A Novel Method for Persian Handwritten Digit Recognition using Support Vector Machines

Mojtaba Mohammadpoor[1*] Abbas Mehdizadeh[2], Hava Alizadeh Noghabi[1]
1- Department of Electrical and Computer Engineering, University of Gonabad, Gonabad, Iran.
Email: mohammadpur@gonabad.ac.ir (Corresponding author)
Email: alizadehn_h@yahoo.com
2- Department of Computing, Nilai University, Negeri Sembilan, Malaysia.
Email: mehdiizadeh@ieee.org

Received: July 2017 Revised: September 2017 Accepted: December 2017

ABSTRACT:
Handwritten digit recognition has got a special role in different applications in the field of digital recognition including; handwritten address detection, check, and document. Persian handwritten digits classification has been facing difficulties due to different handwritten styles, inter-class similarities, and intra-class differences. In this paper, a novel method for detecting Persian handwritten digits is presented. In the proposed method, a combination of Histogram of Oriented Gradients (HOG), 4-side profiles of the digit image, and some horizontal and vertical samples was used and the dimension of the feature vector was reduced using Principal Component Analysis (PCA). The proposed method applied to the HODA database, and Support Vector Machine (SVM) was used in the classification step. Results revealed that the detection accuracy of such method has 99% accuracy with an adequate rate due to existing unacceptable samples in the database, therefore, the proposed method could improve the outcomes compared to other existing methods.

KEYWORDS: Histogram of Oriented Gradients, Principle Component Analysis, Support Vector Machine.

1. INTRODUCTION

Automatic character recognition is widely used as information entry for many applications nowadays [1, 2]. Handwritten digit recognition has a special importance in this way due to its different applications. Some of its different applications are including detecting handwritten addresses, checks and other documents. Peoples' handwritten are varying in different cultures, hence local research is required. Persian is the main language of several countries such as Iran, Afghanistan, and Tajikistan and it is spoken by more than 110 million peoples [3]. Persian handwritten digits have different forms, shapes, and sizes which makes its recognition more challenging. As an example, although the Persian and Arabic digits look each other, peoples write them in different shapes [4].

Automatic Persian handwritten recognition faces different problems coming from their intra-class differences and inter-class similarities. Figure 1 is showing some of them.

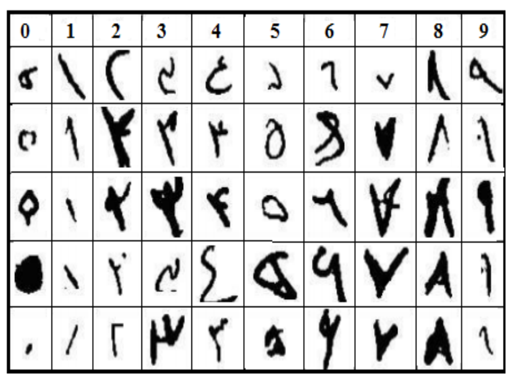

Fig. 1. Some Persian handwritten digits in different classes.

2. REVIEW OF RELATED WORKS

Several types of research have been conducted on handwritten digit recognition in different languages with different results [5-8]. Some researchers have also studied on Persian digits. Samadzade and Rahmati in [9] have proposed a method based on the external profile of the Persian digit image as the main feature, concatenating by a number of the image's cut-offs and also histogram of its projection in different directions. They used a Support Vector Machine (SVM) for classification and achieved 99.5% of recognition rate by applying their method on a self-collected database. Nooraliei was proposed a method based on local features and histogram of image projection and using SVM in [10]. By applying their method to a dedicated database including 8000 train and 600 test data samples which distinguish between different forms of ‘4’ and ‘6’ digits, they’ve reached to 97.83% of accuracy.

HODA is one of the biggest and challenging datasets of Persian handwritten digits and characters which is used in recent research in this area [11]. This dataset includes 60000 training data samples and 20000 test ones. This dataset is used in this paper.

Alaei et al. achieved the accuracy of 98.71% by applying SVM on modified contour features of this database [12]. Ebrahimpour et al. proposed a two-layer Radial Basis Function (RBF) Neural network classification system in [13] where four RBF classifier is applied on the image geometric description and an over-layer RBF neural network makes the final decision by weighing the previous ones. By applying their method to HODA database, they obtained 93.5% of accuracy. Abdi and Salimi have also proposed a similar mixed method in [14], and by applying particle swarm optimization (PSO) method on the same dataset have got 97.1% of accuracy.

