International Journal of Current Engineering and Technology
ISSN 2277 - 4106

Available at http://inpressco.com/category/ijcet

Research Article

SYRIAC Character Recognition System Using Back Propagation Neural Network

Basima Z.Yacob[Á*], Majd S. Mati[Á] and Danny T. Baito[Á]

[Á]Department of Computer Science, University of Duhok –Duhok –Iraq

Accepted 10 December 2013, Available online 25 Dec 2013, **Vol.3, No.5 (December 2013)**

Abstract

The Syriac language is one of the Semitic languages that is being spoken in Iraq, Syria, Turkey and Iran by Assyrians. It's an ancient language, one of the rarest and oldest in the world. In this paper A Syriac character recognition system using Back propagation Neural Network is proposed. A pre-processing step is implemented to separate each character from the others. After that a feature extraction process is applied on each character to obtain the invariant moments. Back propagation neural network is trained on invariant moments of East Syriac alphabet, and tested on these characters to verify each character image belongs to which type of character. This is done by using visual C#.

Keywords: *Neural networks, Syriac alphabet, invariant moments, Back Prorogation Neural Network.*

1. Introduction

Neural networks have been applied to a wide variety of areas including speech synthesis, character recognition, diagnostic problems, medicine, business and finance, robotic control, signal processing, computer vision and many other problems that fall under the category of pattern recognition (Ingrid R., *et al,2005)*. Neural networks have been shown to be particularly useful in solving problems where traditional artificial intelligence techniques involving symbolic methods have failed or proved inefficient, Such networks have shown promise in problems involving low-level tasks that are computationally intensive, including vision, character recognition, speech recognition, and many other problems that fall under the category of pattern recognition (Fakhraddin M., *et al* 2006).

Character recognition is one of the most challenging topics in pattern recognition, In the past several decades, a large number of OCR systems have been developed for natural languages(S. Sardar, A. et al 2010) (M.Soleymani ,*et al* 2003) (B. B. Chaudhuri, *et al* 1998).

Character recognition, both human and machine generated, is a wide and largely studied field in Machine Learning. In fact, nowadays many commercial scanners use Optical Character Recognition (OCR) systems that output a character string having an image of typed text as input(Farah, R.A., *et al* 2005). Back propagation neural network is widely used in the field of pattern recognition because this artificial neural network can classify complex pattern and perform nontrivial mapping function. Neural network are used in pattern recognition because of their ability to learn and to store knowledge e Because of their parallel nature can achieve (Ms Jyoti C Chaudhari 2010).

The Syriac alphabet consists of 22 characters which is written from right to left (Rev. Shlemon I. Khoshaba 2010). Syriac is an ancient Iraqi language, and it is culturally used by human beings in Iraq. It has many religious scripts as well as scientific and literary books which have been completed and achieved throughout the long history and efficient civilization for this language, the problem of Syriac character recognition has been rarely addressed (Abdul Monem S. Rahma, *et al* 2013). In this paper, we propose East Syriac character recognition method that uses a Back-propagation neural net to recognize the characters. The inputs of the Back-propagation neural net are invariants moments features vectors of East Syriac alphabet that were computed in (Abdul Monem S. Rahma, *et al* 2013).

2. Artificial Neural Network

Neural network are simplified models of the biological nervous system and therefore have drawn their motivation from the kind of computing performed by a human brain. An NN in general is a highly interconnected of a large number of processing elements called neurons in an architecture inspired by the brain. An NN can be massively parallel and therefore is said to exhibit parallel distributed processing. Neural Network exhibits characteristics such as mapping capabilities or pattern association, generalization, robustness, fault tolerance, and parallel and high speed information processing. Neural network learn by example. They can therefore be trained

