TELKOMNIKA, Vol.17, No.2, April 2019, pp.620~627
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v17i2.9017

Indonesian license plate recognition based on area feature extraction

Fitri Damayanti[1], Sri Herawati[2], Imamah[3], Fifin Ayu M[4], Aeri Rachmad*[5]
Faculty of Engineering, University of Trunojoyo Madura, Indonesia
Raya Telang Street PO BOX 02 Kamal-Bangkalan, 69192, Indonesia
*Coresponding author, e-mail: fitri2708@yahoo.com[1], aery_r@yahoo.com[5]

Abstract

The main principle of license plate recognition is to recognize the characters in the license plate which indicates the identity of the vehicle. This research will provide a system which can be implemented to the automatic payment in highway. Indonesian license plate consists of two parts, every of which has certain characters. These characters may become problem in the recognition process. Another problem is on the type of the license plate since Indonesia applies different color for every type of vehicle. In this research, different approaches are employed in the recognition of license plate; that is using character area as the feature value, also known as feature area, and K-Nearest Neighbor (KNN) as classification method. In addition, another method that has been used in our previous research is also employed to detect the character of license plate. The result shows very significant accuracy of 99.44%. In the process of recognition, scenario 1 gives the best accuracy at the K-1 value; that is 68.57% on the license plate and 92.72% on the characters of license plate. In the scenario 2 was obtained the license plate accuracy of 52% and license plate character accuracy of 89.36% with K-5. The system ran in a relatively short computational time.

Keywords: *area feature extraction, character recognition of license plate, K-Nearest Neighbor, license plate recognition*

1. Introduction

License plate is an identity for a vehicle which indicates that the vehicle has been registered. With this identity the ownership of the vehicle can be identified. The system of license plate in Indonesia is based on the residency area in which the vehicle is registered and on the registration order. Some countries have many versions on the form of license plate, for example Tunis [1], China [2], Libya [3], and India [4]. These countries have one-line license plate. It is different from Indonesia which applies two-line license plate. The first line contains one or two letters denoting the area registration, one to four digit numbers, and one or two letters denoting the region the vehicle registered. The second line consists of four digits indicating information about month and year when the plate will expire. Besides the form, Indonesia also applies color to the system of license plate, i.e. private vehicle (black), government vehicle (red), public vehicle (yellow), and dealer vehicle (white).

This research is aimed to create a license plate recognition system. This system later can be used as the reference to identify vehicles. The challenge in the process of license plate identification is how to detect and recognize the license plate of vehicles in different form and type like the one in Indonesia. Some previous research in recognition of Indonesian license plate have been carried out [5–7]. A research [5] employing Fourier transform method and hidden markov model obtained an accuracy of 84.38%. Another research [6] resulted accuracy value of 85% on the character recognition and of 97% on the process of detecting the location of license plate. This research uses contour in detecting the location. The process continues by doing figure segmentation utilizing interconnected figure segmentation, and is ended in the process of classification by using static classification technique by which the segmented figure is translated into the character of ASCII. However, a problem aroused when a character cannot be recognized because of its different shape or italic shape.

A high degree of accuracy, that is 96%, is gained in the other previous research [7]. This research implemented a method of discrete cosine transform to extract feature and radial basis function as the classification method. Yet, this research was conducted only on the recognition of one type of license plate, i.e. black license plate (private vehicle).

Received February 22, 2018; Revised November 27, 2018; Accepted December 29, 2018

Considering the above problem, we develop a system to solve the problem related to the system of Indonesia's license plate. The system developed is expected to be able to recognize all types of license plate of vehicles in Indonesia. Moreover, this research will create a system which is able to run in a short computational time with a high degree of accuracy. To gain the two objectives of the research, area feature is chosen as feature extraction method; while K-nearest neighbor, which has been used in some previous research [8–12], is implemented to recognize license plate.

2. Research Method

In order to develop a system of license plate recognition, the first step is identifying the characters on the plate. They are in the forms of letters A to Z and digits 0 to 9. Figure 1 [13] presents the forms and types of license plate used in this research.

