

# Penerapan Metode Long Short-Term Memory Pada Pendataan Warga Berbasis Android

Sri Lestari<sup>1</sup>, Muhammad Fakhri Pratama<sup>2,\*</sup>

<sup>1</sup> Sistem Informasi, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Jakarta, Indonesia

<sup>2</sup> Teknik Informatika, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Jakarta, Indonesia

Email: <sup>1</sup>sri.lestari1203@gmail.com, <sup>2,\*</sup>ari.japindo@gmail.com

Submitted: 24/07/2022; Accepted: 23/08/2022; Published: 30/08/2022

**Abstrak**—Seiring dengan perkembangan teknologi, pemanfaatan dalam pengolahan data baik input maupun output menjadi sangat penting. Dimulai dari penginputan data secara manual, membuktikan bahwa pemanfaatan teknologi masih belum sepenuhnya digunakan. Data yang selama ini diolah secara konvensional terkadang menemui masalah yang disebabkan oleh kesalahan manusia (Human Error). Tak hanya itu, ketika data harus dimasukkan kedalam komputer, membutuhkan tambahan waktu dan ketika terjadinya bencana alam seperti banjir, gempa, tanah longsor. Berkas sensus yang masih berupa konvensional kemungkinan akan rusak dan akan membutuhkan waktu untuk mendata ulang. Hal tersebut menjadi alasan utama bagi RT 004 RW 008 membutuhkan sistem pendataan yang efisien. Penerapan teknologi Optical Character Recognition (OCR) dalam penelitian ini diharapkan mampu menjadi solusi dalam meningkatkan efisiensi proses penginputan data. Pada penelitian ini, proses penginputan dilakukan secara realtime menggunakan metode Long Short-Term Memory (LSTM) untuk memindai data pada e-KTP. Hasil dari pengujian teknologi OCR dengan metode LSTM pada penelitian ini untuk tiga percobaan yang dilakukan adalah 1) Percobaan pendeteksian e-KTP menghasilkan akurasi sebesar 100%, 2) Percobaan pembacaan 19 atribut pada e-KTP menghasilkan akurasi sebesar 98,42%, 3) Pengujian kesalahan pembacaan atribut e-KTP menghasilkan akurasi sebesar 93,56%. Secara keseluruhan akurasi metode LSTM pada pendeteksian kata pada e-KTP menghasilkan akurasi sebesar 92,1%. Dari hasil tersebut menunjukkan bahwa penerapan teknologi OCR dengan metode LSTM merupakan salah satu solusi tepat untuk proses penginputan data.

**Kata Kunci:** LSTM; Optical Character Recognition (OCR); Realtime System; Pendataan Warga; Berbasis Android

**Abstract**—Along with technological developments, utilization in data processing both input and output become very important. Starting from inputting data manually, it proves that the use of technology is still not fully used. Data that has been processed conventionally sometimes encounters problems caused by human error. Not only that, when data must be entered into a computer, it requires additional time and when natural disasters such as floods, earthquakes, landslides occur. Census files that are still in the conventional form are likely to be damaged and will take time to re-data. This is the main reason for RT 004 RW 008 to need an efficient data collection system. The application of Optical Character Recognition (OCR) technology in this study is expected to be a solution in increasing the efficiency of the data input process. In this study, the input process is carried out in real time using the Long Short-Term Memory (LSTM) method to scan data on e-KTP. The results of testing OCR technology with the LSTM method in this study for the three experiments carried out were 1) The e-KTP detection experiment resulted in an accuracy of 100%, 2) The trial of reading 19 attributes on the e-KTP resulted in an accuracy of 98.42%, 3) Testing the error reading of the e-KTP attribute resulted in an accuracy of 93.56%. Overall, the accuracy of the LSTM method in word detection on e-KTP results in an accuracy of 92.1%. These results indicate that the application of OCR technology with the LSTM method is one of the right solutions for the data input process.