*Corresponding author: **Dr. Basima Z.Yacob**

with known examples of a problem to acquire knowledge about it. Once appropriate trained the network can be put to effective use in solving 'unknown' or 'untrained' instances of the problem. Neural network adopt various learning mechanism of which supervised learning and unsupervised learning methods have turned out to be very popular. In supervised learning, a teacher is assumed to be present during the learning process, i.e. the network aims to minimize he error between target (desired) output presented by the teacher and the computed output to achieve better performance. However, in unsupervised learning, there is no teacher present to hand over the desired output and the network therefore tries to learn by itself, organizing the input instances of the problem. NN Architecture has been broadly classified as single layer feed forward networks, multilayer feed forward networks and recurrent networks, over the year several other NN Architecture have evolved , some of the well known NN system include backpropogation network, perceptron, ADALINE, Boltzmann machine, adaptive resonance theory, Self-organized feature map, and Hopfield network. Neural Network has been successfully applied to problem in the field of pattern recognition, image processing, data compression, forecasting and optimization to quote a few (S.P.Kosbatwar ,*et al* 2012).

2.1. Back propagation Network

Back propagation is a common method of training artificial neural networks so as to minimize the objective function. It is a supervised learning method, and is a generalization of the delta. It requires a dataset of the desired output for many inputs, making up the training set. This learning algorithm is applied to multilayer feed-forward networks consisting of processing element with continuous differentiable activation function (networks that have no feedback or simply, that have no connections that loop). The term is an abbreviation for "backward propagation of errors". The aim of neural network is to train the net to achieve balance between the net's ability to respond & its ability to give reasonable responses to the input that is similar but not identical to the one that is used in training. The back propagation algorithm is different from the networks in respect to the process by which the weights are calculated during the learning period of the network. The training of BPN is done in 3 stages:-feed-forward of the input training pattern, the calculation and back propagation of the error & updating of weights. The testing of the BPN involves the computations of feed-forward only, there can be more than one hidden layer. The training process is very slow but once the network is trained it can produce its output very rapidly(S S Sayyad, *et al* 2013).

3. Proposed Syriac Character Recognition System

The construction of East Syriac Character Recognition system consists of several phases are as follows:

3.1Image Acquisition

In Image acquisition, the recognition system acquires a scanned image as an input image. The image should have a specific format such as JPEG, BMT etc. This image is acquired through a scanner, digital camera or any other suitable digital input device.

3.2 Pre-processing

The pre-processing is a series of operations performed on the scanned input image. It essentially enhances the image rendering it suitable for segmentation. The various tasks performed on the image in pre-processing stage for example, Binarization process that converts a gray scale image into a binary image using global thresholding technique.

3.3. Feature Extraction

This is performed in order to capture the vital characteristics of the symbol. Here, the seven moments of East Syriac alphabet are extracted, Table 1 shows the seven moments East Syriac (Abdul Monem S. Rahma, *et al* 2013).

3.4. Classification and Recognition

The seven moments feature vectors computed in the feature extraction stage of East Syriac alphabet is fed into an artificial neural network for character recognition. The stages are summarized in Fig.1 below.

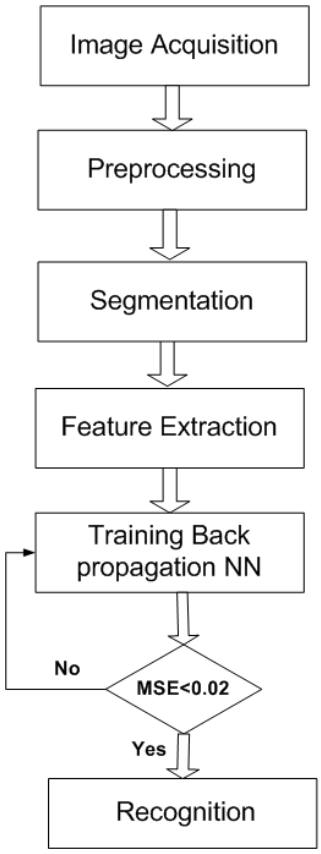

Fig.1 East Syriac Character Recognition System

Table 1: Seven Invariant Moments for East Syriac alphabet (Abdul Monem S. Rahma, et al 2013).

