The pattern recognition problem solution using a convolutional neural network

K E Tokarev¹ ², A F Rogachev¹ ², Yu M Tokareva³, A Yu Rudenko¹, D V Zelyakovskiy¹ and A N Chernyavsky¹

¹Volgograd State Agricultural University, 26, Universitetskiy Avenue, Volgograd, 400002, Russia
²Volgograd State Technical University, 28, Lenina Avenue, Volgograd, 400005, Russia
³Volgograd State Medical University, 1, Pavshikh Bortsov Square, Volgograd, 400131, Russia

E-mail: tke.vgsha@mail.ru

Abstract. The article deals with the issues of theoretical and applied modeling based on convolutional neural network in the field of agricultural production. In particular, the specificity of convolutional neural network is considered, its model, architecture and layers are presented. Possibilities of its use for the solution of problems of recognition and classification of images are revealed. In addition, it was found that the convolutional neural network differs from the usual perceptron in that each neural layer is not connected to all the neurons of the previous layer, but only to a part of it. In networks of this type, a small matrix is used for the convolution operation, which moves along the processed layer with a certain offset step. For the first layer, this matrix moves across the input image. After each shift, a signal is generated to activate the neuron of the next layer and from the corresponding position. As a result of studying the problems of image recognition using electronic computers, the possibilities of machine learning as a method of recognizing an object in an image that characterizes each object with a set of features are revealed. In addition, the author reveals the specifics of using the convolutional neural network mathematical apparatus as a model for image recognition, which consists in using an array of data describing the image and applying the method of error back propagation.

1. Introduction

During the period of active design and implementation of irrigation equipment in the Russian Federation, no special emphasis was placed on such factors as agro-climatic conditions of application of sprinkler machines, the variety of agrotechnical methods of cultivation of agricultural crops, various economic opportunities of farms and, especially, the use of artificial intelligence systems for irrigation management. In this regard, the development of a set of programs based on fuzzy neural network models for optimization of resource-saving parameters of irrigation sprinkler machines is an urgent problem for effective management of water resources and obtaining stable high-yielding crops in the regions of the country.

The cultivation of crops based on the optimization of the management of their irrigation with the use of information technology in terms of differentiated accommodation will provide consistently
high-yielding crops, maintain soil fertility, prevent dryness and increase the level of resursosberegenie in terms of variable climate regions.

Currently, the problems of management of cultivation of agricultural crops in irrigation conditions using convolutional neural networks, mathematical modeling of watering processes and the use of digital technologies remain insufficiently investigated. The complex internal structure and pronounced nonlinearity of organizational and production processes of agriculture, as well as the need for the development of reclamation technologies, require the development of new digital information technologies, modernization of methods of mathematical modeling and structural and parametric optimization of production processes.

2. Materials and methods

One of the main directions of development of digital agriculture is the research, development and implementation of methods for solving problems of pattern recognition. This task is an important part of automated systems, information management and processing systems, as well as decision-making systems. Image processing and classification of visual images are used in security and video surveillance systems, in access control and control systems, in virtual reality and image retrieval systems. The task of image classification is to obtain an image at the input of the classifier, and at the output to obtain a class corresponding to the object in the image, or a probability vector that best classifies the image [1, 2].

This paper presents a study of the problem of solving the problem of recognition and classification of images of aerial photographs based on the mathematical apparatus of convolutional neural network. In machine learning, one of the most common ways to classify an object in an image is a feature description of an object that characterizes each object with a set of features. But it is worth considering that not all data open signs do not give good classification accuracy. Such signs can be a digital sound signal or the color of the dots. It is more difficult for a machine to classify images of cats or machines than for a human, as these data have hidden features that are difficult for a machine to understand, unlike a human.

Deep learning is a set of machine learning algorithms. This set of methods attempts to model high-level abstractions in the data, i.e. tries to isolate hidden features from the data [3].

Autoencoder is one of the algorithms of machine learning without a teacher, in which the size of the output vector is equal to the input. A neural network of direct propagation without feedback is one of the most common autoencoder architectures, and consists of an input layer and an output layer, equal in number of neurons and hidden layers. The data obtained on the input layer is compressed with the help of a hidden layer, after which there is a recovery on the output layer, which highlights the hidden features.

But autoencoder has a significant drawback, namely it should be used on data of small size, as when using big data, the learning process will grow significantly. As an example we can consider the classification of the image size \( m \times n = 8 \times 8 \). We want to allocate 50 signs. Then the size of the input layer will be \( 8 \times 8 = 36 \) neurons, respectively, the size of the output layer will also be 36 neurons (from the autoencoder definition). If we want to isolate 50 features, then the size of the hidden layer will be 50 neurons. Accordingly, the number of weights is equal to \( 2 \times m \times n \times k = 2 \times 8 \times 8 \times 50 = 6400 \). If the size of the images for classification is 100×100, the number of links will increase to 100,000 [3].

To reduce the dependence of the neural network size on the size of the input image, the convolutional neural network method is used. Convolutional neural networks were proposed by Jan Lekun based on the study of the visual cortex of cats. In this study, we discovered simple cells that responded to straight lines at different angles, and complex cells that responded only when several simple cells were combined. Hence came the idea of convolutional neural networks, which consists in the alternation of convolutional layers and subdirectory layers [4, 5].

Convolutional neural network differs from the usual perceptron in that each neuron layer is not associated with all the neurons of the previous layer, but only with a part of it. In this type of networks, a small matrix is used for the convolution operation, which moves through the processed
layer with some offset step. For the first layer, this matrix moves through the input image. After each shift, a signal is generated to activate the neuron of the next layer and from the appropriate position. This matrix is called the convolution kernel or filter, it produces coding of any feature. The next layer thanks to this matrix shows the presence of a feature in the previous processed layer and its coordinates, thereby forming a feature map. Convolutional neural networks use not one matrix of weights, but a set that characterizes a particular trait. Kernels are generated automatically during training, for example, by back-propagation errors. Each set of weights during the passage forms its own sign. Thus the network becomes multichannel [6].

