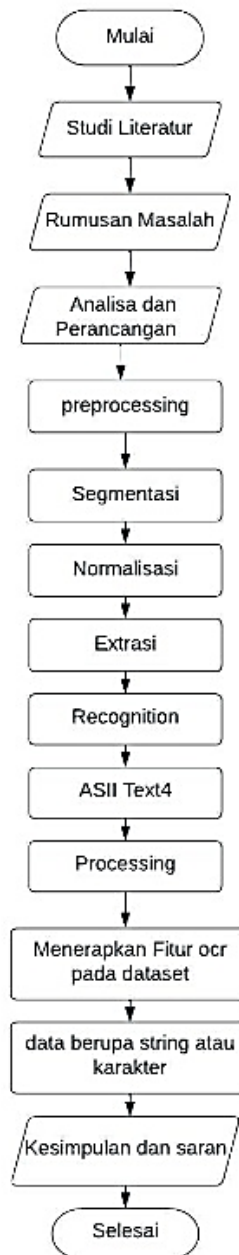
Observasi adalah teknik pengumpulan data untuk mengamati Cara Kerja PendataanWarga di RW 007 Penggilingan Jakarta Timur.

3.1.4 Wawancara

Dalam kegiatan wawancara ini, pewawancara sebagai pihak yang mengajukan pertanyaan dan narasumber menjadi pihak yang memberikan jawaban. Dalam penelitian ini subjek wawancara adalah RW 007 dan Juru Pemantau Jentik atau jumentik.

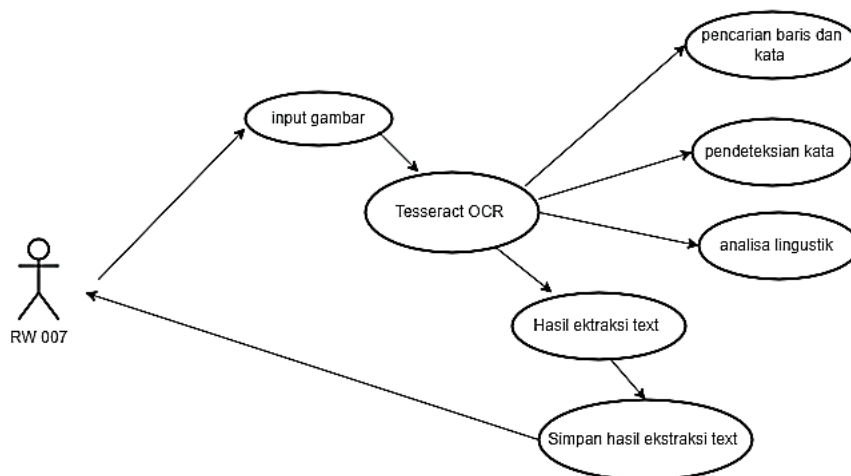
3.2 Penerapan Metodologi

Pada penelitian ini, untuk mencapai tujuan dari suatu penelitian perlu adanya suatu proses atau prosedur sistematis yang digunakan, terdapat beberapa tahapan penelitian yang di jabarkan pada Gambar 1 berikut:



Gambar 1. Tahapan penerapan metodologi

3.3 Rancangan Pengujian
3.3.1 Desain Use Case

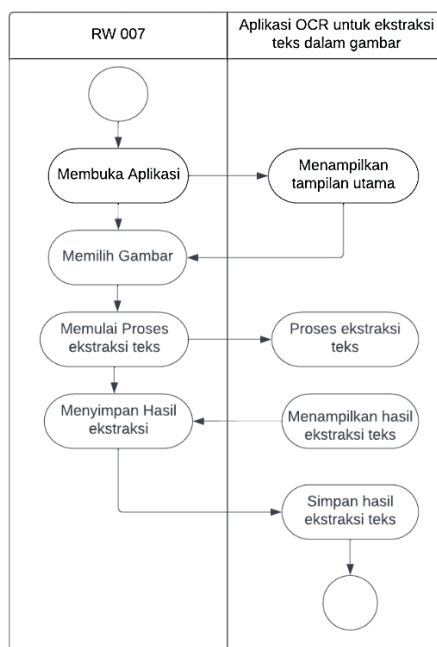


Gambar 2. Desain use case diagram

Gambar 2 diatas adalah *desain use case diagram* aplikasi yang diajukan. Dapat dijelaskan secara singkat sebagai berikut: input gambar ke dalam aplikasi, setelah itu Tesseract-OCR akan mengambil teks yang ada dalam sebuah gambar. dan tulisan dapat disimpan.

