Proceedings of the Fifth International Conference

ICNNAI'2008

NEURAL NETWORKS AND ARTIFICIAL INTELLIGENCE

27–30 May 2008 Minsk, Republic of Belarus

SIG of INNS

БААРИ

Exadel
Extend. Adapt. Deliver.

ГРАНТ БРФФИ

IBA

NovaLab

SKIB

Effective Soft

BELHARD Development

ОИПИ

Minsk, Belarus

5-6

Neocognitron-Like Multilayer Neural Network for Halftone Images Recognition

Aleksandr Doudkin, Maksim Vatkin
United Institute of Informatics Problems of NAS of Belarus, Surganov str. 6
vatkin@lsi.bas-net.by, doudkin@newman.bas-net.by, http://lsi.bas-net.by

Abstract: *Construction peculiarity of neural network "neocognitron" was considered as applied to image recognition task. A new structure of multilayer neural network based on fuzzy logic was developed, that was designed for classification of objects having pixel distortions on halftone images. This neural network allows significantly increase recognition accuracy.*

Keywords: *Neocognitron, Neural networks, Image recognition.*

I. INTRODUCTION

Object recognition is one of the main stages in the image analysis. Object recognition task is stated in the following way. It is necessary to discover a map of an input pattern X to Y, where Y is a class of the input pattern. The map is defined by set of pairs (input pattern, class of pattern). The number of pairs (training samples) is significantly less than the possible amount of pairs (input, output). So training set consists of training samples.

If images are strongly distorted, it is very difficult to find informative features of the object. In common case we have distortions of two types: brightness characteristic or position of a pixel is changed.

Neural network (NN) classifiers give the most robust classification in the case when images are strongly distorted. There is a number of NN suited to this goal: a multilayered perceptron, a network of radial basic function, an ART network (Adaptive Resonant Theory), Hopfield network, Kohonen self-oganasing map.

An image pattern is the structured description of the object on the image. The most appropriate NN classifier of the image patterns is a neocognitron (Fukushima, 1983). It has hierarchical structure of layers and is designed to model human visual path of brain (Hubel, 1962). The neocognitron provides invariant recognition, which is based on successive stages of local template matching and spatial pooling. The following ways are known for development of NN hierarchical image recognition: neocognitron modification to in increase its perfomance (Lovell, D.R., 1992), finding additional feature invariants and changing feature extraction rules to improve recognition accuracy (Wiskott, L., 2003; Satoh, S., 1997), development of unsupervised algorithms for feature hierarchy creation (Pan, Z., 1999; Wersing, H., 2003; Behnke, S., 2003; Scalzo, F., 2005).

Let's generalize main points of structural and algorithmic solutions in image recognition based on necognitron NN.

1. Two types of features are considered for images: simple features that can be found in any image and complex features that are spatial combination of the simple features.

2. The features are organized in hierarchy according to complexity, where the features with the same complexity form a single hierarchical level. The simple features belong to the first hierarchical level.

3. Multi-layer NN is developed, where a number of its layers is equal to a number of the hierarchical levels. Neurons of each layer are grouped, where a number of groups is equal to a number of features of given complexity. Neurons of the group (a sublayer) detects the same local feature on the input image, so a neurons activity represents a map of a feature distribution on the image.

4. Pattern recognition is made using layer-by-layer activation from the input layer that detects the simple features to the output layer that detects all features of the pattern in a whole.

So the following tasks are solves for the neocognitron development:

1. Define feature detection method (the rule of neuron activation).

2. Define a criterion of effectiveness using of set of features on one layer. That criterion allows optimizing a number of neurons groups (sublayers).

3. Develop a hierarchy of the features and a method of combination of the features of the level to form the feature of the next hierarchy level.

In (Sadykhov, 2001) we proposed a neocognitron modification based on the mentioned principles, that classifies objects with distortions of its form. Detection of a local pattern feature is realized by a pair of inhibitory and excitatory neurons. Additional generalizing neurons are used to provide invariance of recognition of a pattern having form distortions. Inputs for generalizing neurons are activities of excitatory neurons of detection layer (fig. 1),

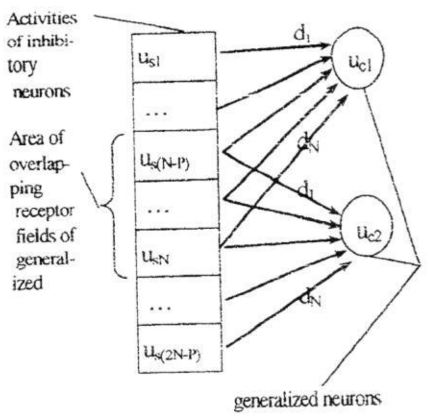

Fig. 1. Connection structure of generalizing layer.

where N is a number of input links of neuron or size of receptor field of the neuron, P is a number of neurons in overlapping area of receptor fields, d_i are unmodified weight of excitatory connections. Values d_i are defined in such a way to decries monotonically form center to the border of neuron reception field. The activation rule acts in such a way: if detected feature lies in overlapping region of two neurons, both neurons are activated. Thus, the neuron activity of previous layer passes into the activity of several neurons of current layer. That rule provides blurring of the feature in the generalization layer.