(a) (b) (c) (d)

Figure 1. Vehicle license plate in Indonesia (a) private vehicle, (b) government vehicle, (c) public vehicle, dan (d) dealer vehicle [13]

After identifying the characters, the next steps are resizing the image, converting the image, detecting the characters of the vehicle's license plate, and the last is recognizing the characters of the plate as shown in Figure 2. The process begins by resizing the images of the plates. The real image is resized into 300x720 pixels in order to ease the next processes. Conversion process is done by converting the captured image which is in the form of RGB into grayscale image, as the following (1)

$$G = (0.2989 \times R) + (0.5870 \times G) + (0.1140 \times B) \quad (1)$$

The process continues detecting the plate by identifying the characters on the plate. By using area feature extraction the characters identified are processed to get the feature value. Lastly, the recognition process is done by utilizing K-NN method to recognize the plate based on the feature value obtained.

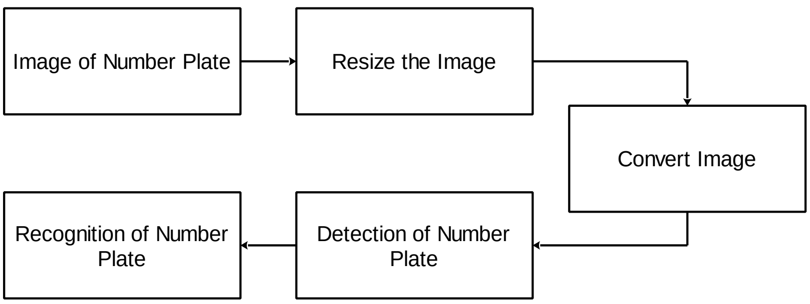


Figure 2. System design

3. Results and Discussion

3.1. Detection of License Plate

Detecting license plate is done to separate the object (the character) from its background and capture the characters which are the vehicle's license number on the first line representing the area registration, license number, and region. This process aims to get the characters on the plate, which can be done using some techniques [14–17]. Based on the types of the plate which have different color of background and characters, a new approach has been developed [14] by utilizing area and ratio image. The approach is developed after conducting an analysis on the characters of the plate. The process of detecting the character of the license plate is explained in Figure 3.

The process of detecting characters begins by processing the grayscale image obtained from the previous process. Then, the process continues converting the grayscale image into two kinds, i.e. binary image and complement drawing. It is done in order to get a white pixel value of 1 on the two images since the plates have different types as shown in Figure 1. The next is adding up the pixel value of 1 on the two images. To get the character on the plate, selection process on the two images is carried out, thus there is only one image selected. This process is done by taking the lowest total pixel value between the two images. The process of selection is illustrated in Figure 4 [14].

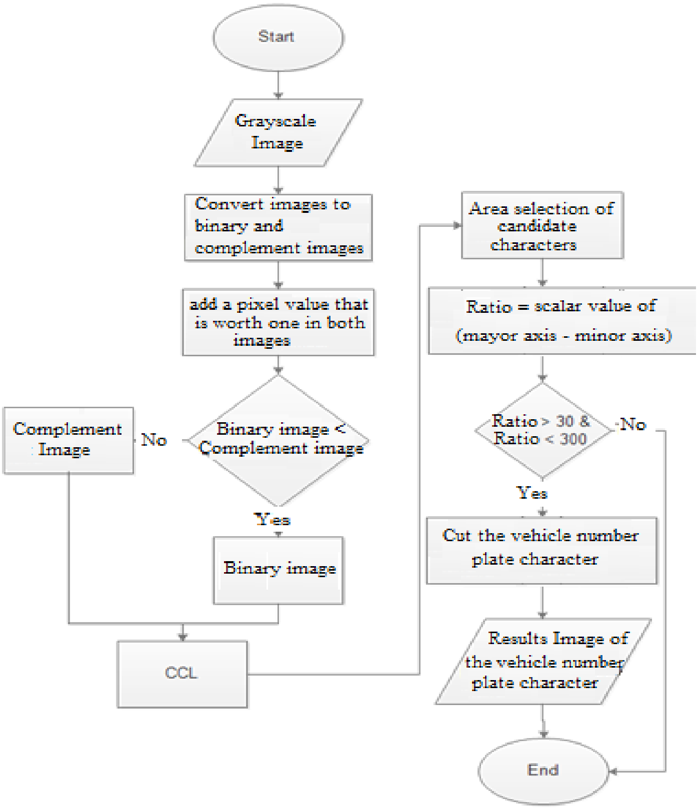


Figure 3. The process of detecting the character of the license plate [13]