3.3.2 Desain Activity Diagram

Activity Diagram berguna untuk memberikan visualisasi alur tindakan dalam sistem, percabangan yang mungkin dapat terjadi, dan bagaimana alur sistem dari awal mula hingga akhir.



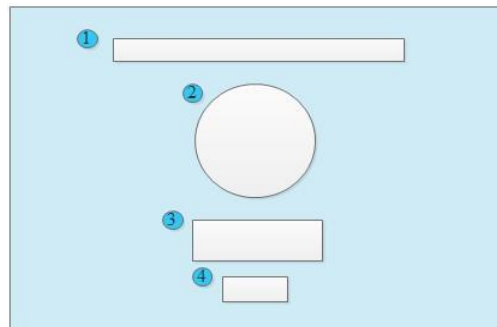
Gambar 3. Activity diagram ekstraksi teks

3.3.3 Desain Rancangan Aplikasi

Rancangan untuk tampilan halaman utama aplikasi dapat dilihat pada Gambar 4.

3.3.4 Evaluasi Prototype

Tahapan pengevaluasian ini dilakukan buat memperjelas spesifikasi kebutuhan pengguna terhadap software.



Gambar 4. Tampilan *prototype*

4. HASIL DAN PEMBAHASAN

4.1 Alat Penelitian

Dalam membuat perancangan sistem, mulai dari analisa, spesifikasi kebutuhan, perancangan, hingga pengujian dan implementasi, dibutuhkan pemilihan software dan hardware yang sesuai (disarankan) dengan spesifikasi yang mumpuni untuk menunjang proses bisnis, seperti berikut:

a. *Software* (perangkat lunak)

Perangkat lunak berisi data yang diprogram atau disimpan dengan fungsi-fungsi tertentu.

Tabel 1. *Software*

No	Kategori Software	Spesifikasi Software
1	Sistem Operasi	Windows 10 Home Single Language
2	Browser	Google Chrome
3	Backend	Express
4	Template Engine	Ejs
5	Programming Language	Javascript
6	Script Editor	Visual Studio Code

b. *Hardware* (Perangkat Keras)

Hardware merupakan komponen yang ada pada komputer yang mana bagian fisiknya dapat terlihat secara kasat mata atau dapat dirasakan secara langsung.

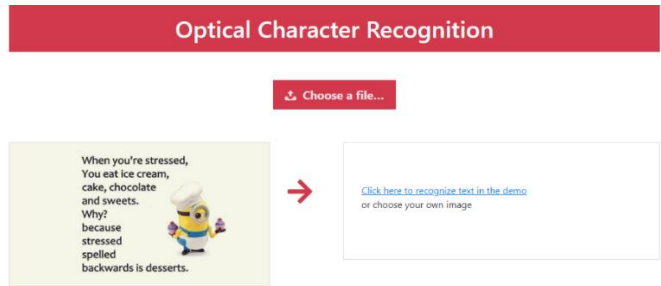
Tabel 2. *Hardware*

No	Kategori Hardware	Spesifikasi Hardware
1	Laptop	ASUS VIVOBOOK MAX
2	Processor	Intel(R) Pentium(R) Silver N5000 CPU @ 1.10GHz 1.10 GHz
3	Motherboard	HP 2192
4	RAM	4GB DDR3-SDRAM
5	Hard disk	Western digital 320 GB
6	Connector	USB 3.0