**Keywords:** LSTM; Optical Character Recognition (OCR); Realtime System; Citizen Data Collection; Android Based

## 1. PENDAHULUAN

Perkembangan teknologi di era terkini ini sudah banyak penemuan berasal dari kalangan ilmuwan, masyarakat serta pelajar buat menginovasikan sebuah temuan yang bertujuan mempermudah aktifitas manusia, yang akan terjadi inovasi ini diperlukan dapat membantu kegiatan warga disetiap harinya [1]. Namun, belum banyak inovasi yg dapat mempermudah pekerjaan manusia, salah satunya dalam bidang pendataan. Pendataan yg sempurna diharapkan akan menghasilkan manfaat yang akbar bagi masyarakat jika diolah dengan benar [2].

Selama ini pendataan warga di wilayah RT 004/08 diolah secara konvensional dengan cara pengurus RT mendatangi rumah warga berdasarkan data sebelumnya dan memberikan formulir kepada warga tersebut, lalu warga mengisi formulir tersebut dan memberikan salinan E-KTP kepada petugas sebagai berkas untuk diarsipkan. Namun dengan cara tersebut terkadang menemui masalah yang ditimbulkan sang kesalahan insan (*Human Error*) [3]-[7]. Tak hanya itu, ketika data harus dimasukkan kedalam komputer, membutuhkan tambahan waktu dan ketika terjadinya bencana alam seperti banjir, gempa, tanah longsor. Berkas berkas yang masih berupa konvensional kemungkinan akan rusak dan akan membutuhkan waktu untuk mendata ulang [8]-[13]. Oleh karena itu, dibutuhkan suatu sistem untuk memudahkan proses input data ke dalam database agar tidak membutuhkan waktu yang lama. Sehingga pengurus RT tidak perlu menulis banyak formulir dan memasukkan data secara manual.

*Optical Character Recognition* (OCR) merupakan prinsip untuk mengenali pola atau pattern recognition, di mana sistem akan mencoba mengenali citra yang dimasukkan. Citra tersebut kemudian dicocokkan dengan citra yang telah ditentukan [4]. Pengenalan pola banyak diaplikasikan untuk pendeteksian tulisan, sidik jari, dan wajah. Ada beberapa penelitian tentang OCR yang sudah dilakukan terkait OCR dan Long Short-Term Memory [15]-[17]. Penelitian oleh (Zhang, Kexin, 2018) mengimplementasikan Long Short-Term Memory

untuk membuat sistem otomatis untuk mengidentifikasi dan mengklasifikasi gambar berdasarkan data fashion-MNIST.