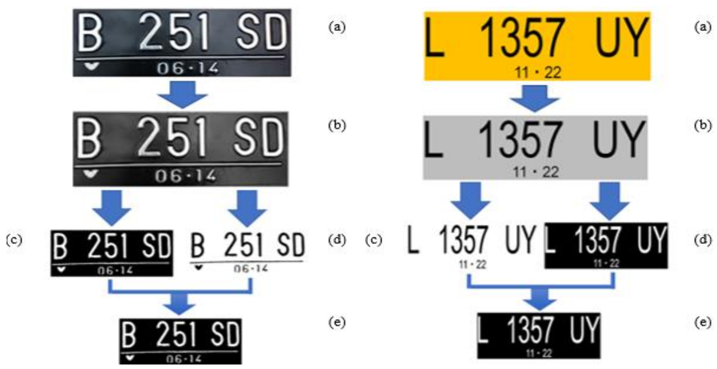


Figure 4. Selection of license plate, (a) original images, (b) grayscale images, (c) binary images, (d) complement drawings, (e) selected image [13]

Then, method of connected component is applied as an algorithm to detect interconnected region blob on a binary image. A previous research by Arai and Tolle [18] uses connected component as a method to detect text in comics. The next process is can it is obtain dilate selection which is the character of the plate by using area of blob. To determine the blob area threshold value of 1500 is used. This process will select when a blob area is higher than or equal to 1500 pixels, then the area is detected as the character, as (2) below:

$$blob_i = \begin{cases} Character & area_i \geq 1500 \\ NotCharacter & area_i < 1500 \end{cases} \quad (2)$$

From the candidate selection process by using area of blob region, the character of the plate is obtained. Nevertheless, a problem aroused, as illustrated in Figure 1 (a), where there is a horizontal line detected as the character for its blob area value $\geq$1500. Regarding to the problem, a calculation on image ratio is carried out. The image ration is gained from the difference of scalar value of major axis length and scalar value of minor axis. From that ratio value, to determine the candidate character a condition is given on the ratio value of the image, like the (3). This process is to eliminate another object whose shape is not like the object or the character of license plate which is in a rectangle.

$$ratio = \begin{cases} Character & ratio > 30 \,\&\, ratio < 300 \\ NotCharacter & otherwise \end{cases} \quad (3)$$

Detected object as the character will be analyzed to find out the area value surrounding the rectangle object, called as bounding box. Then, cutting process is done based on the area of bounding box; that is the area surrounded by four coordinate points ((x_{min},y_{min}), (x_{min},y_{max}), (x_{max},y_{min}), (x_{max},y_{max})). Figure 5 shows the detecting process done started from grayscale process to segmenting license plate's character.

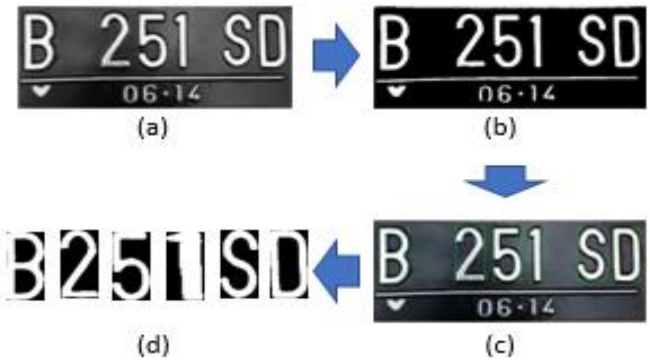


Figure 5. Result of detection process of license plate (a) grayscaleimage, (b) image of selection, (c) detection of character, (d) character of license plate [13]

3.2. Recognition of License Plate

The last process is recognition. The character obtained in the process of detecting the license plate is used as the input. The steps can be seen in Figure 6. Firstly, the process is done by resizing the character into 200x100 pixels, in that the size of every character is same in order to make the process of recognition easy. After that, feature extraction process is carried out to get feature value of the character. By considering computational time, the area of character figure is utilized as the reference in determining feature extraction method. The first step is that from the real image the character image is divided into 10x10 areas (10 rowsx10 columns) as shown in Figure 7, and thus in each area is obtained 20x10 pixel. Then, because the image processed is binary image, addition process is done by adding the value of 1 of each area. The process of saving feature is by placing the value of every area orderly started from area $X_1, X_2, X_3, X_4, \dots\dots X_{100}$. Hence, feature value of 100 is gained in every character as shown in Figure 8 [19].

