UDC 681.327.12 (045)
DOI: 10.18372/1990-5548.56.12937

[1]O. I. Chumachenko,
[2]A. T. Kot

NEURAL NETWORKS MODULE

[1,2]National Technical University of Ukraine "Ihor Sikorsky Kyiv Polytechnic Institute," Kyiv, Ukraine
E-mails: [1]chumachenko@tk.kpi.ua, anatoly.kot@gmail.com

Abstract—*It is considered a basic approach for hybrid neuron network creation. As an example, the counter propagation neural network is analyzed. It is effectively used for image processing. Two modes of this neuron network functioning are considered. They are: accreditation and interpolation. Interpolation mode permits to reveal more complex features and can supply more precise results. Based on this analysis it is developed a new hybrid structure that includes Kohonnen neural network and perceptron. It is proposed a learning algorithm of this hybrid neuron network.*

Index Terms—Hybrid neural networks; Kohonnen neural network; perceptron; learning algorithm.

I. INTRODUCTION

The problem of combining different types of neural structures in a single architecture, which leads to properties that they do not have separately is often discussed [1] – [6]. The example of such combining is counter propagation network. In this article, it is developed a system architecture based on the counter-distribution network, but instead of the Grossberg network, it is taken a single-layer perceptron. Such hybrid neural network, consisting of Kohonen layer and single-layer perceptron has much better characteristics than the network with one hidden layer of neurons. Figure 1 shows a simplified version of a direct action hybrid.

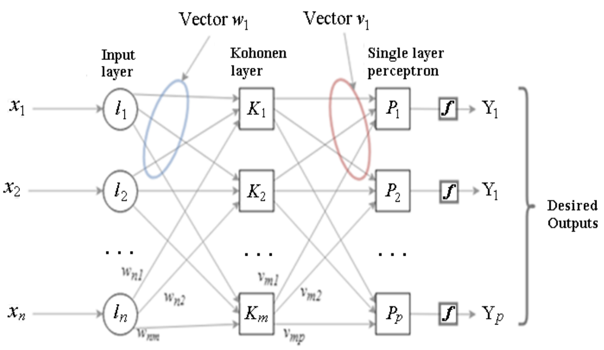

Fig. 1. Simplified structure of hybrid neural network

The ability to generalize, which allows to obtain the correct output at an incomplete input vector is one of the main characteristics of hybrid neural networks. It allows you to effectively use this network for pattern recognition.

II. SOLUTION OF THE PROBLEM OF STRUCTURAL SYNTHESIS OF HYBRID NETWORK

To determine the problem of structural synthesis of a hybrid neural network based on Kohonen and the perceptron, it is necessary to analyze three mechanisms necessary for the operation of this network. The first mechanism is the criterion for selecting "winner" neurons and the frequency of selecting the same neuron as the "winner". This is quite an important question, because there may be a situation when one neuron several times becomes a "winner" and gradually increases its weight, which will lead to the appearance of "dead" neurons that will no longer participate in the training of the network. One of the methods of this problem solution is the method of "justice".

The next mechanism is the strategy of neurons – "winners" choosing. This is done through the interpolation mode in which the Kohonen network operates. Let us consider in more detail the principle of this mode.

The main difference between the accreditation mode and the interpolation mode is the ability to choose not one "winner", but several. The problem is how many "winners" to choose for the pattern recognition problem, in this case – numbers.

The third mechanism is the connection of the Kohonen layer neurons with the layer of the perceptron. The peculiarity of the single-layer perceptron is the absence of "hidden" layers. This means that the input vector immediately forms the output vector depending on the weight on the neuron. The single-layer perceptron does not use the back propagation method because this learning algorithm works for multi-layer networks. Therefore, when training the network, it is used the correction method. Nonlinear activation functions are also used by multilayer networks, so in our case it is more reasonable to use the threshold activation function, which results in 0 or 1.

III. PERCEPTRON. FEATURES AND STRUCTURE

The perceptron is the simplest network and consists of a single layer of artificial neurons connected by weight coefficients with multiple inputs (Fig. 2).

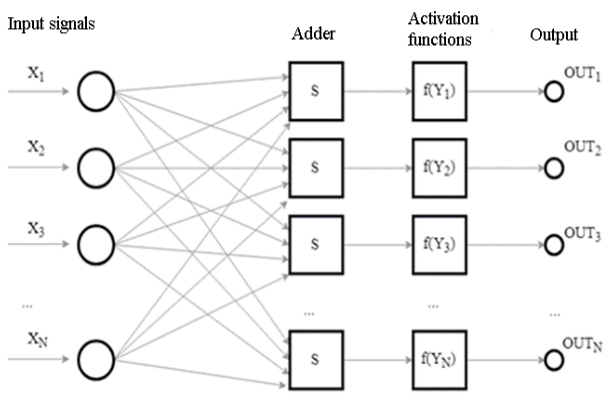

Fig. 2. Structural diagram of the perceptron

The perceptron is trained by feeding a set of images one at a time to the input and adjust its weight until all the elements have reached the desired output.

