

CONSTRUCTION OF AN ADVANCED METHOD FOR RECOGNIZING MONITORED OBJECTS BY A CONVOLUTIONAL NEURAL NETWORK USING A DISCRETE WAVELET TRANSFORM

The tasks that unmanned aircraft systems solve include the detection of objects and determining their state. This paper reports an analysis of image recognition methods in order to automate the specified process. Based on the analysis, an improved method for recognizing images of monitored objects by a convolutional neural network using a discrete wavelet transform has been devised. Underlying the method is the task of automating image processing in unmanned aircraft systems. The operability of the proposed method was tested using an example of processing an image (aircraft, tanks, helicopters) acquired by the optical system of an unmanned aerial vehicle. A discrete wavelet transform has been used to build a database of objects' wavelet images and train a convolutional neural network based on them. That has made it possible to improve the efficiency of recognition of monitored objects and automate a given process. The effectiveness of the improved method is achieved by preliminary decomposition and approximation of the digital image of the monitored object by a discrete wavelet transform. The stages of a given method include the construction of a database of the wavelet images of images and training a convolutional neural network. The effectiveness of recognizing the monitored objects' images by the improved method was tested on a convolutional neural network, which was trained with images of 300 monitored objects. In this case, the time to make a decision, based on the proposed method, decreased on average from 0.7 to 0.84 s compared with the artificial neural networks ResNet and ConvNets.

The method could be used in the information processing systems in unmanned aerial vehicles that monitor objects; in robotic complexes for various purposes; in the video surveillance systems of important objects

Keywords: neural network, discrete wavelet transform, monitored objects, unmanned aircraft system

Vadym Slyusar
Doctor of Technical Sciences, Professor
Research Institute Group*

Mykhailo Protsenko
PhD, Senior Researcher
Office of Special Forces*

Anton Chernukha
Corresponding author
PhD**

E-mail: an_cher@nuczu.edu.ua

Stella Gornostal
PhD, Associate Professor**

Sergey Rudakov
PhD, Associate Professor**

Serhii Shevchenko
PhD**

Oleksandr Chernikov
Doctor of Technical Sciences, Professor
Department of Engineering and Computer Graphics
Kharkiv National Automobile and Highway University
Yaroslava Mudroho str., 25, Kharkiv, Ukraine, 61002

Nadiia Kolpachenko
PhD, Associate Professor
Department of Technological Systems of Repair Production Named after O. Sidashenko
Kharkiv Petro Vasylenko National Technical University of Agriculture
Alchevskih str., 44, Kharkiv, Ukraine, 61002

Volodymyr Timofeyev
Doctor of Technical Sciences, Professor
Department of Automatics and Computer-Integrated Technologies
O.M. Beketov National University of Urban Economy in Kharkiv
Marshala Bazhanova str., 17, Kharkiv, Ukraine, 61002

Roman Artiukh
PhD, Director
State Enterprise "Southern State Design and Research Institute of Aviation Industry"
Sumska str., 130A, Kharkiv, Ukraine, 61023

*Central Scientific Research Institute of Armament and Military Equipment of the Armed Forces of Ukraine
Povitroflotsky ave., 28, Kyiv, Ukraine, 03049

**Department of Fire and Rescue Training
National University of Civil Defence of Ukraine
Chernyshevska str., 94, Kharkiv, Ukraine, 61023

Received date 14.06.2021

Accepted date 10.08.2021

Published date 31.08.2021

How to Cite: Slyusar, V., Protsenko, M., Chernukha, A., Gornostal, S., Rudakov, S., Shevchenko, S., Chernikov, O., Kolpachenko, N., Timofeyev, V., Artiukh, R. (2021). Construction of an advanced method for recognizing monitored objects by a convolutional neural network using a discrete wavelet transform. *Eastern-European Journal of Enterprise Technologies*, 4 (9 (112)), 65–77. doi: <https://doi.org/10.15587/1729-4061.2021.238601>

1. Introduction

The problem of security in recent years is of key importance for the development of mankind. Resolving this issue

is associated with the active evolution of monitoring systems for critical infrastructure [1]. Such facilities include large industrial enterprises, energy plants [2], chemically hazardous industries [3], and other strategic objects [4], the disruption

tion of the normal functioning of which can threaten vital national interests. The main factors threatening the safety of a monitored object (MO) include fires (explosions) [5], emissions of hazardous substances [6], radiation [7], as well as unauthorized entry of persons into the territory of MO. The most needed are systems built on the basis of machine vision and artificial intelligence, including the use of robotic and unmanned aircraft systems (UMAS) [8]. Therefore, devising new methods for recognizing monitored objects by artificial intelligence systems is of particular relevance.