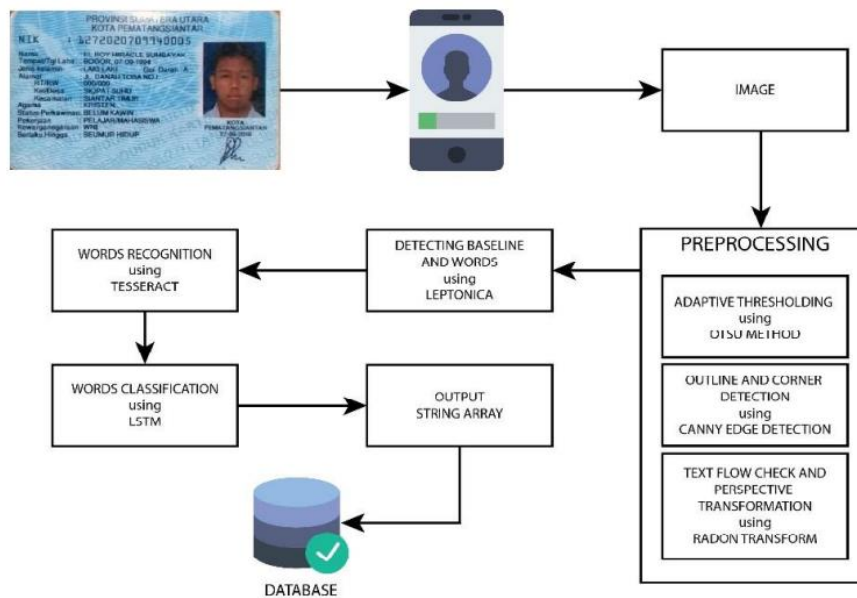
## 2. METODOLOGI PENELITIAN

### 2.1 Teknik Pengumpulan Data

Teknik pengumpulan data digunakan untuk mengumpulkan data sesuai tata cara penelitian sehingga diperoleh data yang dibutuhkan. Teknik pengumpulan data merupakan langkah yang paling strategis dalam penelitian karena tujuan utama dari penelitian adalah mengumpulkan data. Teknik pengumpulan data dalam penelitian ini menggunakan teknik observasi, dan wawancara.

### 2.2 Tahapan Penelitian

Metode LSTM (*Long Short Term Memory*) yang digunakan dalam pembuatan sistem memiliki beberapa tahapan [18]. Tahapan pertama adalah dengan menscan e-KTP untuk mengambil input citra atau *Image*. Sistem akan melakukan tahapan preprocessing yaitu tahapan untuk mengambil atau mengelompokkan bagian citra yang dianggap sebagai e-KTP. Tahapan preprocessing dimulai dengan adaptive tresholding menggunakan metode otsu, kemudian masuk ke tahapan *outline and corner detection* untuk mengambil batas tepi e-KTP dengan metode *canny edge detection*, kemudian masuk ke tahapan *text flow check and perspective transformation* menggunakan algoritma radon transform. Setelah melewati tahapan preprocessing, tahapan berikutnya adalah detecting baseline and words menggunakan leptonica, karakter yang sudah dideteksi kemudian mulai dilakukan pengenalan kata pertama dengan menggunakan tesseract yang kemudian di klasifikasi dengan metode LSTM (*Long Short Term Memory*) yang menghasilkan output dalam bentuk string array dan menyimpannya kedalam database [19]. Arsitektur umum yang digunakan pada sistem dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

Berdasarkan gambar 1 dapat dijelaskan sebagai berikut:

a. Input Citra (*Image*)

Input citra atau *Image* yang diambil dilakukan secara langsung atau realtime dengan cara menscan 3-KTP menggunakan aplikasi yang dibuat.

b. Input Citra (*Image*)

Disaat menscan e-KTP, sistem melakukan tahapan preprocessing. Tahapan preprocessing dimulai dengan proses adaptive tresholding yaitu mengubah *Image* menjadi abu-abu kemudian membedakan object pada *Image* menggunakan metode otsu. Selanjutnya masuk dalam proses outline and corner detection yaitu proses untuk memisahkan object dengan background *Image* dengan mengenali tepi atau batas-batas e-KTP dengan menggunakan metode canny edge detection. Setelah object telah dikenali maka masuk ke proses text flow check and perspective transformation yaitu proses untuk mengenali bentuk dan text yang terdapat pada e-KTP, proses ini dilakukan untuk mengatasi jika *Image* yang diambil dalam keadaan miring.

c. Input Citra (*Image*)

Tahapan ini adalah tahapan di mana sistem yang dibuat mencari object yang dinyatakan sebagai text yang ada pada e-KTP setelah melalui tahapan preprocessing. Tahapan ini menggunakan library atau metode leptonica. Selain itu tahapan ini juga mengelompokkan bagian text pada setiap atribut.