This classification model has following advantages:

- More adequate method of image local feature extraction;
- An automated algorithm of feature extraction for each layer;
- A possibility to process halftone images.

This new modification is intended for recognition of patterns on halftone images that have both form and brightness distortions. The mentioned above rule is close to it biological analog but there is a number of shortcomings in the point of view of software realization:

- A complex rule for detection feature;
- A questionable use of blurring for ensuring of recognition invariance of patterns with form distortions;
- Use of additional layer of generalization neurons to ensure recognition invariance.

As alternative we propose a new neuron structure and an activation rule. That rule is based on a fuzzy difference function of two matrices, that is invariant to value and position changing of matrix elements. As training rule we use a fuzzy mean function that is calculated dynamically. It is also invariant to the mentioned changes.

II. Modification of Rule for Feature Identification

Let's consider two matrices

$$W = \begin{pmatrix} w_{1,1} & \cdots & w_{X,1} \\ \vdots & \ddots & \\ w_{1,Y} & & w_{X,Y} \end{pmatrix} \text{ and} \tag{1}$$

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{X+2R,1} \\ \vdots & \ddots & \\ a_{1,Y+2R} & & a_{X+2R,Y+2R} \end{pmatrix} \tag{2}$$

where W is a matrix of pixel values of sample pattern, A is a matrix of pixel values of pattern that has distortions of pixels of two types: changes in a brightness and a position. Matrix A has bigger size because of the changes in the of pixel positions.

Let's define parameters B and R of a range of pixel distortions. R defines a maximal possible shift of pixel positions in matrix A relatively to equivalent positions in matrix W, and B is a maximal possible interpixel distance, i.e. $\left|w_{i,j} - a_{j+R,i+R}\right| < B$. Let's note, that $|\bar{v}_{x1,y1} - \bar{v}_{x2,y2}| \leq \theta$ for any two neighbouring pixels $a_{x1,y1}$, $a_{x2,y2}$ and their shifts $\bar{v}_{x1,y1} = a'_{x1,y1} - a_{x1,y1}$ and $\bar{v}_{x2,y2} = a'_{x2,y2} - a_{x2,y2}$, where $\theta \leq \max(|\bar{v}_{x1,y1}|, |\bar{v}_{x2,y2}|)$. It specifies uniformity of geometrical distortions, i.e. the neighbouring pixels are displaced almost equally.

Let's construct a matrix

$$D = \begin{pmatrix} d_{1,1} & \cdots & d_{X,1} \\ \vdots & \ddots & \\ d_{1,Y} & & d_{X,Y} \end{pmatrix}, \text{ where} \tag{3}$$

$$d_{x,y} = \begin{cases} 0, \\ \min\limits_{i,j=0\ldots 2R}(a_{x+i,y+j}) \leq w_{x,y} \text{ and} \\ w_{x,y} \leq \max\limits_{i,j=0\ldots 2R}(a_{x+i,y+j}); \\ 0, \\ |w_{x,y} - a_{x+R,y+R}| < B; \\ 1, \quad \text{in other cases.} \end{cases} \tag{4}$$

Now a difference function of matrixes A and W is defined as follows:

$$fd(A,W) = \frac{100}{X \cdot Y} \cdot \sum_{x,y} d_{x,y}\,. \tag{5}$$

If a training image A^t is entered on each training step t, we shall receive a sequence of updating matrices of the reference image $\{W^t \mid t = 0, \ldots, Tn\}$, where T is the number of training images and n is a number of the iteration.

