Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 102 (2016) 588 – 594

Procedia Computer Science

12th International Conference on Application of Fuzzy Systems and Soft Computing, ICAFS 2016, 29-30 August 2016, Vienna, Austria

A hybrid KNN-SVM model for Iranian license plate recognition

Sahar S. Tabrizi[a]*, Nadire Cavus[a]

[a]*Department of Computer Information Systems, Near East University, P.O.BOX:99138, Nicosia, North Cyprus, Mersin 10, Turkey.*

Abstract

This study presents a new method for Iranian License plate recognition systems that will increase the accuracy and decrease the costs of the recognition phase of these systems. In this regard, ahybrid of the k-Nearest Neighbors algorithmand the Multi-Class Support Vector Machines (KNN-SVM) model was developedin the study. K-NN was used as the first classification model as it is simple, robust against noisy data set and effective fora large data set. The confusion among the license plate similar characters problem was overcome by using the multiple SVMs classification model. The SVMs model has improved the performance of the K-NN in the recognition of similar characters. The current study experimental results revealed that there is a significant improvement in the character recognition phase rate compared with a similar study.

Peer-review under responsibility of the Organizing Committee of ICAFS 2016

*Keywords:*KNN-SVM Hybrid Model; Machine Learning; Image Processing; Iranian License Plate Recognition; Feature Extraction;

1. Introduction

Lookingat our daily lives, we canfind the footprints of improving computer science and engineering in every aspect of life such as education, tourism, health and transportation. One of the most rapidly developing areas in the field of engineering is that of Intelligent Transformation System (ITS)[1]. ITS, as an active research area, has started to play an important role in people's lives such as in transport and mobility safety[2]. This phenomenon has encouraged governments and the private sector alike to exploit advanced technologies[3]. Satellite system navigation[4,5,6], road sign automatic recognition[7,8,9] guided parking systems[10,11,12], license plate recognition[13,14,15] and vehicular security[1] are some of the various ITS types.

* Sahar S. Tabrizi; Tel.: +90-392-675-1000. *E-mail address:* sahar.shokouhi@neu.edu.tr

Peer-review under responsibility of the Organizing Committee of ICAFS 2016
doi:10.1016/j.procs.2016.09.447

License Plate Recognition (LPR) plays a crucial role in ITSand hasmainly been used for access control systems. The control systems can be applied in various applications such as automatic parking control, security control in restricted areas, traffic law monitoring, congestion pricing[13], the search for stolen cars or unpaid fees [14,1] and freeway payment[3].

The current paper is organized into five sections.Related works willbe reviewed in Section 2. Afterward, the aim of the study will beexplained in Section 3. The next section discusses the study framework. Experimental results willbe discussed in Section 5 and the last section reveals the conclusion of the study.

2. Related work

Over the last several decades, many research studies have covered the subjectof license plate recognition techniques,wheremost of them were based on grayscale method image processing, character extraction and pattern recognition. Park et al. [15] applied an artificial neural network (ANN) to recognize and extract Korean license platesaccurately, even in bad weather. The Matas et al.[16], study was similar to thePark et al.[11]study. Viand et al. [17] and Kim et al.[18]implemented the Support Vector Machine (SVM) algorithm to determine the license plate. Naito et al.[19] focused on the use of camerasas animportant aspect, in order to obtain abetter image. Cowell et al.[20] explained Arabic and Latin characters' recognition. Yu.M and Kim Y.D [21]underlined various approaches for colorful background license extraction. Chang et al. [22] used fuzzy logic sets to determine plate positions in both vertical and horizontal dimensions. Delforouzi and Pooyan[23]applied bothSVM and OCR to recognize efficiently Iranian license plates with Persian (Farsi) characters. A literature review showed classical image processing, genetic algorithms, artificial neural networks, Fuzzy, Support Vector Machine (SVM), MARKOR Processes, Finite, etc. techniques were all applied in LPR processes.