- d. Input Citra (*Image*)  
Pada tahapan ini object yang dinyatakan sebagai text pada e-KTP akan mulai dilakukan pengenalan dengan tesseract. Object-object yang dinyatakan sebagai text pada tahapan sebelumnya masih dalam berbentuk *Image*, kemudian pada tahapan ini *Image* yang dinyatakan sebagai text tersebut dibaca apakah merupakan sebuah karakter atau tidak dengan menggunakan tesseract.
- e. Input Citra (*Image*)  
Pada tahapan ini, hasil karakter yang dinyatakan pada tahapan sebelumnya mulai di klasifikasi apakah benar sebuah karakter yang dikenali atau tidak dengan menggunakan metode *Long Short Term Memory* (LSTM).
- f. Input Citra (*Image*)  
Output pada sistem adalah string array yang telah di pilih sesuai urutan attribute yang ada pada e-KTP
- g. Input Citra (*Image*)  
String array yang merupakan output dari proses maka akan disimpan kedalam database. Database yang digunakan adalah database mysql.

### 2.3 Rancangan Pengujian

Rancangan pengujian adalah tahap untuk mengetahui bagaimana kinerja dari sistem yang sudah dibangun. Rancangan pengujian ini bertujuan untuk melihat apakah hasil analisa dan rancangan yang dilakukan telah sesuai dengan yang di harapkan. Pada tahap ini, peneliti akan merancang desain aplikasi berbasis Optical *Character Recognition* untuk Ekstraksi Teks pada Gambar yang akan dibuat.

## 3. HASIL DAN PEMBAHASAN

Pada era moderen saat ini di sektor bisnis khususnya pada petugas RT 004/08 Sungai Bambu. sudah banyak yang menggunakan teknologi. Hal tersebut dilakukan karena untuk menunjang proses kerja dari suatu bisnis. Adapun analisa kebutuhan teknologi yang telah dilakukan adalah system aplikasi pendataan penduduk diharapkan dapat online 1 x 24 jam, mempunyai fitur input secara otomatis atau tanpa harus mengetik secara manual dan dapat diakses melalui smartphone sehingga proses pendataan penduduk di RT 004/08 Sungai Bambu Jakarta Utara dapat mempermudah dan berjalan lebih efisien secara waktu dan tenaga.

### 3.1 Pengujian Sistem

Pada tahap ini akan dijelaskan mengenai pengujian pendeteksian kata pada e-KTP. Pengujian sistem yang pertama yaitu untuk akurasi pendeteksian e-KTP, dengan menggunakan 10 data e-KTP akan dilihat berapa banyak e-KTP yang dapat terbaca Text nya. Pengujian kedua yaitu untuk menguji akurasi pembacaan atribut yang ada pada e-KTP, dari 19 atribut yang ada akan dilihat berapa banyak atribut yang bernilai null per e-KTP. Pengujian yang terakhir yaitu untuk menguji kesalahan pembacaan atribut yang ada pada e-KTP, dari 19 atribut yang ada akan dilihat berapa banyak atribut yang terdapat kesalahan dalam pembacaannya per e-KTP.

#### 3.1.1 Pengujian Pendeteksian e-KTP

Pengujian sistem akan dimulai dengan pendeteksian e-KTP. Jika e-KTP berhasil dideteksi, maka proses berikutnya adalah pendeteksian kata e-KTP yang akan memberikan output berupa string array dan akan disimpan kedalam database. Hasil pengujian pendeteksian e-KTP dapat dilihat pada lembar lampiran. Berdasarkan hasil dari penelitian yang telah dilakukan, akurasi dari pendeteksian e-KTP pada sistem dihitung dengan rumus (1), sebagai berikut.

$$\text{Akurasi deteksi ektp} = \frac{\text{jumlah ektp terdeteksi}}{\text{total jumlah ektp}} \times 100\% \quad (1)$$

#### 3.1.2 Pengujian I

Pada pengujian ini, nilai yang ingin dicari adalah akurasi sistem dalam mendeteksi kata pada e-KTP berdasarkan berapa banyak atribut yang dideteksi tidak bernilai null. Hasil dari pengujian ini dapat dilihat pada lembar lampiran. Berdasarkan data diatas, hasil dari pengujian 1 untuk menguji seberapa besar akurasi sistem dalam mendeteksi atribut dan tidak bernilai null dapat dihitung menggunakan rumus (2), sebagai berikut.