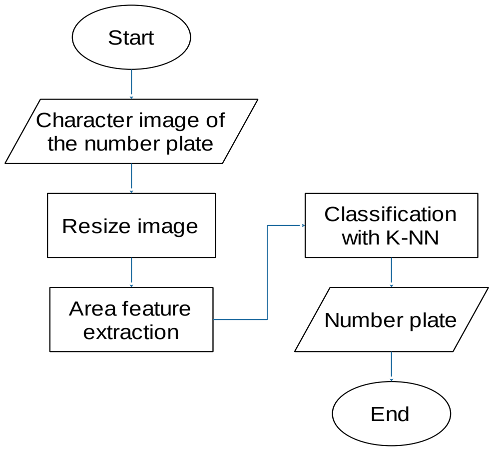


Figure 6. Recognition process

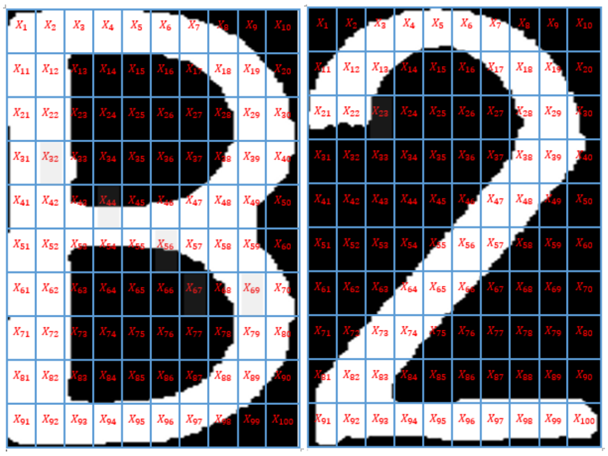
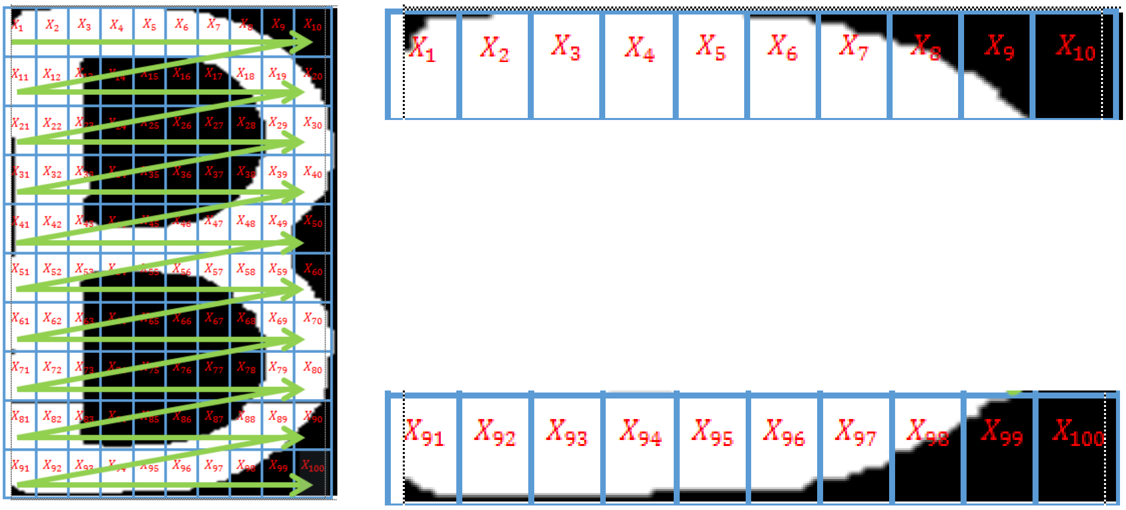


Figure 7. Features with 10x10 areas. (a) the character of the letter B, (b) the character of license



(a) (b)

Figure 8. Feature saving (a) illustration of area-based feature caption, (b) area-based feature value placement in saving process