3. Aim of the Study

The recognition of Iranian License Plates on privately owned cars is the aim of the study. Confusion caused by similarities between the Persian alphabet and numeric characters is one of the problems in the recognition phase ofPersian LPR systems. In this regard, using the benefitsof SVM based on KNN as a Hybrid model was presented in the study to overcome the above-mentioned problem.

4. The Study Frame Work

The main focus of the current is on Iranian license plates on privately owned cars, which consist of a metal plate with standard alphanumeric format and white background, as shown in Fig. 1. This plate consists of 13 lettersand 9 numbers[24] (Table 1).

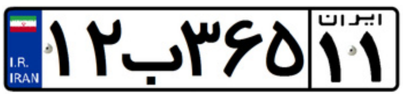

Fig.1. Iranian car license plate sample[24]

Table 1. Licence Plate Letter and Number List.

Letters	Numbers
[illegible]	1 2 3 4 5 6 7 8 9

The system input is an image that is taken by camera in advance. The input image format would be JPG, JEPG, and BMP. This recognition system applies the photos one at a time. In the next step, the inserted image is sent to the image processing phase. The results of step one will be the segmented plates' characters. Afterward, each segmented

character images is converted to the relevant feature vector. These feature vectors are the input data for the recognition phase. Fig. 2. explains the block diagram of the study.

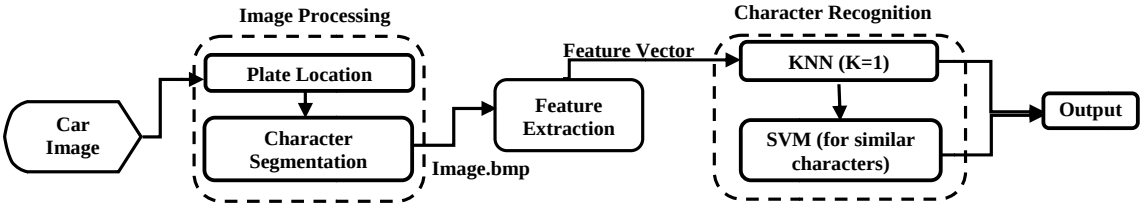

Fig. 2.The Study Block Diagram

4.1. Image processing

As seen in Fig.2, theimage processing phase includes two main sections: plate location identification and character segmentation. In order to obtain the desired result in the study, Morphology was conducted, as it is a robust technique[25,26,27,28]. After converting the colorful images to Grayscale format, edge detection (Prewitt) and Dilate image functions, which are offered by Morphology, are applied. Filling (imfill) and filtering (liner) functions help to fill the images' missed pixels, which are called holes, from the previous step and to delete thenoise in both the vertical and horizontal directions. The final format fromthese processes, as seen in Fig. 3,isconverted to the labeling matrix. The label matrix assigns different values for each object in order to identify them.

In order to identify the license plate location on the image a "for loop" is appliedafter matrix labeling. The spot is deleted if one of the below conditions is TRUE. The extracted license plate is shown in Fig .3(g).

- The width of the spot is less than 3 points;
- The width of the spot is morethan 1/4 of the input image width;
- The height of the spot is lessthan 8 points;
- The height of the spot is morethan 1/3 of the input image height;
- The ratio of the spot's width to its height is less than 0,1;
- The ratio of the spot's width to its height is morethan 4 (because of two or three joined characters);
- The ratio of the spots area to the spot's bounding box area is lessthan 0,15[29];

