On handwritten character recognition through locally connected structural neural networks

Pilar Gomez-Gil[1], Juan Manuel Ramírez[2], William Oldham[1]
1. Computer Science Department, Texas Tech University USA.
2. Electrical Engineering Department. Universidad de las Américas Pueb la, Mexico

Abstract - The problem of handwritten character recognition is still a big challenge for the scientific community. Several approaches have been attempted with that purpose in the last years. In this context, algorithms based on neural networks have been proved to give better results than conventional methods, when applied to problems where the decision rules of the classification problem are not clearly defined. In this paper, results obtained from an off-line handwritten character recognition, writer-independent, are presented. The neural network constructed consists of a locally connected structure based on a multilayer feed-forward perceptron, with a modified back propagation training algorithm, originally proposed by Le Cun. The main idea behind this approach is the generalization obtained when the structure of the network is adjusted to the problem under study, using a knowledge a priori about the general characteristics of the feature space. The neural network is organized through several characteristic map levels. Each characteristic map defines a neuron cluster in the hidden layers, which extract local characteristics from the low levels. The performance of the system is compared with a classical totally connected backpropagation feed-forward neural net. Concluding remarks concerning learning and recognition rates using different network sizes, are presented.

Introduction.

In the last decades there has been strong advances in the field of pattern recognition with applications to a wide variety of knowledge areas, however, there still are perception problems difficult to be solved by computers. Handwritten recognition is one of those problems. Several approaches have been explored [1,2], but in most cases the conditions of the system have to be limited using, for instances, a restricted set of characters, a small database, or a single writer. Neural networks based pattern recognizers have provided better results than conventional methods when they are applied to problems where the decision rules of the classification are unknown. In this paper results from an off-line writer independent handwritten character recognizer are analyzed. The recognition system consists of a locally connected back propagation neural net originally proposed by Yann Lecun [3] The obtained results are compared with a standard backpropagation neural net totally connected.

Artificial neural networks, also known as connectionist models or parallel distributed processors, are mathematical models inspired in biological neurons, adapted and simulated in sequential computers. Even though this is a rough representation, ANN show some characteristics of the brain:

a) Learning. A neural network modify his behavior in response to the environment.
b) Generalization. A trained ANN is tolerant to small input variations, i.e., it can face the ‘imperfections of the real world. This feature is essential to the implementation of a pattern recognizer.
c) Abstraction. Some ANN are able to extract the essence from an input set.

Information processing in an ANN is performed through the iteration of a number of simple elements called neurons. Neurons are connected through some links associated to numerical values or weights, in such a way that each neuron can send excitatory or inhibitory signals to others. This connectivity between neurons is the keystone of an ANN.

Handwritten character recognition can be divided in two areas: on-line recognition, performed at the moment in which the action is executed, and of-line recognition, which involves digitization of some text previously written,

1.Computer Science, Texas Tech University, Lubbock Texas 79409, USA.
2.Electrical Enginering Department, Universidad de las Americas, Cholula Puebla 72820, A.P. 396, Mexico

Published at Procceding of the Second joint Mexico US Internacional Worshop on N. Nets and Neurocontrol. Sian Ka’an 97. pp. 24-228. © August 1997

character segmentation, and the recognition itself. In handwritten recognition we find different levels of complexity. When the recognition is performed off-line, author-independent, and with script lower case writing, the problem is specially difficult. On-line recognition is less complex because additional information about the character to be recognized, can be obtained at the moment in which this is generated. In addition, printed-type characters and digits, tend to be more uniform than script characters, which makes them more easy to handle. In script characters there is a wide variety of fonts, sizes, and leaning, not to mention the additional ornaments introduced by the authors. Several factors such as ethnic origin, social level, education, and age, strongly contribute to the lack of uniformity among characters; even the same person show strong variations due to tiredness, mood, or environmental situations, which, however, are almost completely discriminated by the human brain. It has been shown that human are able to distinguish correctly up to 96% of handwritten characters [4]. An important problem to be solved in the implementation of a recognizer is the selection and design of a vector of characteristics. Learning by samples appear to be an adequate solution to problems where it is not evident to extract the rules of representation of a data set. Furthermore, it is possible to construct a neural network classifier without selecting a vector of characteristics, but feeding it with the original data, and allowing the network to build its own internal representation [5].