Let's define the function of a dynamic average:

$$w_{x,y}(t+1)=\begin{cases} w_{x,y}(t), & \min\limits_{i,j=0...2R}(a_{x+i,y+j}(t)) \le w_{x,y}(t) \ \textit{and} \\ & w_{x,y}(t) \le \max\limits_{i,j=0...2R}(a_{x+i,y+j}(t)); \\ w_{x,y}(t), & |\, w_{x,y}(t)-a_{x+R,y+R}(t)\,| < B; \\ \dfrac{w_{x,y}(t)\cdot t + a'_{x,y}}{t+1}, & \textit{in other cases} \end{cases} \quad (6)$$

where for $a'_{x,y}$ is fulfilled

$$|\, w_{x,y}(t)-a'_{x,y}(t)\,| = \min_{i,j=0...2R}\left(|\, w_{x,y}(t)-a_{x+i,y+j}(t)\,|\right).$$

Thus, the structure of the developed neuron can be represented in in figure 2 (for case of one-dimensional data), where [$\min_k$, $\max_k$], k= 1, …, N, are defined as in the formula (4):

$$\min_k = \min_i(a_{k+i}),\ i = 0...2\cdot R\,;$$

$$\max_k = \max_i(a_{k+i}),\ i = 0...2\cdot R\,.$$

As a result, the performance is raised approximately to kt ≈ 15N times for the activation of one neuron in comparison with the original neocognitron.

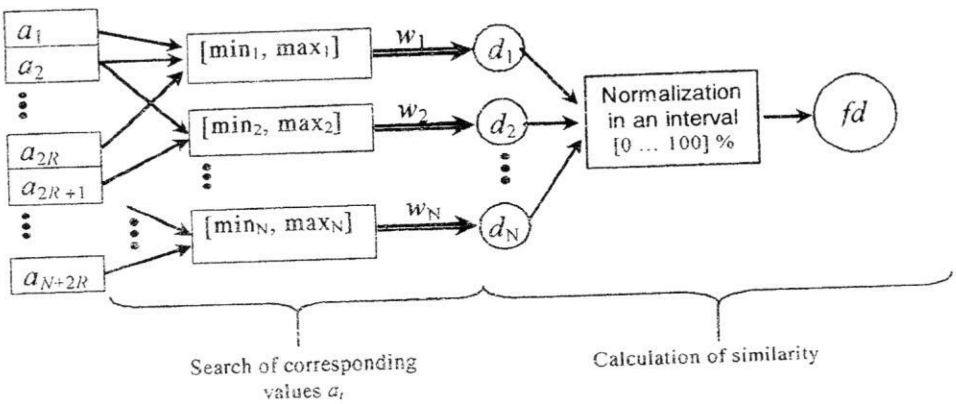

Fig. 2. The neuron structure.

III. NN STRUCTURE

The structure of the developed NN is close to the structure of neocognitron network (Fukushima, 1983), but in contrast to it generalized layers are absent (fig. 3).

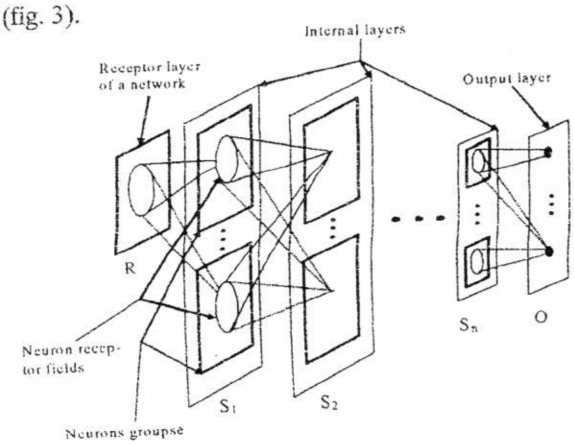

Fig. 3. NN structure.

The receptor layer R transforms pixel brightnesses to values of neuron activity.

Neurons that extract the same feature on the input data are united in one group. Input data for each neuron is the local part of data from the previous layer with displacement that corresponds to neuron indexes inside of it group.

Figure 4 shows the receptor fields of neurons from one group in case of one-dimensional input data.

In the case when input data are bidimentional, neurons also are organized in bidimentional structure. Thus, one group neurons activity forms a distribution of an extracted feature in the input data.

IV. DEFINITION OF DISTORTIONS INVARIANCE RANGES OF NEURONS

Let L be training images of size X on Y, then B is calculated according to formula (7):

$$B=\frac{1}{X\cdot Y}\cdot\sum_{x,y}\left(\frac{1}{L}\cdot\sum_{l}\left|a^{l}_{x,y}-\frac{1}{L}\cdot\sum_{l}a^{l}_{x,y}\right|\right) \qquad (7)$$

Let's define F as an average frequency of a brightness jump on the image, for which the neighbouring pixels belong to one level, if the absolute difference of their brightness is less than B. Now radius R of geometrical distortions we set as

$$R = [0.3/F], \tag{8}$$

where [x] is the nearest integer of x.