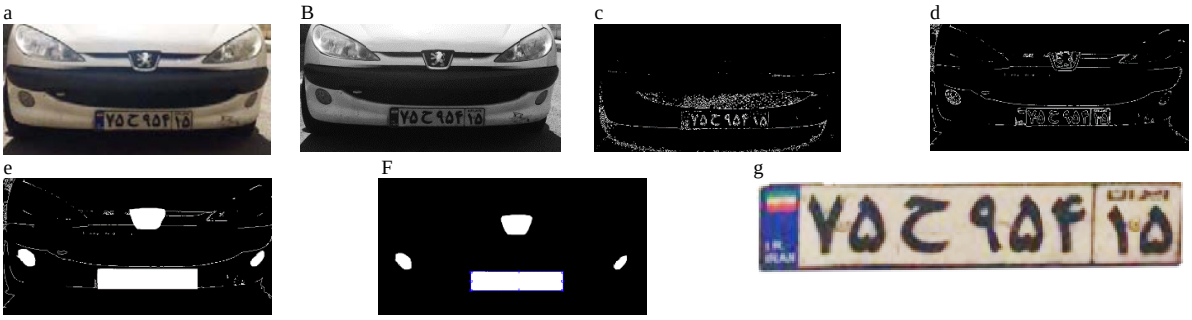

Fig. 3. The License Plate Extraction Phase of the Study (a) Inserted image; (b)Greyscale; (c)Prewitt; (d) Dilation operation;(e) Label matrix;(f) Extracting license plate; (g) Show the license plate.

License plate alphanumeric character segmentation is carried out after extracting the license plate from theinserted cars' image. In this regard, all image processing Morphology steps are applied to segment the license

plates' characters. Furthermore, the license plate rotation over the horizontal vector is applied in advance. Fig.4. shows the steps for character segmentation.

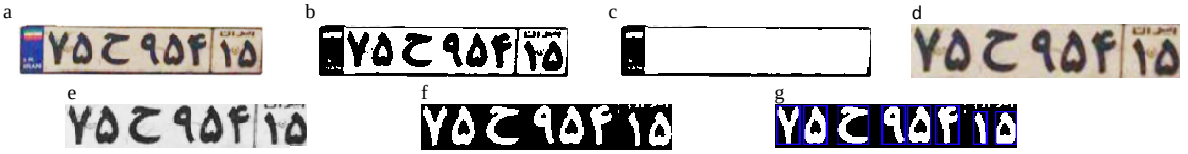

Fig. 4. Character Segmentation Steps of the Study (a) Rotation; (b) Binary; (c) Filling;
(d) Matrix Labelling; (e) Brighten & crop; (f) Grey-scale; (g) Character segmentation and normalization.

4.2. Feature Extraction

Each of the segmented charactersof the license plate which were normalized (32x30 pixel) before are converted to the feature vector in order to feed into the character recognition phase. In order to identify the license plate characters, a hybrid machine was applied in this study; therefore, each machine (KNN and SVM) needs specific feature vectors. Literature reviews reveled that two main features are common in the feature extractions phase for Persian and Arabic based alphanumeric dissimilar characters recognition studies. In this regard, in order to feed and create the KNN feature vector to recognize dissimilar characters (Table2),structural feature and horizontal (H) and vertical (V) crossing count histogram features[1,23] are applied for the KNN machine. The similar characters(Table 2) that were recognized via the SVMsneed another feature vector. The Zoning feature extraction method is applied toeach similar character (Table 2). The similar images are divided into 12 equal zones (8x10 pixels). The Zoning[36] features are extracted from each zone pixels by moving along the diagonals of its respective 10 x 10 pixels. Each divided zone has 17 diagonal lines.In total, 19 featuresare extracted through the Zoning technique, which is used in The SVM. Table 2 explains all the details about both feature vectors.

Table 2. Feature Vector attributes.

Machine	Feature set	Size of feature vector	Total
KNN	Structural features	16	26
	(H&V) Crossing count	10	
SVM	Zoning	19	19

4.3. Character recognition

There are some common problems in the recognition phase of LPR systems. Noisy, dusty or low quality image problems are generallyovercome during the image processing phases. Similarities between the Persian alphabetic and numeric charactersthatare used in the Iranian license plate cause classification confusion problems during the recognition phase. Recognitionof each segmented alphanumeric character thatis already converted to the related feature vector is the last phase of the current study. Previous studieshave shownthat the ANN, K-NN, SVM, Fuzzy logics, etc. methods have beenapplied in LPR systems' character recognition phases. The Hybrid KNN-SVMs model was utilizedin the current study. The main reason behind the choice of this model was the corresponding properties, which use benefits of the combination of K-NN and SVM to decrease the training phase time and increase the efficiency of the LPR system. The K-NN was used as the initial step toclassifyall data set without training[30] and then multiple class SVMs was performed on only the smaller data set with similar characters. Fig.5.explains the proposed hybrid KNN-SVM model of the current study.

