Extraction Method of Handwritten Digit Recognition Tested on the MNIST Database

B. El Kessab[1], C. Daoui[1], B. Bouikhalene[2], M. Fakir[2] and K. Moro[2]

[1]Laboratory of modeling and calculation - Faculty of Science and Technology
PB 523, Beni Mellal, Morocco
[2]Team Information Processing, Faculty of Science and Technology
PB 523, Beni Mellal, Morocco
bade10@hotmail.fr, daouic@yahoo.com, bbouikhalene@yahoo.fr, fakfad@yahoo.fr, kamalmoro@hotmail.com

Abstract

This paper deals with an optical character recognition (OCR) system of handwritten digit, with the use of neural networks (MLP multilayer perceptron). And a method of extraction of characteristics based on the digit form, this method is tested on the MNIST handwritten isolated digit database (60000 images in learning and 10000 images in test). This work has achieved approximately 80% of success rate for MNIST database identification.

Keywords: *Recognition of isolated digit, MNIST digit, neural networks, Extraction of the characteristics, characteristic zones*

1. Introduction

The neural networks are widely used for the recognition of characters [1, 2, 3, 4, 5, 6, 7, and 8]. In this work we train and test a neural network classifier using MNIST database, the important step in the recognition is learning, we used descent of the gradient algorithm. In the training process the synaptic weight of the connections between the neurons is modified. The first layers contain attached binary neurons without editable connections. We have proposed the MLP as a classifier used for the recognition of the binary images (black and white pixels). The MLP contains three layers of neurons, the first layer corresponds to the retina. In technical terms it matches the input image. The second layer (hidden layer) corresponds to the extraction of characteristics subsystems. The third layer corresponds to the output system. Each neuron in this layer corresponds to one of the output classes. In the recognition task of handwritten digits, this layer contains 10 neurons corresponding to the digits 0 ... 9 (Figure 1). The original weights of the network at random connections. The weights are changed during the formation of the perceptron. The rule of change of weight corresponds to the learning algorithm.

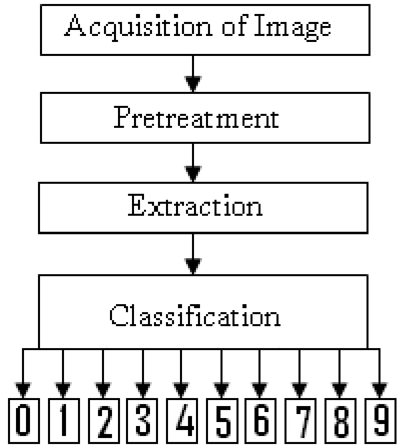


Figure 1. Recognition Process

2. Database

The MNIST database [9] of handwritten digits contains 70000 digits (60000 in learning and 10000 in test) ranging from 0 to 9. The digits have been size-normalized and centered in a fixed-size image with size 28x28. It is free and available on the Web [LeCun (1998)]. Example of the MNIST digit is shown in Figure 2.

Figure 2. Examples of MNIST Database

3. Preprocessing

The pre-processing phase is an important process for recognition digit. In this work, the images of the digit are extracted from the standard MNIST database, and then noise is reduced using a median filter after thresholding process, finally we normalized the images extracted with size 28 x 28.

4. Extraction

The method used for extraction is based on the form of each digit. Dividing the image on five characteristic zones: West zone, East zone, North zone, South zone and Central zone. These characteristic areas are detected by the dilatation of the image processed in four directions.

4.1 Dilatation of Image

The dilation is a transformation based on the form of the image. Each digit is dilated in the four directions.

4.1.1 First Method of Dilatation: Based on the change of pixels values (black or white) of the processed image. Algorithm is shown in Figure 3.

```
1.  Begin
2. For each line
3. For each column
4. {
5. If Image (row, column) = 1
6. Image (row, column + 1) = 1
7. Increment
8. }
9.  Find the Dilatation image to The East.
10. End
```

Figure 3. Dilatation Algorithm to the East

4.1.2 Second Method of Dilatation: Based on the intersection between the object of the image A (white pixels) with a structuring element B. It is defined by the following formula.

$$\textbf{Dilatation (A, B)} = \{x \in \textbf{Image} / B_x \cap A \neq \emptyset\}$$

Where A is the object of the image (the white pixels), B the structuring element which is a particular set of Center x, known size and geometry (in this work is a right half). Example of the dilatation of the digit five to the East

Figure 4. Digit Five and its Dilatation to the East

And it is the same thing for the other directions West, North and South.