2. Literature review and problem statement

Paper [8] reports a method of high-precision geolocation of remote ground MO using the platform of an unmanned aerial vehicle (UAV) equipped with an electro-optical device and laser rangefinder. Using data on UAV position and employing the optical system, the MO coordinates are determined. The proposed method involves multiple angle measurements and range measurements to MO to reduce random measurement errors. The simulation result shows that the MO coordinates were determined with an accuracy of 10 meters for the case when the UAV is at a distance of 4,000 meters from it. The cited paper did not consider the issues related to automating the process of MO image recognition.

Work [9] shows that many applications related to images and videos require high quality. It is proposed to combine a discrete wavelet transform (DWT), a Haar transform, a Kekre transformation, and a cosine transform to compress digital images. A given combination provides for a better compression ratio at high compression ratios from 75 % to 95 %. The disadvantage of a given method is its high computational complexity, instability for different compression ratios.

Study [10] considers the technology of processing a large number of images for 3D reconstructions. However, the amount of these data is quite large while the time is limited. To store high-quality MO images, it is proposed to use a low-rank tensor algorithm based on data compression. The cited study did not consider the issues of automation of the process of MO image recognition.

Analysis of the use of DWT for image compression is carried out in [11]. It is shown that DWT could be used in processes designed to compress images or improve their characteristics. The disadvantage of that approach is the impossibility of selecting image compression parameters for their processing and recognition under an automated mode.

Paper [12] considers the construction of biorthogonal filters based on DWT. Methods using symmetric filters that minimize problems caused by line breaks during conversion are analyzed. The possibility of using biorthogonal filters for the compression of digital images is shown. The disadvantage of a given paper is the lack of practical application of biorthogonal filters for processing digital images in UMAS.

The analysis of medical images is reported in [13]. It is shown that the number of medical images is growing rapidly, so effective image compression algorithms are needed to store them. The cited work proposes a codec for lossless compression of medical images. The disadvantages of that approach are a relatively small compression ratio, the lack

of choice of the optimal basis, the inability to select image compression parameters for their processing and recognition in UMAS.

Study [14] shows that image compression plays a more important role in reducing the size of a graphic file without deteriorating the quality. DWT and hybrid wavelet transform (HWT) have been proven to provide a better quality of compressed images. Experiments were conducted using a bench with ten images. DWT gives better compression quality than orthogonal transformation. The disadvantages of that approach include the impossibility of selecting the parameters of the basic function for their processing and recognition in UMAS.

Paper [15] shows that the growth in software for digital images has increased the need for effective methods of image compression. The HWT performance for the compression of digital images was checked. The experiment involved a set of 20 digital images whose compression ratio was changed. The disadvantages of that approach include the narrow focus of HWT, the lack of choice of the optimal basis during computation, the inability to use a given method for MO recognition.

The personality recognition system by the iris of the eye is tested in [16]. Compression effects for DWT-based images were investigated. In the cited work, the Haar wavelet is used to compress and decompose the image. The results of studies on the indicators PSNR, MSE are reported. It is established that DWT in the Haar basis is effective in compressing the image of the iris. The disadvantage of a given method is the inability to use it for MO image recognition.

It has been shown in [17] that DWT is one of the best compression methods. It provides for a mathematical notation to encode information according to the required level of detail. The Haar wavelet functions are proposed as the DWT basis. Redundancy of DWT detail factors decreases due to threshold values. The quality of compressed images was assessed using compression ratios and PSNR. The experimental results show that the proposed procedure ensures a high enough compression ratio compared to other methods of threshold compression. The disadvantage of a given method is the inability to use it for MO image recognition.

Compression of multicomponent images without loss is considered in [18]. The cited work uses a convolutional neural network (CNN) to select parameters for the DWT wavelet functions. A multicomponent compression system is proposed, which leads to an improvement in the spatial and spectral decorrelation of DWT coefficients. The compression results show a 7.2 % and 23.8 % reduction in bit rates compared to JPEG2000 in the YUV and RGB color spaces, respectively. The disadvantage of a given method is the inability to use it for MO image recognition.

Our review has revealed the following disadvantages of known procedures (methods):

- high computational complexity and instability for different compression ratios of MO images;
- the lack of practical application of mathematical apparatus for processing digital images in UMAS;
- the absence of proven artificial neural networks (ANNs) that solve the task of recognizing MO by classes in UMAS.

Consequently, it is necessary to devise an improved method for recognizing the monitored objects by CNN using DWT.

3. The aim and objectives of the study

The purpose of this study is to improve the efficiency of MO recognition by a neural network by decomposing and approximating the digital image of a monitored object using a discrete wavelet transform.

To accomplish the aim, the following tasks have been set:

- to investigate the quality indicators of MO image recognition;
- to evaluate the effectiveness of the method of recognition of monitored objects by CNN using DWT.