- K-NN

Each plates' characters with the specificstandards has been segmented and resized in the image processing steps. Afterwards, in the feature extraction process, the segmented characters imagesareconverted to the relevant feature vectors.The standard format for license plate uses dissimilar alphanumeric characters. As a simple classifier[30] method, which works well to classifier anomalous data set [31,] ,K-NN was implementedin this study as the first classification step. Furthermore, the K-NN algorithm was chosen because of itszero learning phase costs [32,31], simple implementation[30]and analytical tractability. Thus, K-NN classifies the sample data set for each letter and number

character with minimaltraining costs. The Jaccard and Euclidean distance functions yields revealed that the Euclidean distance is suitable for the study. The results of the KNN classifications show that K-NN is still confused and cannot classify accurately the characters thatare slightlysimilar to each other, as seen in Table 3.

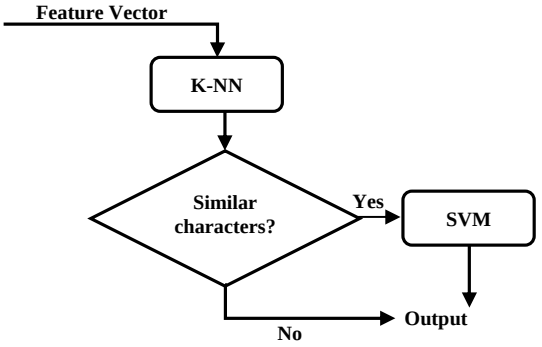

Fig. 5. The Hybrid KNN-SVM Proposed Architecture

Table 3.Similar Characters

Characters types	Confused character
Similar letters	G1(ص, س, ق) G2(ن, ی)
Similar numbers	G1(۲, ۳, ۴)

- SVM

The multiple SVMs classification and recognition model is applied when the K-NN in the first classification step finds one of the similar characters in the recognition phase (Table 3). The SVMs trains only the similar characters sample data sets. Thus, the training rate of the SVM decreases significantly. Various kernel functions test results revealed that the best SVMs function for the current study is the RBF kernel, similar to previous[33,34,35]studies**.**

5. Experimental Results of the Hybrid KNN-SVM

To evaluate the effectiveness of the proposed method in this paper, 257 images of car license plate have been tested as part of the study. These license plate images were taken randomly by normal digital camera with a 1024*768 pixel frame size, in RGB and JPG format. The camera was in the range of 5 to 10 m. from the car. The captured images were divided into2 groups: *T1*- Plate with at least one similar character (n=137) and *T2*- Plate without similar characters (n=120). The experimental results of each group are shown in Table 4.

Table 4. Experimental Tests Results

Applied Test	T1(n=137)	T2(n=120)	Total (n=257)
Plate region extraction	96.01		96.01
Character segmentation	95.24		95.24
Character recognition	96.82	97.31	97.03

Common failures in Plate region extraction and Character segmentation in both groups are caused by characters with dots and image capturing insufficiencies. The K-NN (K=1) with Euclidean distance functions was applied in the first phase of classification and the multiclass SVMs with RBF kernel was conducted as an improved phase. Comparing the character recognition phase results of the current study with asimilar study[23], we can statethat,althoughthe character recognition phase rate of the study is not faster than in the previous study[23], the accuracyof recognition is higher(Table 5). In addition, the lower cost of the proposed system's training phase is incomparable with other models.

Table 5. Feature Vector attributes.