	Φ1	*Φ2*	*Φ3*	*Φ4*	*Φ5*	*Φ6*	*Φ7*
ܐ	0.181384	0.000633	4.61E-05	2.23E-05	8.59E-11	-5.54E-07	-1.32E-10
ܒ	0.220627	5.81E-05	4.07E-06	0.000292	3.82E-08	-5.09E-07	-7.29E-08
ܓ	0.2394	0.000455	0.000249	2.56E-05	-2.01E-09	1.04E-08	1.10E-09
ܕ	0.192234	0.001389	0.000266	6.27E-05	-1.13E-09	-2.32E-06	2.16E-09
ܗ	0.199315	0.000798	0.000133	0.000141	-1.17E-08	-2.48E-06	-1.84E-08
ܘ	0.1943	0.001087	0.000238	8.44E-05	-1.03E-08	-2.28E-06	9.61E-10
ܙ	0.178957	0.001108	0.000288	1.41E-05	-5.39E-10	-4.62E-07	4.49E-10
ܚ	0.193554	0.000209	3.95E-05	0.000129	1.35E-08	-5.04E-07	-1.67E-08
ܛ	0.267302	0.001361	0.000168	3.50E-05	-2.93E-10	-1.06E-06	-1.39E-09
ܝ	0.182153	0.000312	4.99E-05	3.78E-05	-6.46E-10	-2.22E-07	-1.73E-09
ܟ	0.219983	0.000148	3.08E-07	0.000243	2.30E-08	5.10E-07	-5.99E-08
ܠ	0.226912	0.000979	0.000182	0.000178	-1.80E-08	-5.08E-06	4.51E-09
ܡ	0.21297	0.000916	0.000122	0.000119	-1.23E-09	-3.58E-06	4.60E-09
ܢ	0.238187	0.000105	8.90E-05	4.39E-05	2.00E-09	3.96E-07	3.66E-10
ܣ	0.249598	0.000298	2.03E-05	0.000387	9.15E-08	6.50E-06	-5.09E-08
ܥ	0.213008	0.0008	0.000314	0.000215	5.93E-08	-2.52E-06	-6.96E-08
ܦ	0.206189	0.000182	6.30E-07	0.000154	2.67E-09	-3.69E-07	-3.04E-08
ܨ	0.177858	0.000914	0.000233	1.25E-05	-5.07E-10	-2.73E-07	2.40E-10
ܩ	0.17423	0.000278	6.48E-05	2.57E-06	2.92E-11	2.52E-08	-2.55E-11
ܪ	0.223071	0.000203	1.54E-05	0.000352	5.37E-08	-8.49E-07	-1.29E-07
ܫ	0.235799	0.000524	2.74E-07	6.29E-05	1.93E-10	-3.17E-07	-3.28E-09
ܬ	0.235394	0.000238	8.63E-06	0.000149	4.24E-09	-8.52E-07	-1.74E-08
ܐ	0.181221	0.000835	0.000194	1.98E-05	-1.07E-09	-5.35E-07	6.90E-10
ܒ	0.216285	0.000151	5.88E-06	0.000318	3.90E-08	-2.08E-06	-1.09E-07
ܓ	0.191217	0.001322	0.000223	5.73E-05	-4.26E-10	-2.06E-06	1.18E-09
ܕ	0.221562	3.02E-05	9.32E-06	0.000267	3.66E-08	8.72E-07	-3.50E-08
ܗ	0.197171	0.000707	7.80E-05	8.34E-05	-3.64E-09	-1.52E-06	-2.01E-09

4. Experimental results

In this paper ,We presented East Syriac character recognition system trained using back-propagation algorithm, The proposed system is implemented by using Microsoft Visual C# 2012,the BPNN network that is implemented is composed of three layers input layer, output layer and hidden layer. The input layer constitutes of 7 neurons which receive the moments feature vector for each Syriac character. The hidden layer constitutes of 20 neurons. The output layer is composed of 27 neurons number of characters as shown in Fig 2. The learning rate and momentum parameter were set to 0.1 and 0.3 respectively for the backpropagation model with sigmoid as an activation function. Fig 3 shows the window of parameters used for designing and training BPNN, Fig 4 is a plot between performance parameter MSE and its improvement with increased iterations, and Table 2 shows the result of BPNN Training. We test Syriac alphabet with various orientations (with rotation angle 0º,25º, 45º,75º, 90 º,120º, 160º, 180, 200º, 250º, 270º, 300º and 330º) the recognition rates were 100%. Table 3 shows the seven moments for "***Alap***" Syriac character with different rotation angles.

