Evaluation of a neural system for handwritten digits recognition

Isabelle Pottier, Gilles Burel

EVALUATION OF A NEURAL SYSTEM FOR HANDWRITTEN DIGITS RECOGNITION

Isabelle POTTIER & Gilles BUREL
Thomson-CSF Laboratoires Electroniques de Rennes
Avenue de Belle Fontaine, 35510 CESSON-SEVIGNE, France

First European Conference on Postal Technologies, Nantes, France, June 14-16th, 1993

ABSTRACT:

Handwritten digits recognition has been widely studied because of its potential application in automatic sorting of mail pieces. In this paper, we focus on off-line isolated digits with unknown scriptor. TCSF/LER has developed an intermediate approach between classical methods, based on extracting small sets of parameters, and pure neural methods, in which the network is fed with raw image data. The proposed method combines image processing and connectionnist recognition. A vector of 90 parameters consisting in profile curves, measures of density and morphological information is computed from the digit image. Then a multilayer perceptron trained by backpropagation is used to classify. The method has been evaluated on a huge database of real zipcodes, provided by the SRTP Nantes. The database includes around 20000 digits for learning and 12000 digits for testing. The isolated digits come from prefilled or free envelopes. Each digit has two labels provided by two operators: the first one sees the whole address block and the second one is restricted to seeing only the segmented digit. In this paper, we describe our approach and we give many experimental results: recognition rates on prefilled envelopes, on free envelopes, on digits confirmed by the second operator, etc.

1 Introduction

Character recognition has been the subject of many researches over the past thirty years [Govindan,1990]. For handwritten digits, the main applications are recognition of handwritten zip codes [Mitchell,1989], and cheque processing. Here, we focus on "Off-Line" recognition with unknown scriptor. We consider only isolated digits.
In previous works, recognition of handwritten digits by classical methods has been studied, among others, by [Mitchell,1989] (model matching method with morphological features) and by [Shridhar,1986] (rule based system with parameters on profiles). In the neural networks field, many works have been published over the past three years. The neural network is generally fed with an image of the digit, normalized in size [Le Cun,1990].

As shown on the flow chart (fig 1), we propose an intermediate method which combines image processing in order to extract features and connectionist recognition which uses the previous features. Although many recognition rates are published in papers, it is impossible to compare methods which are not working with the same experimental conditions. So, to obtain a meaningful evaluation of the existing methods, the french postal services (SRTP) has set up a validation procedure, using a large digit base which is significant for the mail sorting application. After testing our system with a first base of 19000 digits supplied by the SRTP, it has been validated on a second base of 12000 digits kept secret in SRTP. We give here our experimental results.

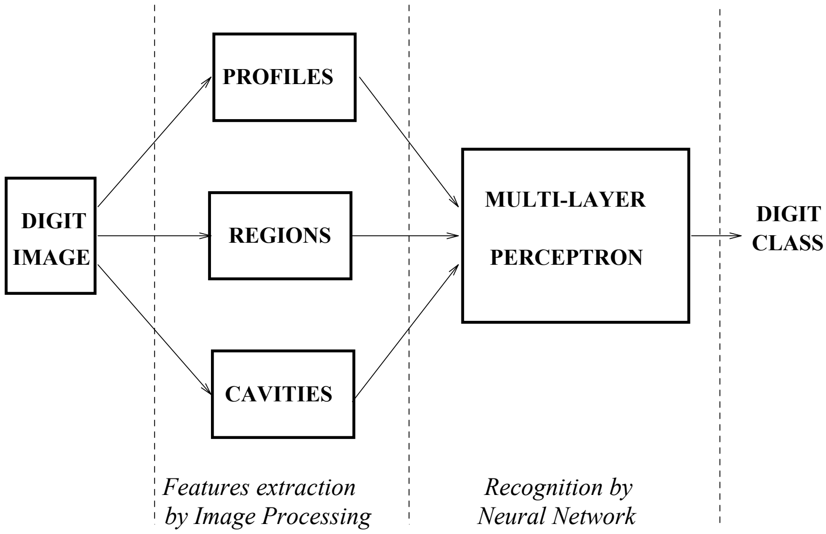

Figure 1: *System overview*

2 Method

2.1 Features extraction by image processing

A vector of 90 parameters is extracted from the digit image. It includes the following features which represent metric, statistical and morphological information.