Verity of expert systems has been used for classification of handwritten characters or digits. Saxena et al. have used neural networks for handwritten digits recognition [20]. Support Vector Machines (SVM) is used frequently for this purpose, and its excellence in handwritten digit recognition is shown by Ebrahimzadeh and Jampour [19]. Therefore, SVM is used in this research.

3. MATERIALS AND METHOD

As other pattern recognition methods, the proposed method can be divided into three major parts, namely pre-processing, feature extraction and classification, which will be discussed in following sections.

3.1. Pre-processing

One of the HODA database problems is the different size of its images. Furthermore, dimension ratio of different images is different inherently. For example, images of ‘1’ or ‘9’ are sketched vertically, while ‘0’ and ‘5’ images are the square shapes. For overcoming this problem, an image squaring algorithm is conducted. The ratio of the length to the width of the image is calculated and some zero rows or columns are padded to it symmetrically, in case that the ratio is less than 0.95 or greater than 1.05.

In the second stage, the size of all images are adjusted, to generate unique block size and feature-length during applying feature extractor algorithm. The final pre-processing part is noise reduction and Binarization of the images. Figure 2 is summarizing the pre-processing parts.

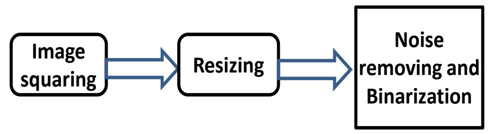

Fig. 2. Pre-processing steps.

3.2. Feature Extraction

Several features are extracted from each image and concatenated together to build feature vector. They are HOG, 4-side profiles of the digit image and some horizontal and vertical samples.

Histogram of oriented gradients (HOG) is a powerful feature extracting method which is frequently used in character and digit recognition algorithms in different languages [15, 16]. It is also used in this research due to its good performance and especially its robustness against image’s dimension changing [17]. This algorithm is counting the number of edges in a local neighbor in the image [18]. Gradient calculated the magnitude and direction of the most changing pixels of the image [19]. It is using Sobel filter, as shown in Figure 3 for horizontal and vertical components.

-1	0	1
-2	0	2
-1	0	1

a)

1	2	1
0	0	2
-1	-2	-1

b)

Fig. 3. Sobel mask used for gradient, a) horizontal, b) vertical.

Vertical and horizontal components can be calculated based on the following formulas:

$$G_x = H * I(x,y) \text{ و } G_y = H^T * I(x,y) \quad (1)$$

Where ‘*’ means correlation.

Gradient of the image is calculating as:

$$G(x,y) = \sqrt{\left(G_x^2(x,y) + G_y^2(x,y)\right)} \quad (2)$$

And its direction as:

$$\theta(x,y) = \tan^{-1}\frac{G_x(x,y)}{G_y(x,y)} \quad (3)$$

HOG of the image is calculated by

$$\Psi = \begin{cases} G(x,y), & if\ \theta(x,y)\ \epsilon\ bin_k \\ 0, & otherwise \end{cases}. \qquad (4)$$

Figure 4 is showing some HOG features of a digit ‘8’ images which are calculated by non-overlapping blocks and different sizes.

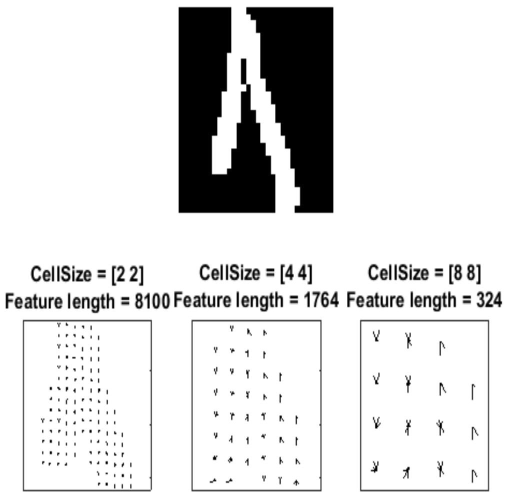

Fig. 4. HOG Features of Persian digit ‘8’ image.

In this research as shown in figure 4, [8 8] HOG block size which produces minimum feature length is used to decrease the processing time. Meanwhile, by increasing HOG bins into 18, the feature-length is duplicated to 648, for ensuring better accuracy.

