Recognition of Handwritten Persian Two-digit Numerals Using a Novel Hybrid SVM/HMM Algorithm

Vahid. Moradi[1], Farbod. Razzazi[2], Alireza. Behrad[3]
1- Electrical and Computer Engineering Department, Science and Research Branch, Islamic Azad University, Tehran, Iran.
Email: v_mo83@yahoo.com
2- Electrical and Computer Engineering Department, Science and Research Branch, Islamic Azad University, Tehran, Iran.
razzazi@srbiau.ac.ir
3-Electrical and Electronic Engineering Department, Shahed University, Tehran, Iran.
behrad@shahed.ac.ir

Received : Oct. 2015 Revised : March 2016 Accepted : June 2016

ABSTRACT:
There is a vast range of potential applications for recognition of handwritten Persian / Arabic digits (e.g. banking transactions, business registration forms and postal code recognition to name a few). In this paper, a new method is presented for automatic recognition of joint two-digit Persian numerals. The proposed method is composed of a combinational structure of Support Vector Machines (SVM) and a Hidden Markov Models (HMM). In this approach, we used SVM and HMM for classification and segmentation goals respectively. Due to the higher performance of SVM in classification with respect to HMM, the main core of recognition is a SVM classifier. In contrast, we used HMM to detect the location of the boundary for two-digit numerals. To evaluate the method, we employed a selection of HADAF Persian isolated characters corpus. We employed a 4 scale Gabor filter bank (24, 12, 6 and 3 scales) in 6 directions (0, 30, 60, 90, 120, 150 degrees) for feature extraction. The results showed the digit recognition rate of about 98.75 percent for the proposed algorithm on Persian two-digit numerals, while the recognition rates were 98.58 and 95.93 for separate SVM and HMM engines on isolated characters respectively.

KEYWORDS: Persian handwritten numeral recognition, SVM/HMM combining classifier.

1. INTRODUCTION

Nowadays, the recognition of handwritten digits has found wide applications. They are applicable to postal zip code reading, handwritten form processing, data entry applications, electronic payment cheques, office automation systems and a large volume of business transactions. The recognition of handwritten Persian isolated numerals, due to high structural similarity of different digits and diversity of handwriting styles faces many challenges. There are valuable glancing results for Persian/Arabic isolated characters. Mahmoud et al address the problem of recognition of Persian/Arabic numerals using SVM and HMM engines separately, the achieved average recognition rates are 99.4%, 97.99% and 94.35% using SVM, HMM and NM classifiers, respectively [1].

Ahranjany et al. used a fusing of two convolutional neural networks (CNN) trained by gradient descent training algorithm for recognizing the handwritten Farsi/Arabic digits. The technique is inspired by the human visual system. This study reports a digit recognition rate of 99.17%. In addition, the recognition rate increases to 99.98% after rejection of ten percent of “hard to recognize” samples [2].

Ching et al used a combining classifier consisting of a CNN and a support vector machine (SVM) for recognition of English digits [3]. In this approach, CNN is employed for feature extraction stage and SVM is used as a classifier for recognition. This algorithm results in an average recognition rate of 99.81%.

Yang Zhang et al used a hybrid SVM/HMM for handwritten chemical symbols recognition [4]. In the first stage of algorithm, they applied SVM classifier to distinguish non-ring structure (NRS) and organic ring structure (ORS) symbols, and then at the second stage, HMM method was employed for fine recognition. The average recognition rate for this algorithm is 98.08%.

Unfortunately, all existing researches have been concentrated on recognizing isolated digits; however there are many applications with inevitable connected handwritten digits and letters written in data entry forms without cellular structure. As an example, it can be referred to numerical bank cheque amount recognition which requires segmentation of the handwritten digits before recognizing them. Although HMM accuracy is very good in sequence classification when there is a huge amount of training data, SVM classifier has a nice record in classification when the number of training data is limited [5]. There is a fundamental difference between SVM and HMM, because SVM is inherently a static classifier which cannot implicitly model temporal evolution of data while HMMs can be used for classification of data that are dynamic and have two certain assumptions including stationary and independence constraints [6].

