

Automatic Iranian Vehicle License Plate Recognition System Based on Support Vector Machine (SVM) Algorithms

Mahdi Aghaie¹, Fatemeh Shokri^{2*}, Meisam Yadolah Zade Tabari³

¹Department of Computer & IT Engineering, Mazandaran Institute of technology
Babol, P.O. Box 47166-96519, IRAN

²Department of Computer & IT Engineering, Mazandaran Institute of technology
Babol, P.O. Box 47166-96519, IRAN – (Corresponding Author) Email: fatemeh.shokri@aol.com

³Lecturer, Department of Computer & IT Engineering, Mazandaran Institute of technology
Babol, P.O. Box 47166-96519, IRAN

ABSTRAKSI

Ada jauh lebih banyak mobil di jalan sekarang daripada dulu. Oleh karena itu, mengontrol dan mengelola lalu lintas besar hampir tidak mungkin tanpa menggunakan teknologi komputer. Makalah ini menyajikan merancang dan mengimplementasikan suatu sistem cerdas untuk pengenalan plat berdasarkan tiga langkah proses. Proses - proses ini meliputi deteksi lokasi plat, segmentasi karakter dan pengenalan karakter. Dalam studi ini, kami menggunakan pengklasifikasian svm untuk mendeteksi karakter. Menurut hasil, kinerja dari sistem yang diusulkan lebih baik dibandingkan dengan algoritma yang sama seperti jaringan syaraf. Perlu disebutkan bahwa pendekatan pengakuan diuji dalam berbagai kondisi dan hasil dijelaskan.

Kata Kunci: pengenalan plat kendaraan bermotor, operasi morfologi, Histogram, batas deteksi, pengklasifikasian svm

ABSTRACT

There are far more cars on the road now than there used to be. Therefore, controlling and managing the huge volume of traffic is virtually impossible without the use of computer technology. This paper represents design and implement of an intelligent system for license plate recognition based on three main steps. This process includes the detection of license plate location, character segmentation and character recognition. In this study, we used Classifier svm to detect the characters. According to the results, the performance of the proposed system is better compared to similar algorithms such as neural network. It is worth mentioning that recognition approach is tested in various conditions and results are described.

Keyword: Vehicle license plate recognition, Morphology Operations, Histogram, The edge detection, Classifier svm.

1. Introduction:

Nowadays, the increasing number of Vehicles in cities has caused problems in the field of traffic management, parking, and toll ways. It is evident that human cannot solve problem alone. As a consequence, we need an intelligent system to recognize the numbers and alphabet of plate. Several methods of image processing are used to extract the license plate. Among these methods, we can mention the use of the color property [2], the edge detection [3], Morphology Operators [4], Hough transform [5], Fuzzy Logic [6]. There is no doubt that each method has its strengths and weaknesses. For example, in the method of the color property when lighting conditions changes, color loses its stability, and it changes. More over, it takes a large time for colored license plate compared to gray ones [2].

On the other hand, Edge detection methods cannot extract candidate regions. Because it is unable to correctly differentiate the license plate regions from other candidates [3]. The advantage of this method is its speed.

Proposed intelligent system for License Plate Recognition (LPR) has three phases which are the recognition of license plate location, character segmentation and character recognition. The correct character recognition is considered as the most crucial step in the ALPR system since there are problems such as poor image quality, image perspective distortion. In this paper, we focus most of their attention on this sector.

Our proposed system in this paper is the combination of edge detection, histogram analysis, and morphological operations. This structure based on researches has been conducted, has more suitable processing time compared to the methods mentioned above.

The remainder of the paper is organized as follows. Section 2 brief descriptions are given about the available methods in recognition license plate. Section 3 presents the proposed system for the detection of license plate location. Section 4 discusses character separation. Section 5 explains character recognition with aid of Classifier svm. Section 6 discussed implementing of this method; it also reviews the results of the implementation of proposed algorithm on images. Finally, a conclusion and discussion are presented in Section 7.

2. Available methods in the field of License Plate Recognition

Current use of methods in the field of the License Plate Recognition depends on the environmental conditions such as light, orientation, noise that could lead to deficiency. Other problems of these methods can be mentioned are the high cost of computation and selection of many candidate regions on the image in the process of Localization. For example, the data obtained from Histogram of the image for finding candidate regions, has the large number of the candidate regions. Furthermore, the vertical mask of Sobel is just used in edge detection, which

has poor performance for skewed plate. In [7], multi-layer Feed Forward Neural Network is used that has its disadvantages such as slow processing time and limit of the license plate size. This method can be implemented only on the plate with particular color. Another point of consideration about this procedure is it is only used on a vehicle without the complex background.