$$\text{Akurasi pengujian I} = \frac{\text{jumlah tidak null}}{\text{total jumlah}} \times 100\% \quad (2)$$

#### 3.1.3 Pengujian II

Pada pengujian 2 ini, dari hasil deteksi atribut yang tidak bernilai null pada pengujian 1, maka akan dicari berapa banyak atribut yang dideteksi benar sesuai dengan e-KTP dan berapa banyak yang miss dalam pembacaan atributnya. Hasil dari pengujian ini dapat dilihat pada lembar lampiran. Berdasarkan pengujian yang dilakukan, maka untuk menghitung akurasi sistem dalam mendeteksi kata setiap atributnya sesuai dengan e-KTP, dapat menggunakan rumus (3), sebagai berikut.

$$\text{Akurasi pengujian II} = \frac{\text{jumlah pendeteksian sesuai ektp}}{\text{jumlah tidak null}} \quad (3)$$

### 3.1.4 Hasil Keseluruhan Pengujian Sistem

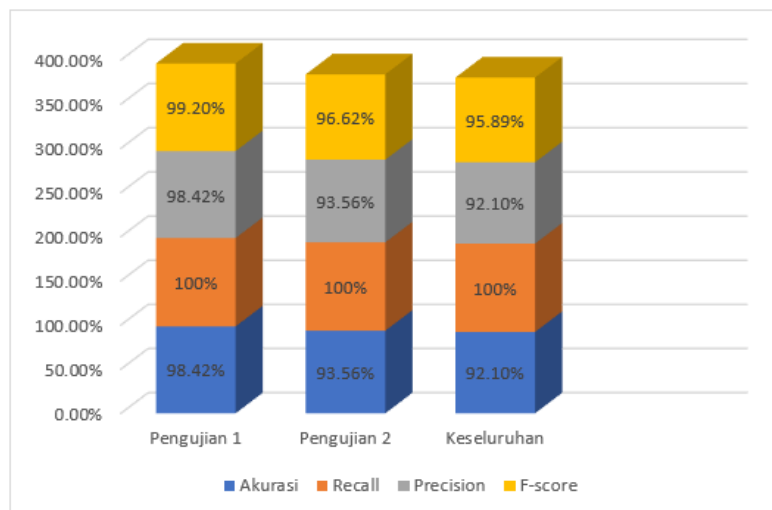
Hasil pengujian yang dilakukan pada pendeteksian kata e-KTP menggunakan metode *Long Short Term Memory* (LSTM) dapat dilihat pada tabel 1 berikut.

**Tabel 1.** Jenis jenis database

No	Pengujian	Jumlah Data Benar	Jumlah data total	Akurasi
1	Deteksi e-KTP	10	10	100%
2	Pengujian 1	187	190	98,42%
3	Pengujian 2	175	187	93,56%
4	Keseluruhan	175	190	92,11

### 3.2 Hasil Metode Evaluasi

Hasil pengujian yang telah didapatkan maka akan dievaluasi dengan menggunakan metode confusion matrix. Hasil evaluasi pada pengujian pendeteksian e-KTP ditampilkan pada gambar 2 berikut.



**Gambar 2.** Confusion Matrix

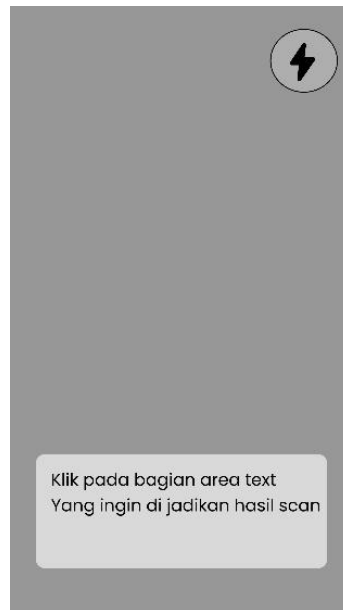
### 3.3 Implementasi Program

Pada bagian ini akan menjelaskan hasil implementasi dari rancangan tampilan antarmuka penerapan metode long *short-term memory* pada pendataan warga berbasis android. Rancangan tampilan beranda berisi 4 *CardView* yaitu Scan OCR, Data warga, bantuan, dan logout. Tampilan beranda bisa dilihat pada gambar 3 berikut.