In this work, we used a hybrid HMM/SVM classifier to recognize connected digit numerals. SVM was invented by Vapnik for the first time based on the idea of minimizing the risk structure [7]. SVM classifier is a basically linear two-class classifier that can be also extended to the nonlinear multi-class classifier by using an appropriate kernel. SVMs have been successfully used in many character recognition tasks in recent years [8]. To train SVM parameters, empirical risk minimization (ERM) can be used to find the optimum hyper plane, although this does not guarantee a unique solution [9]. HMM is one of the recognition methods that have been used to recognize speech by Lawrence Rabiner for the first time [10]. This model is a useful tool for sequential data modeling and has been successfully applied in speech and continuous handwriting recognition [11].

In this paper, we present a hybrid algorithm for recognition of two-digits numerals. In the first stage, SVM and HMM classifiers are trained separately for recognition of Persian digits. Then the two-digit numeral is segmented into digits using HMM classifier, hence we can easily separate features of each digit. After this step, SVM is used for recognition of two separated digits.

The paper is organized as follows. Section 2 describes the proposed algorithm for handwritten Persian two-digit recognition. Experimental results are reported in Section 3 and finally the conclusions are stated in Section 4.

2. THE PROPOSED ALGORITHM FOR TWO-DIGIT NUMERALS RECOGNITION

The architecture of the proposed hybrid model is depicted in Figure 1.

As HMM models are trained by isolated digits, two extend the application of trained HMMs to two-digit numerals, it is necessary to concatenate HMMs to have all possible two digits numerals. Each sub-model should be responsible to represent one digit, while the concatenation of the models should be a good statistical model for transition between two models. Therefore, the parameters of the HMM pairs should be extracted by combining the parameters of HMM sub-models.

The proposed algorithm for handwritten two-digit Persian numerals consists of a concatenation of two HMMs that are pre-trained for isolated digit recognition. Therefore, before the construction of the combined classifier, it is necessary to train SVM and HMM classifiers for isolated digit recognition. The meta-parameters of the trained classifiers were achieved empirically.

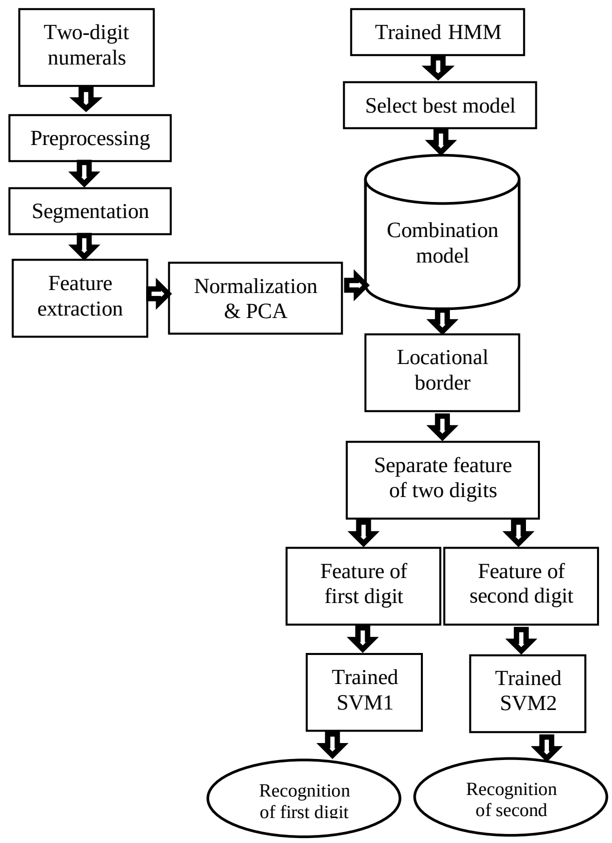

Fig. 1. The Architecture of the Proposed Algorithm.

In the test phase, by using Viterbi algorithm on concatenated HMMs for all possible 2-digit pairs, the location of border of two digits were found by the HMM model whose had the best result in two digits recognition. Then, two trained SVMs for two isolated digits were employed to recognize the segmented input image.

To construct the two-digit HMM models, the parameters of concatenated HMM are determined as full construction of the initial states distributions:

$\pi_c = [\pi_1, \varepsilon, \ldots, \varepsilon]$, (1)

where π_1 is the initial state distribution of the first model and π_c is the initial state distribution of the resulting concatenated HMM model and ε a numeral near to zero.