The method which is proposed in [8], among other methods, can be notable. In this method, Histogram equalization is utilized to find the threshold value in order to enhance the image quality of license plate. Difference between pixels in each region is calculated, and morph metric techniques are used to smooth edges of shapes. Then, the number of pixel points and the aspect ratio of width and length and the number of black pixel points in the image are calculated to eliminate pseudo license plate areas. Finally, plaque image is returned as output. Drawback of this method is that the location of plate should be considered in the middle of image. Otherwise, plate localization is not easily extracted.

In [9], statistical methods and the edges detection of the shapes have been used to find plaque location. Disadvantages of this approach are that, firstly, only one plate is extracted from an image. Secondly, the extraction of plate from images with complex background is quite difficult.

3. Description of Proposed System

In the proposed method, texture information based on the gray-level images is adopted for License Plate Recognition. This approach enables us to identify the plate license with different colors, and it is more resistant to changes in lighting conditions. Different stages of our proposed system are illustrated in Figure 1.

Mahdi Aghaie, Fatemeh Shokri, Meisam Yadolah Zade Tabari
Automatic Iranian Vehicle License Plate Recognition System Based
on Support Vector Machine (svm) Algorithms

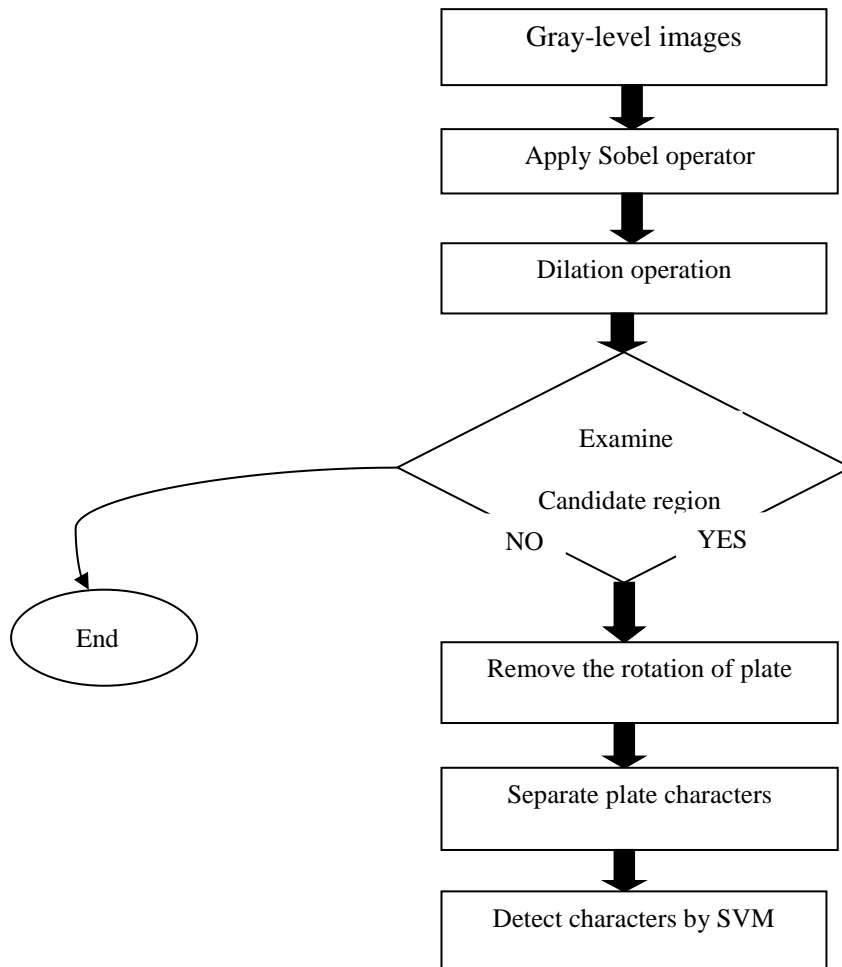


Figure 1: Proposed algorithm in plate recognition system

As can be seen from figure1,the input image is converted to gray-level images. After applying Sobel Operator, the image edge information is obtained. Then, morphological operations are applied to join the image edges in order to create plate candidate regions. These regions which have been identified will be filled with morphology operations. In the next step, candidate regions have been studied with an accurate threshold level, and non-candidate areas will be deleted. Following that, the plate is examined in terms of degree of rotation using the Hough Transform. If a rotation is too much, rotation adjustment will be done. Then plate characters are separated, with the aid of region growing and labeling techniques. Finally, plate characters are recognized using svm Classifier.

