Persian/Arabic Handwritten Digit Recognition Using Local Binary Pattern

Mohsen Biglari*, Faezeh Mirzaei, Jalil Ghavidel Neycharan

biglari@shahroodut.ac.ir, Department of Computer Engineering, Shahrood University of Technology, Shahrood, Iran

fmirzaei@grad.kashanu.ac.ir, Department of Computer Engineering, University of Kashan, Kashan, Iran

jalil.ghavidel@gmail.com, Department of Computer Engineering, Shahrood University of Technology, Shahrood, Iran

ABSTRACT

Automated handwritten character recognition seems to be necessary due to the increasing number of Persian/Arabic handwritten documents. A new approach for Persian/Arabic handwritten digit recognition has been proposed in this paper. This approach employs Local Binary Pattern (LBP) operator as the base feature extraction method. Although this operator has shown great performance in research areas such as context and object recognition, but it has not been used in Persian/Arabic handwritten digit recognition problem. First step in the proposed approach involves smoothing, converting black and white input image to grayscale intensity image and resizing it to a fixed size. In the next step, input image is divided into several blocks. LBP operator is applied to each block to extract features. Finally, these features are used to train a multi-layer perceptron neural network with circular approach. Empirical results shows that the proposed approach has a very good generalization accuracy (99.72%) on Hoda dataset with 60000 train and 20000 test samples. This accuracy is the best among the state-of-the-art methods.

KEYWORDS

Persian/Arabic Handwritten Digit Recognition, Local Binary pattern, Multi-Layer Perceptron, Binary to Grayscale Conversion, Hoda DataSet

1 INTRODUCTION

Automated handwritten character recognition seems to be necessary due to the increasing number of Persian/Arabic handwritten documents. Handwritten character recognition research for Latin script family started over 50 years ago, but it's rather a new emerging research area for Persian/Arabic language with great achievement in the past few years. Persian and Arabic digit handwritings are exactly the same. So, all the descriptions in this text works for both, however for the sake of shortness, only Persian phrase will be used from now on.

The characteristics of Persian digit handwriting are different from the Latin. Some digits in Persian can be written in different forms. For example, five can be written like an empty circle or like 5. Distinguishing the first form from zero is very difficult in some cases. There are more other challenges like this which make Persian handwritten digit recognition harder than the Latin. Good performance of gradient based features in Latin digit recognition [1], has encouraged researchers to consider them in Persian digit recognition. Therefore, although many approaches have been proposed in this area, but our focus was only on gradient based one. For example, authors in [2] achieved 99.15% recognition rate by extracting features from gradient image on Hoda dataset [4]. Authors in [3] have used 16 directions instead of 8 directions which improved recognition rate by 0.32% (99.15% recognition rate) on the same dataset. But it is slower than the base method. A new classifier combining method has been proposed by [5]. This model is composed of four RBF networks as experts and another one as input gate. The gate network learns how to divide input space among experts. This method has reached 95.3% recognition rate and outperforms previous network combination methods. Modified Chain-Code Direction Frequency in surrounding pixels of images has been employed by [6]. This paper achieved 99.02% accuracy using multilayer SVM neural network on Hoda dataset.

This paper presents a new feature extraction approach for Persian handwritten digit recognition problem based on Local Binary Pattern (LBP) operator. LBP has been successfully applied in

context, object and face recognition problems. High success rate of LBP is a good inspiration to use it in Persian handwritten digit recognition. This operator is the main feature extraction method in the proposed approach. Also Multi-Layer Perceptron (MLP) has been used as the base learner. Experimental results show that our proposed approach is more accurate and faster than the previous ones.

The rest of this paper is organized as follows. Section two gives a brief review of LBP operator. The proposed approach is explained in section three. Section four includes the experimental results and the final section concludes the paper.

2 LOCAL BINARY PATTERN

Ojala et.al. [7] proposed Local Binary Pattern (LBP) for texture description. LBP is one of the most popular and successful binary descriptors in context, object and face recognition [8-12]. This operator is also robust to monotonic gray-scale changes.