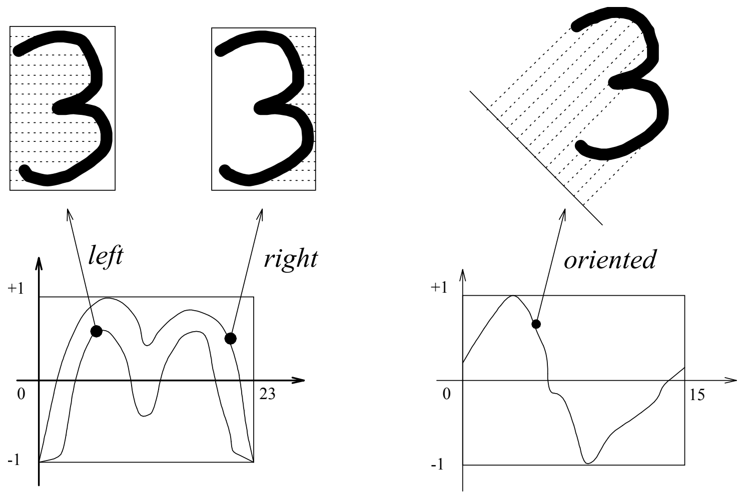

Figure 2: *The left, right, and oriented profiles*

With normalized digit **profiles** (fig 2), we propose to take advantage of metric features which bring almost all shape information about the digit, because it is nearly possible to rebuild the entire image from them. We also add an oriented profile (under 45^o) to reduce ambiguities. To avoid loss of information, our neural net is fed with the complete curves of the profiles and not only with a small set of parameters extracted from them. The curves are represented by 24 samples for the right and the left profiles and 16 samples for the oriented one.

With **regions**, some global information is also used. It consists in characterizing the distribution of the digit pixels inside the including frame. This statistical feature has been proposed by [Beauville,1990]. The including frame is divided in regions in 6 different ways (fig 3). There is a total of 20 regions, which produce 20 features $R_j = n_j/n$, where n_j is the number of digit pixels inside region j, and n is the total number of digit pixels.

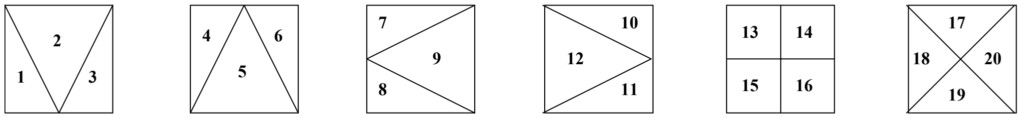

Figure 3: *The 20 regions defined with 6 masks*

Cavities of different types and holes are the last characteristics used, by the means of their relative surfaces. We define 5 cavity types: West, East, North, South and Center. Adding the hole, we obtain 6 morphological features (fig 4). These features were introduced by [Mitchell,1989] to specify a symbolic model of the digit. Each morphological feature could be given to the neural net input as the surface of the feature divided by the total surface of the whole set of morphological features.

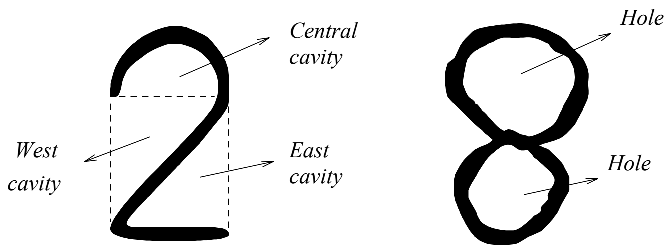

Figure 4: *Some morphological features*

2.2 Recognition by neural network

The neural network is a Multi-layer Perceptron [Rumelhart,1986], trained with the back-propagation algorithm. The neuron model is a weighted adder followed by a non-linear function (hyperbolic tangent). It includes an adaptative bias. Two neural networks have been considered:

5-layer perceptron with local connections

This network (fig 5) has been proposed and evaluated in [Burel,1992]. Instead of a full-connected network, the network structure is adapted to inputs. The proposed network consists of 5 layers where the first ones are devoted to the profiles curves. Local connections and shared weights are introduced to take advantage of neighborhood notion between curve samples. The inter-pattern share of the 2nd layer provides curve analysis. Then, the intra-pattern share of the 3rd layer averages slight local distortions of the profiles. As for the regions and cavities parameters which are independant pieces of information, they feed the neural network from the 4th layer with full-connected neurons.
It can be noticed that this architecture leads to a small size network (130 neurons and 2000 connections) which is advantageous for learning time and implementation reasons. For comparison, on the same application, [Le Cun,1990] uses a network with 4000 neurons and 100000 connections.