Le Cun Neural Network Model.

This model shows characteristics with special relevance to the problem of handwritten recognition. The general architecture is showed in figure 1. It basically consists of a multilayer feed-forward perceptron with partial connections between neurons, and trained with a backpropagation algorithm. This network was used to build a zip code recognizer for the US mail system with excellent results [5]. In that reference, it was shown that a neural network provide a better generalization if its structure is adjusted to the specific problem, using a priori knowledge about it. The objective is to reduce the number of weights to be trained in the learning process, without reducing capacity. The structure of the model is based on several levels of characteristic maps. These maps are neuron clouds in the hidden layers which synthesize local characteristics from the lower levels. This architecture is inspired in the way in which the mammals perform the visual process. The same idea can be found in some other neural networks such as the neocognitron [6] , and the Linsker perceptual network [7]. The extractors consists of neurons which evaluate the same characteristic in different areas of the input image. This is obtained by forcing to the neurons in the same characteristic map to share the same weights. Knowing the precise location of some characteristic in a given character it is not needed, because each neuron extracts information from small overlapping windows in the lower levels. This approach attempts to handle information from the input image in a coarse way, providing some tolerance to rotation and translation of the handwritten symbol. Configuration of the network is as follows: Input level is feed with an NXN matrix with the pixel values corresponding to the digitized input image. The hidden levels are formed by two dimensional arrays with the characteristic maps. Each unit in the level ‘y’ receive an information input from an WiXWi window, where Wi<N, from the immediate lower level i-1. Maps in the i´th level have same size input windows, however to different levels corresponds different window sizes. The connections corresponding to each unit in some specific map are shared by all maps in the same level, but neuron thresholds are not shared. In this way, all the map units extract the same characteristic and the number of weights to be trained is reduced to WixWi plus the number of thresholds. Each neuron in the map takes its input from a different window overlapping each other. In a given map, neurons one pixel apart have their corresponding windows with a separation of two pixels. This situation causes the lost of some position information related to the characteristic learned. This kind of connectivity resembles the convolution operation frequently used in image processing, where the networks weights act as a testing mask applied on an input pattern.

Database Construction.

A database was constructed in order to test the system. This database consists of 7800 characters collected from 61 subjects with different age, origin, and cultural level. The subjects were asked to write the alphabet in his own style on a

Published at Procceding of the Second joint Mexico US Internacional Worshop on N. Nets and Neurocontrol. Sian Ka’an 97. pp. 24-228.

transparent sheet within small squares, using upper and lower cases. 84 % were right handed, and 16% left-handed. 57% of the group used normally script letters, while the rest used printed-type symbols. These sheets were digitized using image processing equipment. The acquired images were binarized, segmented, and normalized to the size of 16X16 pixels. Finally, the letters were alphabetically sorted, and divided in two sets: 75% (5850 characters) as a training set, and 25% (1950) as a testing set.

Results.

Several experiments were carried out using different subsets of letters and the full alphabet. Two cases will be discussed in this section: The first one is formed by the 26 symbols of the alphabet with two representations of the letters ´s´ and ´z´, giving a total of 28 classes. The second one is formed by the letters a, b, c, d, y, r, s, t, u, x, with two representations of the symbol ´s´, giving a total of 11 classes. This set was formed arbitrarily, but aiming to reduce the complexity of the system and compare it with the ten classes digit recognizer of reference 5. Furthermore, results were compared with a totally connected neural network recognizer. The networks were trained using a learning coefficient from 0.7 to 0.05; the criterion used to decrement this parameter was based on the change rate of the error obtained in each sweep on the training set. This reduction provided better performance in recognition and generalization. The scale parameter of the sigmoid function was fixed to 0.2 in the first data set, and 0.08 in the second one. Next, the results obtained in each case are described.