3.1. Edge information extraction

After obtaining data from conducting research on new license plates in Iran, it is clear that this type of plate due to the use of Persian numbers in a dark color with the bright background, edge information seems very strong in plate location.

Various methods and algorithms are proposed to extract the edge information in image processing that the Sobel operators have been more considered among them because of the high speed and low volume processing in the same technique. The Sobel operators use two horizontal and vertical masks for edge detection which are shown in Figure 2.

$$\begin{array}{c}
 \left(\begin{array}{ccc} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{array} \right) \\
 A
 \end{array}
 \qquad
 \begin{array}{c}
 \left(\begin{array}{ccc} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{array} \right) \\
 B
 \end{array}$$

Figure 2. horizontal and vertical masks

The majority of available methods use only the vertical mask for edge detection because vertical edges have further information compared to the horizontal edges. Thus, using one of these masks makes processing speed higher. The demerit of this method is when the input image orientations are too much, perform poorly. The proposed method used both vertical and horizontal mask for edge detection.

3.2. Morphological operations to find region candidate

After obtaining the edge information from Dilation, obtained edges are connected to achieve the range of region candidate.

According to this assumption, with the extension of the A and ingredient B, dilation is defined by Equation (1).

$$1. \quad A \oplus B = \{x \mid [(\hat{B})_x \cap A] \subseteq A\}$$

Extension of A with B is the set of all X which the $(\hat{B})_X$ and A could overlapt least one non-zero elements.

3.3. Filling holes

Since plate region is a connected area, the potential holes in the previous stage are filled using morphological operations.

3.4. License Plate Image Extraction

As it is expected from the previous stage, there are plate candidates in several connected region. To identify plate, the length of connected regions is gained. Then, by considering the two threshold levels T1 (minimum area) and T2 (maximum area), the regions which Smaller than T1 and greater thanT2 are removed and the remaining regions are returned as a plate.

3.5. Rotation Adjustment of the plate candidate

One of the common problems in Intelligent Plate Recognition System is rotation of images which makes recognition of characters on the plate difficult. To solve this problem, the Hough transform is used to smooth the images. Hough transform can be described as a transformation of a point in the x-y plane to the parameter space.

This conversion in the parameter space is defined by the following equation:

$$2. \quad P = x.\cos \theta + y.\sin \theta$$

The Hough transform is a method that, in theory, can be used to find features of any shape in an image. In practice it is only generally used for finding straight lines. Hough transform processes the whole picture in the four directions left, right, up and down and if it has any slight deviation from vertical and horizontal surfaces, image is given the same amount of the rotation in the opposite direction. In this paper, threshold value is considered 3 degree for this operation. In this case for every rotation is less than this amount, there is no need for rotation adjustments.

4. Characters Segmentation

Before character segmentation, the gray-scale image can be converted into a binary image. The obtained binary image has a lot of noise during labeling, are returned as plate characters. To eliminate noise, the region growing and labeling techniques are utilized.

First, binary image of candidate plate should be labeled. Then, using region growing techniques, area of the labeled regions is calculated. Regions whose ratio is less than a threshold T3 or greater than a threshold T4 are removed. The ranges of thresholds T3 and T4 have been selected by trial and error and by learning about the range of region each plate characters.

In this step, image would have tiny noise that is not removed in the previous step. To remove this noise, Persian license plate, which has a certain area, is calculated. Then, with use of the growing techniques that we mentioned above, the image candidate is scrolled. If the regions were equal with the calculated area, they would be isolated and stored as plate characters.

5. Detecting characters based on SVM Classifier

Now that the plate characters have been isolated, it is time to recognize the characters. In this paper, Support Vector Machines Classifier (SVM) is used for detecting the characters. SVM is known for its success in handwritten character recognition equals neural networks which are carefully adjusted. SVM is a linear separator. In SVM data can be seen in the form of P-dimensional vector. SVM separates points with the P-1 hyper plane. What SVM sets it apart from other

separators is how to choose Separator hyper plane. SVM is intended to maximize the margin between the two classes. As result, it chooses a hyper plane that its distance from Linear Separator to the nearest data point on each side is maximized. If there is such a hyper plane, it will be known as maximum-margin hyper plate. Figure 3 illustrates this concept.