The classification process was done from the feature gained in the previous process. In the recognition process the method of K-Nearest Neighbor (KNN) was employed. KNN [20] is a learning method that is good and easily used to solve problem of recognition. Similar to the existing learning method [21-23], KNN works by comparing the feature obtained, (in this research it is feature resulted from license plate character extraction) with (training data) to get the shortest distance. Before the classification process, we did a learning process by training 10 images in every alphabet and in every number, thereby the overall training data was as many as 360 data. Based on the training data, in the process of K-NN we searched the shortest distance between the 360 training data and the trial data. The point K has a function as the class determinant based on the average of the shortest classes. The measurement of the shortest distance used Euclidean distance with (4) [24].

$$d(x,y)=\sqrt{\sum_{i=1}^{i=n}(x_i-y_i)^2} \tag{4}$$

3.3. Analysis Trials

The trials was done by testing 50 data taken randomly. The first trial was detecting the character on the license plate by using a method that has been used previously [13]. As soon as the result was obtained, the next trial was done by using 2 scenarios; as follows:

a) Scenario 1, detected character was processed by using are feature extraction, i.e. 10x10 areas. Classification process used K-NN by applying K-values of 1, of 3, and of 5.
b) Scenario 2, detected character was processed by using are feature extraction, i.e. 20x20 areas. Classification process used K-NN by applying K-values of 1, of 3, and of 5.

The result of detecting process on 50 trial data which were taken randomly from all types of plate as shown in Figure 1 shows accuracy of 99.44%. From all characters on 50 data trial, error occurred only on detecting 2 characters. The error on the first character was that it was not matched with the defined ratio in the detection process, so the character was undetected. This error made the character fail to be recognized in the recognition process. The second error occurred when a noise was detected as a character. It happened because the noise, whose condition matched with the defined condition, passed the blob selection as well as the ratio process. Hence, it can be said that the accuracy obtained by our system was very high.

The result of recognition of the license plate and the plate’s character indicates that two scenarios give different results. In the scenario 1 as shown in Figure 8, the accuracy of the license plate did not change; that is 60% accuracy; whereas the accuracy of the character underwent a reduction when K-value increased. In the scenario 2 as shown in Figure 9, the accuracy of license plate and of the character increased when K-value increased. From the two results, we can conclude that the feature extraction process caused the different accuracy result of the two scenarios. The more we divide the area in the feature extraction process, the higher accuracy cause is when K-value is added. The best accuracy was gained in scenario 1 despite the fact that this scenario does not affect the accuracy gain of the license plate, even decreases the accuracy of the plate’s character.

