

Problems Connected With Application of Neural Networks in Automatic Face Recognition

Rafał Komański, Bohdan Macukow

Faculty of Mathematics and Information Science,
Warsaw University of Technology
00-661 Warszawa, Poland, Pl.Politechniki 1

Abstract. One of possible solutions in creating an automatic system of face recognition is application of auto-associative neural networks for remembering and recognise two-dimensional face images. Experiments with applying the Hopfield network and twolayer perceptron confirmed the possibility of remembering and reproducing face images, even if partially covered or disturbed. Limited technical possibilities enable using only low definition images. It is due to the fact that computations take a long time, and the number of remembered faces is relatively small. Proper working of the network is influenced in a significant way by light and facial expressions.

Introduction

Face is a unique feature of every person, the same applies to fingerprint and pupil. Improvement of computers and of their processor capacity resulted in introducing informatics into this area. Automatic face recognition has been drawing a lot of attention in the recent years. Digital identification possibility has found broad application in safety systems and improvement of communication between a man and digital computers. From among biometric identification systems using unique characteristics of people face recognition systems draw a lot of attention. This results among others from the method specificity. People being identified may not be aware of the fact that there is such a system working next to them and recognising scanning them. Face recognition systems are limited to obtaining a face image, then to analysing the image and deciding whether the person is included in their data base or not. The work concentrates only on the face recognition problem, in particular on using auto-association neural networks for this purpose.

Basic Notions

Automatic face processing includes numerous issues, i.e. face detection in a picture, face recognition, analysis of facial expression, and classification based on facial features. The first problem is the face detection. In some cases conditions in which a picture is taken can be fully controlled. For example the face position can be easily specified in passport photographs. Yet in many situations position, size and orienta-

tion of a face are not known in advance. First step in face detection is checking if a given picture includes any faces. After that the number, situating and size of particular faces are specified. Nevertheless we should differentiate at this point between face detection and face localisation. The second of the problems deals with finding a face situating when it is known that in a given picture there is exactly one face. Figure 1 shows the issue of face detection in a picture [4].



Fig. 1. Face detection in a picture. Detected faces are marked with frames. One of the faces in the picture, tilted right, has not been detected.

The next step is recognising the face found in the picture. The main task here is to compare and to match the unknown face in the picture with the faces in the database. One of possibilities here is the application of neural networks.

Next problems in face processing, although not directly associated with face recognition, are also very important. The first of them is the facial expression analysis. Basing on the information included in the way a given face looks one can specify if a person is e.g. happy, sad, surprised or frightened. Different facial expressions create a difficulty in face recognition.

Another problem is the classification on the basis of facial features. Classifications can be based on the person's sex, age, race, etc. Yet the mechanisms working here are not known.

One of the basic notions is a face representation. All faces must be stored in the basis in the same format. A new face to recognise must be also converted into the same format. Methods based on neural networks directly use two-dimensional face images in grey scale. In studies devoted to face recognition it is assumed that the starting point are pictures with nothing but faces on them, so there is no need for detection. Such a format was used for the tests the results of which are described further in this article. The images can be stored in various resolutions and use different numbers of the grey colour levels.

The application of neural networks solves one of the most important problems, i.e. an economical way of storing faces in memory. An important feature of associative memory is the ability to remember the input data in order to reproduce them even if the input to the network is their disturbed or partially damaged version. Associative memories function as content addressable memories. In such a memory information is remembered in the connections weights. The first person to use auto-associative memory to remember and reproduce face images was Kohonen. He proved that an

auto-associative network can be used as a content addressable memory for storing face images [2].

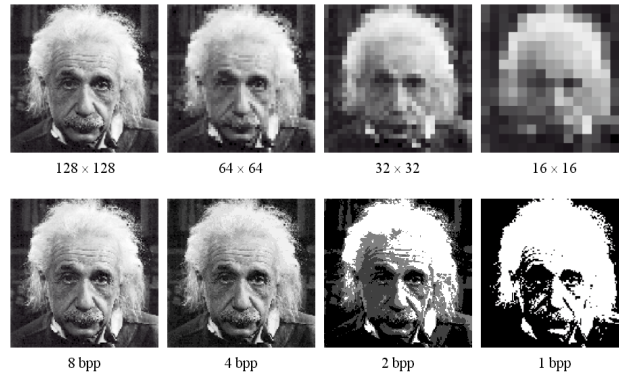


Fig. 2. A face image in various resolutions (scaled to the same size) and in various number of grey colour levels.

