UDC 004.032.26

RATE INCREASE OF THE OBJECTS CLASSIFICATION ON THE CONVOLUTIONAL NEURAL NETWORKS WITH THE SELF-ORGANIZATION MAPS IMPLEMENTATION

Shapovalova S. I., Ph. D.
Moskalenko Yu.V.
National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute" (Kyiv, Ukraine)

Objects and processes classification is a common experimental problem. Its solution, first of all, is needed in automatic diagnosis systems, for example, to determine the equipment operation state through the diagnostic signal or to identify abnormalities with medical images. With the development of convolutional neural networks, new prospects for solving such problems have risen up. However, the classification accuracy that can be achieved on these networks is not sufficient enough for all diagnosis issues. It is subject to, for example, timely diagnosis of the onset of transient phenomena. At the same time, another type of neural network, Kohonen self-organizing maps, has a conceptual property for training on unclassified set of classes, that is, giving the opportunity to solve such an issue. Therefore, the accuracy enhancement of the classification on the basis of Kohonen networks implementation in the architecture of convolutional networks is a relevant objective and has practical significance.

The article analyzes the means of improving the accuracy of convolutional neural networks and arising problems solution. The ways of increasing the proportion of correct clustering on Kohonen networks due to the 'growth' of its grid shape in determining new classes in the learning process are also given. It is this property that makes it possible to recognize transient phenomena. It is determined that the existing solutions of the combination of Kohonen networks and convolutional networks are aimed at improving the efficiency of only self-organizing maps, with the purpose to improve the accuracy of classification by convolutional networks, it became necessary to develop a new architecture. The paper provides a description of this issue. Since the initial information of the Kohonen networks is the weighting matrix values of the grid shape neurons, it was necessary to associate it with the representation of the images in order to process the diagnostic images. The paper proposes the concept of a built-in associative array block based on Kohonen networks.

According to the proposed method, a software implementation of a hybrid neural network is developed. The formulation and results of *computational experiments are presented. The efficiency of the proposed method is experimentally proved.*

Keywords: convolutional neural networks, CNN, Kohonen self-organizing maps, SOM, signal classification, image classification.

Problem statement. Classification based on neural network methods was one of the first objectives of artificial intelligence. With the development of deep learning networks, new perspectives have risen up for the classification of processes by diagnostic signals and images. Such objectives are typical for technical, medical, seismic diagnostics. When solving them, accuracy is crucial. The problem is that on convolutional neural networks, the required classification accuracy may not be achieved for all diagnostic objectives. For example, such networks are not always capable to detect the transition stages of the diagnosed processes. To solve such objectives, there is another kind of neural networks — Kohonen self-organizing map (self-organizing map – SOM). However, SOM is designed to solve the clustering objective and requires the use of additional tools to further classifying of the current example. The use of SOM as a convolutional network unit extends the class of technical diagnostics objectives and increases the accuracy of classification. Therefore, accuracy enhancement of classification through the use of Kohonen self-organizing maps as part of the cohort network is an urgent objective and has practical significance.

Analysis of recent research and publications. Deep learning networks are primarily designed for classification, discovery, segmentation objectives. The current research in deep learning aims to improve accuracy and/or reduce work time.

The results of modern research on machine learning can be tested in special competitions, the most famous of which is ImageNet - competition for the classification of objects in images. The first significant improvement in classification accuracy was achieved by AlexNet [1], which won the mageNet LSVRC-2012 competition. This network was a modification of LeNet [2] and had several main features:

- ReLU activation function, which allowed to increase the learning speed by reducing the gradient vanishing;
- accidental exclusion of the group of connections between neurons (DropOut), which partially solved the problem of overfitting.

The ImageNet LSVRC-2014 competition was won by VGG [3], which was a modification of AlexNet. In the VGG network, the convolutional high-dimensional filters have been replaced by a series of 3*3 filters. With this approach, the receptive field of an individual neuron does not decrease, but the number of network parameters decreases, making the network easier to train.

In 2015, Google's Inception (GoogleNet) network was introduced. [4] When implementing the network, the idea laid down in VGG was developed — replacing the n*n filters with a sequence of n*1 and 1*n filters. Besides, on the basis of Inception, the Inception V2 and V3 networks were developed.