Case I (10 letters). The network was formed by 699 nodes, 15 179 connections, and 1409 training weights. Table 1 shows configuration details. After 65 sweeps this network recognized 91.2% of patterns in the training set and 75.3% of patterns in the test set. The performance was compared with a totally connected network with a 12 nodes hidden level, 256 nodes in the input layer, and 11 nodes in the output layer. This network recognized 91.6% of patterns in the training set, and 70% patterns in the test set. It can be seen that the LeCun network showed a better recognition rate when it was tested with unknown input data.

Case II (alphabet). The network was formed by 716 nodes, 16 012 connections, and 2242 training weights. Table 2 shows configuration details of this case. After 79 sweeps the network recognized 65.3% of patterns in the training set, and 54.6% patterns from the data set, and it stopped at that point. The network was shown to be small for the recognition problem, however, differences in recognition and generalization kept constants through the training process, i.e., the network increases its generalization ability as it goes through the learning process.

TABLE 1. Case I

Network Structure:

Number of levels:	4
Number of levels in two dimensions:	3
Number of levels in one dimension:	1

Level No.	dimension	Number of maps per level	Map size
0	2	1	16x16
1	2	6	8x8
2	2	3	4x4
3	1	11	1x1

Network size:

Level No.	Number of maps or nodes	Total number of nodes	Connections	Weights
0	1	256	0	0
1	6	384	9984	534
2	3	48	4656	336
3	11	11	539	539
	TOTAL:	699	15179	1409

TABLE 2. Case II

Network Structure:

Number of levels:	4
Number of levels in two dimensions:	3
Number of levels in one dimension:	1

Level	dimension	Number of maps	Map size

Published at Procceding of the Second joint Mexico US Internacional Worshop on N. Nets and Neurocontrol. Sian Ka'an 97. pp. 24-228. © August 1997

No.		per level	
0	2	1	16x16
1	2	6	8x8
2	2	3	4x4
3	1	28	1x1

Network size:

Level No.	Number of maps or nodes	Total number of nodes	Connections	Weights
0	1	256	0	0
1	6	384	9984	534
2	3	48	4656	336
3	28	28	1372	1372
	TOTAL:	716	16012	2242

TABLE 3. Comparison of performance.

Network	No. of sweeps	Average error per node	Recognition percentage with training set	Recognition percentage with test set
CASE I LeCun	65	0.0225	91.2%	75.3%
Totally connected	65	0.0221	91.6%	70.0%
CASE II LeCun	70	0.034	65.3%	54.6%

Conclusions.

The obtained results summarized in table 3 confirm the fact that back propagation neural networks with a structure adequate to the problem, provide a better performance that totally connected neural networks. It was shown that the network can make an internal representation based on characteristics maps. An important result of this work is the constructed database, which allows further research in the area of handwritten recognition.

References.

1. Bozinovic R.M., Srihari S.N., ¨Off-line cursive script word recognition,¨ IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, No. 1, pp. 68-83, January, 1989.

2. Mori Y., Yokosawa K., ¨Neural networks that learn to discriminate similar Kanji characters,¨ Advances in Neural Information, Processing Systems, Edit. Touretzki, Morgan Kaufmann Publishers, San Mateo, CA,. 1989.

3. English T.M., Gomez M.P., Oldham W.J.B., ¨A comparison of neural networks and nearest-neighbor classifiers of handwritten lower-case letters,¨ Proceedings of IEEE International Conference on Neural Networks 93, San Francisco, California, March 28 - April 1, 1993.

4. Neisser U., Weene P., ¨A note on human recognition of hand-printed characters,¨ Information and Control, vol. 3, pp. 191-196, 1960.

5. Le Cun Y., Boser B., Denker J.S., Henderson D., Howard R.E., Hubbard W., Jackel L.D., ¨Backpropagation applied to handwritten zip code recognition,¨ Neural Computation, Vol. 1, pp. 541-551, 1989.