The basic LBP operator works with a 3×3 window. Center of this window is placed on each pixel of the image and is compared with other pixels in the window. The results are put along and make a binary string which is interpreted as a decimal number. Histogram of these decimal numbers is the output features. Figure 1 illustrates this process with more details.
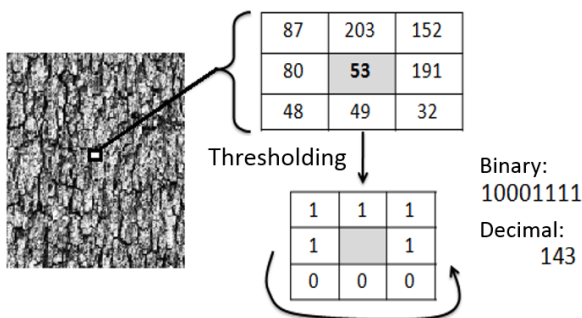

Figure 1: The basic LBP operator

The basic LBP operator has been extended to use circular neighborhoods. This adjustment helps it to catch more structural features of entities like faces, objects, and characters [10]. This extended version's sampling points are on a circle with arbitrary radius. $LBP_{P,R}$ refers to an operator with P

sampling points on a circle with radius R. Figure 2 shows different samples of this operator and formula 1 explains this operator. $LBP_{P,R}(x, y)$ is the interpreted decimal value of pixel (x, y). g_c shows the intensity value of the central pixel and g_i plays the same role for the neighboring pixels. S(x) is calculated as follows:

$$LBP_{P,R}(x, y) = \sum_{i=0}^{P-1} s(g_i - g_c)2^i \tag{1}$$

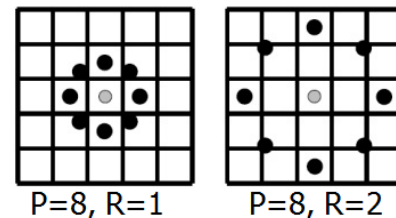$$s(x) = \begin{cases} 1 & x > 0 \\ 0 & x \le 0 \end{cases} \tag{2}$$

Figure 2: Extended LBP with different P and R

Uniform patterns are another extension of the basic LBP operator [10]. A binary pattern is called uniform if it has at most two changes from one to zero and vice versa. This operator is presented with $LBP_{P,R}^{u2}$. Although the number of uniform patterns to the total number of patterns is 58/256 (23%) but it is shown empirically in [10] that $LBP_{8,1}^{u2}$ and $LBP_{8,1}$ have almost the same performance. Histogram of uniform LBP operator has only one bin for non-uniform patterns which leads to total number of 59 bins.

3 THE PROPOSED APPROACH

Our proposed approach is made of two main parts which both of them are very critical. The first part is image normalization that converts binary image to grayscale and normalize its size. The second part is feature extraction. LBP operator is used in a way that both global and local features are extracted. Finally these features are fed to a multi-layer perceptron network. Figure 3 shows the explained process diagram.

The proposed approach has several advantages. The first and the most important one is its high accuracy. Experimental results show that our method is the most accurate among current

methods. The second advantage which make our method easy to use in real application, is its simplicity. This algorithm can be implemented very fast and also has a high speed. Runtime report is presented in the Experimental results section.
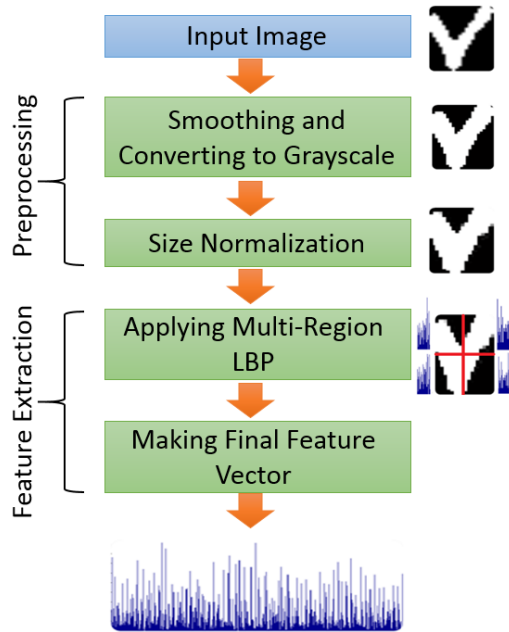

Figure 3: The proposed preprocessing and feature extraction approach diagram

3.1 PREPROCESSING

3.1.1 SMOOTHING AND CONVERTING TO GRAYSCALE

LBP method can only be used for grayscale images, therefore the input binary image should be converted to a grayscale image. Figure 4 shows a filter that can do this conversion. This filter can smooth the input image and convert it to a grayscale image. Figure 5 presents an example image of digit eight (٨). After applying this filter to the input image, all the edges will be blurred and other uniform areas remains the same. So LBP operator will focus only on the edges and the non-edge parts will have no place in the extracted histogram. Any digit has its specific number and type of edges. Therefore by extracting edges information, we can discriminate different type of digits.

