Common Techniques and Bio-Inspired Hierarchical Architecture for Automatic Farsi Handwritten Word Recognition Systems

Reza Ebrahimpour[1], Mona Amini [2], Afra Vahidi Shams [2]

[1]*Department of Electrical Engineering, Shahid Rajaee Teachers Training University, Tehran, Iran, ebrahimpour@ipm.ir*

[2]*Department of Mechatronics Engineering, Islamic Azad University, South Tehran Branch, Tehran, Iran; Mona.amini.mecha86@gmail.com, afra@vahidishams.ir*

Abstract

This paper investigates Farsi handwritten word recognition using common features. Also we applied biologically inspired features (BIFs), derived from a feed forward model of object recognition pathway in visual cortex for Farsi handwritten word recognition problem. Experimental results show that the model achieves high recognition percentage even for large variations and applicability of these features in Small Sample Size problems (SSS).The experiments were achieved using the Iranshahr dataset. This dataset consist of 780 samples of 30 city names of Iran which 600 samples for train and 180 samples for test was used. A set of experiments were conducted to compare Decision Templates with some combination rules. Results show that template based fusion method is superior to the other schemes.

Keywords*: Farsi handwritten word recognition, Feature extraction, Biologically inspired features, Human visual pathway*

1. Introduction

Handwritten Characters Recognition is one of the most remarkable topics in the domain of pattern recognition. During the past decade, an interesting progress has been achieved in the field of handwritten word recognition, and many applications, such as automatic reading of bank checks, postal addresses and forms have been appeared.

Handwritten Recognition is the process of transform scanned images of machine printed or handwritten text into a computer processable format. The process of handwritten recognition can be broadly broken down into three stages:

In the first stage preprocessing has been done then best features must be extracted from the images. Afterward classifiers must select the best class for each input image. In the preprocessing stage, various methods such as Noise Cleaning [1] to reduce noise, Normalizing [2] to remove skewness and Thinning [3-4] to retain skeleton of patterns and binary transformation [3] are used, thereby simplifying the processing of the rest of the stages. The performance of character recognition largely depends on the feature extraction approach. In this paper we focused on comparing various feature extraction methods.

There are various feature extraction methods that are convenient in the handwritten recognition domain such as Discrete Fourier Transform (DCT) [5], Wavelet [6], Linear Discrete Analysis (LDA) [8,13] and Principal Component Analysis (PCA)[16,7,9] , Zoning [10,11], Gradient [12], Histogram [14], and Cross count [15] and Biological feature [17 - 20]

that mainly is used in the face recognition domain and for the handwritten recognition task is new.

Understanding how visual cortex recognizes objects is a critical question for neuroscience. Building a system that emulates object recognition in cortex has always been an attractive idea. Recent works [12, 13, 14] have shown that a computational model based on our knowledge of visual cortex can be competitive with the best existing computer vision systems on some of the standard recognition datasets. Their system follows a recent theory of the feed forward path of object recognition in cortex that accounts for the first 100-200 milliseconds of processing in the ventral stream of primate visual cortex [14].

Next step after feature extraction is the classification that there are various methods for classification stage such as: Multi Layer Perceptron (MLP) [21] that is a kind of feed-forward artificial neural network model that maps sets of input data onto a set of suitable output. Radial Basis Function (RBF)[22-24] that is a kind of supervised neural networks. It is very similar to MLP but connections between hidden layer and output layer are different. RBF has only one hidden layer. Another way for classification is K-Nearest neighbor (KNN) [25-27], that is one of the most popular techniques for classification. This technique is the best in the case of having a small amount of training data. In this method, the distance of an input vector from a set of stored training patterns is computed by suitable distance criteria.

Another way to classify the input image is Support Vector Machine (SVM). The SVM [28, 29] classifier is its basic form implements binary classifications. The objective is to further improve the recognition rate by using support vector machine (SVM) at the segment classification level. In this paper, we use k-NN as a classifier for Farsi handwritten word recognition (FHWR) which is the statistical method and it has the advantages of being a non-parametric classifier. In the papers which worked on Farsi handwritten recognition various feature extraction methods are used. In this paper, we compare prevalent features with the biological.

The rest of this paper is organized as follows. Section 2 describes the feature extraction methods. Section 3 describes the biological features and Section 4 introduces the classification algorithm. In section 5 experimental results are discussed. Finally conclusions are presented in Section 6.

2. Feature Extraction

The most important section for handwritten recognition is feature extraction, which in this section different method are briefly described.

2.1. Histogram

The horizontal and vertical projection histograms of the image are obtained as follows: for each row and column between the first and last non- empty rows and columns of the image, the number of black pixels is counted and these values are placed in separated vectors. By concatenation of the horizontal and vertical projection histogram, the final feature vector is achieved [14].

2.2. Characteristic Loci

The characteristic loci feature is commonly in vertical and horizontal or 45 and 135 degree orientation. The Loci characterization feature vector for each word image is defined by assigning a number for each background point in the image. This number is

dependent on the number of times that the vertical and horizontal lines drawn from the background point toward the four main directions intersected with the body of the word. In these applications, to reduce the feature dimension the maximum number of intersections has been limited to 3. Thus for each point of image background, a four-digit number of base four is attained. These numbers are called locus numbers, which are between 0 and 255. As the result, the loci characterization feature vector consists of 81 elements that show the relative frequency of their corresponding number in the input image. To achieve a normalized feature vector, its values are divided by the total number of white points of the background [30]. For instance, the locus number of point P in Figure 1, is $(2300)_4 = (176)_{10}$.

