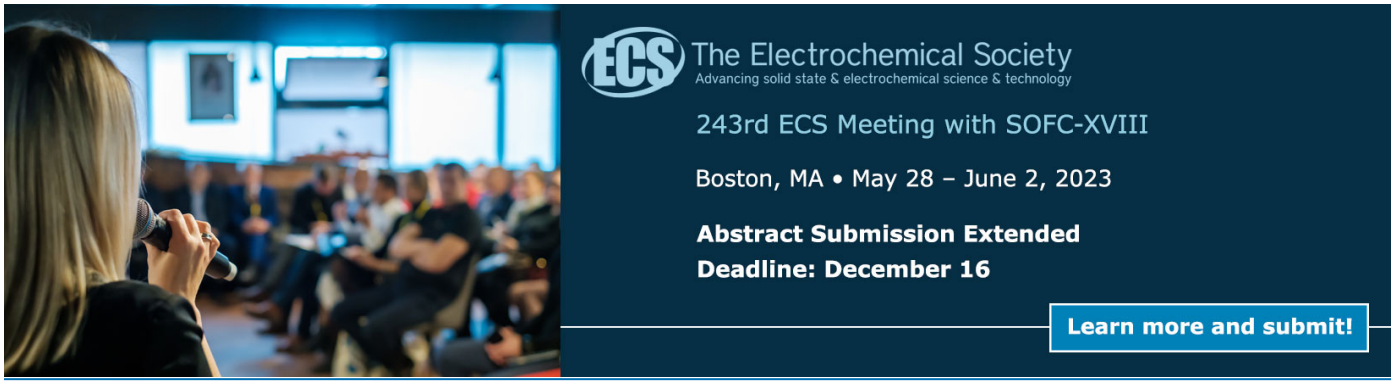
Automatic License Plate Recognition Using Optical Character Recognition



Automatic License Plate Recognition Using Optical Character Recognition

U Salimah, V Maharani and R Nursyanti

Faculty of Technology and Informatics, Indonesian University of Informatics and Business, Soekarno-Hatta Street no. 643, Bandung, Jawa Barat, Indonesia 40285

*Email: u.salimah@mhs.unibi.ac.id

Abstract. Automatic plate recognition is useful for everyday life, for example for law enforcement, traffic control, access to restricted areas, E-Tolls, or for checking in parking areas. The police number plate or TNKB is the main identity of a motorized vehicle that is roadworthy or operating on the highway. Thus, we can use the license plate to identify the vehicle owner. This study aims to recognize vehicle number plates using an Android smartphone. Using a camera on a handphone to scan the plate image, and it is processed to find out the vehicle owner's information from the numbered plate being scanned. This study also uses Optical Character Recognition in processing images scanned by the camera where the steps taken are image acquisition, pre-processing, segmentation, normalization, feature extraction and recognition. The results show that the camera on an Android smartphone can be used to read and display information on number plate owners and also show that vehicle plate recognition using Optical Character Recognition based on the tests conducted shows the percentage of successful character recognition on motorized vehicle plates by 75% while the characters are equal to 97.36%.

1. Introduction

Automatic plate reading is useful for everyday life. There are many reasons why it is important for both individuals and organizations to identify vehicles or their owners, for example for law enforcement, traffic control, access to restricted areas, E-Tolls, or for checking in parking areas. License Plate Recognition (LPR) has the advantage of recording a license plate number to fight crime, if an additional camera is given which is focused on the driver's face so that it can simultaneously recognize the identity of the car driver.

ALPR is a new technology in the transportation system automation world. This technology utilizes image processing to identify vehicles from the license plate image [1], however variations on the license plate and the surrounding environment such as font size, font type, font color, number plate location, and differences in intensity due to lights or environment can cause problems during recognition. license plate [2]. In Indonesia, ALPR is still being developed, especially for the detection of number plate objects. Although in other countries ALPR has been developed a lot, this ALPR system cannot be used directly in Indonesia due to differences in the format or type, color, and structure of number plates.

Research on Automatic License Plate Recognition (ALPR) or automatic number plate recognition in various parts of the world has been carried out, but those who conduct research on number plates in Indonesia are still developing today. In the research conducted by Michael, the method used in image conversion is Optical Character Recognition. Tests carried out resulted in an accuracy of 78.57% if the condition of the number plate was in good condition and for the condition of the number plate in a bad condition it resulted in an accuracy of 57.41% [3]. Referring to further research conducted by Manshur, in 2018, stated that the position of the camera will determine the accuracy of the final result [4]. Budianto, in his research he compared the level of accuracy of various methods on automatic license plate recognition and various factors that influence it, and the results of his research stated that the Neural Network Backpropagation algorithm produced an accuracy rate of 96 percent and there were many factors that influenced this accuracy, one of which was the camera quality and also the distance between the camera and the vehicle [5].