However, these networks were characterized by the degradation problem. The problem is that an increase in classification accuracy can be achieved by adding new layers, but this can be achieved to a certain threshold, after which the accuracy begins to decrease. The main reason for this is the vanishing gradient. On the last layers of the network, the gradient is significant, and on the first — unacceptably small. The degradation problem was solved by ResNet network [5] with new links allowing the input signal to be added to the result of multiple layers of convolution (shortcut connection).

One of the last means of accuracy enhancement was the attention mechanisms [6].

However, all these methods to improve the accuracy of convolutional networks do not always provide the desired accuracy in the diagnosis objectives in cases where it is impossible to introduce a rigid classification of the process due to the emergence of transitional stages, which, in turn, must be recognized. On the other hand, the Kohonen self-organizing map is a kind of neural networks specifically designed to solve the problem of clustering, which allows for training in pre-defined classes. The basis of this work was the assumption that this property SOM will improve convolutional neural networks. SOM properties in combination with deep learning networks should allow for more accurate identification of transient processes with diagnostic signals.

SOM significant advantages include the ability to train on a small number of training sample examples, fast training and resistance to noisy data. However, the SOM network is sensitive to initial initialization of weights, so that the test result may be unstable when conducting a series of identical experiments. Studies presented in [7-9] solve this issue and increase the proportion of correct operation in operating modes. These papers propose an approach aimed at creating a SOM network structure that is changing in the process. The basic idea is to extend the SOM grid shape if the new training sample example differs significantly from the neural network weights. That is, the network can ‘grow’ in the learning process.

Since the definition of the current object class is carried out behind the SOM neuron output matrix, further classification is carried out either on additional neural networks or behind the visualization of a trained map with a dedicated current example. Traditionally, the additional circuitry in combination with the SOM uses a multi-layered perceptron. In modern studies [10-12], SOM has been connected to convolutional networks. These studies aim to improve SOM clustering. In [10], two types of SOM are presented, which differ in application. In the first case — instead of convolutional layer, in the other — instead of pooling. In [11], the concept is to train multiple SOM, each corresponding to a separate area of the input image. In [12], the hidden layers are replaced by modified self-organizing maps. However, all these studies are aimed at improving the efficiency of Kohonen's self-organizing maps. Therefore, it is necessary to conduct research on the concept of improving the accuracy of classification of convolutional neural networks by adding SOM.

The aim of the article is to propose an architectural solution of the neural network complex of the convolution neural network with the Kohonen self-organizing map to improve the classification accuracy. To achieve this object in view the following tasks have been solved:

- to design the concept of SOM integration in the deep learning network;
- to develop a hybrid convolutional network architecture with an integrated Kohonen self-organizing map;
- to carry out the results comparative analysis of the test problems solution on a convolutional neural network and the network modified in a set way.

Implementation of SOM block in deep learning network. It is not possible to directly integrate SOM into convolutional neural networks (CNN) through heterogeneous representations of information transmitted between CNN blocks and outgoing SOM information. The problem lies in the fact that the result of SOM operation in the working mode is a matrix of distance differences between the input vector and the vector of weights of neurons, by which the winning neuron is selected. This result must be compared with the actual display of the input signal, which must be provided further to the convolutional layers. It was proposed to implement the matching of winning neurons to specific fragments of the input signal. To do this, the SOM network is combined with an associative array. This array is implemented as a reflection of the SOM grid shape in the form of a matrix. The elements of such an array are represented as fragments of the SOM input signal. The SOM block implements a direct correspondence between a neuron and a matrix element.

Figure 1 schematically shows a fragment of a two-dimensional SOM grid shape displaying an associative array for the handwriting recognition test. For clarity, the association is presented for the input block, because the following blocks, firstly, many channels, and secondly-less clear contours.

The SOM grid shape can have different dimensions:

– for the input signal as a sequence of values (1D signal) - two-dimensional grid;

– for input signal-images (2D signal) - three-dimensional grid.

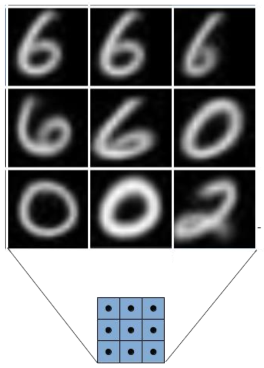

Fig. 1. Associative matching of images on SOM fragment in input block

The associated array has the same dimension as the SOM lattice. The neuron indexes in SOM are the same as the indexes of a cell in an array. However, with such a matching system, the network does not learn universally, but only from the examples of the training sample and is not capable of generalization in the future. That is, the network is also overfitting. Therefore, it is proposed that in the operating mode, the SOM block returns a signal that is obtained by generalizing all associated signals of the winning neuron and its surroundings. The output signal of the SOM block is calculated by the formula:

$$0.5 * I_{winner} + \frac{0.5*1}{n} \sum_{k=0}^{n} I_n \quad (1)$$

where I_{winner} is the signal associated with the winning neuron,

I_n – associated signal of a neighboring neuron,

n – is the number of neighbors.