During training, the so-called Delta rule is used to change the weight.

It is necessary to remark some features:

– single-layer perceptron is represented as a mesh size *a* on *b*. As a result, it gives us *a·b* neurons in the perceptron layer.

– as a result, the output from the networks is compared to the expected value. If the network is wrong, change the weights of those neurons that were part of the neurons with incorrect classified figure.

Let's start with the Kohonen network. The work of this network takes place according to the standard scheme. The only difference is that in combination with the perceptron the network can operate in the interpolation mode, not accreditation. Let us examine this mode in more detail.

In the interpolation mode, several neurons can become "winners". Their number depends on the task.

Algorithm for selecting neurons – "winners" has the following form:

1) Set a certain value C (count), which will be equal to the required number of neurons – "winners" for the task.

2) Calculate the values of the Kohonen layer neurons outputs

$$\delta = T - OUT. \qquad (1)$$

3) After receiving the values of all output signals, select *C* maximum values by format $\max(OUT_k), k = 1...M$, choosing one neuron-"winner" – $OUT_k = 1, C = C - 1$, then repeat the search for the next maximum.

4) Choosing C neurons-"winners", the output signals of other neurons is equal to 0.

5) We are sending output signals to the neuron-"winners" to the input of the perceptron layer.

We will analyze how many neurons are needed for the most optimal network operation for the pattern recognition problem if we use a single-layer perceptron.

If *C* = 2, which is the minimum for the interpolation mode, the network will select two neurons with the highest result value according to the equation (1). Their output will be 1 and the others 0. That is, these are the two figures that are closest to each other. An example is a pair of numbers 0 and 8, 3 and 8, 2 and 7, 7 and 1 and so on. As a result, if the image was not clear enough, there is a possibility of ignoring the required number and not defining its class as a "neuron-winner", which will lead to low accuracy.

If *C* = 3, the network will select three neurons with the highest result value according to the equation (1). Their output will be 1 and the others 0. That is, these are the three figures that are closest to each other. An example is the three digits 3, 8, and 0, and 2, 7, and 1. As a result, if the image was not clear enough, the probability of including the required number and determining its class as a winner neuron is higher than in the case when *C* = 2. But for our case, there is a problem, because the training immediately image of three digits may not lead to a correct change in the weights of neurons. If it is necessary to choose 3 or more neurons-"winners", it is more expedient to use a multilayer perceptron.

If *C* = 4 and 5, the network will select four or five neurons with the highest result value. Since the similarity of numbers as a whole consists of three elements (given in the example C = 3), such a choice is the most optimal.

Since the optimal number of "neurons-winners" is not determined and depends on the task, we can assume that the optimal value is found – *C* = 2 and 3. In our case, we will use 2, because it is necessary to check the quality of the network. In the future, the network can be developed by replacing the single-layer perceptron on the multi-layer.

Let's move on to the key feature and at the same time the problem of the Kohonen network – the probability of choosing the same neuron for the "winner" neuron, which will lead to a primitive ignoring of the training of other neurons. The method of solving this problem is the method of "justice". In this case, it is optimal because of the simplicity of its implementation and, as a result, the homogeneous training of all the neurons of the Kohonen layer. It is determined as:

$$pass = \frac{1}{M}, \qquad (2)$$

where M is the number of iterations for which a neuron can be only once a "winner"; $pass$ is the value of the neural network weight change.

The rule of use of this method is: "for a neuron that has received the status of "winner" more than once in M iterations, the weight decreases, but the activation of the neuron is not canceled." Thus, for the neuron that the second time becomes the winner for a certain number of iterations, the weight decreases, thereby reducing its "rank", which makes it possible to teach other neurons.

$$w_{k(i+1)} = w_{k(i)} - pass, \qquad (3)$$

where $w_{k(i+1)}$ is the new weight of "neuron-winner"; $w_{k(i)}$ the old weight of the "neuron-winner", and is the iteration number. With this method, other neurons on the background of such a "neuron-winner" "rise" in its "rank".

From equation (2), we see that it is not advisable to use a low value of M, since then the weight of the neuron will be severely cut, which dramatically reduces its potential to become a "neuron-winner" in the further training of the network.