The purpose of the study expected by the authors was to produce a transportation automation system capable of recognizing motor vehicle license plates using Optical Character Recognition and recognizing motor vehicle license plate owners in real time using mobile devices as well as knowing the accuracy of the systems tested and creating applications capable of identifying vehicle owners.

2. Material and Methods

2.1 Automatic License Plate Recognition System

Automatic License Plate Recognition or ALPR for short is a method used as vehicle surveillance that uses optical character recognition in images, this method works by reading vehicle license plates. ALPR has an important role in its application in daily life, such as for law enforcement, traffic control, access to restricted areas, E-Tolls, or for checking in parking areas. Vehicle license plates recognition from the ALPR system is taken from an image or image taken through a camera, whether in color or not [6]. Make use of many techniques such as object detection, image processing, and pattern recognition. ALPR is also known as vehicle identification (AVI), car plate recognition (CPR), automatic number plate recognition (ANPR), and optical character recognition (OCR) for vehicles. ALPR system that extracts a license plate number from a given image can be composed of four stages [7].

The first step is to acquire the image using the camera. The second stage is to extract the license plate from the image based on several features such as borders, colors, or the presence of characters. The third stage is to segment number plates and extract characters by projecting their color information, by labeling them, or by matching their position to a template. And the last stage is to recognize extracted characters by matching the template or by using classifiers such as neural networks and fuzzy classifiers [6].

2.2 Optical Character Recognition

OCR stands for Optical Character Recognition. This technology allows the machine to recognize characters automatically via optics. Humans themselves have optics, namely eyes. The image seen by the eye is input to the brain. The ability to recognize this input varies from person to person due to many factors [8]. OCR is a technology that functions much like human reading ability (see Figure 1).

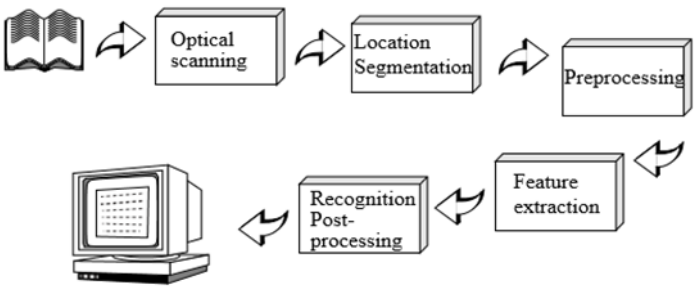


Figure 1. OCR

In OCR the pattern is letters, numbers and some special symbols such as commas, question marks, etc., whereas different classes correspond to different characters. The typical OCR system consists of several components [9]-[10]. In Figure 1, the general configuration is illustrated in OCR. The first step in this process is to digitize analogue documents using an optical scanner. When a region containing text is found, each symbol is extracted through a segmentation process. Extracted symbols can then be processed first, eliminating noise, to facilitate the extraction of features in the next step.

3. Purposed System

Image acquisition is an early stage in the system for acquiring digital imagery. License plate image scan via Vivo 1610 smartphone with camera resolution of 13 megapixels (mp) with sampling distance of 5 – 10 cm and camera height of 150 – 170 cm. For system design can be seen in Figure 2.

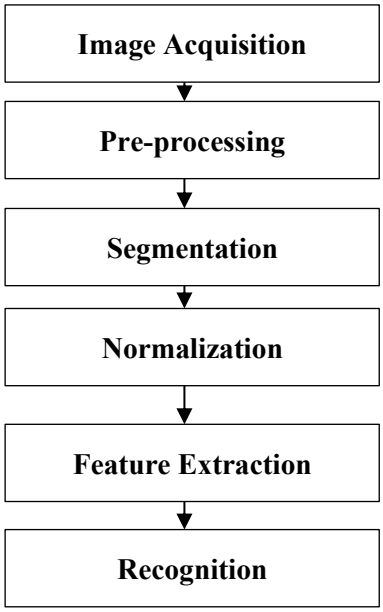


Figure 2. System design

In the pre-processing stage, the system will adjust the image that has been taken to suit the next process. Pre-processing is done in the form of cropping or cutting the plate number area with the background area. This cutting is done by thresholding by setting the threshold value between the

background and the number plate area value. The preprocessing process consists of grayscaling and binaryization processes. Both of these processes are carried out to change the pixel intensity. This process produces an image with binary colors, namely black and white (see Figure 3).