0.1	0.1	0.1
0.1	0.2	0.1
0.1	0.1	0.1

Figure 4: The used filter for smoothing and converting binary image to grayscale image
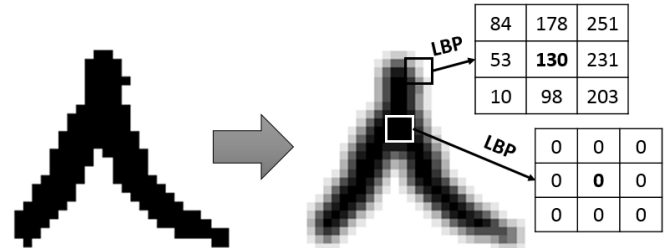

Figure 5: Applying the filter to the left binary image and converting it to a grayscale image

3.1.2 SIZE NORMALIZATION

LBP operator dictates that input images should have the same size. Height to width ratio of Persian alphabet is larger than one. Therefore, empirically, size of input images has been set to 28×20. Figure 6 represents this process in more details.

This method has several advantages over the conventional methods. Persian digits like five and zero have many structural similarities and one of their major differences is their size. The proposed approach preserves their size differences. Moreover, this method preserves height to width ratio that is a critical issue for some digits like one and zero that can fill the whole image after resizing and produce hard cases to classify.
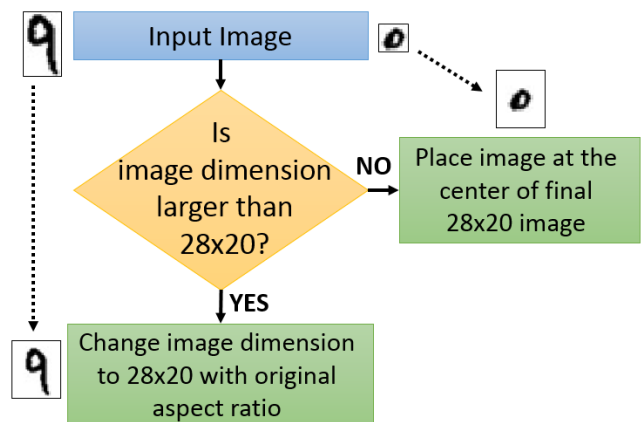

Figure 6: The proposed algorithm's flowchart for image size normalization

3.2 FEATURE EXTRACTION

LBP abilities in Persian handwritten digit recognition have not been evaluated yet. Uniform LBP has been employed to obtain more compact and efficient feature vector. This operator is applied in two forms (figure 7). These two obtained histograms are put along to get the final feature vector which is used in the training process.

1. The uniform LBP operator is applied to whole image to get a histogram with 59 bins.
2. After dividing the image into 4 equal blocks, Uniform LBP is applied to each block and the resulting histograms are concatenated.

Multi-region LBP has been used because, single-region LBP's output histogram is not discriminant enough. For example consider figure 8 which shows two Persian digits, seven (٧) and eight (٨) which have reverse form of each other. Almost all the edges in one image exist in the other one. If we apply LBP to the whole image, it will produce similar histograms. To overcome this problem, the input image is divided into some blocks and LBP is applied to them, therefore the results will no longer be similar (figure 8).
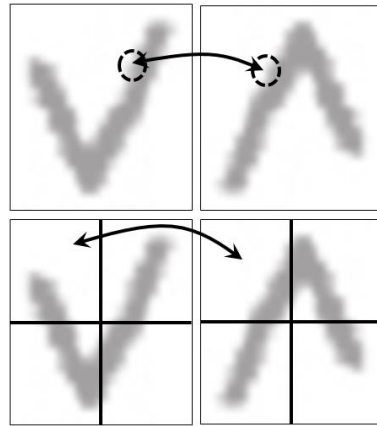

Figure 8: Difference between single-region LBP and multi-regions LBP

4 EXPERIMENTAL RESULTS

Hoda dataset [13] has been used to evaluate and compare the proposed approach to state-of-the-art methods. The proposed approach obtained the first place in the Persian handwritten digit competition, which held during "The First Conference on Pattern Recognition and Image Analysis (PRIA 2013)" [13]. Results of this competition have been reported in this section too. Hoda and Pars [13] datasets have been used to weigh the registered systems in this competition.