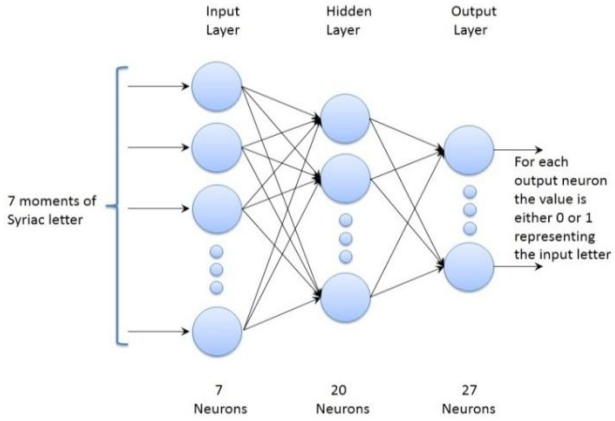

Fig 2: Network architecture for Syriac Characters Recognition System

Fig 3: window of parameters used for designing and training BPNN.

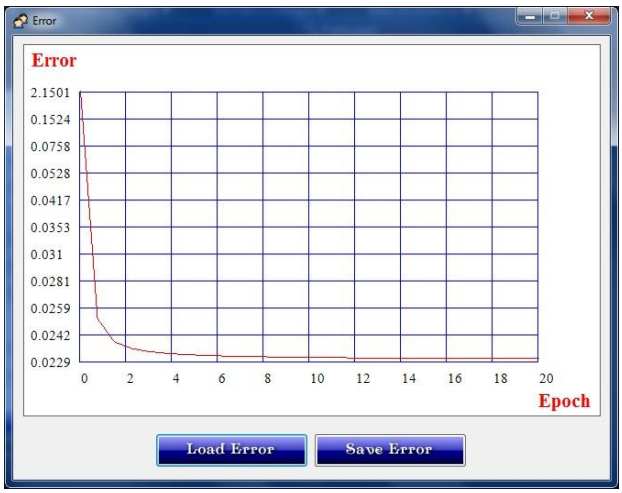

Fig 4: performance parameter MSE and its improvement with increased iterations.

Table 2: The Training BpNN ‘s output of East Syriac Characters .

Letter	Output in binary for each letter used in training																											
ܐ	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܒ	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	s	0	0	0	0	0	0	0	0	0
ܓ	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܕ	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܗ	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܘ	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܙ	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܚ	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܛ	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܝ	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܟ	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܠ	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܡ	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܢ	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ܣ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
ܥ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
ܦ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
ܨ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
ܩ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
ܪ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
ܫ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
ܬ	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
‘	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
‘	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
‘	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
‘	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
‘	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
‘	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Table 3: Seven Invariant Moments for “*Alap*” Syriac character with different rotation angles between 0º and 360º