4.2 Pengujian

4.2.1 Tampilan halaman awal aplikasi

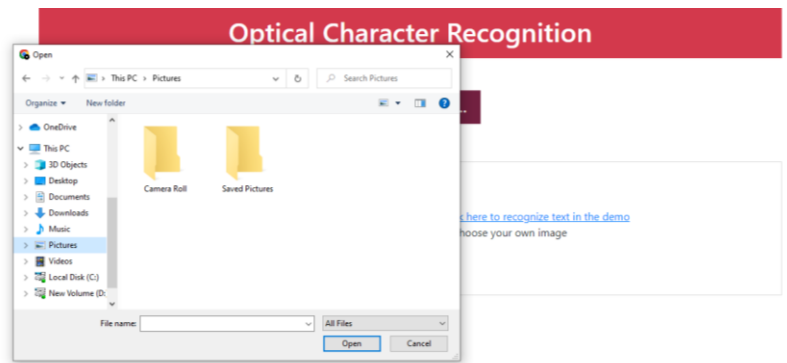
Tampilan halaman ini adalah tampilan dimana user atau pengguna akan memilih file atau gambar mana yang akan di ekstrak. Tampilan halaman pilih file.



Gambar 5. Tampilan halaman awal aplikasi

4.2.2 Tampilan halaman untuk memilih file atau gambar

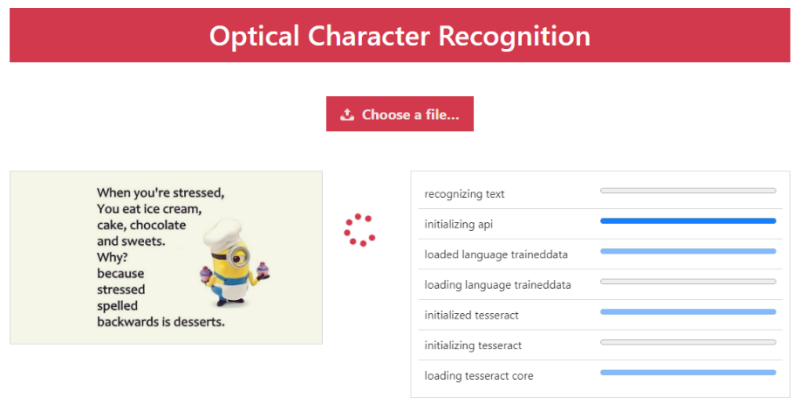
Tampilan halaman ini adalah tampilan dimana user atau pengguna akan memilih file atau gambar mana yang akan di ekstraksi. Tampilan halaman pilih file.



Gambar 6. Tampilan halaman untuk memilih file atau gambar

4.3.3 Tampilan halaman progress

Tampilan halaman ini adalah proses dimana ketika sistem membaca file yang akan di ekstraksi. Tampilan halaman ini dapat dilihat pada gambar 7.



Gambar 7. Tampilan halaman progres

4.3.4 Tampilan hasil akhir

Tampilan halaman ini adalah proses ketika sistem telah selesai meng ekstraksi sebuah gambar menjadi teks atau karakter. Tampilan halaman ini dapat dilihat pada Gambar 8.

Optical Character Recognition



Gambar 8. Tampilan hasil akhir

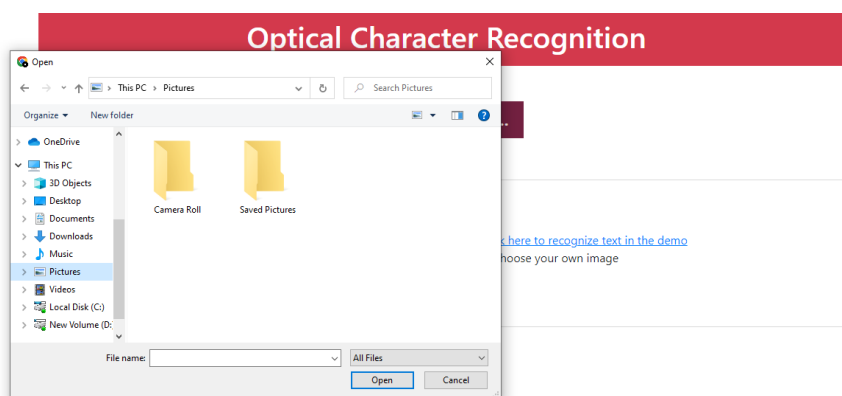
4.3 Hasil Akhir Pengujian

Pada hasil akhir pengujian ini dilakukan pengujian apakah sistem berhasil melakukan proses sosialisasi objek asal gambar ktp serta mengenali kalimat yang ada di KTP tersebut. Gambar yang akan dikenali dapat dilihat di Gambar 9.



