Detection and Identification Indonesia License Plate Using Background Subtraction Based on Area

F Damayanti, W K Dewi, E Rahmanita, A Rachmad

2Departemen of Informatics, Faculty of Engineering, University of Trunojoyo, Madura, Bangkalan, Indonesia

Corresponding author’s e-mail: fitrid@trunojoyo.ac.id, wahyuwkd@gmail.com, eza_rahmanita@yahoo.com, aery_r@yahoo.com

Abstract. The increasing number of vehicles, especially four-wheeled vehicles in big cities, raises traffic problems, including toll roads. With a fast transaction process at the entrance of the toll road, it can reduce the queue of traffic congestion. For this reason, it is necessary to use a vehicle number character recognition image processing application. In this study, an application for license plate character recognition was made on vehicle images starting from the plate detection process using background subtraction. Next is the segmentation process of each character and feature extraction process using area-based features extraction. For the character recognition process, use the K-Nearest Neighbor method. The data used in this study were 50 vehicle images with each of them having 3 distances namely 50 cm, 75 cm and 100 cm. The results of this research trial showed the best level of accuracy for the detection of license plate locations is 94%, the best accuracy of segmentation per character is 71.79% and 82.59% is the best accuracy of character recognition.

Keywords: area-based features extraction, background subtraction, of character recognition.

1. Introduction

Indonesian license plate is the identity of every vehicle that exists. Each vehicle, both two-wheeled and four-wheeled, certainly has a different license plate for each that shows the identity of the vehicle. The increasing number of vehicles, especially four-wheeled vehicles in big cities, creates traffic problems such as traffic jams. One of them is often found at the entrance of the toll road with a very large number of vehicles queuing into the toll road at rush hour and long holidays. As a result the queues are long and create traffic jams. For that, the faster the transaction at the toll gate, will reduce the vehicle queue. Therefore an application is needed to identify the license plate automatically to do a scan that will be recognized and stored. The stored data is then distributed to the toll exit gate so that it can speed up the transaction process on the toll booth.

Digital image processing has various types of fields according to their functions, one of which is pattern recognition. Pattern recognition is a form of the ability of a computer program to analyze the shape of an object from an image. Pattern recognition can be applied to analyze a vehicle license plate from a vehicle license plate image. This character recognition technique is commonly referred to as Optical Character Recognition (OCR) technology. License plate recognition system is one of the smart traffic system technology that can be used widely for many purposes, one of which is for the toll management system automatically [1].

Research on Automatic License Plate Recognition (ALPR) has been carried out with various methods and varying success rates. There are three main processes in the ALPR system, namely license plate localization (LPL), character segmentation, and optical character recognition (OCR) [2]. Several
previous studies have been conducted on Automatic License Plate Recognition (ALPR), namely research [3] [4]. In 2013, M.A. Massoud, M. Sabee, M. Gergais, R. Bakhit in his research "Automated new license plate recognition in Egypt". In this study, morphology was used for plate detection processes, while the introduction process used the Statically Template Matching method. To measure the plate number combination using Euclidian Distance [3].

In 2017, Ibrahim Turkyilmaz and Kirami Kacan in their research entitled "License Plate Recognition System Using Artificial Neural Networks", the character recognition stage use 3 artificial neural network feed-forward layers with backpropagation learning algorithm [4].

In this study, the author made an application for detecting the vehicle license plate on the front image of the car image and character recognition contained in the vehicle number. To detect the location of the plate, the use of the background subtraction method use the opening morphology to look for the background image. The segmentation process of each character uses the otsu thresholding method by finding the optimal threshold value from binarization which can separate the background plate and character on the vehicle license plate [5]. The otsu thresholding method is very good at overcoming disturbances that are not needed in the segmentation process [6]. Area-Based Features Extraction is used as a feature extraction method while K-Nearest Neighbor (KNN) is implemented in recognizing vehicle license plate characters, where the KNN method has been widely used in several studies [7] [8] [9] [10] [11].

2. Methods

The vehicle identity license plate recognition system is an important part of Intelligent Transportation Systems (ITS) [1]. Basically the introduction of license plate characters consists of three main processes, including license plate location detection, then each character's segmentation and the last process is the introduction of vehicle license plate characters. Each process requires an algorithm that is able to meet the needs of the system with fast and accurate performance in various conditions. But before the process, preprocessing stages of image processing are also important in determining the level of accuracy in the next process.