The formula (8) describes the fact, that if the radius of distortion exceeds on 30 % the average diameter of smallest object feature on image, than the image mixes up so that it is impossible to identify the feature and it is impossible to identify the image in the whole. Here the smallest feature of object should be understood as feature of object that cannot be divided into smaller features. The border of image area forming smallest feature is almost convex figure. «Almost convex figure» means that any secant drawn through

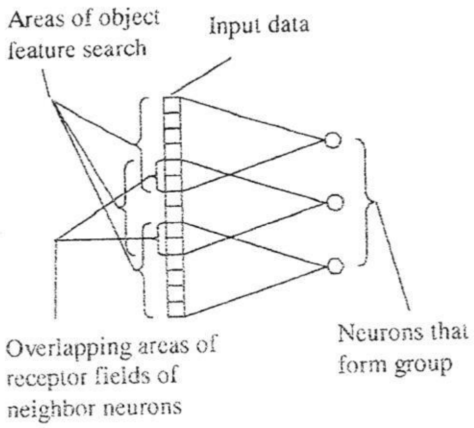

Fig. 4. The scheme neurons association in group.

its centroid crosses a fugure border only twice. Diameter of the smallest feature is defined as an average length of the secants.

Thus, using training we can find neuron parameters R and B.

V. Formation of training set for NN layer

The images for each layer are formed by the network developer according to specificity of recognized objects during neocognitron NN training. It is inefficient for automatic adaptation of recognition system to different training data and we developed an automatic algorithm for training set creation.

For each layer of NN it is possible to define the sizes of a neuron receptor field projection from some layer on an input layer of NN. In figure 5 the receptor field projection from 2-nd layer on an input layer is represented in case of one-dimensional input data.

Thus, for each training image it is possible to obtain K training images for some i-th layer of NN:

$$K = N_{in} - N_i + 1, \tag{9}$$

where N_{in} is the size of an input data and N_i is the size of a projection receptor fields of i-th layer.

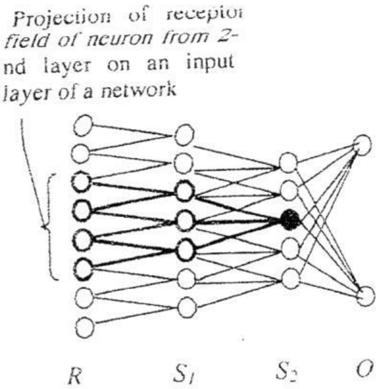

Fig. 5. A method of construction of a receptor field projection.

Training data for i-th layer of NN are considered as informative, if the condition is satisfied:

$$F_{in}/F_k \le 0.8 \text{ or } F_k/F_{in} \le 0.8, \tag{10}$$

where F_{in} is an average frequency of brightness jump on the input image and F_k is an average frequency of of brightness jump on the image received in a receptor field projection window.

Having training set for each layer, it is easy to notice that the problem of training of a layer is reduced to a layer input data clusterization problem.

VI. Algorithm of network layer training

Since neurons from one group have identical weights matrixes, the whole group can be represented in the form of one neuron, and a layer can be represented in the form of a single-layer NN, where each neuron corresponds to the group. For training such network we can use clusterization algorithm, where each cluster corresponds to one neuron. After training a network we use such NN as a prototype for construction of a required layer in a multilayer NN, where the neurons quantity in the prototype corresponds to the quantity of the groups, and the neuron weights of the prototype correspond to neurons weights matrixes of the corresponding groups.

If the training set of a layer is not great and the desirable clusters quantity is known, it is convenient to apply fuzzy C-means clusterization algorithm (Bezdek, 1981).

If the clusters quantity is not known, than it can be used peak grouping clusterization algorithm.

For the case, when the training set of a layer is great and the desirable clusters quantity is not known, we developed clusterization algorithm that is based on modification of multi-dimensional unsupervised FOREL (FORmal Element) algorithm.

VII. NN TESTING

Training set consists from 9 images of "upwards " contact pads and 9 images of "downwards" contact pads. The example of training images is presented on fig. 6.

Fig. 6. An example of training images.

The network has three layers: the receptor layer size is equal to 32x32; the neurons receptor field size of the first layer is equal to 3x3 pixel; the size the neurons receptor field size of output layer is equal to 30x30 pixel.

After training, the values of parameters R and B for each layer are the following: for first layer - R=1, B=28; for target layer - R=1, B=39. The quantity of neurons groups are: for the first layer - 61, for a target layer - 4.