Full-connected 4-layer perceptron

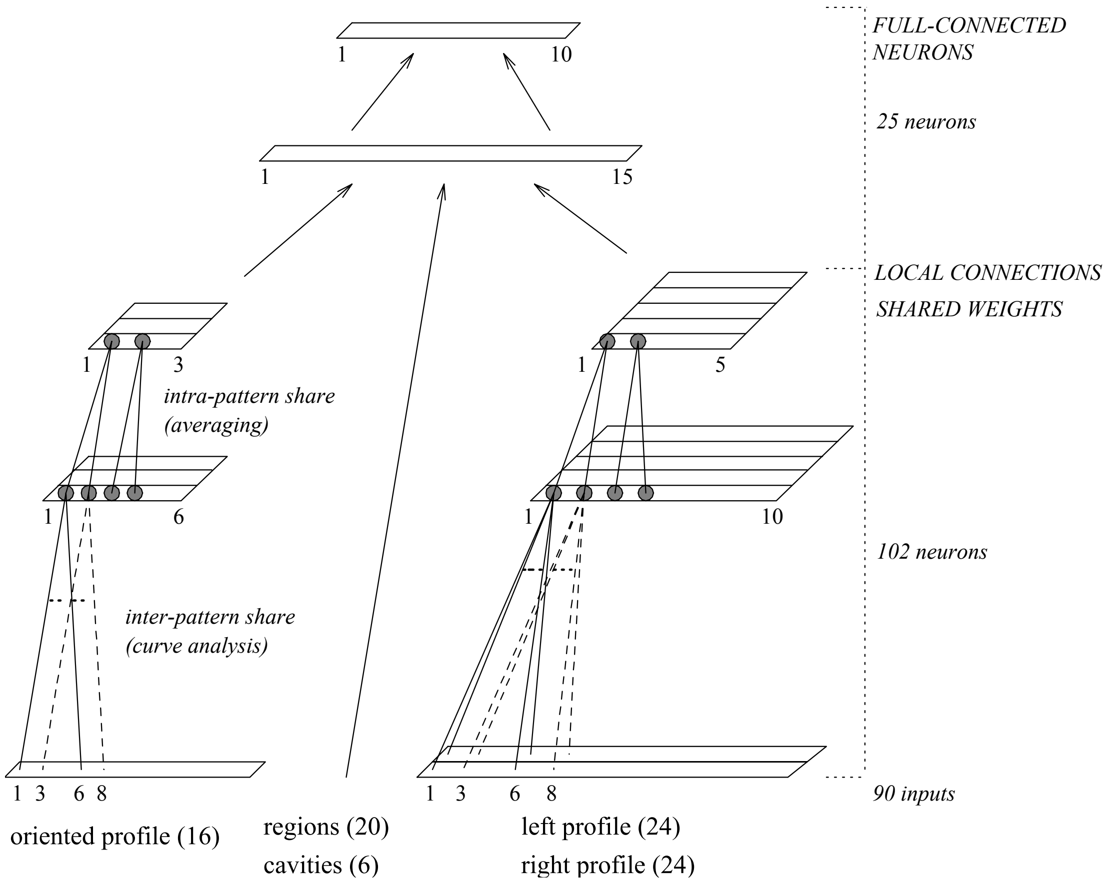

Figure 5: *Structure of the multi-layer perceptron*

This neural net has 90 inputs, 40 and 20 neurons on the hidden layers and 10 neurons on the output layer (one for each digit class). The total number of weights is 4670.
For the learning phase, improvements of the backpropagation algorithm allows to avoid any parameter adjustment. For instance, the learning speed is automatically tuned [Burel,1991].

3 Experiments

3.1 Digit base

The isolated digits come from prefilled or free envelopes. These zip-codes were segmented and binarized. Each digit has two labels provided by two operators: the first one sees the whole address block and the second one is restricted to seeing only the segmented digit. When the two labels are equal, the digit is called "confirmed".

The SRTP has built two huge digit bases :

- The first base (*Public database*) is supplied by the SRTP in order to design at home the recognition system. It includes around 15698 confirmed digits and 19485 all in all.
- The second base (*Secret database*) is kept in SRTP to validate the available methods. It includes 9250 confirmed digits and 2646 unconfirmed digits (either badly segmented or badly labelled).

Furthermore, each base is splitted in 3 categories according to the way of segmenting the digit from the envelope. The zip-codes could come from prefilled envelopes. The digit is normally located in a small box and no segmentation processing is required to isolate it. However, it happens that the writer does not respect this specified location, which gives cut digits. So the bad and the good segmented digits are distinguished by calling them **the bad and the good prefilled digits**. The zip-codes are written on free envelopes. The isolated digits are found by an automatic processing of localization and segmentation. We call them **the free digits**.

