UiTM VEHICLE STICKER SERIAL NUMBER RECOGNITION USING OCR WITH TEMPLATE MATCHING

Nurul Syamimi Mohammod, *Siti 'Aisyah Sa'dan

Faculty of Computer and Mathematical Sciences
Universiti Teknologi MARA (UiTM) Terengganu
23000 Dungun, Terengganu, Malaysia

*Corresponding author's email: sitiaisyah@tganu.uitm.edu.my

Abstract

The level of safety around us has become critical nowadays. There is a need to enhance our security level in order to minimize the risk. In UiTM Terengganu, Polis Bantuan provide vehicle stickers for those who are qualified to bring vehicle to UiTM. These vehicle stickers is being used as an entrance and exit pass. The purpose of this vehicle stickers is to identify the vehicles which are allowed to pass in UiTM. UiTM vehicle sticker serial number (VSSNR) is a system where the serial number is recognized and identified and match to the corresponding plate number to ensure the validity of the sticker and the sticker is genuine. This project focuses on processing the image of UiTM vehicle sticker serial number and is developed to assist Polis Bantuan UiTM in verification of UiTM vehicle sticker. It applies the image processing techniques that manipulate image to text using optical character recogintion (OCR) with template matching. OCR is the process whereby typed or printed pages can be scanned into computer systems, and their contents recognized and converted into machine-readable code. To convert this, several steps are involved which included the process of grayscale, binarizing and slicing the image. This project may ease the Polis Bantuan in performing their task efficiently. In the future work, this project should be implement in mobile application because mobile area is prominent area nowadays.

Keywords: OCR; serial number; template matching

1.0   INTRODUCTION

UiTM established the security unit in 1974 and this unit has continued to safeguard the security of the campus community. The main function of the security unit is to enforce the university's rules and regulations. This unit indirectly responsible for maintaining the security and upholding the reputation of UiTM. This unit have to handle the specific tasks which is traffic section, investigation section, operation section and administration and logistics section (INTEC). To ensure the safety of UiTM Terengganu Campus Dungun, enhancement on verifying UiTM vehicle sticker serial number is required for example, verifying the serial sticker number to ensure sticker validity should be done automatically to save time and cost.

Vehicle sticker is an authorized pass to ensure a vehicle is permitted to enter the campus. There are standard procedure and rules to apply the sticker vehicle. Due to the rigid procedures to ensure the validity of the sticker, few issues arise. One of the problem occurs when the Polis Bantuan are unable to manually inspect each and every sticker on hundreds of vehicles attempt to enter the campus area. This is also time consuming. The situation is denser in the morning where staffs and students are rushing to punch their cards and enter their classes on time. If the sticker checking process is done manually and very particularly every morning, it will cause a huge problem where the vehicles need to queue up at the entrance point and causes road traffic. This could lead to a more risky situation, where irresponsible people might take advantage of this vulnerability.

Image processing is a research to develop the application that can manipulate the image. The data will be extracted from the image during the process of scanning. The ability of computer to read and understand the meaning of the image is different from human being therefore, recognition technique is required to retrieve the data from the image (Hidayatullah, Syakrani, Suhartini, & Muhlis, 2012). In this research, Optical character recognition (OCR) with template matching will be apply on UiTM vehicle sticker serial number to extract the image of the serial number and process it to be readable by the computer. The OCR is a technique of converting image into character from image obtained by scanning a text image into document. OCR has the ability to recognize handwritten or typewritten, but it depends on the quality of that image. OCR has been a well-known technique in area research since the accuracy percentage result obtained by OCR technique is satisfied (Kaur & Banga, 2013).

2.0   RELATED WORK

2.1   Serial number

A serial number is a unique number used for identification and inventory purposes. These serial numbers are similar to the license plate number of vehicles, which represent their unique identification. A serial number allows a company to identify a product and get additional information about it, for replacement, or as a means of finding compatible parts (Computer Hope, 3015).