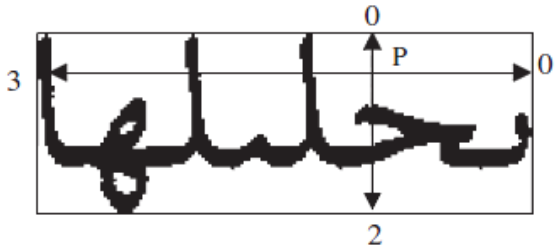


Figure. 1. Calculating Characteristic Loci Features (2300)4 = (176)10

2.3. Zoning

In this method, the word image is divided into the number of zones. In each zone density of object pixels is calculated. Density is calculated by finding the number of object pixels in each zone and dividing it by the total number of pixels having value black as object pixels [31, 32].

2.4. Wavelet Transform

Wavelet transform is a series expansion technique that allows us to represent the signal at different levels of resolution. The wavelet transformation uses the so-called wavelet basis functions (shortly wavelets) and the scaling functions, both forming the orthogonal or biorthogonal family of basis function. The wavelet function has good localization abilities in both time and frequency [31-33].

2.5. Discrete Cosine Transform

The discrete cosine transform (DCT) represents an image as a sum of sinusoids of varying magnitudes and frequencies. This method computes the two-dimensional discrete cosine transform (DCT) of an image. The DCT has the property that, for a typical image, most of the visually significant information about the image is concentrated in just a few coefficients of the DCT [36]. The two-dimensional DCT of an M-by-N matrix A is defined as follows:

$$B_{pq} = \alpha_p \alpha_q \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A_{mn} \cos \frac{\pi(2m+1)p}{2M} \cos \frac{\pi(2n+1)q}{2N} \quad 0 \le p \le M-1, 0 \le q \le N-1$$

$$\alpha_p = \begin{cases} \frac{1}{\sqrt{M}} & , p = 0 \\ \sqrt{\frac{2}{M}} & ,1 \le p \le M-1 \end{cases} \qquad \alpha_q = \begin{cases} \frac{1}{\sqrt{N}} & , q = 1 \\ \sqrt{\frac{1}{N}} & ,1 \le q \le N-1 \end{cases}$$

Where M and N are the row and column size of A, respectively and the values are called the DCT coefficients of A.

2.6. Gradient

To extract gradient features, the gradient operator is the first applied for the gray-scale image of the word to give two gradient components: strength |g(x, y)| and direction g(x, y) at each point (x, y) of the image f. This is done by applying Sobel operator on the image to extract vertical and horizontal gradient components [15]. The gradient features are composed of eight layers; each corresponding to one of the Freeman directions. Each layer is the projection of the gradient vectors of the image into the corresponding Freeman direction [37].

2.7. Cross Count

Cross count is a popular statistical feature, which in this method the number of crossing with a counter by a line segment is calculated. For finding feature vector, the image is the first binarized and then the vertical and horizontal cross counts are calculated by scanning each column and row of the binary image and each transition from 0 to 1 or from 1 to zero increases a counter that has an initial value of zero. Then the scanned column and row are associated with the counter value at the end of the scanning process [38].

2.8. Principle Component Analysis

In some situations, the dimension of the input vector is large, and the components of the vectors are highly redundant. In this situation, it is useful to reduce the dimension of the input vectors that it caused less computational load. An effective procedure for performing this operation is principal component analysis. PCA allows us to compute a linear transformation that maps data from a high dimensional space to a lower dimensional space [39, 40]. The objective of principal component analysis is to reduce the dimensionality of the dataset but retain most of the original variability in the data. PCA projects the data along the directions where the data varies the most. These directions are determined by the eigenvectors of the covariance matrix corresponding to the largest eigenvalues. Figure 2 shows the space transformation by PCA.

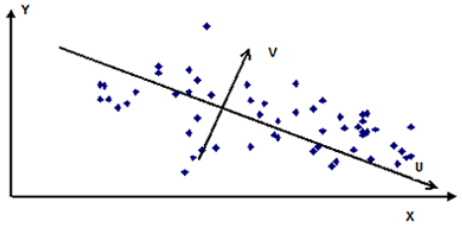


Figure 2. The Space Transformation by PCA.

Applying different feature extraction methods on word images have been shown on Figure 3.

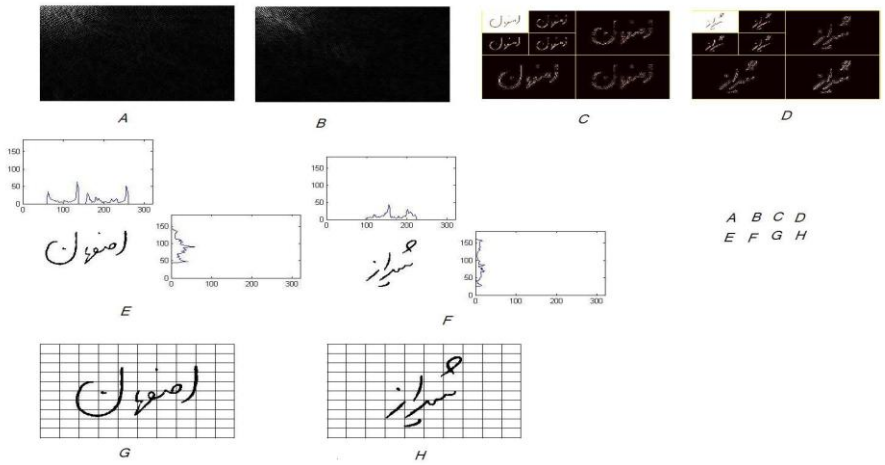


Figure 3. Applying different feature extraction methods on word images of Esfahan and Shiraz.
(A) DCT of Esfahan (B) DCT of Shiraz (C) 2D Wavelet of Esfahan (D) 2D Wavelet of Shiraz (E) Horizontal and Vertical Histograms of Esfahan (F)) Horizontal and Vertical Histograms of Shiraz (G , H) Zoning