2.1. Preprocessing Image

Pre-processing on the image aims to get a better quality image than the previous image by processing image parameters, so that it can produce a form that is more suitable for the pixel values of the image for the next process. The stages of preprocessing for each study vary according to needs. In this study, the preprocessing process was carried out consisting of RGB conversion to grayscale, binaryization and image enhancement which aims to simplify the plate localization process.

In this study the process of improving image quality using background subtraction (image reduction) to get the foreground using morphology to represent the background. This process aims to better show objects in the image in this study are license plates, so that the license plate location can be detected perfectly.

2.2. Image Segmentation

Segmentation is an important process in license plate recognition systems. This process greatly determines the success of the output of the recognition system, because from the input image with a complex background, the system must be able to determine where the location or license plate is in the car. The segmentation process also aims to separate each character on the vehicle license plate. Some methods used in the segmentation process are as follows.

2.2.1 Otsu Thresholding

Otsu Thresholding is an algorithm used to segment gray digital images into foreground and white digital images (background) [12]. This study uses otsu thresholding for the segmentation process of an image. The segmentation stage is the stage of cutting the image which is then processed into several parts. The segmentation process is carried out per character of the vehicle license plate.

2.2.2 Connected Component Labeling

Connected Component Labeling is an approach to grouping pixels that are connected in an image. Pixels will be labeled with gray level or color according to the order so that it produces a labeled
image where the value of each pixel is the label value of the connected component. The image will be separated by cutting the pixel according to the label you have.

2.2.3 Region Properties

Region properties is a default matlab function used in the process of determining the bounding box area. Bounding box is a parameter used to specify objects and mark them with boxes that match the size of the object identified [13]. This function is used to mark each character on the license plate which will be segmented after getting the bounding box properties of each license plate character.

2.3. Feature Extraction

Each object or character must have different special characteristics that are used to distinguish between one pattern with another pattern. The higher the distinguishing power possessed by each feature, the higher the accuracy of grouping patterns based on characteristics. Feature extraction in an image means changing the value of the pixel coordinate intensity contained in the image into a set of value codes in each pixel. Feature extraction aims to capture the characteristics of an object in an image to distinguish one object from another object.

In this study the author uses area-based feature extraction. The method of this method works by dividing the image area into 10x10 areas and adding up the value of each area to get 100 feature values for each character [14].

2.4. Optical Character Recognition (OCR)

Character recognition or commonly called optical character recognition (OCR) in various studies is often used to identify the image of a letter which is then converted into a written file (text) [15]. In this study the character recognition stage uses the k-nearest neighbor method to classify a data based on training or learning data that has the closest distance to the object. Determination of the number of nearest neighbors uses the specified K value. The calculating the distance between the data can be classified with all training data using the euclidean distance.

3. Result

![Diagram of system process and plate localization process]

**Figure 1** (a) System Process and (b) Plate Localisation Process

3.1. Plate Localization

The plate localization stage aims to detect the location of the license plate from the front-facing vehicle image. Figure 1(b) shows the flow of the plate localization process. Inputs in the form of vehicle image appear from the front with a predetermined retrieval distance, which is 50 cm, 75 cm, Figure 1(a) displays the stages of the system work process from input to output. Input in the form of 150 vehicle images consist of 50 vehicle images with each of 3 distances, namely 50 cm, 75 cm and 100 cm. While the output in the form of vehicle license number text results from the introduction of the system.
For 100 cm, The image will be separated between the background and the foreground. Because what you are looking for is the car license plate area, the foreground image is taken, because the car license plate area will be more visible when separated from the foreground background. The foreground image is then binary for further purposes using otsu thresholding. The region's binary image will be measured based on the largest area to get the ROI (Region of Interest), which is the license plate that is considered to be the largest area. After obtaining an area from the ROI, it will be compared to the minimum area of the license plate obtained, if less than the minimum area that has been determined the next process will be carried out. If the minimum area has been met, a process will be carried out to image the car by referring to the ROI area that has been obtained in the previous process. If the minimum area has not been met, the binaryization process will be carried out again. After obtaining a plate image from the image of the car image, the filter region will be maximum from the license plate character. Output license plate image from cropping.

3.2. Character Segmentation

Character segmentation is a process that separates each character's license plate number. For the character segmentation process in vehicle license plate images can be done with several techniques [16] [17] [18] [19]. Figure 2 (a) shows the flow of the character segmentation process carried out in this study. Character segmentation is a process that separates each character's license plate number. For the character segmentation process in vehicle license plates, images can be done with several techniques [16] [17] [18] [19]. Figure 1(b) shows the flow process carried out in this study.