2.2   Optical Character Recognition (OCR)

The OCR is a recognition method in which the input is an image and the output is string of character. OCR is a process which separates the different characters from each other taken from an image. Template matching is one of the approaches of OCR. The cropped image is compared with the template data stored in database. OCR automatically identifies and recognizes the characters without any indirect input. The characters on the number have uniform fonts and the OCR for number recognition is less complex as compared to other methods. The template matching affects the accuracy of number plate recognition. There are some factors which affect the effectiveness of template matching based on OCR technique i.e. font type, noise in image, and tilting (Computer Hope, 2015).

For the OCR processing for this project, it will be applied also using template matching and colour contra since the serial number for the sticker is black and white. This approach is to recognize the number region for black and white plate number. Therefore, it is suitable for UiTM sticker serial number since it is black and white in colour. After number region of interest is extracted, the next process is comparing the neighbouring pixels horizontally and vertically, the summation of difference between neighbouring pixels for each column and row is computed (Simin, Choong, & Mei, 2013). Table 1 shows the different methods and approach in applying the OCR.

Table 1 Method of OCR

| Methods	Description	Disadvantages
Edge detection	High magnitude in edges of number plate (Yasin, Abdullah, & Karim 2009).	Background interference creates high magnitude in edges.
Color contrast	Black and white number plate has high color contrast	Not applicable for number plates with colors
Correlation	Straight forward and reliable	- Not tolerant to font difference and tilt - Huge storage is needed to store all templates
Structure analysis	Tolerant to font difference and tilt	Complex analysis is needed to differentiate all characters
Neural network	Tolerant to font difference and tilt	High investment to train system

3.0 METHODOLOGY

Research methodology is the analysis of method, rules and data that have been collected and needed for the project research. Generally, it is the way to systematically solve the research problem by adopting several steps. It defines an instruction of what the activity research is, way of proceed the project and way to measure progress. In addition, identifying the model or framework used is also being identified in research methodology. Figure 1 is the research framework of this project.
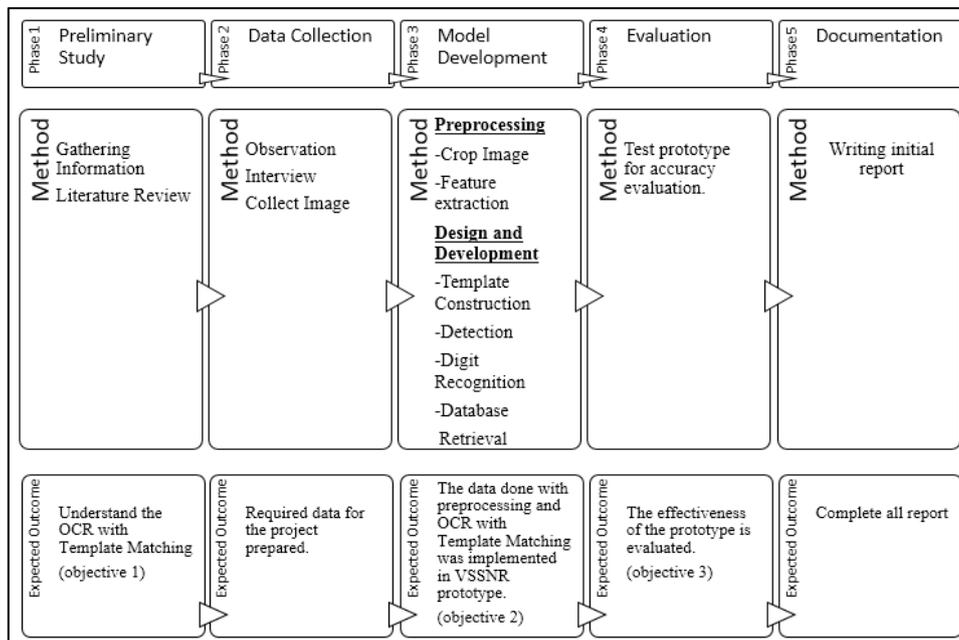