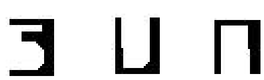


Figure 5. Dilations of Digit Five to the West North and South

With these two methods we found the same results for the dilatation of the image in four directions East, West, North and South. After the Dilatation of the image, we used specific intersections of dilated image in four directions for the detection of characteristic areas of each image.

4.2 Detection of the Characteristic Zones of the Image

The characteristic zones can be detected by the intersections of dilations found to the East, West, North and South. We define for each image five types of characteristic zones: East, West, North, South, and Central zone.

4.2.1. Extraction of East Characteristic Zone: A point of the image (Figure 6) belongs to the East characteristic zone (Figure 7) if and only if:

- This point does not belong to the object (the white pixels in image).
- From this point, moving in a straight line to the East, we do not cross the object.
- From this point, moving in a straight line to the south, north and west one crosses the object. The result of the extraction is illustrated in (Figure 7).

Figure 6. Five Digit Image

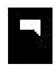


Figure 7. East Characteristic Zone (EZ)

4.2.2. Extraction of West Characteristic Zone: A point of the image (Figure 8) belongs to the West characteristic zone (Figure 9) if and only if:

- This point does not belong to the object (the white pixels in image).
- From this point, moving in a straight line to the West, we do not cross the object.
- From this point, moving in a straight line to the south, north and East one crosses the object. The result of the extraction is illustrated in (Figure 9).

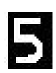


Figure 8. Five Digit Image

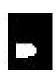


Figure 9. West Characteristic Zone (WZ)

4.2.3. Extraction of South Characteristic Zone: A point of the image (Figure 10) belongs to the South characteristic zone (Figure 11) if and only if:

- This point does not belong to the object (the white pixels in image).
- From this point, moving in a straight line to the South, we do not cross the object.
- From this point, moving in a straight line to the North, East and West one crosses the object. The result of the extraction is illustrated in (Figure 11).

Figure 10. Five Digit Image

Figure 11. South Characteristic Zone (SZ)

4.2.4. Extraction of North Characteristic Zone: A point of the image (Figure 12) belongs to the North characteristic zone (Figure 13) if and only if:

- This point does not belong to the object (the white pixels in image).

- From this point, moving in a straight line to the North, we do not cross the object.

- From this point, moving in a straight line to the East, West and South one crosses the object. The result of the extraction is illustrated in (Figure 13).

Figure 12. Five Digit Image

Figure 13. North Characteristic Zone (NZ)

4.2.5 Extraction of the Central Characteristic Zone: A point of the image (Figure 14) belongs to the Central characteristic zone if and only if:

- This point does not belong to the limit of the object.

- From this point, moving in a straight line to the South, North, East and West we cross the object. The result of the extraction is illustrated in the (Figure 15).

Figure 14. Five Digit Image

Figure 15. Central Characteristic Zone (CZ)

Each digit is characterized by these five characteristic areas. But the problem is it is that the digit two and five for example they also have same characteristic zones West and East (the number of white pixels), even if they are different, the West (East) zone of the digit five is located at the bottom (top) and West (East) zone of the digit two is located at the top (bottom).

Figure 16. Digit Five with West and East Zones

Figure 17. Digit Two with West and East Zones

To remedy this problem. We surround the object of digit and we divided the two zones West and East in two parts the part that is at the top and bottom part. Then we calculate the distance between the first line and the west characteristic zone, and between the last line and the west characteristic zone. If the West zone is near to the first line so this zone belongs to the part who is at the top, if this zone is near to the last line so this zone belongs to the part which is at the bottom, otherwise If the two distances are about equal we divided the zone by the central axis of the lines.