3.2 Tests on the Public Database

We have tested both neural networks, the 5-layer perceptron with local connections (LC5) and the full-connected 4-layer perceptron (FC4). The LC5 with its adapted structure to the inputs does not provide better results than the FC4 network on this large digit base. Yet, on a smaller base of 1000 digits, the selection of a network with as few free coefficients as possible was proved necessary. The recognition rates were: 94% for the LC5, 90% for the FC4, 88% for the k-nearest-neighbour in [Burel,1992] and 83% for a network proposed by [Le Cun,1990] and evaluated by [Auger,1992] on the same data base. As explained in [Lee,1991], when the size of the training database is hugely increased, the recognition rates of the various methods (k-nearest-neighbor, radial basis functions, backpropagation neural networks) converge towards the same asymptotic value. Furthermore, it can be shown [Cover,1967] that the k-nn reaches the theoretical upper bound of recognition rate when the database becomes large. Hence, on a very large digit base, the recognition performances of many methods are very likely to provide the same score. Computational complexity and memory requirements will constrain algorithm selection more strongly than small differences in overall error rate. For instance, the k-nn classifier which is theoretically the best approach on a huge database requires untractable computational power and memory in practise.

The following results are given for the FC4 network. The table below summarizes the results. Inside each box, there is a recognition rate and the number of digits on which it has been computed. The neural network was trained on 75% of the confirmed digits. Its evaluation on the remaining one-fourth provides a recognition rate of 97.7%. On the unconfirmed digits, the recognition rate is 60%. Hence, on average, since the unconfirmed digits represent 19.4% of the database, the recognition rate is 90.4%. For comparison, the nearest neighbour classifier gives 95.8% on confirmed digits and 88.1% when unconfirmed digits are also taken into account.

<u>Results on the public database</u> (19485 digits, 19.4% of unconfirmed digits)

	learning set	test set
Recognition rate for confirmed digits	99.3% *11773*	97.7% *3925*
Recognition rate for unconfirmed digits	*0*	60.0% *3787*
Weighted average	*11773*	90.4% equivalent size:*4870*

Some errors on digits belonging to the test set are shown on (fig 7) where the class provided by the first operator, the found class and the confidence value are written above the digit image. Among the badly recognized digits, we could distinguish two kinds of errors : There are errors due to a bad segmentation or a bad labelling : from prefilled envelopes, there are a lot of cut digits (see 0 6). We notice, as well, strange symbols which do not look like digits and some connected digits. And, there are errors which would not be done by a human being : the digits 1 drawn in the English way are not well classified (learning was performed essentially on digits written in the French way).

If we accept a rejection, it is possible to reduce the error rate by using the confidence which

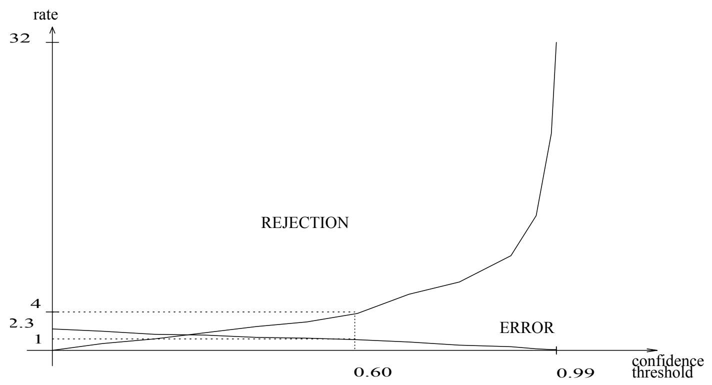

Figure 6: *Error rate and Rejection rate versus confidence threshold*

is here the difference between the two highest network outputs. When the confidence is lower than a threshold, no decision is taken (rejection). An error rate of 1% on confirmed digits may be obtained, at the cost of a rejection rate of 4% (fig 6).

3.3 Validation on the Secret Database

We use here the secret base of 11896 digits. They are splitted in 1223 bad prefilled, 7394 good prefilled and 3279 free. As for the public base, The label provided by the first operator is used as reference when they are different. The table below summarizes the results. On average, we obtain a recognition rate of 96.8% on the confirmed digits, and 87.3% when the unconfirmed digits are also taken into account. These results are comparable with those obtained on the public database, but the secret database seems to be slightly more difficult.