Gambar 9. Foto KTP RW 007

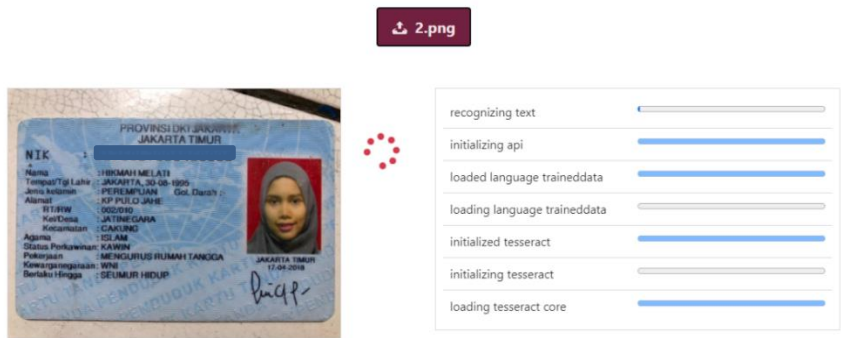
Langkah pertama yang harus dilakukan adalah mengakses halaman web melalui localhost atau komputer, lalu akan ditampilkan halaman utama sistem seperti yang tertera pada gambar foto KTP RW 007. Kemudian pilih tombol choose image buat merogoh gambar yang ada di direktori penyimpanan PC (Personal Computer) kemudian pilih gambar yang akan di proses. Berikut artinya tampilan halaman waktu ingin menentukan gambar yang akan di proses. Setelah memilih gambar, sistem akan langsung memproses gambar tersebut.



Gambar 10. Halaman untuk memilih file

Setelah memilih gambar, sistem akan langsung memproses gambar tersebut.

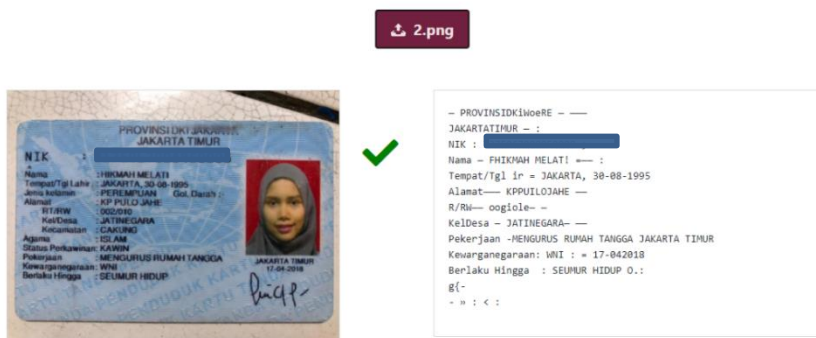
Optical Character Recognition



Gambar 11. Halaman progres

Setelah gambar di proses maka tahap akhir adalah menampilkan teks atau kalimat dari gambar tersebut dan hasilnya dapat dilihat pada Gambar 2 berikut:

Optical Character Recognition



Gambar 12. Tahap akhir

Gambar 12 merupakan hasil dari outputnya yaitu string / karakter dari optical character recognition tesseract.