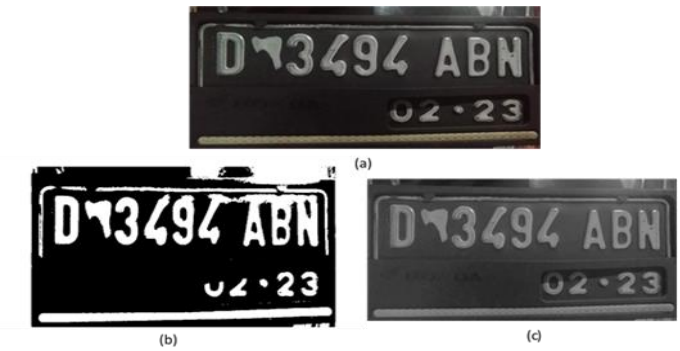


Figure 3. Original Image (a), Binnarizing (b), Grayscaling (c)

At the segmentation stage, the position of each number and character will be determined. In this section, if there are related candidate digits, it is likely that they will be extracted simultaneously. That way, to detect each row, each row and each column will be checked. If the number of consecutive black pixels is less than or more than the specified threshold limit, the pixels will be converted to white. With this process, it can be done by detecting a column containing white and all matrices equal to 0 so it is determined as the initial cutoff limit. This detection is carried out continuously until a black color is found with all matrices equal to 1. If a black color is found, the process will go back one column again and the last white color will be displayed with label number one. the next detection process is carried out the same thing until the last labeling process has been completed. And pixels containing black will be deleted and not saved. That way the characters that have been labeled will be carried out the next process (see Figure 4).

Figure 4. Segmentation

The process that will be done at the normalization stage is scaling against the image so that the image has a fixed resolution. As we know, the camera on the phone has different resolutions so it must be determined the optimal resolution so that the resulting image is not too large, as it will affect the memory and reading process of the image. However, it should not be too small, as it will affect the text reading result because the resolution of the image is too small resulting in the characters in the text being split. In this study, the resolution taken on the camera will be converted into dimensions of 1024 x 960 pixels. After obtaining the edge detection image, each character will be matched to the imagery in the training image database. The process of matching by adjusting the character image size to the character template that's in the library (see Figure 5).

Figure 5. Expresses the number 9 character

After obtaining the image from the edge detection results, each character will be matched with the image in the training image database. The matching process is by adjusting the size of the character image with the character template in the library. The pattern shown in Figure 6 expresses the number 9 character. The black box is worth 1 while the white box is 0. The same is done for the other characters.

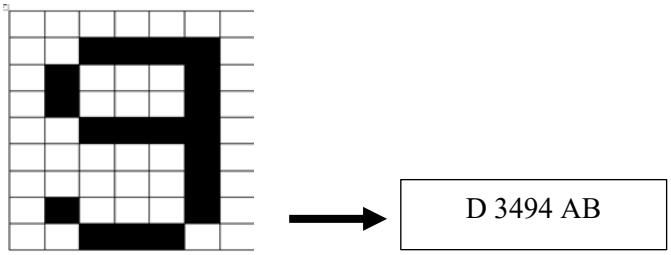


Figure 6. Feature Extraction

On Stage Process Introduction, Collection Fitur Results Extraction Citra On the training process Have Stored Will Used To Process Classification With data input. In This Process, dataset gambar Character To Training (training) Especially First Prepared By Character yang Will Determined.

4. Implementation and Testing

From the research conducted, the implementation for this system uses ML-Kit as an additional library to extract number plate text using an android camera. Which is then connected to the database to access the detected number plate owner data. This implementation is implemented and tested on android devices. The interface for this application can be seen in the image below:

The scanned plate will display the introductory resulting text on the next page, also displaying the vehicle owner's information, if the scanned plate results in a perfect introduction and the license plate is registered on the database.