Figure 18. Digit Five after Surround and its Characteristic Zones

Figure 19. Digit Two after Surround and its Characteristic Zones

We used this method only for the two zones West and East because all digits they have these two zones, then the other digits do not have the characteristic zones north and south except the digit four containing a North characteristic zone. Instead of working with five components we find seven components.

$\mathbf{NWZ_1}$, $\mathbf{NWZ_2}$, $\mathbf{NEZ_1}$, $\mathbf{NEZ_2}$, **NNZ**, **NSZ**, **NCZ**

With,

$\mathbf{NWZ_1}$: Number of pixels of value 1 in the West characteristic zone which is at the top

$\mathbf{NWZ_2}$: Number of pixels of value 1 in the West characteristic zone which is at the bottom

$\mathbf{NEZ_1}$: Number of pixels of value 1 in the East characteristic zone which is at the top

$\mathbf{NEZ_2}$: Number of pixels of value 1 in the East characteristic zone which is at the bottom

NNZ: Number of pixels of value 1 in the North characteristic zone

NSZ: Number of pixels of value 1 in the South characteristic zone

NCZ: Number of pixels of value 1 in the Central characteristic zone.

Therefore the extraction vector will be defined as follows:

$$\mathbf{Vext = [WZ_1, WZ_2, EZ_1, EZ_2, NZ, SZ, CZ]}$$

With, **Npixels**: Number of pixels in the image from the MNIST database size 28 x 28.

$$WZ_1 = NWZ_1 / (Npixels).$$
$$WZ_2 = NWZ_2 / (Npixels).$$
$$EZ_1 = NEZ_1 / (Npixels).$$
$$EZ_2 = NEZ_2 / (Npixels).$$
$$NZ = NNZ / (Npixels).$$
$$SZ = NSZ / (Npixels).$$
$$CZ = NCZ / (Npixels).$$

5. Neural networks

The neural networks [10, 11, 12, 13, 14] based on properties of the brain to build systems of calculation best able to resolve the type of problems as human beings live know resolve. They have several models one of these models is the perceptron.

5.1 The Perceptron (Rosenblatt, 1958-1962) [15, 16] (Figure 20) is the simplest module in neural networks. It is composed of two layers of entry and exit, entry called the retina layer and the layer of the output is called the response. These two layers are connected between them by the coefficients of weights W_{ij}.

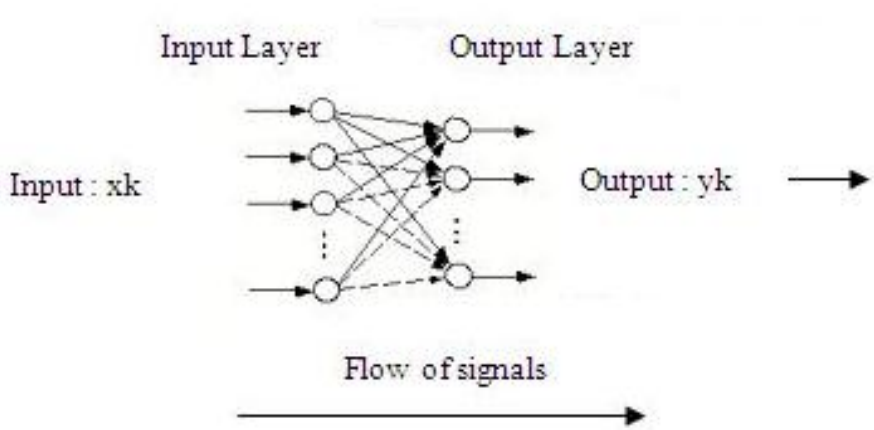


Figure 20. The Perceptron

x: Input layer.
y: Output layer
I: The number of neurons in the input layer x.
J: The number of neurons in the output layer y.
W_{ij}: Weight or synaptic coefficients between the ith entry cell and the jth output cell

5.2 The multi-layer Perceptron (MLP) [17, 18, 19] (Figure 21) is a set of interconnected neurons which has the role of shape recognition, decision, or memory problems. It is also used in signal processing, the treatment of vision, speech, and robotics.