Studies	Feature extraction methods	Models	%
Current Study	KNN "Structural features, (H&V) Crossing count" SVM "Zoning"	KNN-SVM	97.03
Similar	OCR	SVM	94

6. Conclusion

In this study, a new method is presented to apply the LPR systems for Iranian license plates. Increasingthe accuracy of the character recognition phase rate and decreasing the training rate are the main advantages of the new Hybrid model. The K-NN was implemented as the first classification method, as it is simple, robust against a noisy data set and effective in large data sets with zero training cost. The confusion problem related tosimilar characters in the license plates was overcome by using the multiple SVMs classification model. The SVMs has improved the performance of the K-NN in the recognition of similar characters. The SVMs was trained and tested only for the similar characters, thus, the training cost of the SVMs decreased significantly. Comparison resultsbetweenthe current study experimental results of a similar study[23]revealed that that the presented hybrid KNN-SVM model improves the character recognition rate significantly from 94% to 97.03% for all cases tested.

References

1. Nguwi YY, LimWJ.Number plate recognition in noisy image. In.*Proc. 8th International Congress on Image and Signal Processing*;2015. p.476-480.
2. Shaw C.Driving ambition.*New Electronics*2010; 43(18):18-20.
3.Nikolaos E, Ioannis E, Loumos V, Kayafas E. A License Plate-Recognition algorithm for intelligent transportation system applications. *IEEE Transactions on Intelligent Transportation Systems* 2006; 7(3):377-392.
4.Xiaolei Y, Yongrong S, Jianye L, Ling W, Yuan C, Jianfeng M. Key techniques for multi-satellite integrated navigation system modeling and controlling.In. *Proc. 2nd Systems and Control in Aerospace and Astronautics*; 2008. p.1-6.
5.Jianjun Z, Yukui H.Research on mutual interference evaluation method of global navigation satellite system.In.*Proc. International Conference on Information and Automation. IEEE;* 2013. p.291-296.
6.QiaoL, Liu J, Zheng G, Xiong Z.Augmentation of XNAV system to an ultraviolet Sensor-Based satellite navigation system.*J Selected Topics in Signal Processing, IEEE*2009;3(5):777-785.
7. Nguwi YY, Kouzani A.Detection and classification of road signs in natural environments. *Neural Computation and Applications*2008;17(3):265-289.
8.Nguwi YY, Cho SY.Two-tier self-organizing visual model for road sign recognition.In. *Proc. International Joint Conference on Neural Networks, IEEE*;2008. p.1-8.
9. NguwiYY, Kouzani AZ.Automatic road sign recognition using neural networks. In. *Proc. International Joint Conference on Neural Networks*; 2006. p.3955-3962.
10. Lee ChH,Leibiao Lim MS.The optimum configuration of car parking guide system based on wireless sensor network. In. *Proc. International Symposium on Industrial ElectronicsIEEE*;2009. p.1199-1202.
11.Jiang K, Seneviratne LD.A sensor guided autonomous parking system for nonholonomic mobile robots.In. *Proc. International Conference on Robotics and Automation IEEE*; 1999.p.311-316.
12. Rajabioun T, Foster B, Ioannou P. Intelligent parking assist.In.*Proc. 21st Mediterranean Conference Control & Automation*;2013. p.1156-1161.
13. Chang S,Chen L, Chung Y, Chen SW. Automatic license plate recognition.*IEEE Transactions on Intelligent Transportation Systems*2004;5(1):42-53.
14. MohadesKasaei H, MohadesKasaei M. Extraction and recognition of the vehicle license plate for passing under outside environment. In. *Proc. European Intelligence and Security Informatics Conference*; 2011.p.234-237.
15. Park SH, Kim KI, Jung K, Kim HI. Locating car license plates using neural network.*IEEE Electronics Letters*1999; 35(17):1475-1477.
16. Matas J, Zimmermann K. Unconstrained licence plate and text localization and recognition. In.*Proc. Intelligent Transportation Systems, IEEE*; 2005. p.225-230.
17. ViandAS, Seyedjavadi SHH, Rahmani AM.Enhancing automatic speed estimation systems performance using support vector machines.In *Proc. 5th International Conference of Intelligent Computer Communication, IEEE*;2009.p.185-188.
18. Kim KI, JungK, Kim J.Color texture-based object detection: An application to licence plate localization. In Lecture Notes on Computer Science, vol. 2388, S.-W. Lee and A. Verri, Eds. New York: Springer-Verlag; 2002, p.293-309.
19. Naito T, Tsukada T, Yamada K, Kozuka K, Yamamoto S. Robust license-plate recognition method for passing vehicles under outside environment.*IEEE Trans. Veh. Technol.*2000;**49**:2309-2319.