- Constructing the state transition matrix:

$A_c = \begin{bmatrix} A_1 & B \\ \varepsilon_1 & A_2 \end{bmatrix}$, (2)

where A_1 and A_2 are transition matrices of the first and second models, B is the repetition of the scaled version of the initial states distribution of the second model (π_1). The scale is assumed to be 0.1 in this study. ε_1 is square matrix with all elements equal to ε. Finally, A_c is the transition matrix of the concatenated model.

- Building the observation probability distribution matrix:

$E_c = [E_1\ E_2]$ (3)

Where $E_1\ and\ E_2$ are observation probability distribution matrices of the concatenating models and E_c is the observation probability distribution of the concatenated model.

- Determining the number of states in the concatenated model:

$N_c = N_1 + N_2$ (4)

where $N_1 and\ N_2$ are the number of states in concatenating models and N_c is the number of the combinational model. Fig. 2 shows the topology of the concatenated model, the value of P_c is equal to the matrix B.

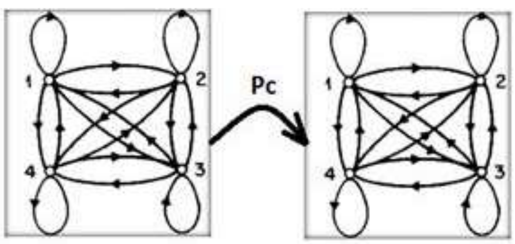

Fig. 2. Topology used in the concatenated model.

By constructing the concatenated model for all two-digit pairs by pre-trained HMM models of isolated digits, we will have 10×10=100 models for two-digit numbers. The selected model among these 100 models to segment the digits into two isolated ones, would be the model which most likely generates the input two-digit combination. Therefore, we put the features of image of two-digit numerals as the input of all concatenated models and trace the path of Viterbi algorithm to find out both the generation probability and the boundary. When the state number changes from N_1 or less to a state greater than N_1, there is a boundary between the two digits in Viterbi path. Therefore the features sequence of the image is separated into two separate feature sequences from this point. For example, if the size of the first digit in the border is m, the size of the second digit feature sequence will be 24-m. Fig. 3 shows the border detection result in a two digit image.

Fig. 3. Segmentation of the image into two isolated digits using concatenated HMM.

Finally, we used the trained SVMs for the recognition of the each segmented images

.

3. EXPERIMENTAL RESULTS

3.1. Dataset Description

In this work, we used a selection of HADAF dataset [12]. Fig. 6 shows samples of this dataset. We used ten classes 1 to 10 (their labels are: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9). The selection of the database consists of 18869 samples of Persian digits, where 60% of the data were used in the training phase of the classifier and the remaining 40% were used for the test. The database for two-digit numerals images are synthetically made by putting together 1000 pairs of isolated digits.

3.2. Preprocessing and Feature extraction

In fig. 4, the block diagram of the proposed algorithm for feature extraction is depicted. In the first stage, it is necessary to perform preprocessing on the test and train images. This task plays a vital role to extract discriminative features which are well invariant against intra-class variation. The images were filtered to reduce salt and pepper noise using a Gaussian filter before binarization. Then, the digits were spatially transferred to the center of the image. The images are binarized and cropped to the outer bounding box of the handwritten digit. Finally, images are normalized to 24×24 pixels. Two-digit numerals are normalized to 24×48 pixels.

In this work, we used Gabor filters for feature extraction. We applied a Gabor filter bank with proper impulse responses for feature extraction. Gabor filters have first been used for handwriting recognition by [13] and have successfully been employed for isolated character recognition in a great set of previous studies. Specifically they showed their good performance in Persian digits recognition in recent papers [14]. Gabor filters' theoretical superiority is due to simultaneous extraction of frequency and spatial information in characters zones. This issue makes the feature vector to be a good representation for character recognition.