Networks Used in the Tests

In the simulation programme used for carrying out the experiments two models of auto-associative memory were applied. The first of them is the model of auto-associative memory proposed by Hopfield in 1982 [1]. The second model of auto-associative memory is a twolayer network, taught with the method of a scaled conjugate gradient algorithm.

A discrete Hopfield network is a single-layer recurrent network. In the learning mode the values of weights are specified in accordance with the Hebbian Learning Rule. In the reproducing mode the values of weights are frozen and the network can work synchronously or asynchronously. The Hopfield network functioning in the synchronous mode and with bipolar neurones has been implemented in the simulation programme. The second type of auto-associative memory used in the tests is a twolayer perceptron. In order to specify the weights of the network's hidden layers a special strategy called a back-propagation algorithm should be applied. The most efficient methods of learning are the optimisation methods based on developing the objective function into the Taylor's series. The process of learning by the network means an iterative choice of minimisation directions, so that the error function is minimised. In order to determine the minimisation direction one can use the algorithm of the highest descent, the algorithm of variable metric, and the most effective for big networks scaled conjugate gradient algorithm. This algorithm was proposed by Møller [3], and it combines determining the minimisation direction with the optimum pace of learning.

Results

The tests have proved that the Hopfield network as well as a twolayer perceptron can be used as a content addressable associative memory. The networks reproduced correctly the face images on which they were taught. This means that when providing a network with the input of any of the learnt images, the original image or one very similar to it was received as the output. Also when the input presentation was disturbed or partial, networks generated correct images with recovered missing parts of faces as the output.

Hopfield network

Enormous demands concerning memory are a serious limitation to the Hopfield network. That is why carrying out the tests with this type of memory has been limited to images not exceeding $32 \times 32 \times 8$ bpp. Exemplary results obtained with the use of Hopfield network are presented in Fig. 3. The network also did very good when partially covered or disturbed images were presented as the input.

Another limitation in case of Hopfield network is the fact that computations during learning and iterative computations in the reproducing mode take a long time.

Even if only several face images were used to teach one network the computation time was counted in minutes and hours. It was necessary to introduce a limitation of the maximum number of iterations. Consequently the obtained results were not always satisfying, because there was not enough time for the network's state to stabilise.

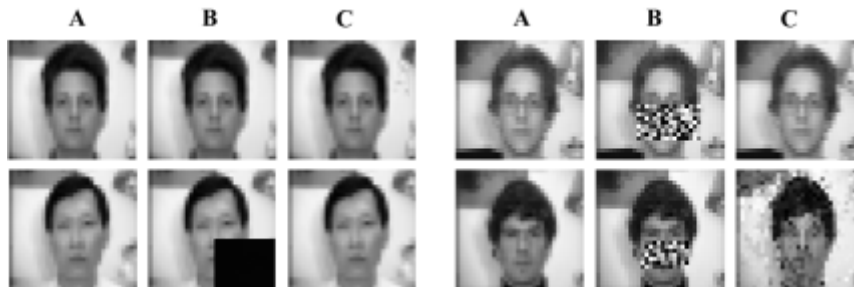


Fig. 3. Exemplary results of operation of the simulation programme for Hopfield network taught $32 \times 32 \times 8$ bpp images. A original images, B input to network, C network response.

Twolayer perceptron

The tests on twolayer network also confirmed possibility of applying it as a content addressable memory for face images. The network reproduced the images from the teaching file correctly. The exemplary results of the operation of simulation programme for the twolayer network are presented in Fig. 4.

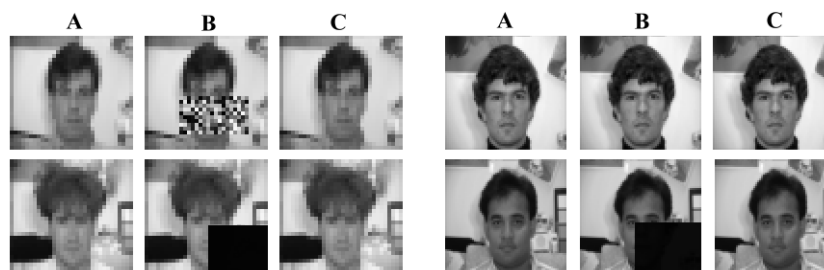


Fig. 4. Exemplary results for twolayer network, on the left $32 \times 32 \times 8$ bpp images, on the right $64 \times 64 \times 8$ bpp images. A original images, B input to network, C network response.

