OCR for Handwriting Characters - A Case Study

Victor Andrei Maiorescu
Technical University Iasi, Faculty of Electronics and Telecommunications, e-mail: mandrei@etc.tuiasi.ro
Tiberiu Dinu Teodorescu
Technical University Iasi, Faculty of Electronics and Telecommunications, e-mail: t-teodor@etc.tuiasi.ro

Abstract

Optical Character Recognition is a major field of signal processing. The problem is almost solved for specific families of character such as Times New Roman, Arial, and so on. Also it was proved that the type of font used influences the results of the classification.

Our study is oriented on handwriting characters. This application may be important, for example in tasks that ensure the interface between a large number of persons and a system that has to process the information sampled from those persons. One example of such application is the working with forms. They can be automatically scanned and introduced in large databases, assuring for the process a certain level of automation.

1 The database with characters

The database with handwriting characters is made of black-white images. From the original images (20x20 pixels), the images used in training and test were obtained. A normalization of the gray levels has been performed. Also, some anti-aliasing techniques have been used. The image is centered relatively to the cipher. The 10 different ciphers are approximately equally represented. The database was downloaded from the Internet and is declared to be a benchmark [7].

The whole database is split in two sets: one is used for training purposes and the other is used in testing. The criteria for the image choosing were choosing typical images and also non-typical images from a very large number of persons. The dimension of the database and the number of exemplars in every class influence the performance [1, 2]. The classes are very good built from the number of exemplars' point of view.

The population used in training and testing is explicitly given below:

Table 1: The population used

	Training set (exemplars)	**Test set (exemplars)**	**Both sets (exemplars)**
cipher '0'	5923	980	6903
cipher '1'	6742	1135	7877
cipher '2'	5958	1032	6990
cipher '3'	6131	1010	7141
cipher '4'	5842	982	6824
cipher '5'	5421	892	6313
cipher '6'	5918	958	6876
cipher '7'	6265	1028	7293
cipher '8'	5851	974	6825
cipher '9'	5949	1009	6958
Total	60.000	10.000	70.000

A typical set of images is given below:

Fig. 1: A typical set of images

Not all the images in the database are so clear as the images represented above.

2 The dataflow

The dataflow in a recognition application is represented below:

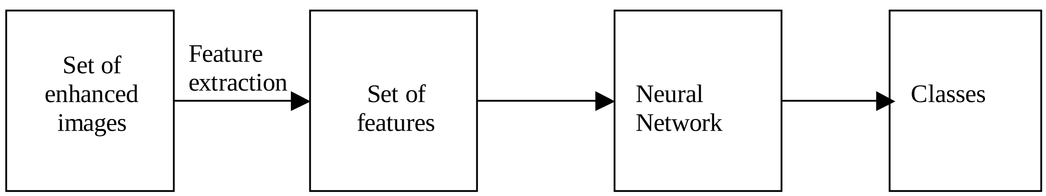


Fig. 2: A typical data-flow in recognition application

3 Feature extraction

Once one has loaded the images, they are pre-processed in order to be suitable for the neural network. One realizes a compression of the data and a de-correlation too. In the particular case of this application one obtain by compressing the data a noise filtering (high frequency noise) that came from the acquisition process. If we keep 5% of the DCT coefficients, a compression of 20:1 is obtained. This high compression rate is used for making easier the job of the neural network [3].

A normalization of the input data is also performed in order to have the data in a certain interval.

Also, as a future direction, a pre-processing by using cellular neural networks can be considered. It is known that by using appropriate convolution masks one can perform the so-called **C**onnected **C**omponent **D**etection [8] that has given good results in applications regarding among other tasks character recognition.

This will be a matter that has to be discussed in a future work. Also, a comparison between those two approaches has to be done in order to choose the one that performs best.

4 User interface and functioning

The interface uses Document/View architecture and it is very useful in order to obtain as much information we can at a certain moment. The user has the ability of browsing the database. At the same moment, he can see the histogram for a certain image and the classification for the same image. Also, he can have an image about the trust degree of the classification (in the bottom right corner of the main window of the application). For the current image in the figure, the degree of trust in the classification is almost 100% (cipher 4).