Analyze the result of the network at various values of M. At $M = 1$, in fact, the network does not have any criterion for lowering the rank, since each subsequent iteration will be new for the entire network and the previous "neuron-winner" will be able to become it again. It is also inappropriate to use this value, since the weight will decrease to a critical value.

At $M = 2$, the network has some control over the situation, but only if two consecutive times one and the same neuron claims to be the "winner". This number of iterations makes sense, but is not very effective. It is also inappropriate to use this value, since the weight will decrease to a critical value.

At $M = 3$, the network more actively analyzes the "winners", but there are still situations where certain neurons will become winners every fourth or fifth iteration, thereby not falling under the method of "justice" and at the same time preventing the training of other neurons. In this case, the weight is reduced to not too large, but still sufficiently large.

At $M = 7...10$ the network maximally controls the moment of determining the neuron-"winner", which negatively affects the learning of the network. If you take the maximum value in 10 iterations, it actually means that for ten iterations, none of the neurons can not be selected as a "winner" twice, because it will lead to weight reduction.

Given this analysis, we can conclude that the optimal value of M is 4.5 or 6. The value of the number of iterations for the "equity" method It's best to test software programmatically and choose one of the three best options that will give the network the highest accuracy.

IV. Parametric Synthesis of Hybrid Neural Network

Let's turn to the parametric synthesis of a hybrid neural network (HNN) based on Kohonen network and perceptron.

As already investigated in structural synthesis, the Kohonen network works for with its standard scheme. When connecting a network with a perceptron is the best the result of the hybrid network will be in interpolation mode.

Hybrid Neural Network Training Algorithm

1) We supply image signals from the input layer to all 10 neurons in the Kohonen network.

2) Normalize the input signals of the image, dividing all the input signals by 255 (since the value of all pixels varies from 0 to 255).

$$x_{ni} = \frac{x_i}{255},$$

where x_i is the value of the ith input signal; x_{ni} is normalized value.

3) For a given number of "neurons-winners" we choose the winners from the biggest value of output.

4) "Neurons-winners" receive at the output – 1, other neurons – 0. If the neuron becomes the winner more than once in M cases – we reduce its weight.

5) Next, the neurons with the output value of 1 transmit it to the inputs of perceptron. The connection Kohonen network and perceptron is executed with help of Drop Connect method.

6) The value of the output signal Y_i of the neurons in the perceptron layer is compared with the reference output corresponding to the given digit.

7) Enter the value of the learning rate $\delta = 0.1$. During training, this value will be reduced by 2 times with each iteration.

8) The output signals of the perceptron neurons will be activated by the threshold activation function $f(Y)$ represented as

$$f(Y)=\begin{cases}0, & Y<\theta,\\ 1, & Y\geq\theta.\end{cases}$$

9) In neurons, the values of which did not coincide with the reference, the threshold and weight of the input is changed. New values can be calculated using expressions

$$w_i=\begin{cases}w_i-\delta, & Y_i=1,\\ w_i+\delta, & Y_i=0.\end{cases}\qquad \theta_i=\begin{cases}\theta_i+\delta, & Y_i=1,\\ \theta_i-\delta, & Y_i=0.\end{cases}$$

10) After changing the weight and threshold values, the value of the learning speed coefficient is changed according to the next formula. That will allow us to get closer to the exact value sooner.

$$\delta=\delta/2.$$

11) After completing the iteration, new output values are compared to the references. The process is repeated as long as the deviation does not become less than two characters.

V. CONCLUSION

The network of counter propagation is analyzed. Thanks to the results of this analysis, the architecture of the hybrid neural network based on the Kohonen network and the perceptron, focused on the task of recognizing handwritten digits, was designed.

In order to maximize the accuracy of the work of the network, two basic mechanisms, the method of "justice" for the ability to teach all neurons of the Kohonen network and the approach to choosing neurons-"winners" in interpolation mode, are analyzed.

As a result of the conducted research, an algorithm for the operation of the hybrid neural network was obtained, allowing to recognize handwritten digits with an accuracy of more than 90%. The developed model can be used for recognition of automobile license plates, or a more global task – *X*-ray detection and diagnosis..

REFERENCES

[1] V. V. Borisov, V. V. Kruglov, and A. S. Fedulov, Fuzzy models and networks. 2 nd ed., The stereotype. 2012. (in Russian)

[2] Khaykin Saymon. Neural networks: full course, 2nd edition. 2006. (in Russian)

[3] A. P. Rotshteyn, Intelligent identification technologies: fuzzy sets, neural networks, genetic algorithms. Monograph. Vinnitsa: "Universum-Vinnitsya," 1999, 295 p. (in Russian)

[4] J.-S. R. Jang, ANFIS: Adaptive-Network-Based Fuzzy Inference Systems, IEEE Trans. Systems, Man & Cybernetics 23 (1993).