4. The study materials and methods

4. 1. Mathematical statement of the problem on image recognition of monitored objects

The MO images were acquired by the UAV optical system and transmitted to the computer of the ground control center. There, they are stored digitally in the form of a matrix $P(x, y)$ whose dimensionality is $M \times N$; they take the following form:

$$P = \| \| r_{x,y} \| \|,$$

where

$$x=1, \dots, M; y=1, \dots, N.$$

In a general case, the problem of MO image processing can be represented in the following form:

$$S[W \| \| r_{x,y} \| \|] = Y, \tag{2}$$

where S is the operator that characterizes ANN performance; W is the DWT operator of the original MO image; Y is the output data matrix.

The task of processing is to choose the operator S , with the help of which the decision is made qualitatively and quickly to categorize an MO image into one of the classes: tank, plane, helicopter.

4. 2. Selecting the quality indicators for the image recognition of monitored objects

The basic quality indicators that characterize the recognition process include the time and probability of the correct recognition of an MO image.

The recognition time T_p is determined from the following ratio [18]:

$$T_p = t_p / N_{np}, \tag{3}$$

where t_p is the time during which all MOs recognition was carried out; N_{np} is the number of correctly recognized MO images.

The probability of correct recognition P_p is estimated by the frequency of correct recognition:

$$P_p = N_{np} / N_o, \tag{4}$$

where N_o is the total number of MO images that are submitted for recognition.

At the stage of approximation of processing quality indicators, the probability of the correct recognition of MO can be determined from the following expression [18]:

$$P_p = \exp \left[- \left(B \frac{R}{L} \right)^2 \right], \tag{5}$$

where B is the shape recognition coefficient; R is the spatial resolution ability (on the ground); L is the maximum geometric size of a simple object on the ground.

The arguments of formula (2) and their ratios make it possible to take into consideration the influence of the most important factors that determine the quality of recognition: the geometric and photometric parameters of objects, the quality of an aerial photograph, and the ability of a person to perceive the image of the object.

The time spent on processing an aerial photograph can be expressed using the following diagram (Fig. 1), where:

- t_0 is the time spent on clarifying the task, analyzing the conditions for acquiring an image, planning the decryption process;
- t_1 is the time spent on the search and identification of a complex object, zones of the location of objects;
- t_2 is the time spent on object recognition, assessment of the state of a complex object;
- t_3 is the time spent on the preparation of conclusions, registration of information and reporting documents.

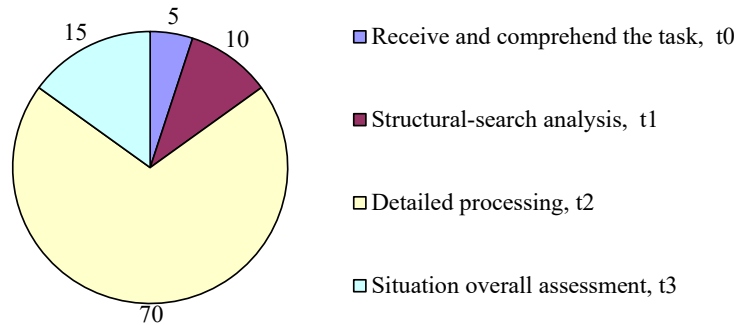


Fig. 1. Aerial image temporal processing diagram

The diagram demonstrates that a significant part of the time is spent on the stage of detailed processing whose main activity is the recognition of MO. Therefore, it is possible to improve the efficiency of the entire processing process by further automating the MO recognition process.

In this regard, it is a relevant issue to devise the basic procedures and algorithm for implementing an improved method of recognizing MO by CNN using DWT.

Our study was carried out under the following assumptions and limitations:

- UAV carries out panoramic aerial photography;
- the camera shoots in the view range, the characteristics of the camera do not change;
- MOs from the aerial photograph are recognized sequentially, one by one;
- information processing is carried out at the ground control point;

– shooting is carried out in the daytime; the season is summer.

4. 3. Devising the basic procedures and algorithm for implementing an improved method of recognizing monitored objects by a convolutional neural network using a discrete wavelet transform

The algorithm for implementing the proposed method of MO recognition is shown in Fig. 2.

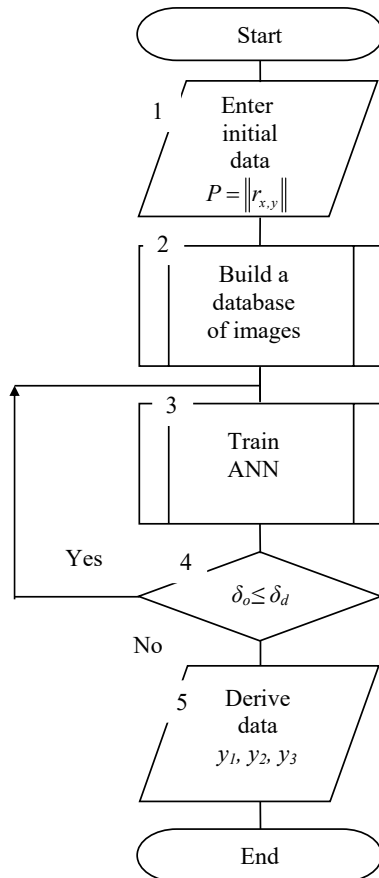


Fig. 2. Algorithm for implementing the proposed method

Step 1. Enter initial data (activity 1 in Fig. 2).