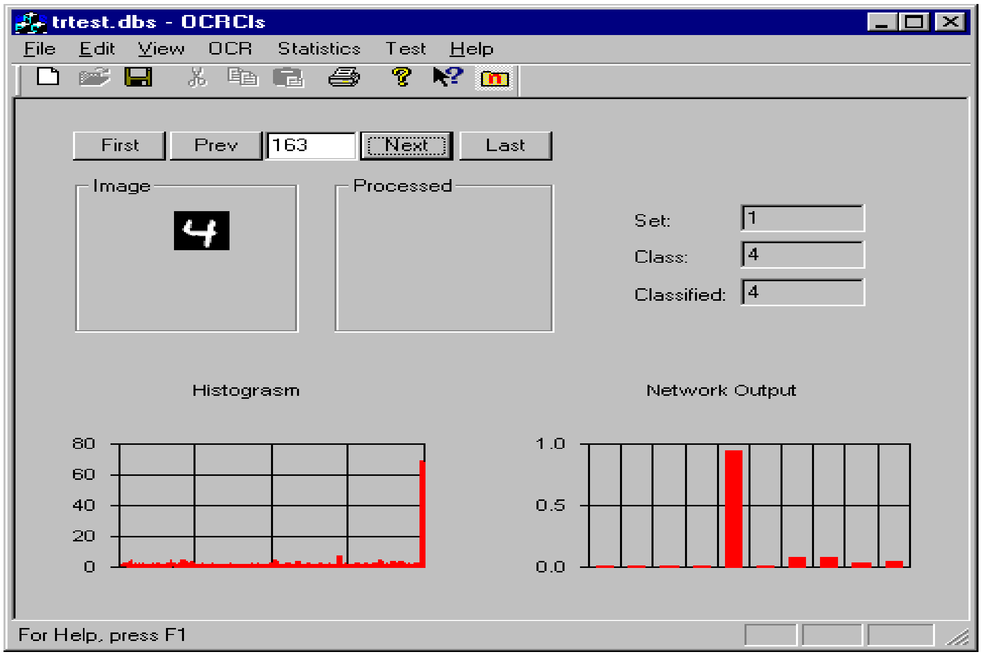


Fig. 3: The main window of the application

In order to train the network one monitor the performance of the system after each epoch. For the training set and test set the error and the performance are displayed. The performance for each class of images is separately displayed:

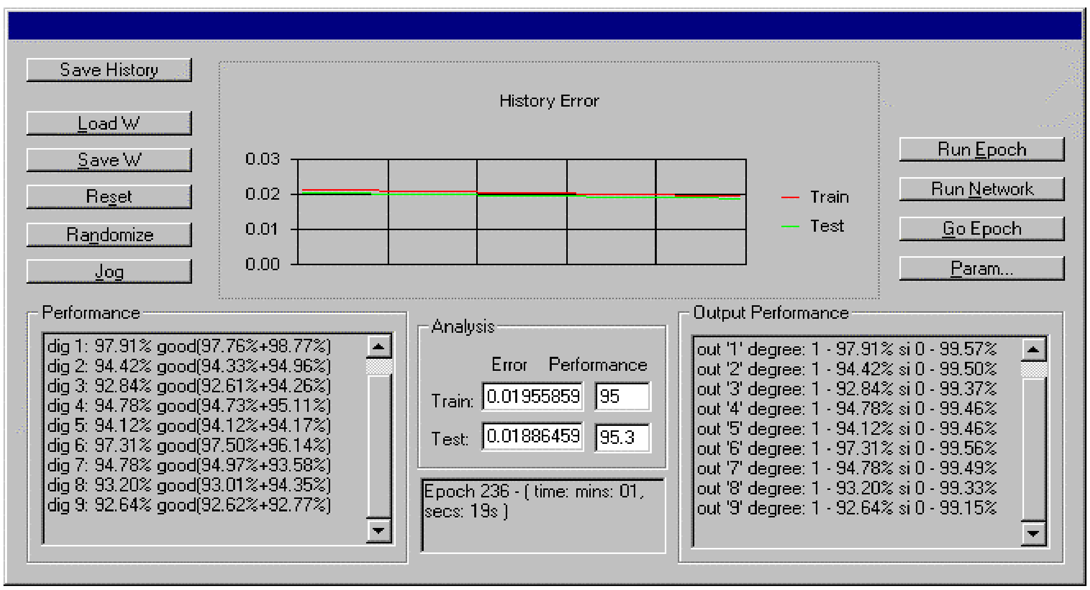


Fig. 4: The performance of the network

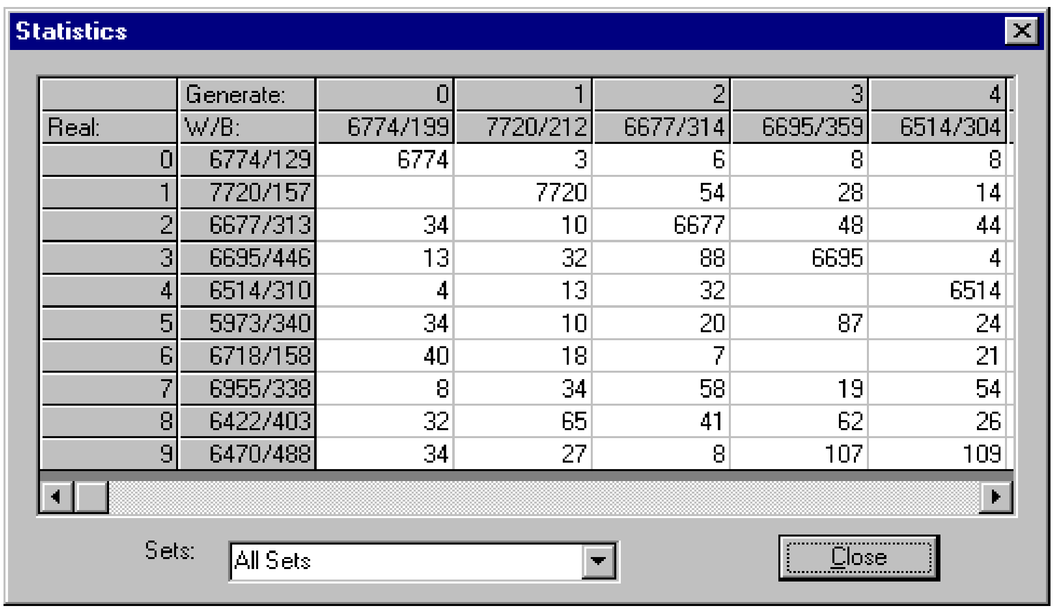
In the center of the dialog window the error history is displayed for the last 50 epochs. The commands for the network such as: reset, weight initialization, exit from local minims, loading and saving the weights, training and online using are buttons on the panel.

In our experiment, a Multilayer Perceptron Network along with the Backprop algorithm has been used [1, 5 ,6] . The results are displayed below:

Table 2: The results of the training for training set (70%), test set (30%) and both sets:

Label of the class	**Training set**	**Test set**	**Overall**
cipher ‘0’	98.1%	95.71%	97.38%
cipher ‘1’	97.76%	98.77%	97.91%
cipher ‘2’	94.33%	94.96%	94.42%
cipher ‘3’	92.61%	94.26%	92.84%
cipher ‘4’	94.73%	95.11%	94.78%
cipher ‘5’	94.12%	94.17%	94.12%
cipher ‘6’	97.5%	96.14%	97.31%
cipher ‘7’	94.97%	93.58%	94.78%
cipher ‘8’	93.01%	94.35%	93.2%
cipher ‘9’	92.62%	92.77%	92.64%

A detailed analysis of the classification results is made:


Statistics

	Generate:	0	1	2	3	4
Real:	W/B:	6774/199	7720/212	6677/314	6695/359	6514/304
0	6774/129	6774	3	6	8	8
1	7720/157		7720	54	28	14
2	6677/313	34	10	6677	48	44
3	6695/446	13	32	88	6695	4
4	6514/310	4	13	32		6514
5	5973/340	34	10	20	87	24
6	6718/158	40	18	7		21
7	6955/338	8	34	58	19	54
8	6422/403	32	65	41	62	26
9	6470/488	34	27	8	107	109

Sets: All Sets

Close

Fig. 5: Statistics for the network performance

The table above should be interpreted in the following way: the first column has listed the real classes (from "0" to "9", the second presents the number of the well-classified versus bad-classified exemplars in each class. For example, class corresponding to cipher "0" has 6903 exemplars. Within these exemplars, 6774 exemplars were well-classified, and 129 were taken as exemplars that belong to other classes: 3 "belong" to class corresponding to cipher "1", 6 exemplars "belong" to class corresponding to cipher "2", 8 exemplars "correspond" to class labeled with "3" and so on.