4.3.1 Tingkat Akurasi

Tabel 3. Tingkat akurasi

KTP	Pre processing	Hasil OCR	Klasifikasi
	PROVINSI DKI JAKARTA JAKARTA TIMUR NIK: 337506xxxxxxxxxx Nama: HIKMAH MELATI Tempat/Tgl Lahir = JAKARTA, 30-08-1995 Jenis Kelamin: PEREMPUAN Alamat: KP PULO JAHE RT/RW: 002/010 Kel/Desa: JATINEGARA Kecamatan: CAKUNG Agama: ISLAM Status Perkawinan: KAWIN Pekerjaan: MENGURUS RUMAH TANGGA Kewarganegaraan: WNI Berlaku Hingga: SEUMUR HIDUP	— PROVINSIDKIhOeRE — JAKARTA Timur — : NIK : ██████████ Nama — FHIKMAH MELATI — : Tempat/Tgl Ir = JAKARTA, 30-08-1995 Alamat— KPPUILOJAHE — R/RW— oogiole— KelDesa — JATINEGARA— Pekerjaan -MENGURUS RUMAH TANGGA JAKARTA TIMUR Kewarganegaraan: WNI : = 17-042018 Berlaku Hingga : SEUMUR HIDUP 0.: g(- - : < :	Terdeteksi



PROVINSI DKI JAKARTA
JAKARTA TIMUR
NIK: 317506xxxxxxxxxx
Nama: NURKAMILAH
Tempat/Tgl Lahir = JAKARTA, 12-10-1986
Jenis Kelamin: PEREMPUAN
Alamat: KP.RAWA BADUNG
RT/RW: 011/007
Kel/Desa: PULO GEBANG
Kecamatan: CAKUNG
Agama: ISLAM
Status Perkawinan: KAWIN
Pekerjaan: MENGURUS RUMAH TANGGA
Kewarganegaraan: WNI
Berlaku Hingga: SEUMUR HIDUP

Kewarganegaraan: WNI : =
17-042018
Berlaku Hingga : SEUMUR
HIDUP O.:

g{-
- » : < :
PROVINSI DKi JAKARTA
JAKARTA TIMUR
NIK : 317506xxxxxxxxxx
Nama *NURKAMILAH
empat/Tgl Lahir : JAKARTA,
12-10-1986
Jenis kelamin :PEREMPUAN
Gol. Darah :-
Alamat : KP RAWA
BADUNG
RT/RW : 011/007
Kel/Desa 1 JATINEGARA
Kecamatan :CAKUNG i
Agama (ISLAM :
Status Perkawinan: KAWIN

Terdeteksi



PROVINSI DKI JAKARTA
JAKARTA TIMUR
NIK: 317506xxxxxxxxxx
Nama: ALFIA WIDYASARI
Tempat/Tgl Lahir = JAKARTA, 21-02-2000
Jenis Kelamin: PEREMPUAN
Alamat: KP.PISANGAN II
RT/RW: 006/004
Kel/Desa: PENGGILINAN
Kecamatan: CAKUNG
Agama: ISLAM
Status Perkawinan: BELUM KAWIN
Pekerjaan: PELAJAR/MAHASISWA
Kewarganegaraan: WNI
Berlaku Hingga: SEUMUR HIDUP

\ ae IS =. -
PROVINSIDKI JAKARTA
JAKARTA TIMUR
NIK 317506xxxxxxxxxx
Nama ALFIA WIDYASARI
amy AKARTA, 21-02-2000
Je "EREMPUAN Gol. Darah
Alama KP. PISANGAN ii i .
RT/RW 006/004 i '
Kel/Desa PENGGILINAN
9
Kecamatan CAKUNG
Agama ISLAM
Status Perkawinan: BELUM
KAWIN
Pekerjaan
PELAJAR/MAHASISWA
JAKARTA TIMUR
Kewarganegaraan: WNI 31-
08-2018
Berlaku Hingga : SEUMUR
HIDUP 1

Terdeteksi



PROVINSI DKI JAKARTA
JAKARTA TIMUR
NIK: 317206xxxxxxxxxx
Nama: RIFQAH LUTHFIYYAH
Tempat/Tgl Lahir = JAKARTA, 16-10-2000
Jenis Kelamin: PEREMPUAN
Alamat: KANNDANG BESAR
RT/RW: 012/004
Kel/Desa : UJUNG MENTENG
Kecamatan: KELAPA GADING
Agama: ISLAM

PROVINSI DKI JAKARTA
JAKARTA TIMUR
: 31750b5L10000013 oe
SRIEQAH LUTHFFYY AH
FAERIE
Izhir * JAKARTA 16-10-
2000 ps oy
5 PEREMPUAN — Gol
Darah — A A
KANDANG BESAR / f a
1 012/004 h F
> SUJUNG MENTENG EB
“CAKUNG a