3. Biological Feature Extraction Method

The printing Object recognition in cortex is mediated by the ventral visual pathway running from primary visual cortex , V1, through extrastriate visual areas V2 and V4 to inferotemporal cortex , IT, and then in order to prefrontal cortex (PFC) which is involved in linking perception to memory and action [43-45]. In this section, we will briefly describe feed-forward model of the primate visual object recognition pathway. The standard model, in its simplest form, consists of four layers of computational units where simple S units alternate with complex C units. The S units combine their inputs with Gaussian-like tuning to increase object selectivity. The C units pool their inputs through a maximum operation, thereby introducing invariance to scale and translation [51]. The sketch of the model is depicted in Figure 4.

3.1. Image Layer

The input for the model is an image and we convert the input image to gray-value image. In this layer an input image is transformed into an image-pyramid with ten scales. For creating the pyramid, the shorter edge is scaled to 140 pixels while maintaining the aspect ratio, then using bucolic interpolation, an image pyramid of ten scales, each a factor of 21/4 smaller than the previous is created.

3.2. S1 and C1Layers

Image layer is first analyzed by a multi-dimensional array of simple S1 units which correspond to the classical V1 simple cells of Hubel & Wiesel [46]. Mathematically, the weight vector w of the S1 units takes the form of a Gabor function which has been shown to provide a good model of simple cell receptive fields [47] and can be described by the following equation:

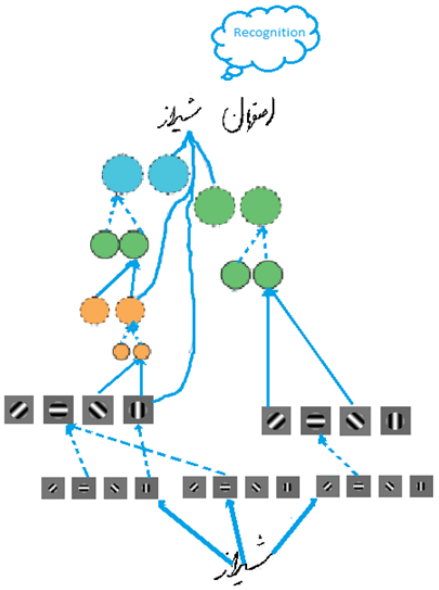
$$G(x,y) = \exp\left(-\frac{X^2+\gamma^2 Y^2}{2\sigma^2}\right) \cos\left(\frac{2\pi}{\lambda}\right)$$

Where $X = x\cos\theta - y\sin\theta$ and $Y = x\sin\theta + y\cos\theta$ and θ varies between 0 and π. Each of three parameters γ (aspect ratio), σ (effective width), and λ (wavelength) are set to γ =0.3, σ = 0.0036 * RF size+ 0.35 *RF size + 0.18 and λ = σ/0.8 (σ= 3.5 λ =2.8). Finally, the components of each filter are normalized so that their mean is 0 and the sum of their squares is 1. The response of a patch of pixels X to a particular S1 filter G is given by:

$$R(X,G) = \left|\frac{\sum X_i G_i}{\sqrt{\sum X_i^2}}\right|$$

The next C1 level corresponds to striate complex cells of Hubel and Wiesel [48]. This layer pools nearby S1 units (of the same orientation) to reach position and scale invariance over larger local regions, and as a result can also subsample S1 to reduce the number of units. The value of a C1 unit is simply the value of the maximum S1 units (of that orientation) that falls within a max filter. Similarly, position invariance is increased by pooling over S1 cells at the same preferred orientation but whose receptive fields are centered on neighboring locations.

The resulting C1 layer is smaller in spatial extent and has the same number of feature types (orientations) as S1. This layer provides a model for V1 complex cells.As a result, the size of the receptive fields increases from S1 to C1 layer (0.2-1.0 to 0.4-2.0 degree). Similarly the effect of the pooling over scales is a broadening of the frequency bandwidth from S1 to C1 units which is in agreement with physiological evidences [43].



Model Layer	RF Size	Description
Classification Units		
S4	7°	
C3	7°	6 Bands (1 to 6), 2000 Features
C2b	7°	Pooling over bands (1,2), (3,4) & (5,6)
S3	1.2°-3.2°	2000 3-by-3 Patches
S2b	0.9°-4.4°	2000 3-by-3, 5-by-5, 7-by-7, 9-by-9 Patches (500 per size)
C2	1.1°-3°	Bands (1 to 6), 2000 Features
S2	0.6°-2.4°	3-by-3 Filters in -45°,0°,45°,90°
1	0.4°-1.6°	6 Bands (1 to 6)
S1	0.2°-1.1°	Gabor Orientations: -45°,0°,45°,90°
Input Image		

Figure 4. Sketch of the model. There is an increase in invariance to position and scale, and in parallel, an increase in the size of the receptive fields as well as in the complexity of the optimal stimuli for the neurons

3.3. S2 andC2 Layers

At the S2 layer, units pool the activities of retinotopically organized complex C1 units at different preferred orientations over a 3×3 neighborhood of C1. As a result, the complexity of the preferred stimuli is increased: At the C1 level units are selective for single bars at a particular orientation, whereas at the S2 level, units become selective to more complex patterns (such as the combination of oriented bars to form contours or boundary) conformations. Receptive field sizes at the S2 level range between 0.6−2.4 degrees.