During SOM training, the winning neuron is selected. After that, the signal associated with it is mapped to the signal received from the previous block. Signal fragments are also updated in the same way as weights in the Kohonen neural network are updated:

$$Inew = (1 - t) * I_{old} + t * I, \quad (2)$$

where I – is the input signal,

I_{old} – signal - associated with this cell,

t – is a function that vanishes away from the winning neuron.

The proposed SOM block is attached to the CNN base blocks.

The architecture of the hybrid network. The convolutional network contains the following base layers:

– convolution layer - ***(Conv)*** basic component of convolutional neural networks;

– pooling layer - ***(Pool)*** to reduce the dimension of the input signal;

– activation layer - ***(Activation)*** to introduce nonlinearity into the network;

– full connection layers ***(Dense)*** - for direct classification-usually used in the last layers of the network.

In addition, support layers are used, for example, normalization layers

(Norm), link exclusion layers, global pooling ***(Global pooling)***, uniting layers ***(Concatenate)*** and others.

In this research, SOM was added to the traditional layers. Figure 2 shows the convolutional network block architectures: for the traditional case (2a) and the two proposed architectures using the SOM block.

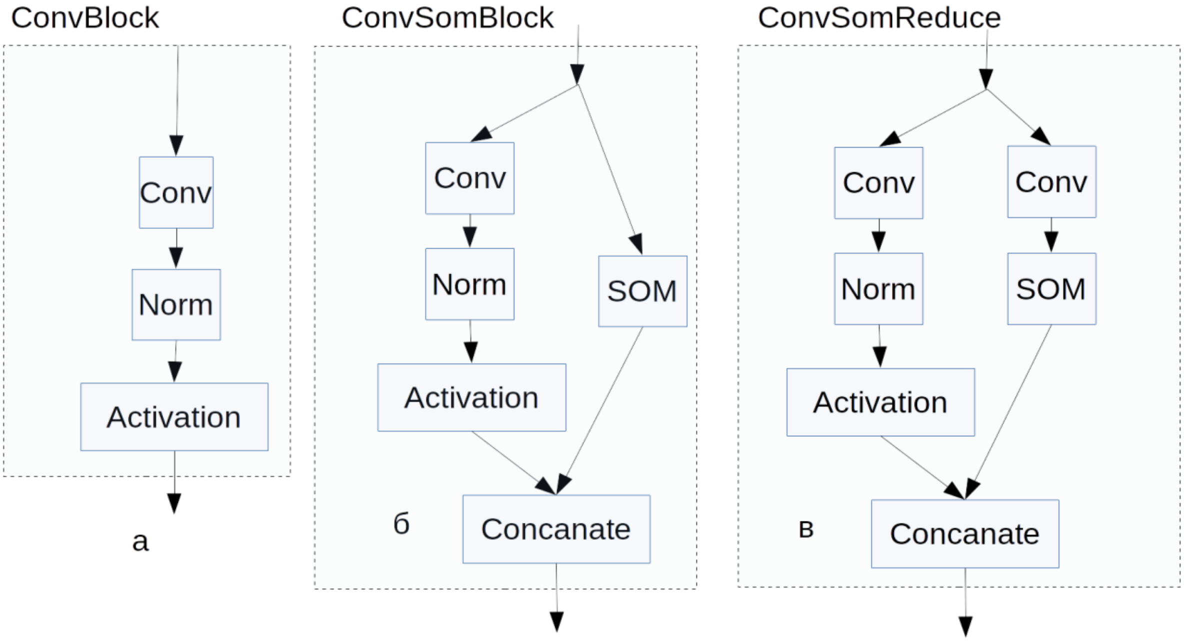

Fig. 2 – convolutional network base unit: a) currently existing, b) using the SOM block, c) using the SOM block with the previous dimension reduction

The input signal is fed in parallel to the convolutional layer and the SOM layer (Figure 1,b). After processing, the results are concatenated (combined) and then fed to the next layers. The unit is the SOM architecture (2,b), (2,c) applied starting from the second block from the input layer of the SOM network. This is necessary for the first SOM network to be signaled after the convolution layer. Otherwise, the input signal will be almost unchanged throughout the network. Thus, in the first (input) CNN block, the SOM layer is not applied.