Results on the secret base (11896 digits, 22.2% of unconfirmed digits)

	bad prefilled	good prefilled	free	weighted average
Recognition rate for confirmed digits	74.8% *330*	98.4% *6372*	95.9% *2548*	96.8% *9250*
Recognition rate for unconfirmed digits	19.9% *893*	74.5% *1022*	66.3% *731*	54.1% *2646*
Weighed average	34.7% *1223*	95.1% *7394*	89.3% *3279*	87.3% *11896*

Here is the confusion matrix from all the digits extracted from free envelopes. We remember that a recognition rate of 89.3% is obtained on these 3279 digits. The values are expressed in percent except for the classes distribution:

classes	C(0)	C(1)	C(2)	C(3)	C(4)	C(5)	C(6)	C(7)	C(8)	C(9)	error	size
C(0)	**96**	0	1	0	1	0	0	0	1	0	4	838
C(1)	0	**71**	11	1	3	1	0	7	6	1	29	338
C(2)	1	1	**93**	1	0	0	0	0	1	1	7	408
C(3)	0	2	3	**89**	0	1	0	1	0	3	11	204
C(4)	2	2	3	0	**89**	1	2	1	1	0	11	186
C(5)	0	1	4	4	0	**90**	1	0	0	0	10	212
C(6)	1	1	1	0	3	1	**92**	0	0	0	8	143
C(7)	0	3	1	0	1	1	0	**92**	0	2	8	212
C(8)	2	1	5	1	0	1	0	1	**88**	1	12	359
C(9)	5	1	2	3	1	0	0	2	1	**85**	15	379

This matrix reminds that the digit 0 represents, in french mail application, around 25% of the classes distribution and it is well recognized and seldom confused with another digit (sometimes the 9).

We check that our method has trouble with the digit 1 drawn like a stick. This kind of 1 is not well represented in the learning examples because it is not often confirmed by the second operator. So, the network had not the ability to generalize correctly these 1. This digit image creates a very small including frame in which the extracted features are very sensitive. For instance, all density parameters are almost the same, the profiles are unstable according to the digit skewness and every cavity parameter is equal to 0. Therefore, they exist heterogeneous clusters in the parametric space. To solve this problem, we could either take into account relationship between height and width along the main axis or split the digits 1 into 2 classes, the French and the English ones. The other kind of errors are not too significant due to the images quality, mainly in the prefilled envelopes.
Since the validation base is not at our disposal, we can not give more precise results like recognition rates using the second label or introduction of rejection.