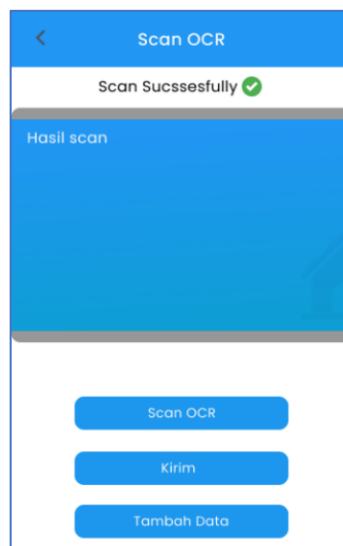
**Gambar 3.** Tampilan beranda

Tampilan scan e-KTP merupakan halaman untuk menambahkan data warga yang akan digunakan sebagai data latih. Halaman ini akan muncul ketika pengguna menekan tombol scan OCR. Tampilan scan dapat dilihat pada gambar 4 berikut.



**Gambar 4.** Tampilan scan

Tampilan hasil scan merupakan tampilan di mana hasil Scan e-KTP ketika sudah diklik pada tampilan scan OCR, bisa dilihat pada gambar 5 berikut ini.



**Gambar 5.** Hasil scan

## 4. KESIMPULAN

Pengujian dan penelitian yang telah dilakukan pada implementasi metode *Long Short Term Memory* (LSTM) dalam pendeteksian kata pada e-KTP menghasilkan tingkat akurasi secara keseluruhan adalah 92,1% dengan nilai *F-Score* sebesar 95,89%. Hasil dari pengujian untuk setiap atribut yang tidak bernilai null sebesar 98,42% dengan nilai *F-Score* sebesar 99,2%. Hasil dari pengujian untuk setiap atribut yang tidak bernilai null dan benar sesuai dengan e-KTP sebesar 93,56% dengan nilai *F-Score* sebesar 96.62%.

## REFERENCES

- [1] E. Utami, O. D. Nurhayati, and K. T. Martono, "Aplikasi Penerjemah Bahasa Inggris – Indonesia dengan Optical Character Recognition Berbasis Android," *J. Teknol. dan Sist. Komput.*, vol. 4, no. 1, p. 167, 2016, doi: 10.14710/jtsiskom.4.1.2016.167-177.