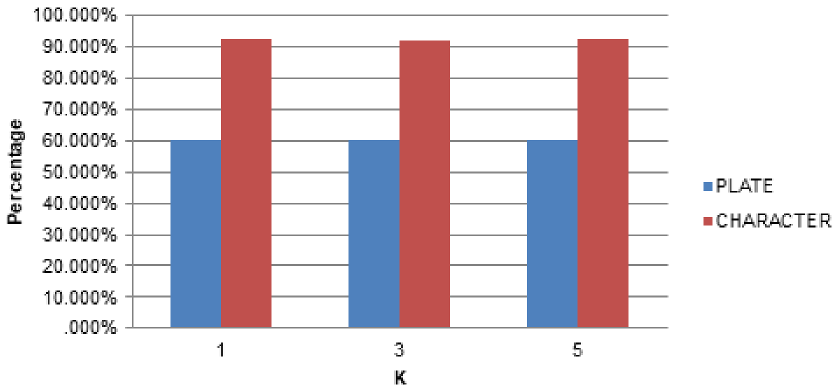


Figure 8. Recognition results of scenario 1

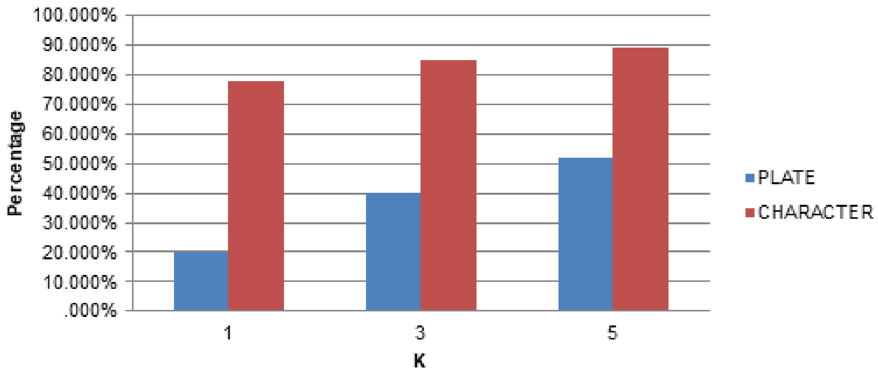


Figure 9. Recognition results of scenario 2

The detailed result is presented in Tables 1 and 2. From Table 1 it can be seen that the best result is K-1, whose accuracy is 92.72%. In Table 2, the best result is K-5 with the accuracy of 89.36%. Seen from the computational time needed during the recognition process, scenario 1 takes shorter time that scenario 2 does, that is 0.616125 second dan 1.307945 second respectively as shown in Figure 10.

Table 1. Recognition Results of Scenario 1

The value of K	Amount of License Plate	Recognized of License Plate	Accuracy of License Plate	Number of Characters	True Character	False Character	Accuracy of Character
1	50	30	60%	357	331	26	92.72%
3	50	30	60%	357	329	28	92.16%
5	50	30	60%	357	330	27	92.44%

Table 2. Recognition Results of Scenario 2

The value of K	Amount of License Plate	Recognized of License Plate	Accuracy of License Plate	Number of Characters	True Character	False Character	Accuracy Of Character
1	50	10	20%	357	278	79	77.87%
3	50	20	40%	357	304	53	85.15%
5	50	26	52%	357	319	38	89.36%

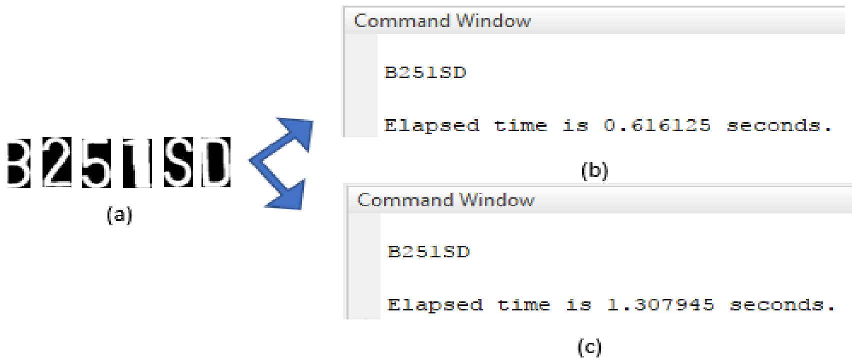


Figure 10. The result of the process of recognizing the character of the vehicle license pelate (a) features 10x10 area (b) features 20x20 area

Table 3 presents the comparison of the trial result using the method applied in this research and the trial result using other method in the previous research. In the recognition of the plate's character the method we propose is better than the method of feature from interconnected image applied in the research [6]. Despite the fact that the accuracy we gain in this research cannot exceed the accuracy gained in the research [7] which applied DCT, we applied the method to all types of vehicle license plate; while the research [7] applied its method only to one type of license plate as shown in Figure 1.

Table 3. Comparison of Trial Results

Methods	Accuracy
Feature from interconnected image segmentation [6]	Character : 85% Localization plate (Detection) : 97%
Discrete Cosine Transform [7]	Character : 96%
Area Feature Extraction (Research conducted)	Character : 92.72% Detection : 99.44% License Plate : 60%

4. Conclusion

From the research that has been done, the method of recognizing the character of vehicle license plate is able to recognize the character very well which is proven with the result of accuracy reaching 92.72%. The feature and value areas of K, both have a role in recognizing the character of the vehicle license plate so it is necessary to test the feature area and the value of K to get the most accuracy. In this study, the computational time required by the system to recognize the character of the average vehicle license plate is 0.62 seconds in the area of 10x10 and 1.31 seconds in the 20x20 area. The computational time difference is caused by the number of feature values on each processed image, thus causing the 20x20 area to be longer than the 10x10 area.