To reduce the dimension of the card signs used operation subdisciplinary, and the network layer in this case is called a layer of subdisciplinary or pooling. For the image classification task, information about the presence of a trait is more important than its position, so the pooling layer of several neighboring neurons selects the maximum or finds their average weight. The resulting neuron is substituted in a new compaction feature map [7].

We consider the convolutional neural network architecture in more detail (figure 1). It consists of many layers. Initially, there is an input layer, which is the images to classify and is a matrix of numbers describing each pixel of the image. After the input layer, the signal passes alternately through the layers of convolution and subsampling, which alternate. The presence of convolution layers allows you to make new feature maps already available, and the peeling layer allows you to reduce the size of these layers.

The alternating layers enables recognition of complex hierarchical features. After passing through many layers of convolution and subsampling, usually a map of signs is transformed into a feature vector, but such vectors becomes hundreds.

Before the output layer in the convolutional neural network there is a conventional multilayer perceptron, which is fed to the input feature map. This perceptron solves the problem of feature map classification.

Convolutional neural network usually consists of the following layers [8]:
- Convolution layer (figure 2), which is used for feature mapping.

![Figure 1. Convolutional neural network model](image-url)
Figure 2. Layer convolution
- ReLU layer - block linear rectification, used as the activation function after the convolutional layer and is a function of the form \( f(x) = \max(0, x) \).
- Layer subsampling or pooling (figure 3). It serves to reduce the dimension of feature maps.

Figure 3. Layer subsampling
- Fully connected layer (figure 4) is a layer for the map features defines what classification type of the observed object and the output produces a vector of probabilities of belonging to a particular class:

Figure 4. Fully connected layer
- Dropout layer. It is used to combat retraining, when learning stochastic gradient descent. The work of this layer is to modernize the structure of the network, by ejecting a neuron with some probability.

Speaking about the advantages of convolutional neural network, we can highlight the following [9]:
- Reducing the number of parameters and increasing the learning speed of the network, in contrast to the multilayer perceptron.
- Resistance to noise, distortion, shifts and rotations of the input image.
- Ability to parallelize calculations and implement network learning algorithms on graphics processors (GPUs), which reduces the learning time of the neural network.
- At the moment there is one of the best and proven algorithms for classification of images.

The disadvantages of the algorithm include only a large number of configurable parameters.

When solving the problem of classification of symbols in the image, an array describing the image of the symbol is fed to the input of an ultra-precise neural network. After the signal passes through the network, the output vector is formed; if the recognition is carried out only in the English alphabet and numbers, the size of the output vector is 62. The output vector shows the probability of a symbol's relation to any class. As mentioned above, the method of back propagation of errors is usually used for training [3, 10].

3. Conclusion
As a result of the study of the problem of image recognition using computers, the possibilities of machine learning as a way of recognizing an object in the image, characterizing each object with a set of features, are revealed. In addition, the specificity of convolutional neural network as a model of image recognition is revealed, which consists in the use of an array describing the image and the application of the method of error back propagation.

4. Acknowledgments
The reported study was funded by RFBR and Volgograd region according to research project № 19-416-343006. The reported study was funded by RFBR and Volgograd region according to research project № 19-416-340014.

References
[1] Rogachev A F, Medvedeva L N, Timoshenko M A, Tokarev K E and Shiro M S 2019 Using computer modeling during evaluation of socio-labor potential of the region Advances in Intelligent Systems and Computing 726 989–1004
[2] Tokarev K E, Rogachev A F, T V Pleschenko, A Yu Rudenko and V A Kuzmin 2019 Economic and mathematical modelling of regional market of grain crops IOP Conf. Series: Earth and Environmental Science 341(2019) 012212
[3] Nikitin N A, Rozaliev V L, Orlova Y A and Alekseev A A 2018 Automated sound generation based on image colour spectrum with using the recurrent neural network CEUR Workshop Proceedings 2212 399–408
[4] Bannikova N V, Baydakov A N and Vaytskekhovskaya S S 2015 Identification of strategic alternatives in agribusiness Modern Applied Science 9(4) 344–353
[5] Gagarin A G and A F Rogachev 2012 Yield prediction based on cross-regional data analysis Proc. of Lower Volga Agrouniversity Complex: Science and Higher Vocational Education 4 339–346
[6] Ovchinnikov A S and Bocharnikov V S 2012 New technical solutions improve the efficiency of resource-saving irrigation methods Proc. of Lower Volga Agrouniversity Complex: Science and Higher Vocational Education 1(25) 119–124
[7] Skiter N N and Mazaeva T I 2015 Modeling ecological security of a state Mediterranean J. of Social Sci. 6(3) 192–195
[8] Shokhnek A V, Skiter N N, Rogachev A F and Pleschenko T V 2017 Features of optimal
modeling of tax mechanisms in the leveling system of environmental and food security taking into account inter-industry externalities *J. of Environmental Management and Tourism* **1**(17) 100–110

[9] Tokarev K E, Orlova Yu A, Rogachev A F, Kuzmin V A and Yu M Tokareva 2020 Crops reclamation management based on hybrid neuro-fuzzy systems *IOP Conf. Series: Earth and Environmental Science* 421**(2020)** 042015

[10] Tokarev K E 2015 Development of a system for decision support in the field of ecological-economic security *Mediterranean J. of Social Sci.* 6 415–420