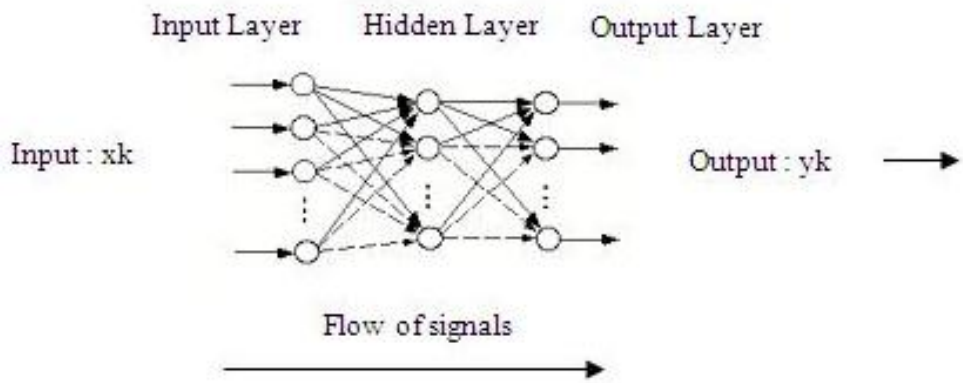


Figure 21. The Multi-layer Perceptron

The MLP is composed:

- With a layer of input x: the input neurons contains I neurons.
- With a layer of output y: output neuron contains J neurons.
- Hidden layers (intermediate): located between the input layer and the layer of output, which having no contact with the outside world. Each hidden layer is composed of a number of neurons. Connections are all oriented the input layer to output layer, i.e: each neuron is connected to all the following layer neurons. The number of neurons in the network is:
- Seven neurons in the input layer (the number seven corresponds to the values found in the vector of extraction).
- Ten neurons in the output layer (the number ten corresponds to the numbers of digit using in recognition).

The number of neurons in the hidden layer is chosen after its three conditions:

- Equal the number of neurons in the input layer.
- Equal 75% of number of neurons in the input layer.
- Equal the square root of the product of two layers of exit and entry.

We followed these three conditions; we varied the number of neurons in layer hidden between four and nine neurons. The method used for learning is gradient back-propagation algorithm [20, 21, 22, 23, and 24].

Table 1. Details of the Neural Networks

Layers	Neurons
Input	7
Hidden	7
Output	10
Constant of learning	
$\alpha = 0.9$	
Squared error	
$E = \frac{1}{2} * (t - o)^2$ t: the theoretical output. o: the desired output.	
Activation function	
$F(x)=\frac{1}{(1+e^{-x})}$	

5.3 Learning Algorithm

1. Initialization of the network settings.
2. Calculation of the State of the network
3. Calculation of error = fct (output - output desired)
4. Calculation of gradient back-propagation algorithm
5. Modification of synaptic weights
6. Stopping criterion, on the error, and the recognition rate of validation database (we stop the algorithm if the rate of recognition of validation it does not increase). After obtaining the optimal parameters of the network, we used a database validation consists 25% of learning entry database initially to validate the results.

5.4 Methodological Issues

- Choice of learning function?
- When stop learning?
- Validation of the results?
- Choice of structure: number of layers, cells, numbers of neurons?
- Quality of the base of learning?

6. Classification

After extracting the characteristic data of the input image, we will classify its data with neural networks (the multilayer perceptron). We must compute the coefficients of weight and desired outputs.

6.1 Experimental Results

For the classification by neural networks, we put values of characteristic vectors in the neurons of input layer, and we known the desired output, we forced network to converge towards a final specific State (supervised learning). Each digit is characterized by a vector of seven components extraction. For the training of the network (multi-layer perceptron MLP), 1 set of 10 images, and 100 sets of 1000 images, We started by a set of ten images, to find the best parameters that maximizes network (Figure 21).