- [2] A. Mulyanto, E. Susanti, F. Rossi, W. Wajiran, and R. I. Borman, “Penerapan Convolutional Neural Network (CNN) pada Pengenalan Aksara Lampung Berbasis Optical Character Recognition (OCR),” *J. Edukasi dan Penelit. Inform.*, vol. 7, no. 1, p. 52, 2021, doi: 10.26418/jp.v7i1.44133.
- [3] A. S. Kusnanto, “IMPLEMENTASI OCR (OPTICAL CHARACTER RECOGNITION) MENGGUNAKAN METODE STRING MATCHING UNTUK Anggi Sanjaya Kusnanto *Jurnal Ilmiah Komputer dan Informatika (KOMPUTA)*,” Implementasi OCR (Optical Character Recognition) Menggunakan Metod. String Matching Untuk Mendeteksi Obat dan Makanan Berbas. Android, 2014.
- [4] A. S. Rizal, “Analisis Dan Implementasi Aplikasi Penerjemah Dan Penambah Harakat Kitab Klasik / Kitab Kuning,” *Pendidik. Agama Islam*, vol. IX, no. 2, pp. 85–95, 2011.
- [5] A. Solichin and Z. Rahman, “Aplikasi Identifikasi Nomor Kendaraan Berbasis Android Dengan Metode Learning Vector Quantization,” *Tek. Inform.*, vol. 3, no. 3, pp. 216–222, 2015.
- [6] A. Sonita and Khairunnisyah, “Aplikasi Pendeteksi Obat dan Makanan Menggunakan OCR (Optical Character Recognition),” *J. Inform. UPGRIS*, vol. 4, no. 1, pp. 111–116, 2018.
- [7] A. W. Hamdani and A. Prapanca, “Sistem Deteksi Plat Kendaraan pada Parkiran Rumah Pribadi dengan Metode Background Subtraction dan Optical Character Recognition,” vol. 03, pp. 250–257, 2022.
- [8] C. Wibisono and S. Budi, “Form Recognition dan Character Mapping Menggunakan *Image* Segmentation dan Optical Character Recognition,” *J. Tek. Inform. dan Sist. Inf.*, vol. 7, no. 1, pp. 154–164, 2021, doi: 10.28932/jutisi.v7i1.3340.
- [9] D. Aditama, H. Rosyid, and N. Fahriani, “Aplikasi Pembelajaran Alfanumerik Untuk Anak Usia Pra-Sekolah Berbasis Android Menggunakan Metode Tesseract-Ocr,” *J. Comput. Sci. Informatics Eng.*, vol. 2, no. 1, pp. 70–76, 2018, doi: 10.29303/jcosine.v2i1.128.
- [10] D. Z. Putri, D. Puspitaningrum, and Y. Setiawan, “Konversi Citra Kartu Nama ke Teks Menggunakan Teknik OCR dan Jaro-Winkler Distance,” *J. Teknoinfo*, vol. 12, no. 1, p. 1, 2018, doi: 10.33365/jti.v12i1.35.
- [11] F. P. Putra, O. N. Huda, and J. Jaroji, “Aplikasi E-Voting Berbasis Mobile Menggunakan Optical Character Recognition,” *J. Inform. Polinema*, vol. 4, no. 3, p. 235, 2018, doi: 10.33795/jip.v4i3.199.
- [12] H. Diwanti, I. S. Sumaryo, and C. Setianingsih, “Real Time Smart CCTV Untuk Mendeteksi Plat Nomor Kendaraan Menggunakan Optical Character Recognition,” *e-Proceeding Eng.*, vol. 6, no. 2, pp. 2–9, 2019.
- [13] J. Sahertian, M. I. Khotmuniza, and R. Helilintar, “Sistem Parkir Menggunakan Ocr (Optical Character Recognition) Plat Nomer Dan Iot (Internet of Things),” *Joutica*, vol. 5, no. 2, p. 363, 2020, doi: 10.30736/jti.v5i2.443.
- [14] K. D. Laksana, M. G. Sunarya, and G. M. Darmawiguna, “Aplikasi Ocr (Optical Character Recognition) Tulisan Tangan Aksara Bali Dengan Metode Contour Analysis,” *Kumpul. Artik. Mhs. Pendidik. Tek. Inform.*, vol. 4, no. 1, pp. 2252–9063, 2015.
- [15] 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.
- [16] M. P. Damayanti et al., “Data Penulis :,” 2018.
- [17] S. Komputer, F. Teknik, E. Univesitas, and T. Jln, “Perancangan Dan Implementasi Directional Feature Extraction Dan Indonesia Berbasis Android,” vol. 3, no. 2, pp. 2194–2202, 2016.
- [18] S. S. Nurhaliza and L. ETP, “Sistem Pengenalan Karakter Dokumen Secara Otomatis Menggunakan Metode Optical Character Recognition,” *Petir*, vol. 15, no. 1, pp. 166–175, 2022, doi: 10.33322/petir.v15i1.1610.
- [19] Y. R. Prayogi and S. N. Budiman, “Color Grading Systems to Classify Ripeness of Apple Mango Fruit,” *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 3, no. 2, pp. 57–61, 2018, doi: 10.25139/inform.v3i2.1010.