[5] Y. Bengio, A. Courville, and P. Vincent, Representation Learning: A Review and New Perspectives. Department of computer science and operations research, U. Montreal. 2014.

[6] X. Glorot and Y. Bengio, Understanding the difficulty of training deep feedforward neural networks. 2010.

Received February 12, 2018

Chumachenko Olena. Candidate of Science (Engineering). Assosiate Professor.
Technical Cybernetic Department, National Technical University of Ukraine “Ihor Sikorsky Kyiv Polytechnic Institute,” Kyiv, Ukraine.
Education: Georgian Politechnic Institute, Tbilisi, Georgia, (1980).
Research area: system analysis, artificial neuron networks.
Publications: more than 60 papers.
E-mail: chumachenko@tk.kpi.ua

Kot Ananatoliy. Post-graduate student.
Technical Cybernetic Department, National Technical University of Ukraine “Ihor Sikorsky Kyiv Polytechnic Institute,” Kyiv, Ukraine.
Education: National Technical University of Ukraine “Ihor Sikorsky Kyiv Polytechnic Institute,” Kyiv, Ukraine, (2017).
Research area: artificial Intelligence.
Publications: 2.
E-mail: anatoly.kot@gmail.com

О. І. Чумаченко, А. Т. Кот. Модуль нейронних мереж

Розглянуто базовий підхід для побудови гібридної нейронної мережі. Як приклад, проаналізовано нейронну мережу зустрічного поширення. Розглянуто два режими функціонування цієї нейронної мережі. Це режими акредитації та інтерполяції. Режим інтерполяції дозволяє виявляти більш складні особливості і дає більш точні результати. На основі цього аналізу розроблена нова гібридна структура, що включає нейронну мережу Кохонена і персептрон. Запропоновано алгоритм навчання цієї гібридної нейронної мережі.

Ключові слова: гібридні нейронні мережі; нейронна мережа Кохоннена; персептрон; алгоритм навчання.

Чумаченко Олена Іллівна. Кандидат технічних наук. Доцент.
Кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», Київ, Україна.
Освіта: Грузинський політехнічний інститут, Тбілісі, Грузія, (1980).
Напрям наукової діяльності: системний аналіз, штучні нейронні мережі.
Кількість публікацій: більше 60 наукових робіт.
E-mail: chumachenko@tk.kpi.ua

Кот Анатолій Тарасович. Аспірант.
Кафедра технічної кібернетики, Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», Київ, Україна.
Освіта: Національний технічний університет України «Київський політехнічний інститут ім. Ігоря Сікорського», Київ, Україна, (2017).
Напрям наукової діяльності: штучний інтелект.
Кількість публікацій: 2.
E-mail: anatoly.kot@gmail.com

Е. И. Чумаченко, А. Т. Кот. Модуль нейронных сетей

Рассмотрен базовый подход для построения гибридной нейронной сети. В качестве примера, проанализирована нейронная сеть встречного распространения. Рассмотрены два режима функционирования этой нейронной сети. Это режимы аккредитации и интерполяции. Режим интерполяции позволяет выявлять более сложные особенности и дает более точные результаты. На основе этого анализа разработана новая гибридная структура, включающая нейронную сеть Кохоннен и персептрон. Предложен алгоритм обучения этой гибридной нейронной сети.

Ключевые слова: гибридные нейронные сети; нейронная сеть Кохоннена; персептрон; алгоритм обучения.

Чумаченко Елена Ильинична. Кандидат технических наук. Доцент.
Кафедра технической кибернетики, Национальный технический университет Украины «Киевский политехнический институт им. Игоря Сикорского», Киев, Украина.
Образование: Грузинский политехнический институт, Тбилиси, Грузия, (1980).
Направление научной деятельности: системный анализ, искусственные нейронные сети.
Количество публикаций: более 60 научных работ.
E-mail: chumachenko@tk.kpi.ua

Кот Анатолий Тарасович. Аспирант.
Кафедра технической кибернетики, Национальный технический университет Украины «Киевский политехнический институт им. Игоря Сикорского», Киев, Украина.
Образование: Национальный технический университет Украины «Киевский политехнический институт им. Игоря Сикорского», Киев, Украина, (2017).
Направление научной деятельности: искусственный интеллект.
Количество публикаций: 2.
E-mail: anatoly.kot@gmail.com