4.1 HODA DATASET

Hoda is the largest available dataset of handwritten Persian digits. It is composed of 60000 train and 20000 test samples. For each digit, varying from zero to nine, there is 6000 train and 2000 test samples. This dataset can be obtained from [4]. Figure 9 shows some sample images from this dataset.
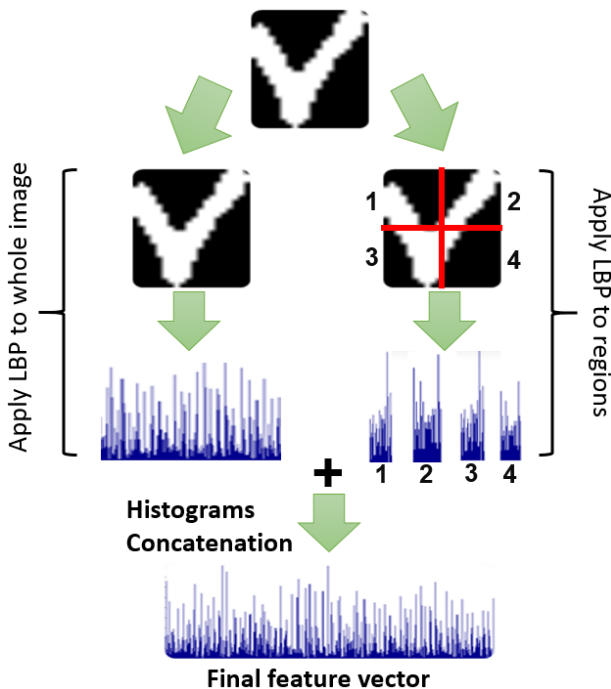

Figure 9: Sample images from Hoda dataset

Figure 7: The proposed approach for feature extraction using LBP operator

4.2 PARS DATASET

This dataset gathered by previously mentioned competition [13] and has not been officially released yet. It has 81863 samples which are divided into 59451 train and 24412 test samples. Figure 10 shows some sample images of this dataset.

Figure 10: Sample images from Pars dataset

4.3 TRAINING PROCESS

A multi-layer perceptron, consisting of one hidden layer with 30 neurons is used in the training process. 295 inputs (feature vector length) and 10 outputs (0 to 9) are other specification of this network.
Hoda dataset has 60000 training samples. Because of this large number of training samples, two different approaches have been considered for training process:

1. Training at once: In this approach, the whole dataset is fed to network in training process.
2. Circular training: In this approach, training set is divided into some subsets. All subsets are used for training in successive turns. This process is repeated several times, which prevents the network from forgetting the first subsets.

The proposed method accuracy has been compared with several other methods. The results have been summarized in table 1. The results indicate several points. Circular training is almost as good as training at once and running times of these two approaches are very close. In some cases, training with the whole dataset is not feasible. For example, sometimes training dataset is relatively too large to fit in the RAM. In these cases Circular training can be a good alternative.

Table 1: Comparison of the proposed method and the best previous methods

Method	Precision (%)
Ref. [5]	95.30
Ref. [14]	97.99
Ref. [15]	98.71
Ref. [6]	99.02
Ref. [15] (5-fold)	99.37
Proposed Method (Circular training)	99.59
Proposed Method (At once training)	99.72

Although the proposed method has the best accuracy among the other methods, it is much simpler than the others and has much less features which makes it very fast. Table 2 and table 3 report the results of the competition [13]. These tables compare the methods proposed in [13] on Hoda and Pars datasets. Name of the universities indicates only their representatives. As the results show, the proposed method outperforms all other methods on both datasets. Although these systems are only trained on Hoda dataset, the proposed method still has a better result than the others on Pars dataset. Experimental results of [13] shows that our method can recognize a digit in 180 milliseconds while it takes 400 milliseconds for the second place team.

Table 2: Comparison of the proposed method on Hoda dataset in competition [13]

Rank	Team	Precision (%)
4	Shahid Beheshti University	94.63
3	Amirkabir University	95.94
2	Birjand & Hormozgan University	97.39
1	Proposed Method	99.59

Table 3: Comparison of the proposed method on Pars dataset in competition [13]

Rank	Team	Precision (%)
4	Shahid Beheshti University	96.07
3	Amirkabir University	97.81
2	Birjand & Hormozgan University	98.94
1	Proposed Method	99.53

5 CONCLUSION AND FUTURE WORK

In this paper, a new feature extraction method for Persian handwritten digit recognition has been

proposed. Experimental results show that LBP operator can extract very discriminating features for handwritten digit recognition. Extraordinary recognition rate (99.72%) of the proposed approach makes it one of the best methods that has been proposed in this research area.