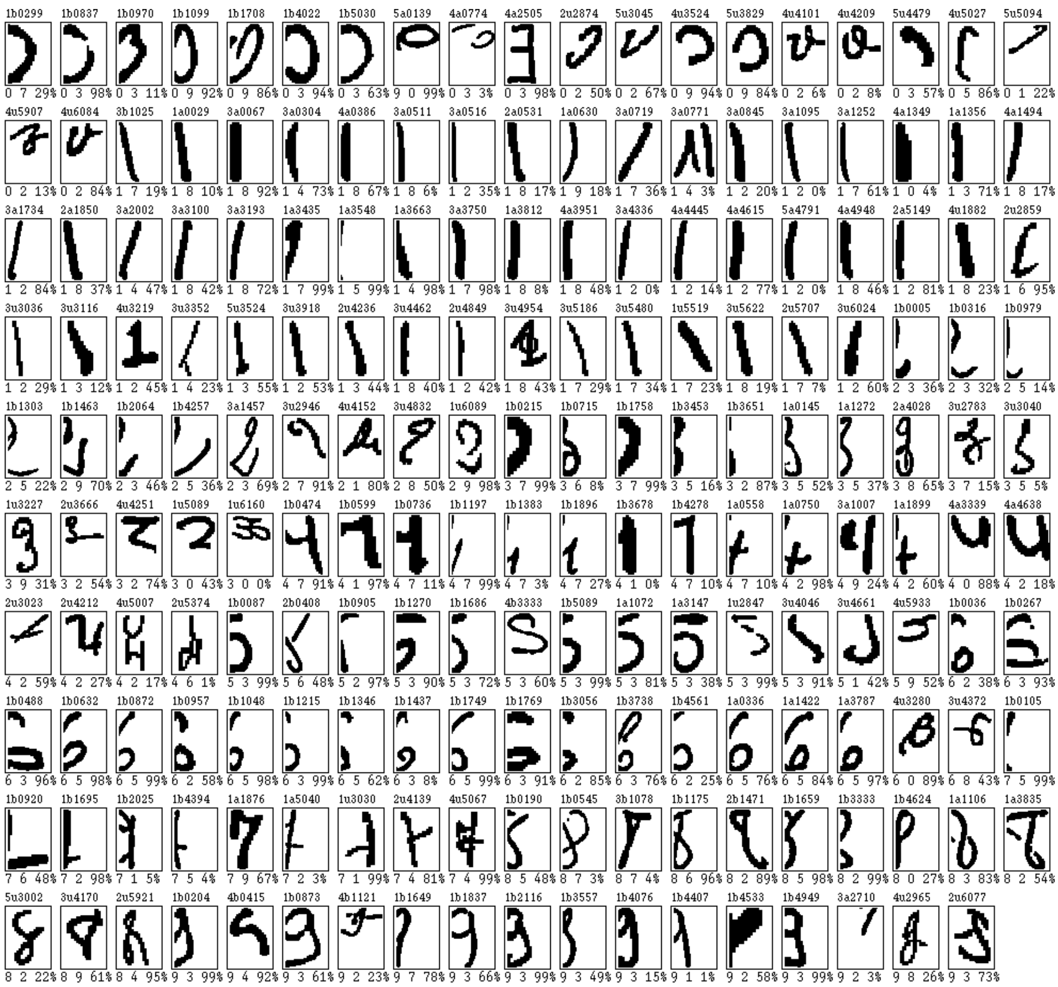

Figure 7: Some errors in the complete base

4 Conclusion

An intermediate approach between classical handwritten digit recognition methods, and pure neural methods has been proposed. Compared with classical methods, this approach has the advantage to avoid the difficult task of selecting a right parameter set and decision rules. Compared with pure neural methods, it has the advantage of a lower complexity. The simplicity of the presented approach and the good performances obtained on a very large digit base provided by the French postal services are really strong advantages in comparison with previous methods. In the near future, we plan to enhance our method by investigating three axis : understanding of errors by synthesis of the digit image from the extracted features, refinement of preprocessing step by adding new features, improvement of recognition step by mixing several kinds of neural networks.

ACKNOWLEDGEMENT:
The authors would like to thank Mr Gilloux and Mr Lemarié of SRTP/Nantes for providing the database of french ZIP-codes and for their interest and support during the evaluation.

REFERENCES:

Auger, J.M., Idan, Y., Chevallier, R., Dorizzi, B., 1992. "Complementary aspects of topological maps and time delay neural networks for character recognition", IJCNN Baltimore, vol IV, pp 444-449, June 1992

Beauville, J.P.A., et al., 1990. "Reconnaissance de chiffres manuscrits saisis sans contrainte" BIGRE n^o68 (Reconnaissance automatique de l'écrit), Le Havre - France, may 18th, 1990, pp 89-101

Burel, G., 1991. "Réseaux de neurones en traitement d'images: des modèles théoriques aux applications industrielles", Ph.D. Thesis, University of Brest, December 6th, 1991

Burel, G., Pottier, I., Catros, J.Y., 1992. "Recognition of handwritten digits by neural network", Proceedings of IEEE-IJCNN92, June 8-11th, 1992, Baltimore

Cover, T.M., Hart, P.E., 1967. "Nearest neighbor classification", IEEE Trans. on Information Theory, vol 3, n^o1, pp 21-27, January 1967

Govindan, V.K., Shivaprasad, A.P., 1990. "Character Recognition - A Review", Pattern Recognition, vol 23, n^o7, pp 671-683, 1990

Le Cun, Y., et al., 1990. "Optical Character Recognition and Neural-Net Chips", International Conference on Neural Networks, Paris, july 9-13th, 1990, pp 651-655

Lee, Y., 1991. "Handwritten digit recognition using k-nearest-neighbor, radial basis function, and backpropagation neural network", NEURAL COMPUTATION, 3, pp 440-449, 1991.

Mitchell, B.T., Gillies, A.M., 1989. "A model based computer vision system for recognizing handwritten ZIP codes", Machine Vision and Applications, vol 2, n^o4, Fall 1989, pp 231-243

Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. "Learning internal representations by error backpropagation", Parallel Distributed Processing, D.E. RUMELHART and J.L. Mc CLELLAND, Chap8, Bradford book - MIT Press - 1986

Shridhar, M., Badreldin, A., 1986. "Recognition of isolated and simply connected handwritten numerals", Pattern Recognition, vol 19, n^o1, pp 1-12, 1986