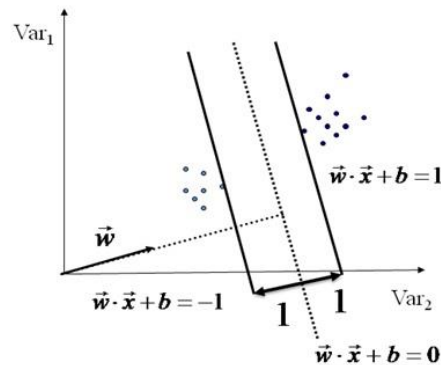


Figure 3: Linear Separator

To construct the maximum margin, two parallel boundary planes and separator plane are drawn. We separate them until they encounter the data. A separator plane that is the farthest from the boundary planes would be the best separator.

Decision function is determined for separating data with a subset of training data which is called the support vectors. In fact, the optimal hyper plane in SVM is separator between the support vectors.

In case of proper use of SVM, this algorithm has good generalization ability. Despite the large size, it avoids from over fitting. Due to use of support vectors instead of all the data, this algorithm also performs data compression. Consider a set of points in the following form:

$$\{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{1, -1\}\} i = 1, \dots, n$$

According to equation, y_i Can be 1 or -1, and it determines that point x_i belongs to what class. Each x_i is a p-dimensional vector.

We want to find the maximum margin Separator. Each hyper plate can be a set of points x that satisfy the following relation which is written below:

$$3. \quad W \cdot X - b = 0$$

The vector w is a normal vector. This vector is perpendicular to the hyper plate. Parameter b/w determine the offset of the hyper plane from the origin along the

normal vector W . Thus, W and b should be chosen till the distance between parallel hyper plates to be maximized. Moreover, parallel hyper plate shave to be far apart from each other. As a consequence, the data can be separated.

These parallel hyper plates can be described by the following equation:

$$W \cdot X + b \geq 1 \quad \forall x_i \text{ with } y_i = 1$$

$$W \cdot X + b \leq -1 \quad \forall x_i \text{ with } y_i = -1$$

In this regard, if the training data can be linearly separated, so two Separator plates can be selected that any data is not among them. And then try to maximize their distance. Data sets that are linearly separable with some noise work out great. But if the Data sets are very complicated, they cannot be simply classified by a line, to address these cases, mapped into a higher dimensional space can be used for classification. SVM algorithm has a high efficiency that makes application very suitable for data classification.

6. Implementation

In many cases, there are problems such as poor image quality, image perspective distortion, other disturbance characters or reflection on vehicle surface, and the color similarity between the license plate and the background vehicle body, the license plate is often difficult to be located accurately and efficiently. To overcome this problem, a database full of dirty and damaged plates is created. The desired Classifier is trained to deal with.

The images are captured in different positions and different angles such as: Left, right and direct, near and far distance and different lighting conditions, like day, night. Furthermore, it includes plates which are dirty and their characters extremely have been corrupted.

One of the notable advantages of this method over conventional methods Such as neural networks MLP, RBF, HOPFIELD is that is exempt from extraction of image feature. This makes the characters easy to identify. In the first stage, the isolated plate characters are converted into the same size. Then, the plate characters are selected in 5 different positions from the opposite perspective, the left perspective with angles greater than 30 degrees, right perspective with angles greater than 30 degrees, damaged characters and characters with different lighting conditions. Database built for the training algorithm for each character is consists of 30 photos. That is a total of 240 images taken by a digital camera. Some of these images can be seen in Figure 4.



Figure 4. a number of samples image

For example, the proposed method was applied to one of the images which are in the database and it can be seen from Figure 4, and the result is shown in Figure 5. The final results also are shown in Table 1. Due to drop accuracy, the infrared camera was used because there is not good and adequate lighting at night. This method has been applied to the target database.

Results of accurate license plate localization and the accurate characters identification are shown in Table 1 and Table 2 respectively. As it can be seen from tables, the proposed method has high efficiency both in terms of plate recognition and character detection.