Another feature used in this paper is 4-side profiles. For this purpose, as shown in Figure 5, the distances of the image from all borders are calculated and padded to the feature vector as a 128-bit new feature.

The final image features are some vertical and horizontal samples of the image. In this way, the image edges are extracted using morphological operators and then its samples across vertical and horizontal lines in 5, 15, 25 coordinates are padded to the feature vector.

By applying the previous feature extracting methods, the length of the feature vector is reached to 1160 bits. Principle component analysis (PCA) algorithm is used for reducing its dimension, to improve the total performance. In order to find the optimum PCA length, the entire algorithm is done several times over 2% of training data by different PCA length. Figure 6 is showing the final classification accuracy versus different PCA length. As expected by increasing the feature vector length, the classification accuracy is improved, but after reaching its maximum values, it remains adequately constant. Therefore, the maximum point is selected as the optimum PCA length, which is 763 bit for this experiment. This optimum length is used for other experiments which will be performed over the entire training dataset.

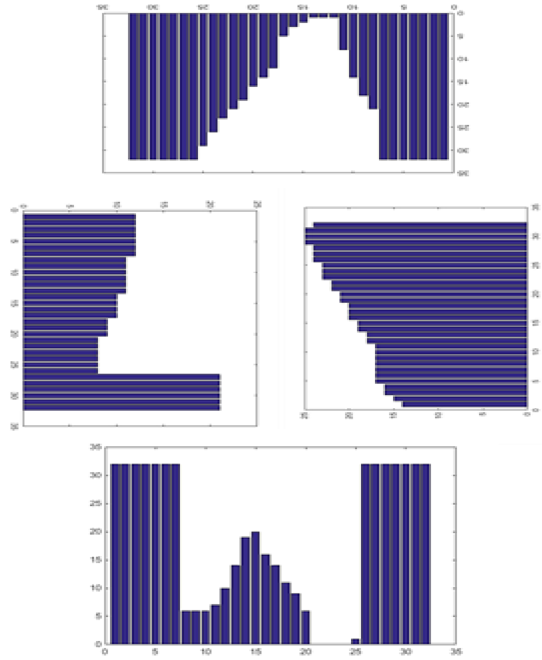

Fig. 5. 4-side Profiles of the digit ‘8’ image.

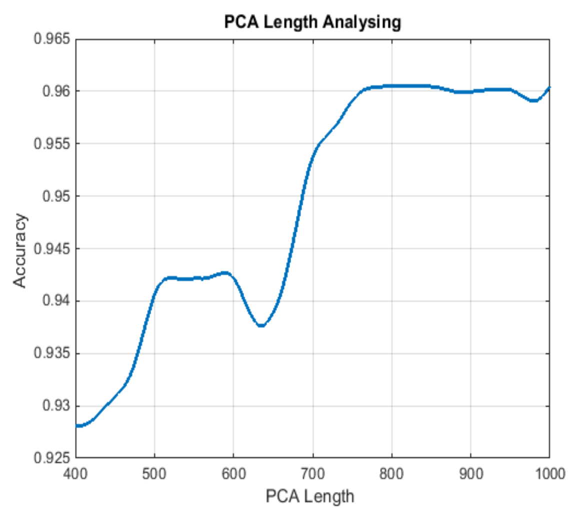

Fig. 6. PCA length analysing.

3.3. Classification

Different experts may be used for classification. Support vector machines in its simplest form, the linear SVM, consist of a hyperplane that separates the set of positive and negative samples with maximum distance.

In the training step, each pixel of a learning set is assigned a class tag. The training algorithm tries to find the optimal separation hyperplane that maximizes the margin between the nearest pixels [21]. The boundary pixels, called support vectors, are used to create a decision surface. In the prediction step, each unlabeled pixel receives a label based on the relative distance to the hyperplane. The decision function is given by

$$\mathrm{Lbl}(x) = \mathrm{sign}\left(\sum_{i=1}^{N}(a_i * y_i * K(x_i, x) + b)\right) \qquad (5)$$

Where x is the input pixel vector, x_i is the i^{th} support vector, N is the number of support vectors, a_i and y_i are the i^{th} coefficient of Lagrange and the corresponding classification tag, respectively, and finally b is the decision shift coefficient. K is a kernel function used to convert the original data to the feature space. Different kernels can be used in SVM such as linear, polynomial, and Gaussian (RBF). The RBF kernel can be defined as

$$K(x_i, x) = \exp(-\gamma * \|x_i - x\|^2) \qquad (6)$$

Where γ is the RBF coefficient, determined after scanning the data, x is the input pixel vector, and x_i is the i^{th} support vector.