Determining the optimal number of layers is a separate issue. To avoid this, often use proven architectures such as ResNet34, ResNet50, VGG16 and others.

Because input signals, such as graphics, are noisy in applications, it is necessary to reduce their impact on SOM operation. For this purpose, batch normalization is used. Thus, the SOM layers will be less sensitive to noise and changes in the amplitude of the input signal. If batch normalization is used in the convolution block, the use of normalization before the SOM layer is not required.

In the concept of convolutional networks has been widely used to reduce the dimension of the signal with a simultaneous increase in the number of channels. The SOM layer must be used independently for each channel. Since the neural network on the last layers often has many channels, to save computational resources and memory on the lower layers before the SOM block used convolution dimension 1*1, which reduces the number of channels (Fig. 1, c). The block in Figure 2,b it would be advisable to use if the number of channels is less than 8. Otherwise, it is advisable to use a block with architecture 2.c.

Research results. To experimentally confirm the effectiveness of the proposed method of embedding SOM in CNN, the following test issues were solved:

– classification by diagnostic signal, which is set as a time series of values of characteristic indicators;

– handwritten digit recognition on MNIST dataset.

For realization of the first issue the example on recognition of fitting curves of the second order on similar fragments, upper halves of a circle, an ellipse, a parabola is chosen. The problem is provided in [13]. The characteristic signal is calculated by second-order fitting curve equations: circle, ellipse, parabola. For each discrete value, a noise curvature in a given range is randomly generated. Thus, the input signal is a list of discrete values of the function. The data set consists of train and validation samples, each containing 999 original examples: 333 examples for each fitting curve. The length of the input signal vector is 100.

The test problem of recognition is solved on a dataset MNIST - a database of handwritten digits. Every sample is represented as an image with a scale of 28*28 pixels. The dataset also consists of train and validation samples. Train one contains 60000 single-channel (black and white) images, validation-10000. The task is to classify the numbers in the current image.

2 neural networks were constructed for computational experiments. The first one uses traditional convolutional ***ConvBlock*** blocks (Fig. 2,a), in the second - the proposed convolutions using SOM - ***ConvSomBlock*** and ***ConvSomReduce***. Simple models were chosen for test neural networks. For correctness of experimental set up both networks, both traditional and modified, have identical architecture. The block and layer designations match the descriptions provided when defining the hybrid network architecture.

To solve the first issue, all convolutions were one-dimensional (1D), and the SOM matrices were two – dimensional (2D), in the second issue, respectively, 2D and 3D. In the experiments, the maximum number of SOM channels was 4, 8, 16 and 32.