Figure 1 Research Framework

System design is a process of defining the architecture, components, modules, interfaces, and data for a system to satisfy specified requirements. The input and output will be determine based on data collection. Then the process is proceed to designing the interface. Figure 2 show the process design of the project. This process involves the use of a database of characters or templates and there is a template for all possible input characters. For recognition to occur, the current input character is compared to each template to find the exact match or the closest representation of the input character.
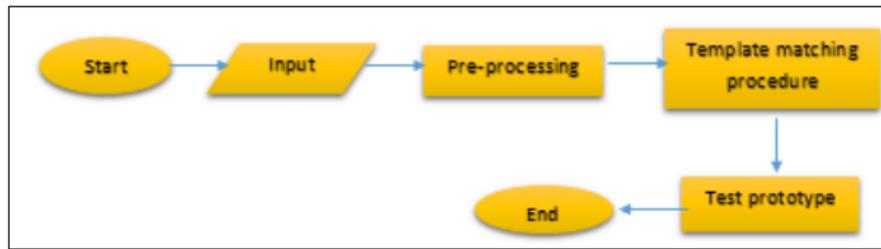
Figure 2 Project Design

The method can only be successful if the input character can match the templates of the same or similar font. Template matching can be performed on binary, threshold characters or on grey-level characters. For grey-level characters, it is more common for Normalized Correlation to be used as it provides improved resistance to variations in OCR method with the template matching. After serial number image is converted into binary, the process is proceed to comparing between input images and template images to find the best match. The character's identity is assigned to be the one with the most similar template. The highest similarity will considered as the character. Figure 3 show the template matching algorithm implements the following step.

1) The character image from the detected string is selected using the previous pre-processing.
2) The input character image size is rescale with the first template.
3) After rescale the input image to the size of the first template, the matching metric is computed.
4) Then, the highest match found is stored. If the image is not match repeat again the third step.
5) The index of the best match is stored as the recognized character.
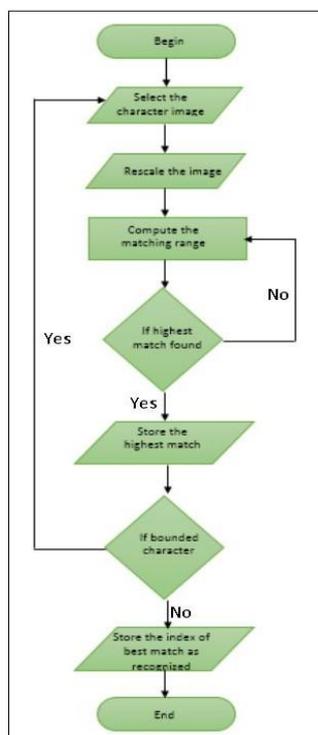
Figure 3 Template Matching Algorithm

The next step is the evaluation. The main purpose of testing the project is to make sure that the objective three that is to evaluate the effectiveness of the prototype is achieved and also to trace the error. If error is detected in the project, then correcting and debugging the error is required. From the result of testing, percentage of the correct and incorrect output will be calculated. The result will identify the system performance.

4.0   RESULT AND DISSCUSSION

The results include in any outcomes such as in diagram, table, interfaces, analysis and evaluation. The results is on the ability of UiTM vehicle sticker serial number recognition using OCR with template matching system.

OCR with the template matching method is one of the effective ways to achieve the classification of discrete input mode. Its essence is to measure a certain similarity between the input mode and the sample, then take maximum similarity as the category of the input mode. OCR with template matching is the straightforward method to recognize object or characters and digits. Therefore understand how this technique implement in the program is important.

After going through the pre-processing of the data, the data will then be recognized. The prototype of this project will recognize the serial number from the input image that had been cropped. The detail regarding the vehicle and the owner will be display. The detail will be display based on recognized output. If the process of recognized produce the incorrect result, the detail regarding owner and vehicle displayed will be incorrect.