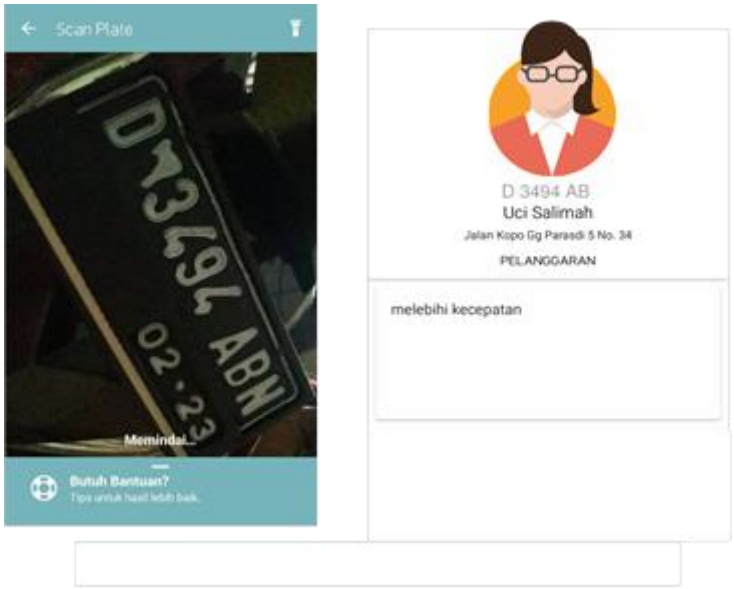
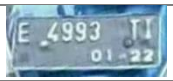
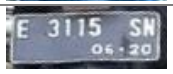
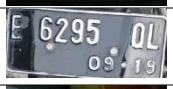
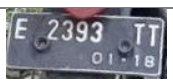


Figure 7. Interface implementation

Testing has been conducted to test the accuracy of this system. Test images are taken during the day under various conditions as well as shooting angles. As a result of the proposed implementation, it can be seen in Table 1.

Table 1. Test Results in this table are displayed 10 test results from 100 tests.

No.	Citra Flat	No. Plat	j. experiment	Introduction Results
		E 3996 R	3	E 3996 R
		E 5142 TW	3	142 TU
		E 4993 TI	3	E 4993 H
		E 3115 SN	3	E 3115 SN
		E 4564 SZ	3	E 4564 S
		E 1245 RG	3	E 1245 RG
		E 6295 QL	3	E 6295 OL
		E 5361 TI	3	E 5361 TI
		E 2956 RP	3	E 2956 RP
		E 2393 TT	3	E 2393 TT

After testing carried out on 100 number plate samples, it can be concluded with the percentage of success in the recognition process from 100 number plates, 75 number plates can be recognized, namely 75%. Character recognition accuracy with 722 total characters and 703 number of successful character recognition, amounting to 97.36%.

5. Conclusion

In this study, the program was able to recognize license plates using Optical Character Recognition (OCR) techniques. The application of the OCR method on the android operating system results in good accuracy in number plate readings, where from the test results are obtained 75% accuracy, where the accuracy for the reading of each character is 97.36 %. The system is able to display vehicle owner data that is successfully recognise correctly. Recognition failures are due to the resemblance of characters to each other, such as B and 8, 2 and Z, 1and I, D and 0, A and 4. Imperfections in character segmentation can cause characters to be truncated that can change the perception of character recognition.

References

[1] Singh R V and Randhawa N 2014 Automobile Number Plate Recognition AndExtraction Using Optical Character Recognition *International Journal of Scientific & Technology Research* **3** 37-39

[2] Ramachandran R 2015 Multi Depot Vehicle Routing Problem using ODV EV Based Genetic Algorithm *International Journal Of Engineering And Computer Science* **4** 10779-10786

[3] Michael A 2016 Pengenalan Plat Kendaraan Berbasis Android menggunakan Viola Jones dan Kohonen Neural Network *Ilkom Scientific Journal* **8** 95-102

[4] Manshur Y 2018 *Optical Character Recognition For Detection of Car Plates And Motor Vehicles On Gowa Engineering Campus* Bachelor Degree 54 – 55

[5] Budianto A 2018 Automatic License Plate Recognition: A Review with Indonesian Case Study *Scientific Journal of Informatics* **5** 258-270

[6] Du S, Ibrahim M, Shehata M and Badawy W 2012 Automatic license plate recognition (ALPR): A state-of-the-art review *IEEE Transactions on circuits and systems for video technology* **23** 311-325

[7] Sarfraz M, Ahmed M J and Ghazi S A 2003 Saudi Arabian license plate recognition system *International Conference on Geometric Modeling and Graphics, 2003. Proceedings* 36-41 IEEE.

[8] Chiou Y C, Lan L W, Tseng C M and Fan C C 2011 Optimal locations of license plate recognition to enhance the origin-destination matrix estimation *Proceedings of the Eastern Asia Society for Transportation Studies* **8** 297-297

[9] Sobh T and Elleithy K 2010 *Innovations in Computing Sciences and Software Engineering* Berlin: Springer Science & Business Media

[10] Shreya S, Upadhyay Y, Manchanda M, Vohra R and Singh G D 2019 Optical character recognition using convolutional neural network *2019 6th International Conference on Computing for Sustainable Global Development* 55-59