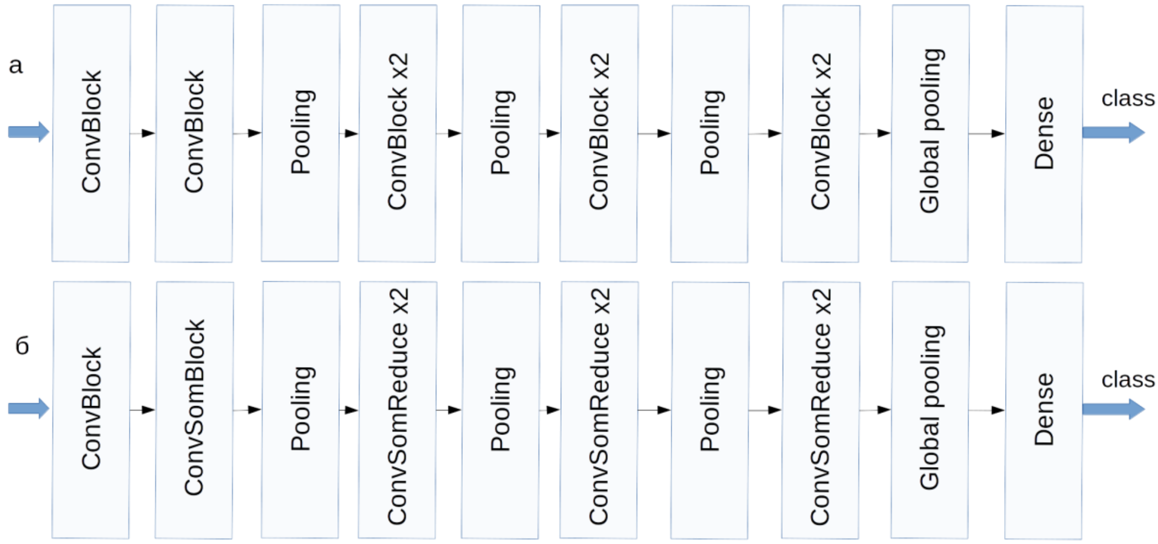

Fig. 3. Network architectures: a) a convolutional network, b) hybrid network

The accuracy metric in all experiments was the ratio of the number of correctly defined classes to the total number of examples.

The results of the experiments are shown in Table 1.

Table 1.

Results of computational experiments

Network	Number of SOM channels	Accuracy metric
The classification problem for time series		
Convolutional	-	0.9629
Hybrid	4	0.9709
Hybrid	8	**0.9719**
Hybrid	16	0.9709
Hybrid	32	0.9709
Image classification task		
Convolutional	-	0.9937
Hybrid	4	0.9945
Hybrid	8	**0.9949**
Hybrid	16	0.9948
Hybrid	32	0.9945

The results of the experiment proved that the proposed method of modifying the signal network increases the accuracy of classification for the problem of signal recognition over a time series in the range of 0.9629 to 0.9719; for the problem of image recognition – from 0.9937 to 0.9949.

Conclusions.

1. A method of association of the self-organizing grid shape reflection with fragments of the input signal for embedding the SOM block in a hybrid convolutional neural network is proposed.

2. The architecture of hybrid convolutional network with SOM block is proposed.

3. The efficiency of the proposed method of convolutional network modification is experimentally proved.

Prospects for further research are the improvement of SOM block algorithms, in particular the implementation of vanishing depending on the distance to the winning neuron, as well as the study of the influence of the number of channels on the accuracy of classification for complex images.

References

1. Krizhevsky A., Sutskever I. and HintonG. Imagenet classification with deep convolutional neural networks. In *NIPS*. 2012.
2. Le Cun Y., Boser B., Denker J. S., Henderson D., Howard R. E., W. Hubbard, and Jackel L. D.. Backpropagation applied to handwritten zip code recognition. Neural computation. 1989.
3. Simonyan K. and Zisserman A. Very deep convolutional networks for large-scale image recognition. In ICLR. 2015.
4. 4.Szegedy C., Reed S., Erhan D., and Anguelov D. Scalable, high-quality object detection. arXiv:1412.1441v2. 2015.
5. He K., Zhang X., Ren S., and Sun J. Deep residual learning for image recognition. In CVPR. 2016.
6. Hu J., Shen L., Sun G. Squeeze-and-excitation networks. arXiv preprint arXiv:1709.01507. 2017. 5, 7.
7. Cao M., Li A. Growing self-organizing map approach for semantic acquisition modeling Proceedings of the 4th IEEE International Conference on Cognitive Infocommunications (CogInfoCom). 2013. P. 33–38.
8. Furao G., Hasegawa O. An incremental network for on-line unsupervised classification and topology learning Neural Netw N19. 2006. P. 90–106.
9. Furao S., Ogura T., Hasegawa O. An enhanced self-organizing incremental neural network for online unsupervised learning Neural Networks N20. 2007. P. 893–903.
10. Hiroshi D., Gen N., Satoru A. Convolutional self-organizing map. International Conference on Computational Science and Computational Intelligence (CSCI). 2016.
11. Liu N., Wang J., Gong Y. Deep Self-Organizing Map for visual classification. In Neural Networks (IJCNN). International Joint Conference. 2015. P. 1-6.
12. Trappenberg T., Hollensen P., Hartono P. Classifier with hierarchical topographical maps as internal representation. Intelligent engineering

systems (INES), IEEE 19th international conference. 2015. P. 341–345.
13. Shapovalova S., Moskalenko Yu. Increasing the share of correct clustering of characteristic signal with random losses in self-organizing maps. *Eastern-European Journal of Enterprise Technologies*. 2019. № 2/4(98). P. 13-21.

КЛАСИФІКАЦІЯ ОБ'ЄКТІВ НА ЗГОРТКОВИХ МЕРЕЖАХ З ВИКОРИСТАННЯМ КАРТ САМООРГАНІЗАЦІЇ

Шаповалова С.І., Москаленко Ю.В.