![Figure 2 (a) Character Segmentation Process and (b) Character recognition process](image)

Binary image output results from the plate localization process are used as input to the segmentation process for each license plate character. To segment each character license plate, the region is measured first to get bounding box properties of each license plate character. After that, the process of cropping and resizing will be carried out to 200 * 100 on binary plate images with bounding box for each region that has been obtained in the previous process. Each license plate character will be marked with rectangle with bounding box for each region obtained in the previous process. Then, a filter process is carried out to dispose of regions that are not considered license plate characters by measuring the height of each segmented character image as a minimum average. So, if a character does not meet the requirements / height below the specified one, it will be discarded. The results obtained are a number of images for each license plate character and also a license plate image with a rectangle marker.

3.3. Character Recognition

Character recognition is the process of changing the image of the characterization of the character into a license plate character text. Figure 2(b) shows the flow of the character recognition process.

In the process of character recognition, the input is in the form of segmented images per character on the license plate. Then for each segmented image per character, the license plate will be extracted to feature an area of size 10 * 10. After that, it will be classified using KNN between the area features.
of the license plate and the area features in the database, and the measurement of similarity will be measured using the euclidean distance. KNN [20] is an easy-to-use and good learning method to overcome recognition problems. The same is the case with existing learning methods [21] [22] [23]. The K data feature area in the database that is most similar to the area feature on the license plate will be taken as a reference; and then the most similar class will be converted into alphanumeric characters (A-Z and 0-9). Once it is obtained in alphanumeric form, the license plate character text will be obtained.

Digital image processing of license plate character recognition system is implemented using Matlab R2017a. Figure 3(a) is a GUI (Graphical User Interface) in the Matlab program. System testing scenarios can be seen in table 1(a). Experiments carried out with the amount of data used are 50 vehicle image data for each scenario.

![Table 1](a) Trial Scenario and (b) Accuracy Analysis

| Scenario | Distance | Total image |
|----------|----------|-------------|
| 1        | 50 cm    | 50          |
| 2        | 75 cm    | 50          |
| 3        | 100 cm   | 50          |

| Scenario | Plate Lokalisation | Character Segmentation | Character Recognition |
|----------|--------------------|------------------------|-----------------------|
| 1        | 94.00%             | 70.21%                 | 82.59%                |
| 2        | 78.00%             | 71.79%                 | 79.77%                |
| 3        | 30.00%             | 60.00%                 | 78.04%                |

From each subsequent, scenario calculates plate localization accuracy, segmentation per character and character recognition. After calculating the accuracy of each trial scenario, then it is analyzed from the accuracy produced. The accuracy of the results of testing each image will be averaged so that it gets the average accuracy for each scenario. Table 1(b) shows the results of calculating accuracy.

Graph of average results from calculation of plate localization accuracy, segmentation per character and character recognition can be seen in Figure 3(b).

From the graph of Figure 3(b) it can be seen that:

- The highest plate localization accuracy is generated in scenario 1 of 94%. In scenario 2, the accuracy is 78%, while the lowest accuracy is generated in scenario 3 by 30%. From these data it indicates that the shooting distance is very influential on the results of plate localization which will also affect the process afterwards. The farther the shooting distance, the more difficult the system to detect the license plate location.

- The highest character accuracy of segmentation is in scenario 2 of 71.79%, then in scenario 1 it is 70.21% and the last is 60% in scenario 3. Some things that affect the success of character segmentation when shooting are lighting that is sometimes not sunlight so that the results of the image on the character plate are bright and dim. Another cause is the license plate print that is already blurry. This makes the system fail to read the characters on the vehicle license plate during the character segmentation process.

- The highest character recognition accuracy is found in scenario 1 of 82.59%, then in scenario 2 it is 79.77%, and scenario 3 is 78.04%. The results of character segmentation affect the level of success of the application in character recognition. The clearer the character generated from the results of character segmentation, the higher the accuracy of character recognition produced.
4. Conclusion
After a trial analysis, conclusions that can be taken from the results of the research that have been done are as follows:

1. The best plate localization accuracy is generated in scenario 1 of 94%, while the best character segmentation accuracy is generated in scenario 2 of 71.79%, the best accuracy of character recognition is produced in scenario 1 of 82.59%.
2. The farther the shooting distance, the vehicle license plate is increasingly difficult to detect by the system. The success of character segmentation is influenced by the light and dim lighting when shooting. The presence of noise also affects the level of accuracy. The imperfect character segmentation results in character recognition failure so that the results of character recognition are not appropriate.