In case of more than two classes, SVM could works in different ways, for example one-versus-one and on-versus-all. Results show that selecting any of these methods had no strong effect on the overall accuracy.

4. RESULTS AND DISCUSSION

In order to see the performance of them in Persian handwritten digits classifying, some famous algorithms are examined separately. In order to ensure the results, some of the experiments are performed several times by random train data samples. The results are shown in Table 1. As shown, SVM has the best classification accuracy over other methods.

Table 1. Some classification methods comparison.

Learner	knn	Tree	discriminant	SVM
Detection Accuracy	0.8875	0.8875	0.9125	0.9725

All types of SVM kernels are examined in this research using the whole training data to find the best one. Table 2 is showing the results. As shown, the polynomial kernel has the best classification accuracy in this usage, as discussed and proof in [19].

Table 2. Different SVM kernels comparison.

SVM Kernel Function	Gaussian	Linear	polynomial
Detection Accuracy	0.9868	0.9855	0.99025

As shown in Table 2, the best classification accuracy is achieved by applying polynomial SVM which is 99.025%. Table 3 is showing the percentage confusion matrix calculated by applying it. The values of the main diagonal are showing the true positive values of detection for each digit, and the average of them is implying the overall classification accuracy. For example, the value for Digit "1" is one, which means all test candidates are predicted true. While the value for the digit "2" is 0.99, due to one percent miss-classification to digit "3' which is very similar to "2" in Persian writing.

Table 3. The classification confusion matrix.

		Predicted classes									
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Real classes	**0**	0.9875	0	0	0.0025	0	0.005	0	0.0025	0	0.0025
	1	0	1	0	0	0	0	0	0	0	0
	2	0	0	0.99	0.01	0	0	0	0	0	0
	3	0	0	0.0225	0.9675	0.0025	0	0	0.0025	0.0025	0.0025
	4	0	0	0	0.005	0.9925	0	0	0	0	0.0025
	5	0.005	0	0	0	0.0025	0.9925	0	0	0	0
	6	0	0	0	0	0.005	0.005	0.985	0	0	0.005
	7	0	0.005	0.0025	0	0	0	0.0025	0.99	0	0
	8	0	0	0	0	0	0	0	0	1	0
	9	0	0	0	0	0	0	0.0025	0	0	0.9975

5. CONCLUSION

Regarding the importance of Persian handwritten digits recognition, a method for classifying them is introduced in this paper. The proposed algorithm is based on applying SVM over a publicly available dataset name HODA. In this way, a feature vector is extracted from each image using a mixture of different features. By reducing the dimension of the feature vector using the PCA algorithm, its optimum length is calculated. Polynomial SVM is found to be the best classifier for this goal. By training it using train data and evaluating it, the overall classification accuracy is measured around 99% which shows improvement over the existing methods.

6. ACKNOWLEDGEMENT

The authors would like to thank the University of Gonabad for their financial support through grant number 93-1.

REFERENCES

[1] Chaudhuri, Arindam, Krupa Mandaviya, Pratixa Badelia, and Soumya K. Ghosh, **"Optical Character Recognition Systems,"** *In Optical Character Recognition Systems for Different Languages with Soft Computing*, pp. 9-41. Springer International Publishing, 2017.

[2] Sawant, Sudarshan, and Seema Baji, **"Handwritten Character And Word Recognition using Their Geometrical Features Through Neural Networks,"** *International Journal of Application or Innovation in Engineering & Management (IJAIEM)*, 5, 2016.

[3] W.M. Pan, T.D. Bui, and C.Y. Suen, **"Isolated Handwritten Farsi Numerals Recognition Using Sparse and Over-Complete Representations,"** *in 10th International Conference on Document Analysis and Recognition*, pp. 586 – 590, 2009.

[4] Hamid Salimi, DavarGiveki, **"Farsi/Arabic Handwritten Digit Recognition Based On Ensemble of SVD Classifiers And Reliable Multi-Phase Pso Combination Rule,"** *International Journal on Document Analysis and Recognition (IJDAR)* , Vol. 16(4), pp. 371-386, 2013.