6. Wasserman Phillip D., ¨Neural Computing: Theory and Practice¨, Van Nostrand Reinhold, New York, 1989.

7. Linsker R., ¨Self-Organization in a Perceptual Network,¨ Computer, pp. 105-117, 1988.

Published at Procceding of the Second joint Mexico US Internacional Worshop on N. Nets and Neurocontrol. Sian Ka'an 97. pp. 24-228. © August 1997

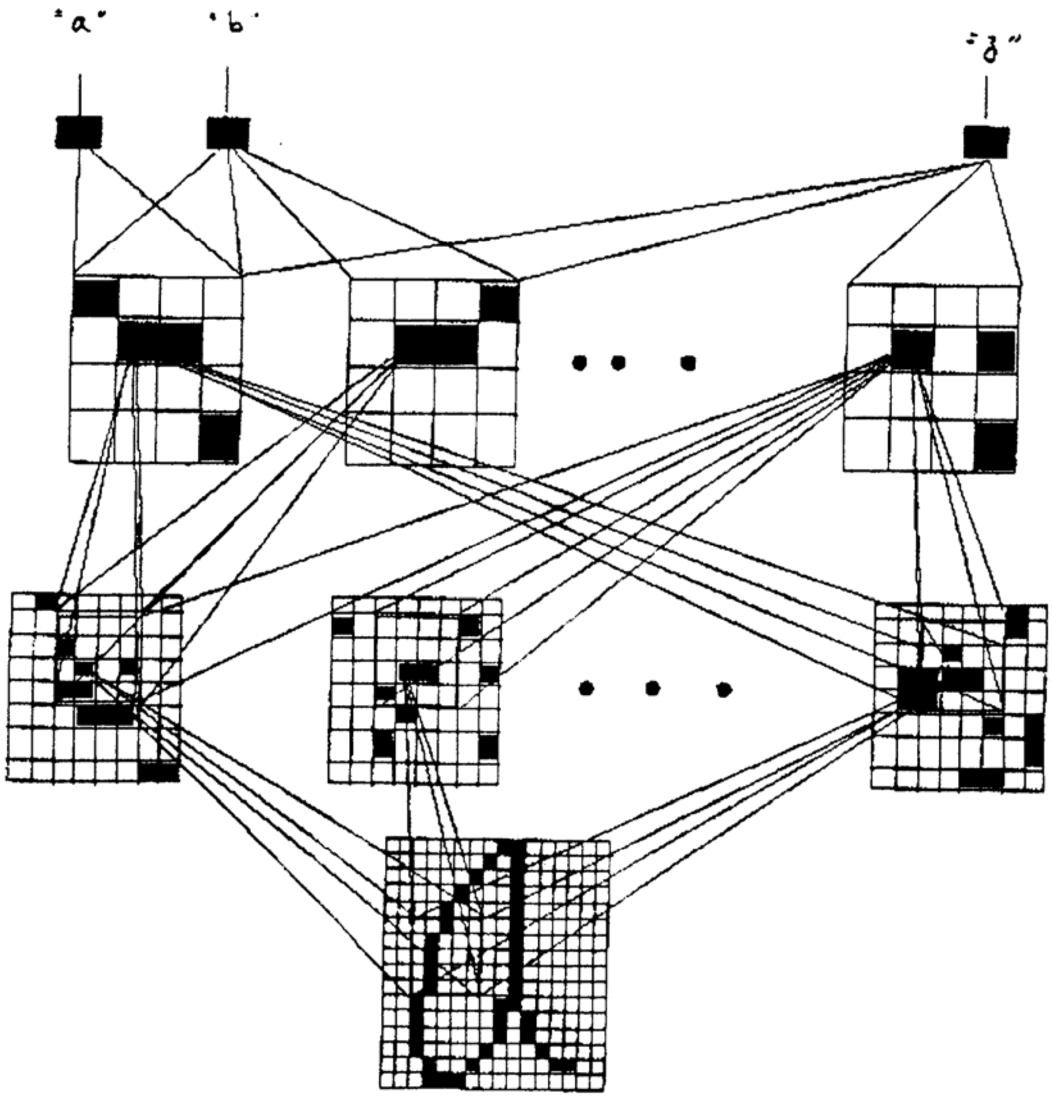

Figure 1. LeCun Neural Network Structure

Published at Procceding of the Second joint Mexico US Internacional Worshop on N. Nets and Neurocontrol. Sian Ka'an 97. pp. 24-228. © August 1997