At the C2 Layer, units pool over S2 units that are tuned to the same preferred stimulus (they correspond to the same combination of C1 units and therefore share the same weight vector w) but at slightly different positions and scales. C2 units are therefore selective for the same stimulus as their afferent S2 units. Yet they are less sensitive to the position and scale of the stimulus within their receptive field. Receptive field sizes at the C2 level range between 1.1−3.0 degrees.

3.4. C3 andS3 Layers

Beyond S2 and C2 stages, the same process is iterated once more to increase the complexity of the preferred stimulus at the S3 layer (possibly related to Tanaka's feature columns in TEO [45]), where the response of C2 units with different selectivities are combined with a tuning operation to yield even more complex selectivities. In the next stage (possibly overlapping between TEO and TE), the complex C3 units, obtained by pooling S3 units with the same selectivity at neighboring positions and scales, are also selective to moderately complex features as the S3 units, but with a larger range of invariance. The S3 and C3 layers provide a representation based on broadly tuned shape components.

3.5. S2b and C2b Layers

S2b and C2b Layers may correspond to the bypass routes that have been found in visual cortex, e.g., direct projections from V2 to TEO (bypassing V4) [48] and from V4 to TE (bypassing TEO) [49]. S2b units combine the response of several retinotopically organized V1-like complex C1 units at different orientations just like S2 units. Yet their receptive field is larger (2-3 times larger) than the receptive fields of the S2 units. The tuning of the C2b units agrees with the read out data from IT [50].

4. K-Nearest Neighbor Classification

For the classification stage, we used the K-Nearest Neighbor (KNN). In this method, there is no training stage and training samples used as prototypes [19]. For each of test sample Euclidean distances with the all of training samples are calculated and unlabeled testing samples assigned to the classes based on their similarity with samples in the training set.

5. Experimental Results

The efficiency of the Biology method for feature extraction has been well tested via several experiments using a dataset of city names of Iran. The dataset which was used in this paper consists of 780 samples of 30 city names of Iran. For each city name 26 samples were available, which were written by 26 different persons. All of the samples are scanned at 96 dpi resolution in the gray scale format. Figure 5 shows some examples of this dataset.

For all the experiments, train and test samples have been chosen randomly.

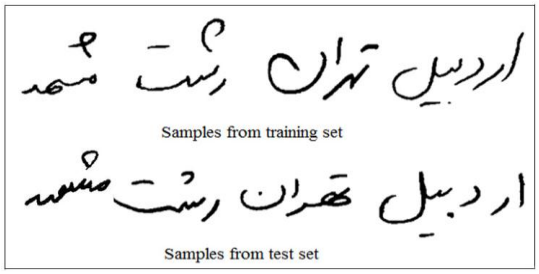


Figure 5. Samples of City Names of Iran from Training and Test Set.

By applying different experiments, the results of eight feature extraction method in the Farsi Hand Written Recognition (FHWR) domain via KNN classifier are comprised.

For many of the feature extraction methods which were used in this paper, there are some parameters that need to be adjusted. These parameters and how the values have been set are discussed here:

(i) Size of zones in Row Image Zoning (Z)
For word images with 184×320 pixels, Row Image Zoning was implemented with the following size of zones Z=20,40,64,80,128,160 and the 128 zones which gives the highest accuracy on the training set is chosen.

(ii) Size of sub image in DCT (D)
Different experiments were implemented and the 6*7 sub image that leads to an acceptable accuracy was the best choice.

(iii) Kind of wavelet basic function and number of levels in Wavelet (W)
Here the DWT2 function from Matlab toolbox has been used. The dwt2 command performs single-level two-dimensional wavelet decomposition and low frequency, level 4 and Haar wavelet analysis are being specified by the user in this work.

4.1. Experiment I

Classification results of the Biology feature extraction method have been compared with other seven feature extraction methods being described in section II, are summarized in Table1 and shown in Figure 6. In this experiment, feature spaces have been compared with different dimensions, due to high dimensionality of input space, KNN is used as a classifier. K domain changes in the range of 1,3,5,7,9,11. To increase the accuracy of the experiment, it had been repeated for 50 times and the average of these 50 times is shown as the result.

Table1. The average of Recognition rate of different features extraction methods with different K on our dataset

	Dimension	1	3	5	7	9	11	Average Accuracy	Rank
Wavelet	42	0.10407	0.06653	0.04856	0.04351	0.03864	0.03603	0.05622	8
DCT	42	0.14906	0.11898	0.11774	0.11387	0.11022	0.10739	0.11954	3
Gradient	640	0.12657	0.09221	0.09076	0.08779	0.08517	0.08285	0.09423	6
Histogram	504	0.31910	0.27836	0.29322	0.28908	0.28208	0.27793	0.28996	2
Loci	256	0.12782	0.09475	0.09700	0.09350	0.08936	0.08596	0.09807	5
Zoning	128	0.15090	0.11708	0.11408	0.11041	0.10684	0.10286	0.11703	4
CrossCount	504	0.12540	0.09123	0.08831	0.08470	0.08039	0.07759	0.09127	7
Biology	600	0.62956	0.56637	0.56034	0.54277	0.52437	0.50722	0.555109	1
Average Accuracy		0.21656	0.17819	0.17625	0.17070	0.1646	0.15973		
Rank		1	2	3	4	5	6		

The column named "average accuracy" in Table1 shows the average accuracy of each feature set. These values evaluate robustness of Biology feature extraction method with respect to other methods. Similarly, the row named "average accuracy" shows the average accuracy for different Ks in KNN classifier. For each value of K the recognition rate of feature extraction methods has been calculated, and then the highest recognition rate is chosen. In the "Rank" column the feature sets are put in order according to their average accuracies. The experimental results show that the Biology method outperforms other suggested methods .As sees in table1, K=1 leading to highest accuracy.