Gabor filters is a plane wave that it is modulated by a Gaussian function. Two-dimensional Gabor filter in the direction θ is defined as follows:

$$g(x, y) = y.z, \tag{1}$$

where y and z are described as:

$$\mathrm{y} = \exp\left(\frac{-1}{2\sigma^2}(x^2 + y^2)\right), and \tag{2}$$

$$\mathrm{z} = \exp(j\omega(x cos\theta + y sin\theta)), \tag{3}$$

where σ is the standard deviation of the Gaussian function along x and y axes, and ω is the spatial frequency. In addition, σ determines the bandwidth of the filter and θ is the rotation angle of the Gaussian major axis of the plane wave. For constructing filter bank, we used 6 orientations θ from (0, 30, 60, 90, 120 and 150) and 4 scales λ from (3, 6, 12 and 24) [9]. The impulse response of the filter to be applied on the image i(x,y) according to the convolution relation is computed as follows:

$$\mathrm{G}(\mathrm{x}, \mathrm{y}; \lambda, \theta) = \mathrm{g}(\mathrm{x}, \mathrm{y}; \lambda, \theta) * \mathrm{i}(\mathrm{x}, \mathrm{y}), \tag{4}$$

$$\sigma = \lambda.k, \tag{5}$$

where k is a factor for determination of filter scale.

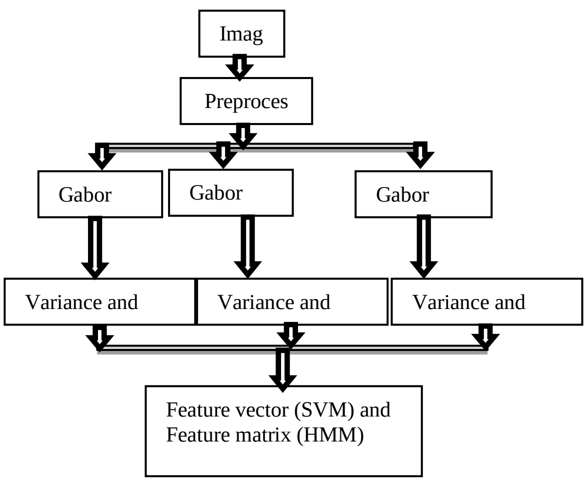

Fig. 4. Block diagram of the proposed feature extraction algorithm.

The mean and the variance of each segment are taken as the features of the segment.

After determining the filter bank impulse responses, we employed two types of zoning strategies to extract the features for each zone of the image separately. The zoning strategies were static zoning and sliding window strategies. In static zoning, the image is segmented into predefined checker type blocks. In sliding window zoning, a vertical strip is slid on the image from left to right with appropriate overlap to model the non-predefined length images. Fig. 5 shows the sliding windows strategy on an image containing the numeral 94. To extract the features using the static zoning method, the digit images and two-digit numerals are segmented into 3 × 3 and 3×6 segments respectively. Our filter bank had 24 filters and hence, the feature vector for each digit has 432 components while this number is 864 for two-digit numerals.

In sliding window zoning approach, each window is slid into vertically aligned segments. Each segment has 4×3 pixels. The window is slid from left to right horizontally with two pixels overlapping. In HMM approach; the final feature vector sequence would be 12 vectors with 384 components for isolated digits and 24 vectors for two-digit numerals. Fig. 5 shows this case for digit 94.

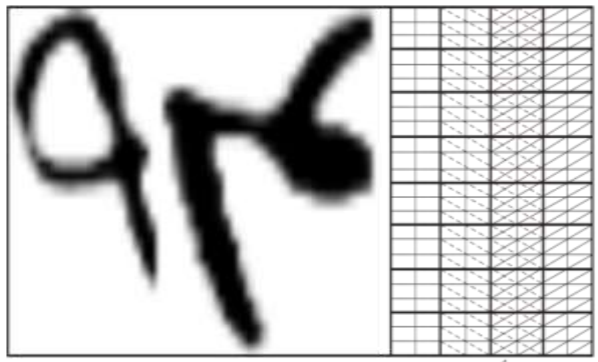

Fig. 5. Right-to-left sliding window strategy to extract features from the numeral images.

Fig. 6. Handwritten samples of Persian digits [11].

We used principal components analysis (PCA) for features compaction.

3.3. Experiments

To evaluate the system, in the first stage, the performance of the proposed algorithm for isolated digit recognition was investigated in terms of recognition accuracy. The recognition results for this algorithm are presented in table 1.