[5] Wang, Jeen-Shing; Fang-Chen Chuang, **"An Accelerometer-Based Digital Pen with A Trajectory Recognition Algorithm for Handwritten Digit and Gesture Recognition,"** *IEEE Transactions on Industrial Electronics*, Vol. 59, No. 7, pp. 2998-3007, July 2012.

[6] Labusch, K.; Barth, E.; Martinetz, T., **"Simple Method for High-Performance Digit Recognition Based on Sparse Coding,"** *IEEE Transactions on Neural Networks*, Vol. 19, No. 11, pp. 1985-1989, Nov. 2008.

[7] Cireşan, D., Meier, U., Gambardella, L., Schmidhuber, J., **"Deep, Big, Simple Neural Nets for Handwritten Digit Recognition,"** *Neural Computation,* Vol.22, No.12, pp. 3207-3220, Dec. 2010.

[8] Sarkhel, Ritesh, Nibaran Das, Aritra Das, Mahantapas Kundu, and Mita Nasipuri, **"A Multi-Scale Deep Quad Tree Based Feature Extraction Method For The Recognition of Isolated Handwritten Characters of Popular Indic Scripts,"** *Pattern Recognition*, 2017.

[9] Hasan Soltanzadeh, Mohammad Rahmati, **"Recognition Of Persian Handwritten Digits Using Image Profiles of Multiple Orientations,"** *Pattern Recognition Letters*, Vol. 25, Issue 14, pp. 1569-1576, 2004.

[10] Nooraliei, A., **"Persian Handwritten Digits Recognition by using Zoning And Histogram Projection,"** *3rd Joint Conference of AI & Robotics and 5th RoboCup Iran Open International Symposium (RIOS)*, pp. 1-5, April 2013.

[11] http://farsiocr.ir/farsi-digit-dataset, retrieved at 3rd, march,2015

[12] Alaei, Alireza, Umapada Pal, and P. Nagabhushan, **"Using Modified Contour Features And Svm Based Classifier for The Recognition Of Persian/Arabic Handwritten Numerals,"** *Seventh IEEE International Conference on Advances in Pattern Recognition*, ICAPR'09, 2009.

[13] Ebrahimpour, Reza, Alireza Esmkhani, and Soheil Faridi, **"Farsi Handwritten Digit Recognition Based on Mixture of RBF Experts,"** *IEICE Electronics Express 7.14*, pp. 1014-1019, 2010.

[14] Abdi, Mohammad Javad, and H. A. M. I. D. Salimi, **"Farsi Handwriting Recognition With Mixture of RBF Experts Based on Particle Swarm Optimization,"** *International Journal of Information Science and Computer Mathematics*, 2.2, pp. 129-136, 2010.

[15] Kamble, Parshuram M., and Ravinda S. Hegadi, **"Handwritten Marathi Character Recognition using R-HOG Feature,"** *Procedia Computer Science*, 45, pp. 266-274, 2015.

[16] Iamsa-at, S.; Horata, P., **"Handwritten Character Recognition using Histograms of Oriented Gradient Features in Deep Learning of Artificial Neural Network,"** *International Conference on IT Convergence and Security (ICITCS)*, pp. 1-5, 2013.

[17] Lowe, David G., **"Distinctive Image Features from Scale-Invariant Keypoints,"** *International Journal of Computer Vision*, Vol. 60.2, pp. 91-110, 2004.

[18] Dalal, Navneet, and Bill Triggs, **"Histograms of Oriented Gradients for Human Detection,"** *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, Vol. 1, 2005.

[19] Ebrahimzadeh, Reza, and Mahdi Jampour, **'Efficient Handwritten Digit Recognition Based on Histogram of Oriented Gradients and SVM,"** *International Journal of Computer Applications*, Vol. 104.9, 2014.

[20] Neera Saxena, Qasima Abbas Kazmi, Chandra Pal and O.P. Vyas, **"Employing Neocognitron Neural Network Base Ensemble Classifiers to Enhance Efficiency of Classification in Handwritten Digit Datasets,"** *D.C. Wyld, et al. (Eds): CCSEA 2011*, CS & IT 02, pp. 408–416, 2011.

[21] Wu, Yuanfeng, Xinghua Yang, Antonio Plaza, Fei Qiao, Lianru Gao, Bing Zhang, and Yabo Cui. **"Approximate Computing of Remotely Sensed Data: SVM Hyperspectral Image Classification as a Case Study."** *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9*, No. 12, pp. 5806-5818, 2016.