Advantages of this method can be categorized as follows:

- High discriminative power of the extracted features
- Easy implementation
- Very low feature extraction time
- Very low computational overhead

On the other hand, one of the drawbacks of this system can be the number of its features. Feature vector length can be reduced to 50% of the current length without noticeable performance fall. This helps the system to operate about two times faster. Methods like PCA, LDA and evolutionary algorithms can be considered as good feature reduction and selection methods. Also LBP operator can be modified to be more suitable for Persian handwritten digit recognition. We think it will be a good idea to extend LBP and combine it with a feature selection method.

7 REFERENCES

[1] C.-L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, "Handwritten digit recognition: benchmarking of state-of-the-art techniques," Pattern Recognition, vol. 36, pp. 2271-2285, 2003.

[2] M. Nahvi, K. Kiani, R. Ebrahimpour, "Gradient feature extraction improvement based on discrete cosine transform for handwritten Persian digits recognition", 18th Iranian Conference on Electric Engineering (ICEE2010), pp. 3066-3071, 2010.

[3] H. Khosravi, A. A. Kabir, "Introduction of two fast and efficient features for handwritten Persian digits recognition", 4th Iranian conference on Machine Vision and Image processing, pp. 2023-2029, 2007.

[4] H. Khosravi, E. Kabir, "Introducing a very large dataset of handwritten Farsi digits and a study on their varieties", Pattern Recognition Letter, vol. 28, pp. 1133-1141, 2007.

[5] R. Ebrahimpour, A. Esmkhani, and S. Faridi, "Farsi handwritten digit recognition based on mixture of RBF experts," IEICE Electronics Express, vol. 7, pp. 1014-1019, 2010.

[6] A. Alaei, P. Nagabhushan, and U. Pal, "Fine classification of unconstrained handwritten persian/arabic numerals by removing confusion amongst similar classes," in Document Analysis and Recognition, 2009. ICDAR'09. 10th International Conference on, 2009, pp. 601-605.

[7] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern Recognition, vol. 29, pp. 51-59, 1996.

[8] H. K. Ekenel, M. Fischer, E. Tekeli, R. Stiefelhagen, and A. Ercil, "Local binary pattern domain local appearance face recognition," 2008, pp. 1-4.

[9] D. Maturana, D. Mery, and Á. Soto, "Face recognition with local binary patterns, spatial pyramid histograms and naive Bayes nearest neighbor classification," 2009, pp. 125-132.

[10] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 24, pp. 971-987, 2002.

[11] L. Sánchez López, "Local Binary Patterns applied to Face Detection and Recognition," 2011.

[12] H. M. Vázquez, E. G. Reyes, and Y. C. Molleda, "A new image division for LBP method to improve face recognition under varying lighting conditions," 2008, pp. 1-4.

[13] Birjand University, "Competitions Report," The First Iranian Conference on Pattern Recognition and Image Analysis (PRIA 2013), 2013.

[14] R. Ebrahimpour and S. Hamedi, "Hand written digit recognition by multiple classifier fusion based on decision templates approach," in The International Conference on Computer, Electrical, and Systems Science, and Engineering, 2009, pp. 245-250.

[15] A. Alaei, U. Pal, and P. Nagabhushan, "Using modified contour features and SVM based classifier for the recognition of Persian/Arabic handwritten numerals," in Advances in Pattern Recognition, 2009. ICAPR'09. Seventh International Conference on, 2009, pp. 391-394.

[16] G. E. Hinton, M. Revow, and P. Dayan, "Recognizing handwritten digits using mixtures of linear models," Advances in neural information processing systems, pp. 1015-1022, 1995.

[17] T.-F. Gao and C.-L. Liu, "High accuracy handwritten Chinese character recognition using

LDA-based compound distances," Pattern Recognition, vol. 41, pp. 3442-3451, 2008.

[18] L. S. Oliveira, R. Sabourin, F. Bortolozzi, and C. Y. Suen, "Feature selection using multi-objective genetic algorithms for handwritten digit recognition," in Pattern Recognition, 2002. Proceedings. 16th International Conference on, 2002, pp. 568-571.