Table 1. Results of the SVM and HMM classifier for isolated digit recognition.

Digit	SVM	HMM
0	%99	%99
1	%98.2759	%97.985
2	%96.625	%91.75
3	%97.4969	%91.239049
4	%97.625	%92.625
5	%99.75	%99.375
6	%99.125	%93.5
7	%99.5	%97.375
8	%99	%97.625
9	%98.875	%98.75
Average	%98.5	%95.9

Recognition results show that the worst performance is achieved in classes 3, 4 and 5 (i.e. digits 2, 3 and 4); which may be explained by their intrinsic high inter-class geometric similarity. Both Support Vector Machine and Hidden Markov Model classifiers have taken highest misclassification for these 3 classes. The class 3 (digit 2) has the lowest recognition rate.

In the second experiment, we explored the effect of the proposed architecture for two digit pairs.

Table 2. Recognition rate of two trained SVM for two digits.

Digit	First digit In proposed architecture	Second digit In proposed architecture	First Digit in two digit HMM classifier	Second Digit in two digit HMM classifier
0	%99	%99	%98	%97
1	%98	%97	%97	%88
2	%98	1	%95	%78
3	%98	%98	%95	%81
4	%98	%97	%98	%84
5	%98	1	%97	%87
6	1	1	%92	%71
7	%99	%99	%98	%97
8	1	%98	%97	%91
9	%99	1	%96	%87
Average	%98.7	%98.8	%96.3	%86.1

According to table 2, recognition results of two-digit numerals show that the proposed system is better than simple methods of Hidden Markov Model classifier.

The SVM classifier alone cannot be applied to two digits directly. The obtained average recognition rates was 98.58% and 95.96% for digits using SVM and HMM and 98.7% and 98.8% for first digit and second digit of numerals using the proposed architecture respectively.

Three important parameters are effective in the SVM recognition rate. The first parameter is the value of gamma which is used in radial basis function (RBF) kernel, which can be written as:

$$K(x, y) = e^{\frac{-|x-y|^2}{2\sigma^2}} \quad (5)$$

This equation is a Gaussian function where the gamma value is $\frac{1}{2\sigma^2}$. The second parameter C is the regularization parameter which controls the tradeoff between the margin and the misclassification. Third is the effect of Principal Component Analysis (PCA) on features. Although it can be observed that PCA is not effective on the average recognition rate significantly, it increases the computational performance of the algorithm. The optimum point of gamma and C were achieved by grid search. The highest recognition rate is acquired using gamma equal $4.47 * 10^{-6}$ and C equal 10^6.

The effective parameters on the performance and the recognition rate of HMM are the number of states and the dimension of features obtained by PCA. We explored the effect of two parameters in fig. 7 and table 3. It should be noted that the aforementioned results are reported in the empirically optimum point of this investigation.

Table 3. The recognition rate for different dimension of feature matrix.

Matrixdimension (Effect of PCA)	Recognition rate
4	%78.8128
9	%91.6281
17	%94.66
25	%95.82
38	%95.82
52	%95.92
67	%95.32
78	%94.97

Although the performance was slightly degraded after employing the optimum value, table 3 indicates that the average recognition rate increases by increasing the dimensions of features. The next parameter is the state number of HMM, which saturated the performance in about 30 states.

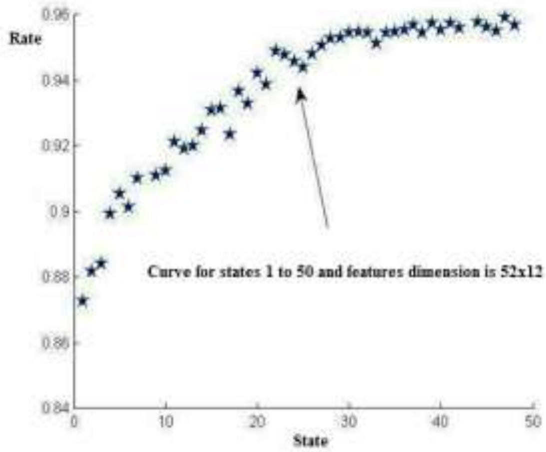

Fig. 7. Effect of states on recognition rate.