References
[1] P. Kulkarni, A. Khandebharad, D. Khope, and P. U. Chavan, “License Plate Recognition: A review,” 2012 Fourth Int. Conf. Adv. Comput., pp. 1–8, 2012.
[2] J. Yepez and S. Ko, “Improved license plate localisation algorithm based on morphological operations,” IET Intelligent Transport Systems, vol. 12, Issues 6, 2018.
[3] M. A. Massoud, M. Sabee, M. Gergais, and R. Bakhit, “Automated new license plate recognition in Egypt,” Alexandria Eng. J., vol. 52, no. 3, pp. 319–326, 2013.
[4] T. Ibrahim, “License Plate Recognition System Using Artificial Neural Networks,” ETRI Jurnal, vol. 39, Issue 2, 2017.
[5] Y. Ma, J. Chi, R. Hu, and G. Yang, “A new algorithm for characters segmentation of license plate based on variance projection and mean filter,” Proc. IEEE 5th Int. Conf. Cybern. Intell. Syst. CIS 2011, pp. 132–135, 2011.
[6] J. G. Park, “An intelligent framework of illumination effects elimination for Car License Plate character segmentation,” Int. Conf. Mach. Learn. Cybern. ICMLC, vol. 3, no. July, pp. 1268–1272, 2010.
[7] P. A. Charde, S. D. Lokhande, ”Classification Using K Nearest Neighbor for Brain Image Retrieval,” International Journal of Scientific & Engineering Research, vol. 4, Issue. 8, pp. 760–765, 2013.
[8] Y. H. Sharath Kumar, N. Manohar, H. K. Chetan, ”Animal Classification System : A Block Based Approach,” Procedia Computer Science, vol. 45, pp. 336-343, 2015.
[9] S. S. Sarikan, A. M. Ozbayoglu, O. Ziceli, ”Automated Vehicle Classification with Image Processing and Computational Intelligence,” Procedia Computer Science, vol. 114, pp. 515-522, 2017.
[10] S. Letkovits, L. Letkovits, ”Distance Based k-NN Classification of Gabor Jet Local Descriptors,” Procedia Technology, vol. 19, pp. 780-785, 2015.
[11] F. A. Mufarohra and F. Utaminingrum, ”Hand Gesture Recognition using Adaptive Network based Fuzzy Inference System and K-Nearest Neighbor,” Int. J. Technol., vol. 3, pp. 1117–1125, 2017.
[12] B. Bhusan, ”License Plate Recognition System using Neural Networks and Multi thresholding Technique,” Int. J. Comput. Appl., vol. 84, no. 5, pp. 45–50, 2013.
[13] W. Jia, H. Zhang, and X. He, ”Region-based license plate detection,” J. Netw. Comput. Appl., vol. 30, no. 4, pp. 1324–1333, 2007.
[14] T. Panchal, H. Patel, and A. Panchal, ”License Plate Detection Using Harris Corner and Character Segmentation by Integrated Approach from an Image,” Procedia Comput. Sci., vol. 79, pp. 419–425, 2016.
[15] V. Tadic, M. Popovic, and P. Odry, ”Fuzzified Gabor filter for license plate detection,” Eng. Appl. Artif. Intell., vol. 48, pp. 40–58, 2016.
[16] M. K. Saini and S. Saini, ”Multiwavelet transform based license plate detection,” J. Vis. Commun. Image Represent., vol. 44, no. January, pp. 128–138, 2017.
[17] C. M. Bishop, Pattern Recognition and Machine Learning. 2006.
[18] F. Camasta and A. Vinciarelli, ”Cursive character recognition by learning vector quantization,” vol. 22, pp. 625–629, 2001.
[19] T. Munisami, M. Ramsurn, S. Kishnah, and S. Pudaruth, ”Plant leaf recognition using shape features and colour histogram with k-nearest neighbor classifiers,” Second Int. Symp. Comput. Vis. Internet, vol. 58, pp. 740–747, 2015.
[20] F. Damayanti, W. Setiawan, S. Herawati and A. Rachmad, ”Comparing The Similarities Measurement Of Face- Expression Based On 2DLDA Modification Method,” Journal Of Theoretical And Applied Information Technology, vol. 94, no. 2, pp. 484 - 489, 2016.