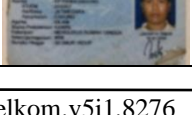
Terdeteksi








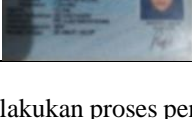
	Status Perkawinan: BELUM KAWIN Pekerjaan: PELAJAR/MAHASIWA Kewarganegaraan: WNI Berlaku Hingga: SEUMUR HIDUP	ISLAM RR inan: BELUM KAWIN Bare = : PELAJAR/MAHASISWA JAKARTA The ps : WNi 804 32: SEUMUR HIDUP > : ~~~ PROVINSIDKIJAKARTA = JARRE ——— NIK ~~~ : 317506b502820014 ——— ee = = — — _ ——— { Kecamatan CAKHUNG— \$A So Perkawinan: KAWIN — Pekerjaan : MENGURUS RUMAH TANGGA JAKARTA TIMUR Kewarganegaraan: WNI 29- 96-2018 Berlaky Hingga : SEUMUR HIDUP Ff	Terdeteksi
	PROVINSI DKI JAKARTA JAKARTA TIMUR NIK: 31750xxxxxxxxxx Nama: ALFIYAH Tempat/Tgl Lahir = JAKARTA, 08-01-1990 Jenis Kelamin: PEREMPUAN Alamat: UJUNG KARAWANG RT/RW: 012/005 Kel/Desa: PULO GEBANG Kecamatan: CAKUNG Agama: ISLAM Status Perkawinan: KAWIN Pekerjaan: KARYAWAN SWASTA Kewarganegaraan: WNI Berlaku Hingga: SEUMUR HIDUP		

Tujuan pengujian ini adalah untuk melihat seberapa besar kemampuan sistem buat mengenali objek dari sebuah gambar serta buat melihat apakah penggunaan metode ocr ini berpengaruh terhadap kinerja Tesseract Engine. Proses pengujian taraf akurasi sistem dapat dicermati di tabel 4 tingkat akurasi berikut.

4.3.2 Hasil Data Uji pada KTP

Tabel 4. Hasil data uji pada KTP

No.	KTP	Klasifikasi	Keterangan
1		Terdeteksi	Sesuai tapi belum sempurna
2		Terdeteksi	Sesuai
3		Terdeteksi	Sesuai
4		Terdeteksi	Sesuai tapi belum sempurna
5		Terdeteksi	Sesuai tapi belum sempurna
6		Terdeteksi	Sesuai tapi belum sempurna
7		Terdeteksi	Sesuai

No.	KTP	Klasifikasi	Keterangan
8		Terdeteksi	Sesuai
9		Terdeteksi	sesuai
10		Terdeteksi	Sesuai tapi belum sempurna
11		Terdeteksi	Sesuai
12		Terdeteksi	Sesuai tapi belum sempurna
13		Terdeteksi	Sesuai tapi belum sempurna
14		Terdeteksi	Sesuai
15		Terdeteksi	Sesuai tapi belum sempurna

Setelah dilakukan proses pengujian sistem menggunakan gambar inputan, OCR Tesseract Engine sudah bisa berjalan baik tanpa adanya tambahan pada proses sebelum sosialisasi. taraf akurasi sosialisasi OCR juga relatif baik. Perhitungan taraf akurasi sistem menggunakan jumlah data uji sebesar 15 data uji, maka tingkat akurasi pada sistem dapat dilihat menjadi berikut.

$$\text{Akurasi} = \frac{\text{Jumlah data yang benar}}{\text{banyak jumlah data}} \times 100 \% \quad (1)$$

$$\text{Akurasi} = \frac{7}{15} \times 100\%$$

$$= 46,7$$

5. KESIMPULAN

Berdasarkan penelitian yang peneliti lakukan, dengan adanya sistem informasi pendataan warga di RW007 Maka dapat disimpulkan beberapa hal sebagai berikut:

- Sulitnya mendapatkan data valid karena penduduk yang masih nomad (berpindah-pindah tempat tinggal karena persoalan kerjaan).
- Dengan adanya aplikasi Optical Character Recognition untuk ekstraksi teks pada gambar memudahkan para warga RW 007 Penggilingan Jakarta timur untuk mengambil tulisan yang ada dalam sebuah gambar.
- Tidak dapat mengenali karakter dengan sempurna jika memiliki background atau latar belakang yang gambar atau beberapa warna.
- Aplikasi yang dibangun masih memiliki beberapa kelemahan, buat itu masih dibutuhkan adanya perbaikan maupun pengembangan buat penelitian selanjutnya.