3.4. Error analysis

In this section, the recognition results are evaluated and the reasons for wrong classifications of numerals are investigated. This is due to several factors such as wrong labeling in the database, intrinsic ambiguity in handwritten numerals, non-descriptive features, and poor classification. Some images of numerals and digits that have been classified incorrectly and their real labels are shown in fig. 8 and fig. 9.

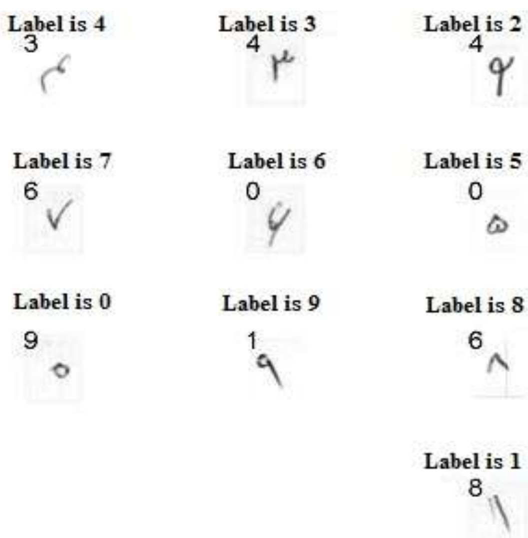

Fig. 8. Digits samples that are labeled incorrectly.

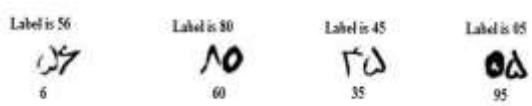

Fig. 9. Numeral samples that have been labeled incorrectly.

The confusion matrixes of the final proposed system are presented in table 4 and 5. As it can be observed, the most frequent errors are confusion of two, three and four which are intrinsically misclassified. However, the recognition result can be regarded as acceptable due to state of the art Persian character recognition systems, even in these confusing digits.

Table 4. Confusion matrix of the SVM1 of the final proposed system.

SVM 1	1	2	3	4	5	6	7	8	9	10
1	**0.99**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0.01**
2	**0**	**0.98**	**0**	**0**	**0.01**	**0**	**0**	**0**	**0.01**	**0**
3	**0**	**0**	**0.98**	**0**	**0**	**0**	**0**	**0**	**0**	**0.02**
4	**0**	**0**	**0**	**0.98**	**0.02**	**0**	**0**	**0**	**0**	**0**
5	**0**	**0**	**0**	**0.02**	**0.98**	**0**	**0**	**0**	**0**	**0**
6	**0.01**	**0**	**0**	**0**	**0**	**0.98**	**0**	**0.01**	**0**	**0**
7	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**
8	**0**	**0**	**0**	**0**	**0**	**0.01**	**0**	**0.99**	**0**	**0**
9	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**
10	**0**	**0**	**0**	**0**	**0**	**0**	**0.01**	**0**	**0**	**0.99**

Table 5. Confusion matrix of the SVM1 of the final proposed system.

SVM 2	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	**0.99**	**0**	**0.01**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
2	**0.01**	**0.97**	**0.01**	**0**	**0**	**0**	**0**	**0**	**0.01**	**0**
3	**0**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
4	**0**	**0**	**0**	**0.98**	**0.02**	**0**	**0**	**0**	**0**	**0**
5	**0**	**0**	**0**	**0.02**	**0.97**	**0.01**	**0**	**0**	**0**	**0**
6	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**
7	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**
8	**0**	**0**	**0**	**0**	**0.01**	**0**	**0**	**0.99**	**0**	**0**
9	**0**	**0**	**0**	**0**	**0**	**0**	**0.01**	**0**	**0.98**	**0.01**
10	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**

4. CONCLUSIONS

In this paper, a novel system for classification of writer independent off-line handwritten Persian two-digit numerals is proposed based on a combination of HMM and SVM. Average recognition rate of 98.58% and 95.93% was achieved for digits with SVM and HMM separately. These results indicated that SVM yields a better performance compared to HMM, hence we have used SVM as the main core of numerals recognition, while we applied HMM for determination of the border location between two digits. The average recognition rate of this approach is assessed as 98.75% for two-digit numerals. Recognition results show that the classification error due to factors such as ambiguity of images, inter-class overlap of handwritten classes and classification errors are investigated. Future work will be concentrated on the possibility of extending this approach to the recognition of handwritten multi-digit numbers and handwritten Persian cursive words. Although knowing the number of digits makes the model directly applicable to three or more digits by concatenating the previously trained single digit HMMs, the problem is still open for unknown number of digits.