Moment / Rotation	$\phi 1$	$\phi 2$	$\phi 3$	$\phi 4$	$\phi 5$	$\phi 6$	$\phi 7$
0º	0.181384	0.000633	4.61E-05	2.23E-05	8.59E-11	-5.54E-07	-1.32E-10
25º	0.181658	0.000611	4.41E-05	2.21E-05	-1.05E-09	-5.36E-07	1.99E-10
45º	0.180647	0.000496	4.85E-05	2.23E-05	-9.77E-10	-4.79E-07	6.15E-10
75º	0.181432	0.000636	4.05E-05	2.20E-05	-4.00E-10	-5.49E-07	4.77E-10
90º	0.181384	0.000633	4.61E-05	2.23E-05	-4.80E-10	-5.54E-07	9.91E-10
120º	0.18108	0.000659	4.99E-05	2.08E-05	4.68E-11	-5.20E-07	-7.59E-10
160º	0.182086	0.000649	4.28E-05	2.42E-05	-5.75E-10	-6.11E-07	6.33E-10
180º	0.181384	0.000633	4.61E-05	2.23E-05	8.59E-11	-5.54E-07	-1.32E-10
200º	0.181248	0.000624	4.62E-05	2.14E-05	-7.87E-10	-5.29E-07	-4.19E-11
250º	0.182086	0.000649	4.28E-05	2.42E-05	-4.37E-10	-6.11E-07	3.32E-10
270º	0.181384	0.000633	4.61E-05	2.23E-05	-4.80E-10	-5.54E-07	9.91E-10
300º	0.18108	0.000659	4.99E-05	2.08E-05	4.68E-11	-5.20E-07	-7.59E-10
330º	0.180859	0.000595	4.18E-05	2.06E-05	-6.93E-10	-4.98E-07	2.23E-10

Conclusions

The features of each East Syriac character are extracted by using moments to build database and using this data input nodes to the network. The Back Prorogation Neural Network is able to train the character data to verify that each character image belongs to which character. The system is able to recognize the samples of characters in training with 27 character images of 100% training recognition rate and testing with 351 Syriac character images(with rotation angle 0º,25º, 45º,75º,90 º,120º,160 º,180,200º,250º,270º,300ºand 330º)of 100% testing recognition rate.

References

Ingrid R., Zdravko M., Todd N.(2005) ,Character Recognition and Learning with Neural Networks ,*UMAP Journal*, Vol. 14, No.1

Fakhraddin M., Jamal F.(2006), Character Recognition using Neural Networks, Near east university, North Cyprus, Turkey via Mersin-10,KKTC, *http://ww1.ucmss.com/books/LFS/CSREA2006/ICA5025.pdf*

S. Sardar, A. Wahab.(2010), Optical Character Recognition system for Urdu, *International conference on Information and Emerging Technologies* .

M.Soleymani and F.Razzazi(2003) ,An efficient front-end system for Isolated Persian/Arabic character Recognition of handwritten data-Entry Forms, *International Journal Of Computational Intelligence*, vol 1, pp.193-196.

B. B. Chaudhuri and U. Pal, (1998),A Complete Printed Bangla OCR System, *Pattern Recognition.* Vol.31, pp.531-549.

Farah, R.A., M.T. Khadir and M. Sellaim, (2005),Artificial Neural Network Fusion: Application to Arabic words Recognition, *ESANN 2005 Proc., Bruges (Belgium)*, pp. 27-29.

Ms Jyoti C Chaudhari ,(2010), Design of Artificial Back Propagation Neural Network for Drug Pattern Recognition, *International Journal on Computer Science and Engineering (IJCSE)*, Special Issue.

Rev. Shlemon I. Khoshaba,(2010) ,Lessons in the teaching of the Syriac language*”, AL-Mashrq House Cultural, Duhok – Iraq.*

Abdul Monem S. Rahma, Basima Z.Yacob and Danny T. Baito ,(2013),Moment invariants Based Features Extraction For Classification of Syriac Alphabet Language, *International Journal of Advances in Engineering and Technology*, Vol. 6, Issue 4, pp. 1442-1451.

S.P.Kosbatwar and S.K.Pathan(2012),Pattern Association for Character Recognition by Back-Propagation Algorithm Using Neural Network Approach, *International Journal of Computer Science & Engineering Survey (IJCSES)* Vol.3, No.1, pp 127-134.

S S Sayyad, Abhay Jadhav, Manoj Jadhav, Smita Miraje, Pradip Bele, Avinash Pandhare(2013) ,Devnagiri Character Recognition Using Neural Networks, *International Journal of Engineering and Innovative Technology (IJEIT)* Vol 3, Issue 1,pp 476-480.