20. CowellJ, Hussain F. A fast recognition system for isolatedArabic characters.In.*Proc. 6tl' international conference on Information and Visualization, IEEE*; 2002, p.650-654.
21. YuM, Kim YD. An approach to Korean licence plate recognition based on vertical edge matching.In *Proc.International conference on Systems, Man and Cybernetics*;2001. p.2975-2980.
22. Chang SL, Chen LSo, ChungYC, Chen SW.Automatic license plate recognition.*IEEE Transactions on Intelligent Transportation Systems* 2004;**5**(1):42-53.
23. Delforouzi A, Pooyan M. Efficient Farsi license plate recognition. In.*Proc. 7th international Conference on Information, Communications and Signal Processing, IEEE*; 2009. p.1-5.
24. Vehicle registration plates of Iran. <https: //en.wikipedia.org/wiki/Vehicle_registration_plates_of_Iran/>; 2016 [accessed 12.05.16].
25. Sulehria HK, Zhang Y, Irfan D. Mathematical Morphology methodology for extraction of vehicle number plates. *Int J Computers* 2007;3(1):70-73.
26. Bhardwaj D, Mahajan S. Review Paper on Automated Number Plate Recognition Techniques.*IntJ Emerging Research in Management &Technology* 2015;4(5):319-324.
27. Yan-qing Z, Dai-ping L, Shu-wen Z. License plate location based on combinatorial feature. In. *Proc. Microwaves, Radar and Remote Sensing Symposium, IEEE*; 2011. p.165-168.
28. Nguwi YY, Lim WJ.Number plate recognition in noisy image. In.*Proc. 8th International Congress on Image and Signal, IEEE*; 2015. p.476-480.
29. KwaśnickaH, Wawrzyniak B. License plate localization and recognition in camera pictures. In.*Proc. 3th Symposium on Methods of Artificial Intelligence*; 2002. p.243-246.
30. Bhatia N, Vandana. Survey of Nearest Neighbour techniques.*Int J Computer Science and Information Security* 2010;8(2):302-305.
31.Imandoust SB, Bolandraftar M. Application of K-Nearest Neighbor (KNN) approach for predicting economic events: theoretical background.*Int J Engineering Research and Application* 2013;3(5):605-610.
32. Kim J, Kim BS, Savarese S. Comparing image classification methods:K-Nearest-Neighbour and Support-Vector-Machines. *In. Proc. 6th WSEAS international conference on Computer Engineering and Applications*; 2012. p.133-138.
33. Mozaffari S, Parnia B. Farsi/Arabic handwritten from Machine-Printed words discrimination.In.*Proc. International Conference on Frontiers in Handwriting Recognition, IEEE*; 2012.p.698-703.
34. Zanchettin C, Bezerra BLD, Azevedo WW. A KNN-SVM hybrid model for cursive handwriting recognition. In.*Proc. International Joint Conference on Neural Networks, IEEE*; 2012. p.1-8.
35. Karamizadeh S, Abdullah SM, Halimi M, Shayan J, Rajabi MJ. Advantage and drawback of Support Vector Machine functionality.In.*Proc. International Conference on Computer, Communications, and Control Technology, IEEE*; 2014. p.63-65.
36. Kumar G, Bhatia PK. A Detailed Review of Feature Extraction in Image Processing Systems.In.*Proc. 4thInternational Conference on Advanced Computing & Communication Technologies, IEEE*; 2014. p.5-12.