A fragment of testing image is represented on figure 7. Areas of the image, which have been classified, are led round by a black square.

As a result from 57 contact pads it has been correctly found 46. The quantity misclassified pads was 3. Quantity not found pads was 11. Total the recognition percent was 80 %, percent of mistakes was 25 %.

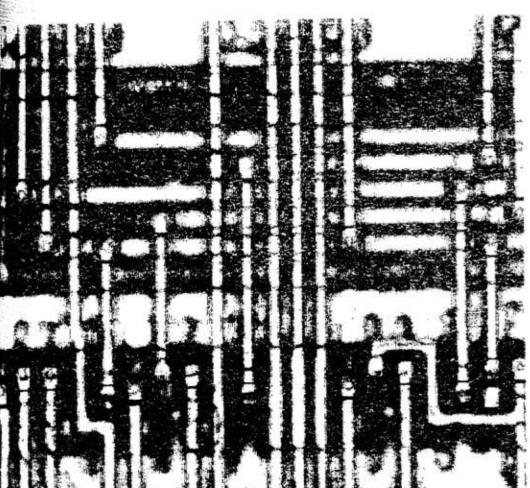

Fig. 7. Fragment of testing image.

As the volume of training sample was small (on 4 typical samples), the objects having very strong distortions, have not been recognized (fig. 8).

The recognition accuracy was increased up to 90% due to adding not classified images into old training set.

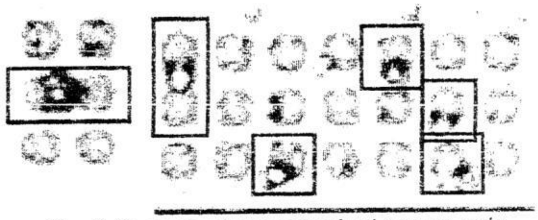

Fig. 8. Example of not recognized contact pads.

CONCLUSIONS

In the paper two problems were considered related to hierarchical NN neocognitron design: performance increasing and creation of unsupervised algorithm for layers training that determines weights of neurons and a number of sub-layers. The new fast activation and training rule of neuron were introduced. The rules are based on distortions invariant difference function that allows comparing matrices, which have changes of matrix element position and value. Algorithms of layer training set creation and layer training of network was developed. On the base of the algorithms full unsupervised algorithm of NN training was developed. Developed NN was approved on task of topological elements recognition.

REFERENCES

1. Fukushima, K., Miyake, S., Ito, T., 1983. Neocognitron: A neural network model for a mechanism of visual pattern recognition. In IEEE Trans. Syst., Man, Cybern., SMC-13, 826-834.
2. Hubel, D.H., Wiesel, T.N., 1962. Receptive fields, binocular interaction and functional architecture in cat's visual cortex. In Jornal of Physiology, 160, 106-154.
3. Lovell, D.R., Tsoi, A.C., 1992. The performance of the Neocognitron with various S-cell and C-cell transfer functions. In Intelligent Machines Lab., Dept. of Elec. Eng., Univ. of Queensland.
4. Wiskott, L., 2003. How does our visual system achieve shift and size invariance? In J. L. van Hemmen and T. J. Sejnowski, editors, Problems in Systems Neuroscience. Oxford University Press.
5. Satoh, S., Kuroiwa, J., Aso, H., Miyake, S., 1997. Recognition of rotated patterns using neocognitron, In Proc. Int. Conf. Neural Information Processing, 1, 112-116.
6. Pan, Z., Sabisch, T., Adams, R., Bolouri, H., 1999. Staged Training of Neocognitron by Evolutionary Algorithms. In Proceedings of the Congress on Evolutionary Computation, 3, 1965-1972. IEEE Press.
7. Wersing, H., Krner, E., 2003. Learning Optimized Features for Hierarchical Models of Invariant Object Recognition. In Neural Comp., 15(7), 1559-1588.
8. Behnke, S., 2003. Meter Value Recognition using Locally Connected Hierarchical Networks. In Proceedings of 11th European Symposium on Artificial Neural Networks, 535-540.
9. Scalzo, F., Piater, J., 2005. Statistical learning of visual feature hierarchies In IEEE Workshop on Learning in CVPR.
10. Sadykhov, R., Klimovich, A., Vatkin, M., 2001. The models of neural networks for videoimage processing in industrial applications. In NIMA-SC2001. NATO Advanced Study Institute on Neural Networks for Instrumentation, Measurment, and Related Industrial Application,. 78-84.
11. Bezdek, J. C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum Press, New York.