Table 2. Experimental Results

Handwritten digit sets	Numbers of digits	Validation Database	Test Database
1 Set	10	61.50	40.22
10 Sets	100	71.90	68.98
50 Sets	500	82.00	77.43
100 Sets	1000	82.00	79.28
Numbers of images for test		1.000 Images	60.000 Images

Table 3. Recognition Rate for each Digit with 100 Sets of Images

Types of digits	Recognition rate of test Database	Error rate	Reject rate	Execution Time (s)
0	86.45	13.55	00.00	91.64
1	94.39	05.61	00.00	112.58
2	88.73	11.27	00.00	95.63
3	77.02	22.98	00.00	76.71
4	76.12	2388	00.00	84.29
5	84.10	15.90	00.00	70.77
6	78.81	21.19	00.00	85.91
7	77.12	22.88	00.00	90.33
8	79.03	20.97	00.00	85.37
9	49.64	50.36	00.00	93.90

Table 4. Recognition Rate for All Digits

Digits	0	1	2	3	4	5	6	7	8	9
0	**86.45**	00.81	01.10	00.01	01.40	00.27	00.22	00.17	06.36	03.21
1	00.02	**94.39**	01.00	00.84	00.02	00.48	00.44	00.18	01.03	01.59
2	00.09	01.42	**88.73**	04.33	00.82	00.65	00.03	01.00	02.00	00.93
3	00.06	00.52	00.96	**77.02**	01.18	00.45	00.00	14.53	04.95	00.32
4	01.21	02.92	00.34	01.44	**77.94**	02.85	03.30	00.58	03.59	05.83
5	00.03	01.37	01.27	01.03	00.74	**84.10**	06.73	00.22	02.58	01.92
6	02.29	02.41	00.81	00.10	06.21	02.43	**78.81**	00.00	06.78	00.15
7	00.04	03.47	06.66	03.83	01.50	00.04	00.00	**77.12**	04.84	02.48
8	04.82	02.58	01.50	02.03	04.08	00.22	00.70	03.16	**79.03**	01.86
9	04.85	09.89	08.34	00.65	02.70	08.14	02.51	01.90	11.37	**49.64**

7. Conclusion

In this work, the neural network (the multilayer perceptron and back propagation) is proposed for the classification of the standard base MNIST isolated digit. An extraction technique is used in the phase of extraction of characteristics before implementing the classification of the digits. The recognition rate is 80.00% with a Test database containing 60,000. The method of extraction shows enough good results.

Acknowledgements

We are grateful to Yan LeCun for the MNIST database, which is a standard used by anyone in the world, who wants to test its classifier.