The scroll-bar permits the user to see the miss-classifications corresponding to the other classes labeled with "4", "5", "6", "7", "8" and "9". All sets have been considered for the experiment.

5 Results' interpretation

The classification process is made roughly. The best output gives the winner in the classification process. Such a decision is motivated through the hypothesis that only one output is activated at a certain moment of time. The drawback is that when one have two outputs significantly activated only one is taken into account (the biggest)

Such a situation is presented in the figure below, for the digit 5.

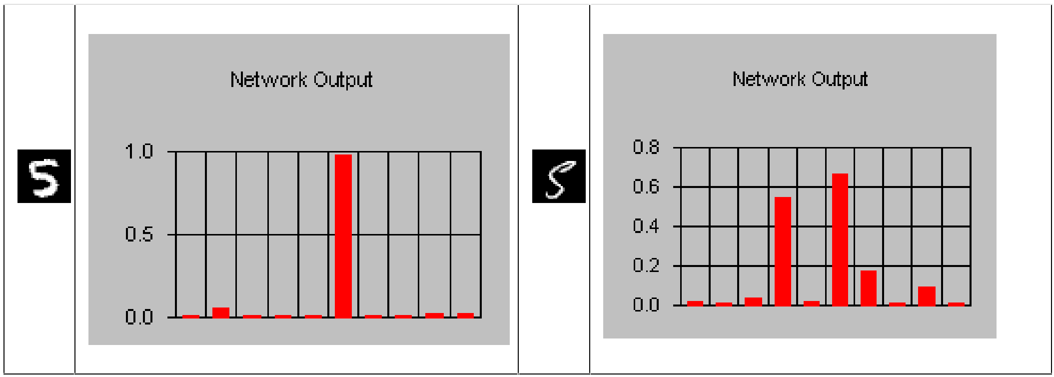


Fig. 6: Ambiguous decision case

One can easily see that for the second case of digit 5, two outputs are activated, one corresponding to digit 3 and the other corresponding to digit 3.

6 Performance evaluation

The performance is separately calculated for the training set and for the test set. A complete view of the classification process can be obtained if one can see the classification performance for each class. This is linked with the outputs of the neural network as well. The classification process is very important, but at least the same importance has the decision making process. To adopt a trust degree is very appropriate for cases similar to that one presented above.

7 Conclusions and further work

We can conclude based on the results obtained in the experiment that one can approach an application for automatic recognition of handwriting characters in forms because of the good results obtained in classification (online functioning). The application can help the post services to split the correspondence that is to be sent to different locations, and so on. Additionally, some algorithms can be used to increase the good classification rate. For the post service, the additional information can be for example the finite number of possible names of the destination cities, streets, and so on.

In order to make the application suitable for Romanian language and for other languages, a database with the specific handwriting characters should be built. This database should be large enough in order to obtain good results. This is a time consuming task and the results are to be sensitive to the number of different persons who give their handwriting samples for the database. We have validated here only the method and we expect to obtain worst results when increasing the number of classes.

Further work should focus especially on pre-processing tasks. We have studied in this article mainly the methods that work in frequency domain. There are also other methods that work in the space domain (the image itself) as methods that extract geometrical shapes from the image. A powerful tool for this kind of tasks is the cellular neural network, which can be useful, as stressed in the "Feature extraction" section. A comparative approach has to be done in order to establish the best method.

Another problem that is maybe the most important for practical validation is the extension of the database used for the recognition with other characters than ciphers (letters, characters that belong to other languages, marks, etc).

Once the recognition provides good results, the integration into bigger applications can be performed successfully.

References

[1] S. Haykin, Neural Networks – A Comprehensive Foundation, Prentice Hall, 1994
[2] C. M. Bishop, Neural Networks for Pattern Recognition, Clarendon Press, Oxford, 1995
[3] ***, Image Processing Library Reference Manual, Intel, 1998
[4] ***, Neurosolutions Reference Manual, Neurodimensions, 1998
[5] ***, Fuzzy Systems and Neural Networks, course notes, Technical University of Iasi, Romania
[6] ***, Signals, Circuits and Systems, course notes, Technical University of Iasi, Romania
[7] ***, AT&T Laboratories benchmark set of handwriting characters, 1999
[8] ***, CNN Software Library (Templates and Algorithms), version 7.3, Analogical and Neural Computing Laboratory, Computer and Automation Institute, Hungarian Academy of Sciences, Budapest, 1999