Mahdi Aghaie, Fatemeh Shokri, Meisam Yadolah Zade Tabari
 Automatic Iranian Vehicle License Plate Recognition System Based
 on Support Vector Machine (svm) Algorithms



D) Image after Morphology Operations

E) Image of Isolated plate

F) Plate image after rotation



G) Image after removing additional regions

H) Binary image

I) Image after removing noise



L) Characters segmentation.

Figure 5 the implementation of the proposed algorithm

Table 1. Implementation results of the proposed method

The total number of experimental data	Total=300	The number of accurate localization	Percentage of efficiency
Plates with rotation from left to right	(40%) 120	107	89/16%
Photographed Plates from an direct angle	(20%) 60	56	93/34%
Photographed Plates with more than 3 meters	(13/34%) 40	34	85%
Plates were taken at night by infrared cameras	(12/34%) 37	30	81/08%
dirty or damaged Plates	(14/34%) 43	34	79/06%

Table 2. Implementation results of the proposed method

The total number of experimental data	Total=240	The number of accurate localization	Percentage of efficiency
Plate image without distance	(73/75%)177	170	96/04%
dirty or damaged Plate image	(13/75%) 33	25	75/75%
Plates image with more than 3 meters	(12/5%) 30	28	93/4%

According to the Table 1, the percentage of plate *recognition* that is directly photographed is approximately 93/34%. In comparison with the method proposed in [1] and [2] are respectively 81/3 % and 82/44 %. It is clear that there is a significant improvement. The percentage of efficiency for proposed method, providing plates photographed from opposite angle, is 96.04% given in Table 2, Compared to

Mahdi Aghaie, Fatemeh Shokri, Meisam Yadolah Zade Tabari
Automatic Iranian Vehicle License Plate Recognition System Based
on Support Vector Machine (svm) Algorithms

methods presented in [3] and [4], which account for about 93% and 91% respectively that have higher efficiency.

7. Conclusion

In this paper, it is intended to propose a method for Iranian license plate recognition. Proposed method uses the Sobel mask followed by morphological operations for LPR. Furthermore, SVM Classifier is utilized in order to recognize characters and numbers. Proposed method was applied to a large number of plates with different angles and distances, different lighting conditions and when the plate is smeared with mud or the characters are damaged. The results show a significant improvement over previous methods.

8 – Reference

- [1] Farhad Faraji, Reza safe bakhsh, "Novel approach to license plate recognition system based on Morphology Operations" National Conference on Signal Processing, Iran, February 2007.
- [2] X.Shi, W.Zhao, Y.Shen, "Automatic license plate recognition system based on color image processing", In lecture notes on computer science, Vol.3483, O.Gervasi, et al, Eds. New York: Springer-Verlag, pp.1159– 1168, 2005.
- [3] W.Jia, H.Zhang, X.He, M.Piccardi, "Mean shift for accurate license plate localization", IEEE/ITS, pp.566-570, 2005.
- [4] D.Zheng, Y.Zhao, J.Wang, "An efficient method of license plate location", Pattern recognit. Lett., Vol.26, No.15, pp.2431–2438, Nov. 2005.
- [5] T.D.Duan, T.L.Hong Du, T.V.Phuoc, N.V.Hoang, "Building an automatic vehicle license plate recognition system", Int. Conf. Comp. Sci. RIVF, pp.59–63, 2005.
- [6] J.A.G.Nijhuis, M.H.ter Brugge, K.A.Helmholt, J.P.W.Pluim, L.Spaanenburg, R.S.Venema, M.A.Westenberg, "Car license plate recognition with neural networks and fuzzy logic", IEEE /ICNN, Vol.5, pp.2232-2236, 1995.
- [7] Sung Han Park, Kwang In Kim, Keechul Jung, and Hyung Jin Kim, "Locating Car License Plates Using Neural Network", IEEE, Vol.35, no.17, Aug 1999.
- [8] D.S.Gao and J.Zhou, "Car License Plates Detection from Complex Scene", International Conference on Signal Processing, Beijing, August 2000.

- [9] I. Maro, M. Chacon, and S. Alejandro Zimmerman, "License Plate Location based on a Dynamic PCNN Scheme", IEEE International Symposium on Computational Intelligence in Robotics and Automation, July 2003.

Mahdi Aghaie, Fatemeh Shokri, Meisam Yadolah Zade Tabari
Automatic Iranian Vehicle License Plate Recognition System Based
on Support Vector Machine (svm) Algorithms