References

[1] A. A. Desai, "Gujarati handwritten numeral optical character reorganization through neural network", Pattern Recognition, vol. 43, **(2010)**, pp. 2582–2589.
[2] S. S. Baboo, P. Subashini and M. Krishnaveni, "Combining Self-Organizing Maps and Radial Basis Function Networks for Tamil handwritten Character Recognition", ICGST-GVIP Journal, vol. 9, Issue 4, **(2009)**.
[3] Z. Cheng and Y. Ma, "A research about pattern recognition of control chart using probability neural network", In: Proc. ISECS, **(2008)**, pp. 140-145.
[4] N. P. Banashree, D. Andhre, R. Vasanta and P. S. Satyanarayana, "OCR for script identification of Hindi (Devnagari) numerals using error diffusion Halftoning Algorithm with neural classifier", Proceedings of World Academy of Science Engineering and Technology, vol. 20, **(2007)**, pp. 46–50.
[5] J. Cao, M. Ahmadi and M. Shridar, "A Hierarchical Neural Network Architecture For Handwritten Numeral Recognition", Pattern Recognition, vol. 30, **(1997)**.
[6] K. Fukushima, "Recognition of partly occluded patterns: A neural network model", Biol. Cybernet., vol. 84, **(2001)**, pp. 251–259.
[7] S. Sagiroujlu, E. Besdoc and M. Erler, "Contro chart pattern recognition using artificial neural networks", Turkish Journal of Electrical Engineering, vol. 8, **(2000)**, pp. 137-147.
[8] B. Gabrys and A. Bargiela, "General fuzzy min–max neural network for clustering and classification", IEEE Trans. Neural Networks, vol. 11, no. 3, **(2000)**, pp. 769–783.
[9] Y. LeCun, "THE MNIST handwritten digit database", **(1998)**.
[10] C. Bishop, "Neural networks for pattern recognition", Clarendon Press – Oxford, **(1995)**.
[11] Haykin, "Neural Networks", Prentice Hall, **(1998)**.
[12] Hertz, Krogh and Palmer, "Introduction to the theory of neural computation", Addison Wesley, **(1991)**.
[13] Thiria, Gascuel, Lechevallier and Canu, « Statistiques et méthodes neuronales", Dunod, **(1997)**.
[14] S. Haykin, "Neural Networks: A Comprehensive Foundation", Second Edition, Pearson Education Asia, **(2001)**, pp. 208-209.
[15] F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain", **(1958)**.
[16] F. Rosenblatt, "Principles of Neurodynamics", Spartan, **(1962)**.
[17] D. Rumelhart, E. Hinton, McClelland and R. J. Williams, "Learning Internal Representations by Error Propagation, **(1986)**.
[18] B. B. Chaudhuri and U. Bhattacharya, "Efficient training and improved performance of multilayer perceptron in pattern classification", Neurocomputing, vol. 34, no. 1–4, **(2000)**, pp. 11–27.
[19] E. M. Kussul, T. N. Baidyk, *et al.*, "Rosenblatt perceptrons for handwritten digit recognition", Proceedings of International Joint Conference on Neural Networks IJCNN 2, **(2001)**, pp. 1516–1520.
[20] R. Flitcher, "Practical methods of optimization", New York: Wiley, **(1987)**.
[21] P. G. Ciarlet, "Introduction to numerical linear algebra and optimization", Cambridge: C.U. **(1989)**.
[22] Y. Le Cun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard and L. Jackel, "Handwritten digit recognition with a back-propagation network", Advances in neural information processing systems, San Mateo, Morgan Kaufmann, **(1990)**, pp. 396-404.
[23] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-Based Learning Applied to Document Recognition", Intelligent Signal Processing, **(2001)**, pp. 306-351.
[24] J. A. Snyman, "Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms", Springer Publishing, **(2005)**.

Authors

B. El Kessab

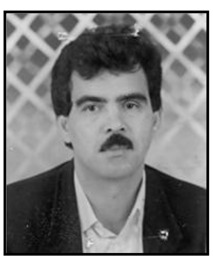
Received his Master's degree in 2009 from Faculty of Sciences and Technology University Sultan Moulay Slimane Beni Mellal Morocco, currently working on his Ph.D. in the modeling and calculation Laboratory at Sultan Moulay Slimane University. His current research interests include pattern recognition, image analysis, document processing and automatic processing of natural languages using hidden Markov models and neural networks.

C. Daoui

Received his Ph.D degree on mathematics in 2009 from Mohamed V Univesity Rabat Morocco. Currently is a professor in Faculty of Sciences and Technology, University Sultan Moulay Slimane Beni Mellal Morocco. His research topics are: the mathematics, operational research and pattern recognition.

B. Bouikhalene

Received his Ph.D degree on mathematics in 2001 and Master's degree on Science of Computer and Telecommunications in 2007 from the University Ibn Tofel Kenitra. Currently is a professor in the Sultan Moulay Slimane University Beni Mellal Morocco. His research topics are: the pattern recognition, artificial intelligence and mathematics and its applications.



M. Fakir

Received his Master's degree on Electrical engineering from Nagaoka University of Technology in 1991 and Ph.D. degree on Electrical engineering Cadi Ayyad University Morocco. it is a team Hitachi Ltd., Japan between 1991 and 1994. Currently is a professor Faculty of Sciences and Technology, University Sultan Moulay Slimane Beni Mellal Morocco. His research topics the recognition and artificial intelligence.

K. Moro

Received his Master's degree in 2009 from Faculty of Sciences and Technology University Sultan Moulay Slimane Beni Mellal Morocco, currently working on his Ph. D in Sultan Moulay Slimane University. His current research interests include pattern recognition, image analysis, document processing and automatic processing of natural languages using hidden Markov models and neural networks.