Also in cases when the images presented as the input to network were slightly modified, i.e. disturbed or partially covered, the network reproduced the images in a satisfactory way.

Differences and similarities

There are differences in functioning of both the types of tested memories. The two-layer network is a feedforward network. Hence the computing time during reproduction is relatively short. There are no such problems with huge memory demand as in the case of Hopfield network. Yet the time of computation during learning is much longer. What is more the process of learning is not always successful. It depends on the network's structure and initial values of weights. It was indispensable to introduce a limit for the maximum number of iterations during learning. Thus it happened that the network finished learning with a high level of energy. In such a case the network responded with a noise. Better results were achieved when the level of energy after finishing learning was lower. The number of elements in the hidden layer can be set in a simulation programme. The appropriate tests were carried out with/ a few to a few dozen neurones in the hidden layer.

Both the types of the tested networks did very well during tests with images saved in smaller number of the grey colour levels. Yet with lower resolutions of picture cutting down on the quality of photos resulted in making them hard to distinguish even by a human. The images had a sensible quality only in the minimum mode 4 bpp.

The common problem for both types of memory was an incorrect reproduction of face images produced with the use of a different light than in the images on which the network was taught. To give an example, a partially shadowed face was confused with a bearded person's face. It could be observed that the network simply aimed at finding the closest picture in terms of Hamming's distance. Nevertheless in such a situation even a person could make a mistake.

Another problem worth paying attention to is a wrong recognition in case of different position of a head in the picture. Although it is known in advance that the picture presents one face, the face can be in frontal position, in semi-profile, or in profile. Besides the head can be turned up, down, to the front or can be tilted to the side. Networks aiming at finding the closest picture in terms of Hamming's distance. They were more willing to present rather other faces in the given position than the correct

faces as the output. There is yet another problem when comparing faces. It is different appearance of the same person's face depending on the person's emotional condition.

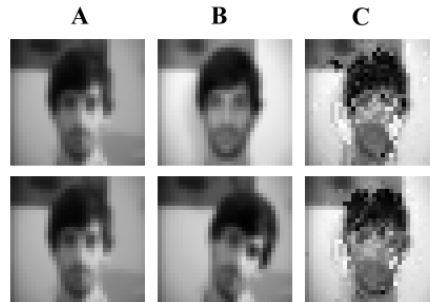


Fig. 5. Exemplary results of the operation of a simulation programme for images with various light (top) and with changing position of head (down). **A** original images, **B** input to the network, **C** network response.

Exemplary results with the changing light problem and various position of head in photos are presented in Fig. 5.

Summary

Face recognition is a difficult and complex problem. It can be divided into a few sub-problems: detection, identification, analysis of the facial expression and categorisation on the basis of facial expressions. The article presents a discussion on the possibilities of using neural networks to recognise faces. Experiments carried out with the use of Hopfield network and a twolayer perceptron taught with the method of a scaled conjugate gradient algorithm have proved that the networks can be used as a context addressable memory for the images of faces. The main two problems that drew the author's attention were the facts of incorrect recognition when the light and the position of head in photographs are variable. A serious problem is also different appearance of faces depending on people's emotional states. Additional limitations are relatively small capacities of network, slow speed of computations and high demand for memory.

References

1. Hopfield J.J, Neural network and physical systems with emergent collective computational abilities, Proc. National Academy of Science, USA, 79:2554-2558, 1982
2. Kohonen T., Associative Memory: A System Theoretic Approach, Springer, Berlin 1977
3. Møller M.F., A scaled conjugate gradient algorithm for fast supervised learning, Neural Networks, Vol. 6, pp. 525-533, 1993
4. Samal A., Iyengar P.A., Automatic recognition and analysis of human faces and facial expressions: A survey, Pattern Recognition, Vol. 25, no. 1, pp. 65-77, 1992