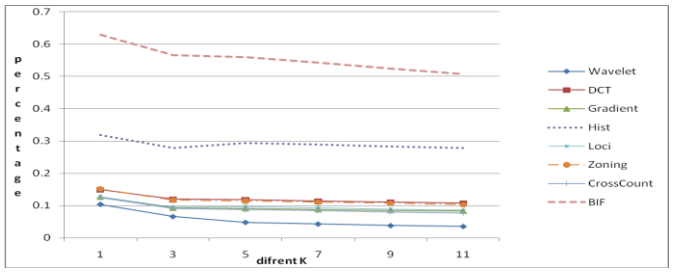


Figure 6. Shows the Recognition Rate of different Features with different K

At first, for all feature methods experiment has been done on 1 to 20 training samples with different Ks. As an example, the results of Biology feature are shown in Table2.

Table2. Recognition rate of Biology feature extraction method for different training sets from 1 to 20 using different Ks

K / No. of Samples	1	3	5	7	9	11
1	0.3286	0.1719	0.1199	0.0970	0.0836	0.0722
2	0.4218	0.3036	0.2795	0.2512	0.2238	0.1943
3	0.4958	0.3852	0.3618	0.3300	0.3024	0.2797
4	0.5258	0.4416	0.4270	0.4054	0.3804	0.3586
5	0.5655	0.4857	0.4799	0.4568	0.4328	0.4060
6	0.5973	0.5300	0.5170	0.4970	0.4757	0.4521
7	0.6130	0.5506	0.5387	0.5237	0.5030	0.4808
8	0.6379	0.5754	0.5683	0.5511	0.5320	0.5135
9	0.6433	0.5851	0.5858	0.5685	0.5504	0.5324
10	0.6670	0.6116	0.6076	0.5884	0.5736	0.5564
11	0.6694	0.6198	0.6147	0.6003	0.5865	0.5723
12	0.6804	0.6369	0.6414	0.6234	0.6100	0.5937
13	0.6952	0.6468	0.6499	0.6326	0.6187	0.6029
14	0.6972	0.6592	0.6647	0.6492	0.6316	0.6244
15	0.7003	0.6609	0.6664	0.6505	0.6405	0.6312
16	0.7062	0.6707	0.6699	0.6598	0.6460	0.6330
17	0.7264	0.6845	0.6964	0.6810	0.6685	0.6542
18	0.7336	0.6922	0.6964	0.6837	0.6635	0.6513
19	0.7387	0.7062	0.7039	0.6970	0.6768	0.6659
20	0.7480	0.7095	0.7177	0.7088	0.6877	0.6695

According to the table, it is obvious that as the training samples are increased, the recognition rate has been improved gradually. It can also be inferred from the average accuracy row, that K=1 is gives the highest performance result.

4.2. Experiment II

As explained in previous experiment the best result is from K=1. To show the robustness of biology feature, all experiments have been done with 1 to 20 train samples and K=1. The results are shown in Table 3 and Figure 7. It is obvious from Table 3 that Biology has the best result in average accuracy and even with 1 training sample Biology has the most recognition rate.

Table 3. Recognition Rate of Different Features with Different Training Samples

	DCT	Gradient	Hist	Loci	Wavelet	Zoning	CrossCount	BIF
1	0.1693	0.1399	0.1769	0.1547	0.0863	0.1626	0.1283	0.3286
2	0.1529	0.1233	0.2179	0.1375	0.0821	0.1523	0.1136	0.4218
3	0.1422	0.1138	0.2495	0.1116	0.0799	0.1412	0.1077	0.4958
4	0.1363	0.1130	0.2724	0.1061	0.0797	0.1391	0.1088	0.5258
5	0.1324	0.1041	0.2773	0.1048	0.0778	0.1313	0.1013	0.5655
6	0.1153	0.0964	0.2895	0.0867	0.0748	0.1230	0.0931	0.5973
7	0.1299	0.1015	0.3039	0.1035	0.0796	0.1280	0.1029	0.6130
8	0.1083	0.0927	0.3124	0.0907	0.0757	0.1095	0.0907	0.6379
9	0.1185	0.0998	0.3213	0.0941	0.0815	0.1186	0.1005	0.6433
10	0.1244	0.1041	0.3305	0.1083	0.0871	0.1307	0.1045	0.6670
11	0.1384	0.1121	0.3374	0.1111	0.0925	0.1383	0.1128	0.6694
12	0.1110	0.0976	0.3473	0.0952	0.0867	0.1135	0.0970	0.6804
13	0.1178	0.1066	0.3521	0.1051	0.0917	0.1200	0.1046	0.6952
14	0.1289	0.1142	0.3608	0.1194	0.1011	0.1300	0.1148	0.6972
15	0.1412	0.1234	0.3645	0.1273	0.1109	0.1436	0.1270	0.7003
16	0.1550	0.1369	0.3653	0.1533	0.1199	0.1517	0.1383	0.7062
17	0.1719	0.1545	0.3704	0.1519	0.1357	0.1773	0.1514	0.7264
18	0.2004	0.1708	0.3720	0.1667	0.1553	0.2032	0.1763	0.7336
19	0.2286	0.2008	0.3800	0.1952	0.1758	0.2299	0.2000	0.7387
20	0.2586	0.2260	0.3809	0.2333	0.2073	0.2742	0.2344	0.7480
Average Accuracy	0.1491	0.1266	0.3191	0.1278	0.1041	0.1509	0.1254	0.6296
Rank	4	6	2	5	8	3	7	1