References

[1] Ktata S, Khadhraoui T, Benzarti F, Amiri H. Tunisian License Plate Number Recognition. *Procedia Computer Science.* 2015; 73: 312–319.
[2] Tian J, Wang R, Wang G, Liu J, Xia Y. A two-stage character segmentation method for Chinese license plate. *Computers & Electrical Engineering.* 2015; 46(C): 539-553.
[3] Ansari NN, Singh AK, Student MT. License Number Plate Recognition using Template Matching. *International Journal of Computer Trends and Technology (IJCTT)*. 2016; 35(4): 175–178.
[4] Kodabagi MM, Kanavi VS. License Plate Recognition System For Indian Vehicles. *International Journal of Advanced Research in Engineering and Technology*. 2013; 4(2): 295–304.
[5] Hermawati FA, Koesdijarto R. A Real-Time License Plate Detection System For Parking Access. *TELKOMNIKA Telecommunication Computing Electronics and Control*. 2010; 8(2): 97–106.
[6] Tedjojuwono SM. Fast Performance Indonesian Automatic License Plate Recognition Algorithm Using Interconnected Image Segmentation. In: Intan R, Chi CH, Palit H, Santoso L. *Editors*. Intelligence in the Era of Big Data. Communications in Computer and Information Science. 2015: 516.
[7] Yugopuspito P, Lukas S, Sutrisno D. Krisnadi. Recognition of Indonesian Vehicle Registration Plate by Discrete Cosine Transform and Radial Basis Function Network. *International Journal of Information and Electronics Engineering*. 2015; 5(4): 275-279.
[8] Charde PA, Lokhande SD. Classification Using K-Nearest Neighbor for Brain Image Retrieval. *International Journal of Scientific & Engineering Research*. 2013; 4(8): 760-765.
[9] Sharath Kumar YH, Manohar N, Chetan HK. Animal Classification System: A Block Based Approach. *Procedia Computer Science*. 2015; 45: 336-343.
[10] Sarikan SS, Ozbayoglu AM, Zilci O. Automated Vehicle Classification with Image Processing and Computational Intelligence. *Procedia Computer Science*. 2017; 114: 515-522.
[11] Lefkovits S, Lefkovits L. Distance Based k-NN Classification of Gabor Jet Local Descriptors. *Procedia Technology*. 2015; 19: 780-785.
[12] Bhuvaneswan P, Therese AB. Detection of Cancer in Lung with K-NN Classification Using Genetic Algorithm. *Procedia Materials Science*. 2015; 10: 433-440.
[13] Damayanti F, Herawati S, Setiawan W, Rachmad A. Segmenting the Character of Plate Vehicle Number Using Connected Component Analysis Method. *Int. Journal of Engineering Research and Application.* 2017; 7: 78-82.
[14] Jia W, Zhang H, He X. Region-based license plate detection. *Journal of Network and computer Applications*. 2007; 30(4): 1324–1333.
[15] Panchal T, Patel H, Panchal A. License Plate Detection Using Harris Corner and Character Segmentation by Integrated Approach from an Image. *Procedia Computer Science.* 2016; 79: 419–425.
[16] Tadic V, Popovic M, Odry P. Fuzzified Gabor filter for license pelate detection. *Engineering Applications of Artificial Intelligence*. 2016; 48: 40–58.
[17] Saini MK, Saini S. Multiwavelet transform based license plate detection. *Journal of Visual Communication and Image Representation*. 2017; 44: 128–138.
[18] Arai K, Tolle H. Automatic E-Comic Content Adaptation. *International Journal of Ubiquitous Computing*. 2010; 1(1): 1–11.
[19] Mufarroha FA, Utaminingrum F. Hand Gesture Recognition using Adaptive Network based Fuzzy Inference System and K-Nearest Neighbor. *International Journal of Technology*. 2017; 8(3).
[20] Bishop CM. *Pattern Recognition and Machine Learning*. 2006.
[21] Camastra F, Vinciarelli A. Cursive character recognition by learning vector quantization. *Pattern Recognition Letters*. 2001; 22: 625–629.
[22] Munisami T, Ramsurn M, Kishnah S, Pudaruth S. *Plant leaf recognition using shape features and colour histogram with k-nearest neighbor classifiers*. Second Int. Symp. Comput. Vis. Internet. 2015; 58: 740–747.
[23] Damayanti F, Setiawan W, Herawati S, Rachmad A. Comparing The Similarities Measurement Of Face- Expression Based On 2DLDA Modification Method. *Journal of Theoretical and Applied Information Technology*. 2016; 94(2): 484 - 489.
[24] Massoud MA, Sabee M, Gergais M, Bakhit R. Automated New License Pelate Recognition In Egypt. *Alexandria Engineering Journal*. 2013; 52: 319-326.