- e. Saran yang dibutuhkan buat pengembangan perangkat lunak ini lebih lanjut adalah:
 - i. Meningkatkan tingkat akurasi dan pendeteksian teks dalam aplikasi sehingga tulisan yang diambil sesuai dengan tulisan apa yang ada dalam sebuah gambar.
 - ii. Menambahkan fitur lainnya ke dalam aplikasi ini, misalnya yaitu pengaturan ukuran tulisan saat ditampilkan, dukungan bahasa yang lebih luas, dan dapat memperbesar sendiri gambar yang digunakan sehingga dapat meningkatkan akurasi pendeteksian teks.

REFERENSI

- [1] N. Rahmawati, S. A. Wibowo, and U. Sunarya, "Analisis Sistem Optical Character Recognition (Ocr) Pada Dokumen Digital Menggunakan Metode Tesseract Performance Analysis of Optical Character Recognition (Ocr) System on Digital Documents Using Tesseract Method," *e-Proceeding Eng.*, vol. Vol.8, no. 5, pp. 4777–4785, 2021.
- [2] K. Ibnutama and M. G. Suryanata, "Ekstraksi Karakter Citra Menggunakan Optical Character Recognition Untuk Pencetakan Nomor Kendaraan Pada Struk Parkir," *J. Media*, vol. 4, pp. 1119–1125, 2020, doi: 10.30865/mib.v4i4.2432.
- [3] M. M. Effendi, A. Y. Permana, and I. Nawangsih, "Penerapan Ekstraksi Image ke TXT dengan Optical Character Recognition untuk Otomatis Data Kependudukan," *Simp. Nas. Ilm. Simponi*, no. November, pp. 496–502, 2019, doi: 10.30998/simponi.v0i0.484.
- [4] S. Supriadi, "Aplikasi Kalkulator Tulisan Tangan Sederhana Menggunakan Optical Character Recognition (OCR)," *Appl. Technol. Comput. Sci. J.*, vol. 3, no. 2, pp. 103–116, 2021, doi: 10.33086/atcsj.v3i2.1867.
- [5] K. Ibnutama and M. G. Suryanata, "Ekstraksi Karakter Citra Menggunakan Optical Character Recognition Untuk Pencetakan Nomor Kendaraan Pada Struk Parkir," *J. Media*, vol. 4, pp. 1119–1125, 2020, doi: 10.30865/mib.v4i4.2432.
- [6] K. Santoso, I. S. Nasution, and D. S. Jayanti, "Sistem Informasi Geografis Berbasis Web Untuk Pemetaan Potensi Pembangkit Listrik Tenaga Mikro Hidro (PLTMH) di DAS Krueng Aceh Bagian Hulu (Web-Based Geographic Information System for Mapping the Potential of Micro Hydro Power (MHP) in the Krueng Ac," *J. Ilm. Mhs. Pertan.*, vol. 4, no. November,
- [7] N. Mamuriyah and J. Jacky, "Perancangan dan Pembuatan Alat untuk Mendeteksi Teks Hangul dan Inggris pada Menu Makanan Menggunakan metode OCR (Optical Character Recognition)," *Telcomatics*, vol. 6, no. 1, pp. 1–10, 2021, doi: 10.37253/telcomatics.v6i1.5054.
- [8] H. Rizqi, G. E. Setyawan, and W. Kurniawan, "Sistem Kendali Navigasi Ar Drone Berbasis Pengenalan Teks Dengan Menggunakan Metode Optical Character Recognition," vol. 2, no. 11, pp. 5277–5284, 2018, [Online]. Available: <http://j-ptiik.ub.ac.id>.