Based on image 25 images tested, 18 is accurate, and 7 is inaccurate. The space between each number contribute to the incorrect output. If the serial number too close among the other number, the prototype will recognize as a number. To evaluate the effectiveness of the prototype, the accuracy have to be analysed. The accuracy of the prototype is 72% as shown in Figure 4.
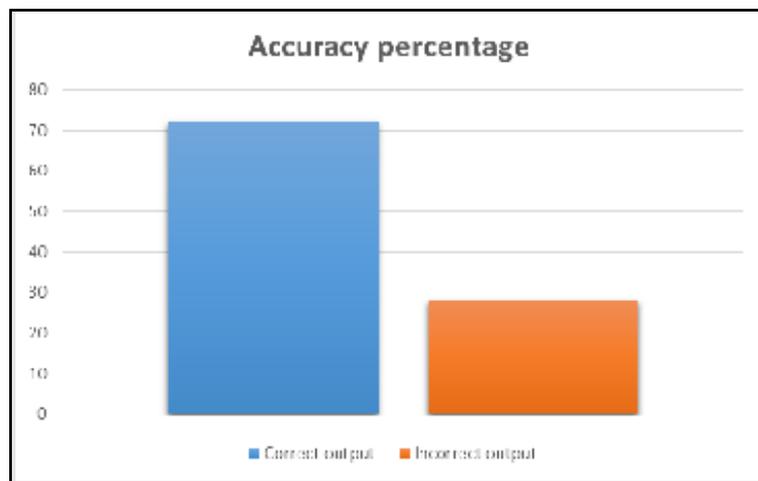

Figure 4 Accuracy Percentage

5.0   CONCLUSION

After the prototype is developed, the ability to recognize numbers can be tested. It begins with the input image so that the interface can display the required image, then the image have to pre-process, which is convert the image to greyscale, binary, and remove the noise and unrequired area by

cropping method. The process of recognizing the serial number is taking place after the pre-processing successfully done. The output from recognized serial number will be compared with the serial number in the database. The detail of vehicle and owner will be display if the serial number has recorded in the database. In case the serial number does not contain in the database, then the validity of the sticker should be hesitate. The main idea about this research is to show OCR with template matching to recognize the characters and digits. From many research made, OCR with template matching is one of the well-known study in image processing filed.

6.0   FUTURE WORK

For future works, the study may use different angles of images, capture the image in the different lighting condition and at a different distance. In UiTM Terengganu, the time taken should be reduced in order to increase the security level at UiTM Terengganu and to make the traffic running smoothly. This prototype also should be implement in mobile field and make it more user friendly and attractive. So that in the future, it is better if all the varieties are used in the research. A new method for recognizing the vehicle serial number might be expected for next for appears as the technology and demand increase.

References

Banga, V. K. & Kaur, E. K. (2013). Number plate recognition using OCR technique. *International Journal of Research in Engineering and Technology, 2*(9), 286–290. Retrieved from http://esatjournals.net/ijret/2013v02/i09/IJRET20130209042.pdf

Computer Hope. (2015). Serial number. Retrieved from http://www.computerhope.com/jargon/s/serinumb.htm.

Hidayatullah P., Syakrani N., Suhartini I., & Muhlis, W. (2012). Optical character recognition improvement for license plate recognition in Indonesia. Paper presented at 2012 UKSim-AMSS 6th European Modelling Symposium. pp. 249–254. doi: 10.1109/EMS.2012.95

Yasin A.M., Abdullah A.S., & Karim, M. R. (2009). Travel time measurement in real-time using automatic mumber plate recognition for Malaysian enviroment. *Journal of the Eastern Asia Society for Transportation Studies, 8,* 1738-1751.

Simin N., Choong F., & Mei, C. (2013). Automatic car-plate detection and recognition system. Paper presented at Engineering Undergraduate Research Catalyst Conference (EURECA). pp. 113–114. Retrieved from http://www.taylors.edu.my/EURECA/2014/downloads/41.pdf