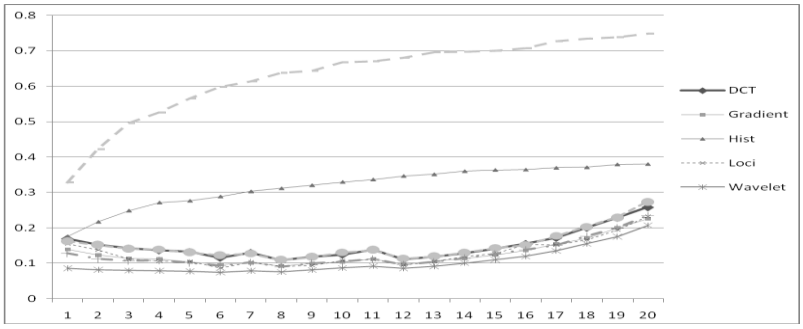


Figure 7. Shows the recognition rate for different feature extraction methods with different number of training samples. As the figure shows, Biology has the highest recognition rate even with 1 training sample and after that Histogram has the highest rate.

4.3. Experiment III

Some of the feature extraction methods that were used in this paper have high dimensions which lead to increase the computational load. For a better comparison between explained methods, Principal Component Analysis (PCA) is used to reduce the size of feature vectors then the classification stage is applied. The Results of different features with training samples from 1 to 20 with K=1 have been shown in table4. Different PCAs have been applied and the results showed that PCA=40 has the best performance. The results of this experiment have been shown in Table4 and Figure 8.

Table 4. The Result of recognition rate of different feature extraction methods with 1 to 20 training samples and K=1, using PCA to reduce the dimension.

	CrossCount	Gradient	Histogram	Loci	Wavelet	Zoning	DCT	BIF
1	0.1283	0.1399	0.1267	0.1357	0.0857	0.1357	0.1363	0.3227
2	0.1118	0.1234	0.1149	0.1219	0.0683	0.1219	0.1245	0.4245
3	0.1101	0.1137	0.1092	0.1140	0.0752	0.1140	0.1234	0.4953
4	0.1093	0.1130	0.1081	0.1088	0.0820	0.1088	0.1145	0.5332
5	0.1029	0.1041	0.1041	0.1008	0.0827	0.1008	0.1098	0.5751
6	0.0948	0.0965	0.1004	0.0921	0.0817	0.0921	0.1031	0.5954
7	0.1044	0.1016	0.1028	0.1042	0.0914	0.1042	0.0956	0.6103
8	0.0936	0.0927	0.0946	0.0899	0.0832	0.0899	0.0832	0.6290
9	0.1019	0.0998	0.0982	0.0976	0.0943	0.0976	0.0845	0.6481
10	0.1059	0.1040	0.1061	0.1013	0.1010	0.1013	0.0923	0.6705
11	0.1162	0.1121	0.1156	0.1103	0.1109	0.1103	0.1176	0.6709
12	0.0994	0.0976	0.0982	0.0936	0.0956	0.0936	0.1025	0.6818
13	0.1058	0.1066	0.1069	0.1023	0.1041	0.1023	0.0954	0.6904
14	0.1157	0.1142	0.1171	0.1123	0.1163	0.1123	0.1032	0.7028
15	0.1267	0.1235	0.1261	0.1244	0.1290	0.1244	0.1265	0.7107
16	0.1405	0.1369	0.1345	0.1332	0.1410	0.1332	0.1287	0.7152
17	0.1550	0.1545	0.1584	0.1486	0.1573	0.1486	0.1534	0.7199
18	0.1792	0.1708	0.1765	0.1690	0.1808	0.1690	0.1976	0.7321
19	0.2042	0.2006	0.2002	0.1989	0.2093	0.1989	0.2009	0.7319
20	0.2353	0.2260	0.2427	0.2256	0.2459	0.2256	0.2398	0.7440

As table shows Biology feature even by reducing the dimension to 42 has the highest rate.

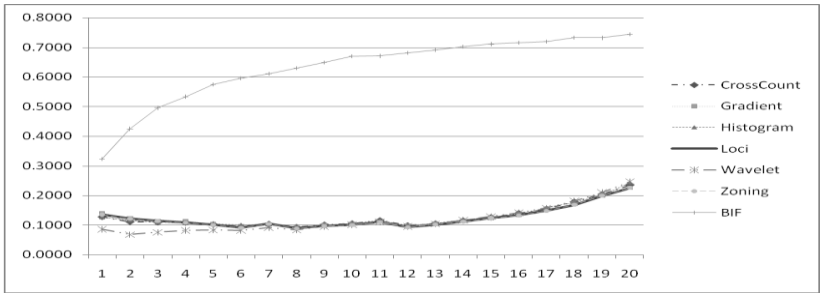


Figure 8. Recognition rate of different features by using PCA 40 on different number of training samples.

Where $d_{y,z}(x_i)$ is the degree of support given by the y'th classifier for the sample x_i of the class z. $Ind_c(x_i)$ has a value of one if the class of x_i is c, and zero, otherwise [41].

5. Conclusions

In this paper, various feature extraction methods for handwritten recognition were briefly described. Classification results of the Biology feature extraction method have been compared with other seven common feature extraction methods which being described in section II. The experiments were conducted on Iranshahr dataset and The experiments show that, even only as a computational tool to solve problems such as the handwritten word recognition problem, BIFs can boost the performance of Farsi handwritten word recognition system. The results also show the applicability of these features in Small Sample Size problems (SSS) and even "one sample per subject" problems.

References

[1] D. Deodhare, R. Amit, Preprocessing and Image Enhancement Algorithms for a Form-based Intelligent Character Recognition System, *International Journal of Computer Science and Application*, 2005, pp.131-144.

[2] P. plomondom, N. Srihari, On-line and Off-line Handwriting Recognition: A Comprehensive Survey, *IEEE PAMI, vol.1, n. 22*, 2005, pp. 63-84.