REFERENCES

[1] Mahmoud ,S.A, Olatunji,S.O., "**Handwritten Arabic Numerals Recognition Using Multi-span Features & Support Vector Machines**",*Information Sciences Signal Processing and their Applications (ISSPA),10th International Conference on*, pp. 618–621,2010.

[2] Ahranjany, S.S; Razzazi, F, Ghassemian, M.H, "***A very High Accuracy Handwritten Character Recognition System for Farsi/Arabic Digits Using Convolutional Neural Networks**", *Bio-Inspired Computing: Theories and Applications (BIC-TA), IEEE Fifth International Conference on*, pp. 1585 –1592, 2010.

[3] Ching Y.Suen, Xiao-xiao Niu," **A novel Hybrid CNN-SVM Classifier for Recognizing Handwritten Digits**", *Pattern Recognition*, Volume 45, pp. 1318-13-25, 2012.

[4] Yang Zhang, Guangshun Shi, Kai Wang, "**A SVM-HMM Based Online Classifier for Handwritten Chemical Symbols**". *Pattern Recognition (ICPR), 2010 20th International Conference on*, pp: 1888 – 1891(2010).

[5] Hejazi, S.A; Sharif Univ. of Technol., Tehran; Kazemi, R. ; Ghaemmaghami, S. "**Isolated Persian Digit Recognition Using a Hybrid Hmm-Svm**". *International Symposium on Intelligent Signal Processing and Communications Systems*, pp: 1-4, 2008.

[6] Shao Qiang ; Inst. of Adv. Control Technol., Dalian Univ. of Technol., Dalian, China ; Shao Cheng ; Kang Jing. "**Hybrid SVM/HMM Method for Tool Wear Intelligence Measure in Cutting Process**", *Information Management, Innovation Management and Industrial Engineering (ICIII), International Conference on Volume 1*, pp. 525-529, 2010.

[7] Usama Fayyad," **A Tutorial on Support Vector Machine for pattern Recognition**", *Data Mining and Knowledge discovery*, PP. 121-167, 1998.

[8] Abdul Rahim Ahmad, C. Viard-Gaudin, Marzuki Khalid, Emilie Poisson,"**Online Handwriting Recognition Using Support Vector Machine**", TENCON 2004, 2004 IEEE Region 10 Conference, pp. 311-314, 2004.

[9] Ganapathiraju, A; Conversay, Redmond, WA, USA; Hamaker, J.E; Picone, J, "**Applications of Support Vector Machines to Speech Recognition**" Signal Processing, IEEE Transactions on, Volume 52, PP. 2348 - 2355, 2004.

[10] L.R. Rabiner, "**A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition**", Proceedings of IEEE, Vol.77 (2) PP. 257-286, 1989.

[11] M. Dehghan, K. Faez, M. Ahmadi, M. Shridhar," **Handwritten Farsi (Arabic) word recognition:a Holistic Approach Using Discrete HMM**". *Pattern Recognition*, Volume 34, Issue 5, pp. 1057-1065, 2001.

[12] Sara Khosravi, Farbod Razzazi, Hamideh Rezaei, Mohammad Reza Sadigh. "**A Comprehensive Handwritten Image Corpus of Isolated Persian/Arabic Characters for OCR Development and evaluation**". Signal Processing and Its Applications .pp. 1-4, 2007.

[13] Xuewen Wang, Xiaoqing Ding, Changsong Liu. "**Gabor Filters-Based Feature Extraction for Character Recognition**", *Pattern Recognition*, Volume 38, Issue 3, pp. 369–379, 2005.

[14] Hedieh Sajedi,Mehran Bahador. "**Persian Handwritten Number Recognition Using Adapted Framing Feature and Support Vector Machines**", *International Journal of Computational Intelligence and Applications*, Volume 15, No.01, 2016.