Класифікація об'єктів і процесів є розповсюдженою прикладною задачею. Її розв'язання, насамперед, потрібно в системах автоматичного діагностування, наприклад, для визначення стану роботи обладнання за діагностичним сигналом або виявлення патологій за медичними знімками. З розвитком згорткових нейронних мереж відкрились нові перспективи розв'язання таких задач. Однак не для всіх випадків діагностування достатньо точності класифікації, яку можна досягти на цих мережах. Це стосується, наприклад, своєчасного діагностування початку перехідних процесів. Водночас інший різновид нейронної мережі – карти самоорганізації Кохонена – мають концептуальну властивість для навчання за невизначеною множиною класів, тобто дають можливість розв'язувати таку задачу. Тому підвищення точності класифікації на основі вбудування мереж Кохонена в архітектуру згорткових мереж є актуальною задачею і має практичну значущість.

В статті проведено аналіз засобів підвищення точності згорткових нейронних мереж та розв'язання проблем, які при цьому виникають. Також наведені способи підвищення частки коректної кластеризації на мережах Кохонена за рахунок 'зростання' її решітки при визначенні нових класів в процесі навчання. Саме ця властивість дає можливість розпізнавання перехідних процесів. Визначено, що існуючі рішення поєднання мереж Кохонена та згорткових мереж спрямовані на підвищення ефективності тільки карт самоорганізації, тому для підвищення точності класифікації згортковими мережами виникла необхідність розробки нової архітектури. В роботі наведено її опис. Оскільки вихідною інформацією мереж Кохонена є матриця вагових значень нейронів решітки, для обробки діагностичних знімків необхідно було пов'язати її з представленням зображень. В роботі запропоновано концепцію вбудованого блоку асоціативного масиву на основі мереж Кохонена. За запропонованим способом розроблено програмну реалізацію гібридної нейронної мережі. Наведено постановку та результати обчислювальних експериментів. Експериментально доведено ефективність запропонованого способу.

Ключові слова: згорткові нейронні мережі, CNN, карти самоорганізації Кохонена, SOM, класифікація сигналів, класифікація зображень.

КЛАССИФИКАЦИЯ ОБЪЕКТОВ НА СВЁРТОЧНЫХ СЕТЯХ С ИСПОЛЬЗОВАНИЕМ КАРТ САМООРГАНИЗАЦИИ

Шаповалова С.И., Москаленко Ю.В.

Классификация объектов и процессов является распространённой прикладной задачей. Решение данной задачи необходимо для систем автоматической диагностики, например, для определения состояния работы оборудования, для анализа диагностических сигналов, или для выявления патологий на основе медицинских снимков. Однако данный метод подходит не для всех задач диагностики, поскольку в данных сетях невозможно достичь требуемой точности. В первую очередь это касается своевременной диагностики начала переходных процессов. В то же время другая разновидность нейронной сети — карты самоорганизации Кохонена SOM (self-organizing map) имеют концептуальное свойство для обучения на неопределённом множестве классов, и таким образом дают возможность решить такую задачу. Поэтому, повышение точности классификации на основе встраивания сетей Кохонена в архитектуру свёрточных сетей является актуальной задачей и имеет практическое значение.

В статье проведён анализ способов повышения точности свёрточных нейронных сетей, а также решение возникающих при этом проблем. Соответственно, приведены способы повышения доли корректной кластеризации на сетях Кохонена за счёт «возрастания» её решётки в процессе определения новых классов во время обучения. Именно это свойство делает возможным распознавание переходных процессов. Определено, что существующие решения объединения сетей Кохонена и свёрточных сетей направлены на повышение эффективности только карт самоорганизации, ввиду этого для повышения точности классификации свёрточными нейронными сетями возникла необходимость разработать новую архитектуру. В работе приведено ее описание. Поскольку выходной информацией сетей Кохонена является матрица весовых значений нейронов решётки, для обработки диагностических изображений её необходимо связать с представлением изображений. В работе предложено концепцию встроенного блока памяти на основе сетей Кохонена.

Согласно предложенному способу разработано программную реализацию гибридной нейронной сети. Приведено постановку и результаты вычислительных экспериментов. Экспериментально доказано эффективность предложенного способа.

Ключевые слова: сверточные нейронные сети, CNN, карты самоорганизации Кохонена, SOM, классификация сигналов, классификация изображений.