[3] T. Y. Zhang, C. Y. Suen, A fast parallel algorithm for thinning digital patterns, *Communications of the ACM, vol. 27, n. 3*, 1984, pp. 236 – 239.

[4] L. Lam, S. Lee, C.Y., Suan, Thinning Methodologies – A Comprehensive Survey, *Transactions on Pattern Analysis and Machine Intelligence (IEEE), vol.9, n. 14*, 1992, pp. 869–885.

[5] K. Cabeen, P. Gent, Image Compression and the Discrete Cosine Transform, College *of the Redwoods.*

[6] D.J. Romero, L.M. Seijas, Directional Continuous Wavelet Transform Applied to Handwritten Numerals Recognition Using Neural Networks, *JCS&T, Vol. 7 No. 1*, 2007.

[7] P. Moallem, B. Mirzaeian Dehkordi, S.M. Shirvani Boroujeni, Determination of Number of Broken Rotor Bars in Squirrel-Cage Induction Motors Using Wavelet, PCA and Neural Networks, *IREE, Vol. 4. n. 1*, 2009, pp. 242-248.

[8] M. Wienecke, G. A. Fink, G. Sagerer., Experiments in Unconstrained Offline Handwritten Text Recognition, *Workshop on Frontiers in Handwriting Recognition,. IEEE*, 2002.

[9] Y. Wen. Y, Lu. P, Shia, Handwritten Bangla numeral recognition system and its application to postal automation, *Pattern Recognition, n.40* , 2007, pp.99–107.

[10] - M. Hanmandlu, J. Grover, V. K. Madasu. S. Vasikarla, Input fuzzy for the recognition of handwritten Hindi numeral, *International Conference on Informational Technology, vol. 2*, 2007, pp. 208-213.

[11] S.V. Rajashekararadhya, P. V. Ranjan, V.N. Manjunath Aradhya, Isolated handwritten Kannada and Tamil numeral recognition: A novel approach, *First International Conference on Emerging Trends in Engineering and Technology ICETET*, n.8, 2008, pp.1192-1195.

[12] Ch.Lin. Liu, Ch. Y. Suen, A new benchmark on the recognition of handwritten Bangla and Farsi numeral characters, *Pattern Recognition*, 2008. 666666

[13] L-F Chen, H-Y Mark Liao, M-T Ko, J-C Lin, and G-J Yu, A New LDA-based face recognition system which can solve the small sample size problem, *pattern recognition, vol. 33*, 2000, pp. 1713-1726.

[14] Sh., Suen, Ch.Y, Yamamoto k, Historical review of OCR research And developing, *IEEE.* 1992, pp.18-9219.

[15] Sh. Abdleazeem, E. EL-sherif, Arabic handwritten digit recognition, *IJDAR, vol.*11, pp.127-141, (2008).

[16] V.N. Manjunath, A. Aradhya, G.H. Kumara, S. Noushatha., Multilingual OCR system for South Indian scripts and English documents An approach based on Fourier transform and principal component analysis, *Engineering Applications of Artificial Intelligence, Vol. 21, n. 4*, 2008, pp 658-668.

[17] T anaka, K.: Inferotemporal cortex and object vision. Annu. Rev. Neurosci. vol. 19, pp. 109--139 (1996)

[18] Hubel, D.H., Wiesel, T.N.: Receptive fields of single neurons in the cat's striate cortex. J. Physiol. vol. 148 pp.574--591 (1959)

[19] Riesenhuber, M., Poggio, T.: Hierarchical models of object recognition in cortex.Nat.Neurosci.vol. 2, no. 11, pp. 1019--1025 (1999)

[20] Serre, T., Wolf, L., Poggio, T.: Object recognition with features inspired by visual cortex. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). vol. 2, pp. 994--1000 (2005)

[21] Serre, T., Kouh, M., Cadieu, C., Knoblich, U., Kreiman, G., Poggio, T.: A Theory of Object Recognition: Computations and Circuits in the Feedforward Path of the Ventral Stream in Primate Visual Cortex. AI Memo 2005-036/CBCL Memo 259, Massachusetts Inst. of Technology, Cambridge (2005)

[22] Ni. D. Xiao, Application of Neural Networks to Character Recognition, Pace University, 2007.

[23] Z. Liu, H. Bozdogan, RBF Neural Network for classification using new Kernel functions, Advances in intelligent systems and applications.

[24] A. Baraldi, N. Borghese, Learning from data: general issue and special applications of Radial Basis Function Networks, *Technical Reports, CS department, UC Berkeley*, 1998.

[25] H. Bozdogan, Z. Liu, RBF Neural Networks for classification using new kernel functions, *Second international workshop on Intelligent systems design and application*, 2002, pp.17-22.

[26] V. L, Lajish, T.K.K, Suneesh, N.K. Narayanan, Recognition of Isolated Handwritten Character Images using Kolmogrov- Smirnov Statistical Classifier and k-nearest Neighbour classifier, *Prof of the international Conference on COGNITION AND RICOGNITION ICCR*, 2005.

[27] B.V. Dhandra., M. Hangarge., On Separation of English Numerals from Multilingual Document Images, *JOURNAL OF MULTIMEDIA, VOL. 2, NO. 6* ,2007.

[28] G.D. Joshi., S. Garg., J. Sivaswamy , Script Identification for Indian Documents , *In. Pro. of 7th IAPR workshop on Document Image Systems, (DAS),* 2006,pp.255-267,

[29] B. Philip., R. D. Sudhaker Samue **,** An Efficient OCR for Printed Malayalam Text using Novel Segmentation Algorithm and SVM Classifiers, *International Journal of Recent Trends in Engineering, Issue. 1, Vol. 1*, 2009.

[30] S. Hua., Z. Sun., Support vector machine approach for protein subcellular localization prediction. *Bioinformatics, 17:721–728*, 2001.

[31] R.Ebrahimpour ,M.R.moradian, A.Esmkhani,F.M.Jafarlou, Recognition of Persian handwritten digit using characterization loci and mixture o experts, *International journal of Digital content Technology and its Applications, Vol. 3, September* 2009, n. 3.

[32] O.D. Trier, A.K. Jain , T. Taxt , Feature extraction methods for character recognition—a survey. *Pattern Recognition, 29,* 1996, pp. *641–662*).

[33] M. Bokser., Omnidocument Technologies. *Proceedings of the IEEE, vol.80, July 1992,pp. 1066-1077.*

[34] R.C. Gonzales, R.E. Woods, *Digital image processing*.(2nd end. Addison-Wesley, Reading, MA, USA ,2002*).*

[35] P. Zhang, T. Bui, C. Suen , Hybrid feature extraction and feature selection for improving recognition accuracy of handwritten numerals. *Proc. 8th ICDAR,* 2005.

[36] G.Li. Ashfaq, A. Khokhar, Content-based Indexing and Retrieval of Audio Data using Wavelets. *IEEE international conference on MULTIMEDIA AND EXPO , vol.2,*2000, pp.884-888.

[37] N. Ahmad, T. Natarjan , K,Roa, Discrete Cosine Transform. *IEEE. Trans Compute., Vol C-23,1974, PP 90-93.*

[38] C.L. Liu, Y.J. Liu, R.W. Dai, Preprocessing and statistical/structural feature extraction for handwritten numeral recognition, *In: A.C. Downton, S. Impedovo (Eds.), Progress of Handwriting Recognition, World Scientific*, 1997, pp. 161–168.

[39] H. Soltanzadeh, M. Rahmati, Recognition of Persian handwritten digits using image profiles of multiple orientations, *Science direct. Pattern Recognition Letters. 25*,2004, pp.1569–1576.

[40] A. Webb, *Statistical Pattern Recognition*, (Oxford university, 1999).

[41] P.N. Belhumeur, Eigenfaces vs. Fisherfaces , recognition using class special linear projection , *IEEE Trans. Pattern Anal. Mach. Intell,* 1997, 19 (7) 711–720.

[42] H. Yu, J. Yang, A direct LDA algorithm for high-dimensional data—with application to face recognition, *Pattern Recognition 34 (11),* 2001, pp.2067–2070.

[43] P.N. Belhumeur, recognition using class speci9c linear projection. *IEEE Trans. Pattern Anal. Mach. Intell. 19 (7),* 1997, pp.711–720.

[44] . Hubel, D.H., Wiesel, T.N.: Receptive fields and functional architecture of monkey striate cortex. J. Physiol. vol. 195, no. 1, pp. 215--243 (1968).

[45] 16. Perrett, D.I., Oram, M.W.: Neurophysiology of shape processing.Img. Vis. Comput. vol. 11, pp. 317--333 (1993).

[46] 17. Tanaka, K.: Inferotemporal cortex and object vision.Annu. Rev. Neurosci. vol. 19, pp. 109--139 (1996)

[47] 19. Jones, J.P., Palmer, L.A.: The two-dimensional spatial structure of simple receptive fields in cat striate cortex. J. Neurophysiol. vol 68, no. 6, pp. 1187--1211 (1987)

[48] 20. Boussaoud, D.R., Ungerleider, L.G., Desimone,R.: Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. J. Comp. Neurol.vol. 296, pp. 462--495(1990)

[49] 21.Desimone, R., Flemig, J., Gross, C.G.:Prestriate afferents to inferior temporal cortex: an HRP study. Brain Res.vol. 184, pp.41--55(1980)

[50] 22. Hung, C.P., Kreiman, G., Poggio, T., DiCarlo, J.J.: Fast readout of object identity from macaque inferior temporal cortex. Science vol. 310 pp. 863—866

Authors

Reza Ebrahimpour is an Assistant Professor at the Department of Electrical and Computer Engineering, Shahid Rajaee Teacher Training University, Tehran, Iran. He is also a research scientist at School of Cognitive Sciences, Institute for Research in Fundamental Sciences, Tehran, Iran. He received the B.S. degree in Electronics Engineering from Mazandaran University, Mazandaran, Iran and the M.S. degree in Biomedical Engineering from Tarbiat Modares University, Tehran, Iran, in 1999 and 2001, respectively. He obtained Ph.D. degree in the field of Computational Neuroscience from the School of Cognitive Science, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran in July 2007. Dr. Ebrahimpour is the author or co-author of more than 60 international journal and conference publications in his research areas, which include Human and Machine Vision, Computational Neuroscience, Neural Networks, Pattern Recognition and Multiple Classifier Systems.

Mona Amini was born in Tehran Iran, in August 1984. She received the BS degree in Electronic engineering from Azad University, Qazvin, Iran, in 2007. She is student in MS degree of Mechatronic Engineering in Azad University, Tehran, Iran. Her research interests include human and machine vision, neural networks and pattern recognition.

Afra Vahidi Shams was born in Tehran, Iran (Octobr 1984). He received the B.Sc. degree in Electronics engineering from Mazandaran University (Noshirvani University of Technology), Babol, Iran. In sep. 2007 and his M.Sc. degree in Mechatronics engineering from Islamic Azad University of Tehran (south branch), Tehran, Iran in sep. 2010. His research interests include Intelligent Robotic Systems and Humanoids, Human and Machine Vision, Biologically Motivated Object Recognition, Computational Neuroscience.