At this stage, the initial data are entered. Type: image; dimensionality: 768×768×3; RGB type; JPEG format.

Step 2. Build a database of MO images (activity 2 in Fig. 2).

Images whose dimensionality is 768×768×3 are saved in the following classes: 1 – tank; 2 – airplane; 3 – helicopter.

Step 3. Train ANN (activity 3 in Fig. 2).

Input data type: 768×768 images divided into classes: 1 – tank; 2 – airplane; 3 – helicopter.

In the proposed method, the LeNet-5 CNN architecture was taken as an ANN. A given CNN of the static architecture of direct propagation has shown high efficiency for image processing [18]. CNN training was conducted according to the error backpropagation algorithm (EBPA), which refers to the supervised methods of learning with a trainer [18]. EBPA is currently considered one of the most effective algorithms for training a CNN and determines the strategy for selecting the weights of a multilayer neural network using gradient optimization methods [18]. EBPA

advances the generalized delta rule and is a gradient descent algorithm that minimizes the total RMS error. In accordance with the delta rule, the adjustment of the weights at the current training step is carried out in the direction of the anti-gradient of the error function [19].

Fig. 3 shows a diagram of an artificial neuron.

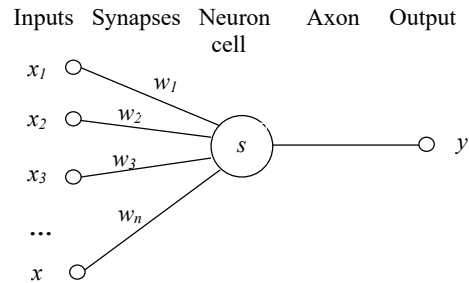


Fig. 3. Mathematical model of an artificial neuron

Fig. 3 shows that an artificial neuron consists of synapses that connect the neuron’s inputs to the nucleus. The nucleus of the neuron processes the input signals and the axon that connects the neuron to the neurons of the next layer. Each synapse has a weight that determines how much the corresponding input of a neuron affects its state. The state of the neuron is determined from the following formula [19]:

$$s = \sum_{i=1}^n x_i w_i, \tag{6}$$

where n is the number of neuron inputs;

x_i is the value of the i -th input of the neuron;

w_i is the weight of the i -th synapse.

Then the value of the axon of the neuron is determined from the following formula:

$$y = f(s), \tag{7}$$

where $f(s)$ is the activation function.

By substituting the value of s in formula (7), the following expression is obtained:

$$y = f\left(\sum_{i=1}^n x_i w_i\right). \tag{8}$$

Expression (8) is a mathematical model of the formal McCulloch-Pitts neuron. For this model, threshold activation functions were used.

For the next neuron, the output signal can be written as:

$$y_k = f\left(\sum_{j=1}^k w_j \cdot f\left(\sum_{i=1}^n x_i w_i\right)\right), \tag{9}$$

where k is the number of neuron inputs; w_j is the weight of the j -th synapse.

The development of the McCulloch-Pitts formal neuron model led to the emergence of new (more efficient) activation functions. As an activation function for the convolutional layer, a positively linear one is chosen – ReLu (10). The activation function ReLu returns 0 if it accepts a negative argument, and in the case of a positive argument, the function returns the number itself.

$$f(s) = \max(0, s), \quad (10)$$

where $f(s)$ is the activation function; s is the value of the argument.

The SoftMax function is the logistic function for a multi-dimensional case and is used in the CNN last layer. The function converts a vector s of dimensionality K to a vector f of the same dimensionality, where each coordinate of the resulting vector is represented by a real number in the interval $[0, 1]$. Coordinate values are calculated from the following formula:

$$f(s) = \frac{e^s}{\sum_{k=1}^K e^{s_k}}, \quad (11)$$

where $f(s)$ is the activation function; $k=1, \dots, K$ is the number of classes.

The SoftMax function is applied not to a single value but to a vector. It is used in the case of the multiclass classification problem. The network is built in such a way that on the last layer the number of neurons is equal to the number of classes sought. In this case, each neuron must give the value of the probability of belonging to the class of the object, that is, the value between zero and unity, and all neurons in the sum must give unity.

The scheme of the proposed CNN is shown in Fig. 4. The problem for the CNN to solve is to categorize images by classes: 1 – tank; 2 – airplane; 3 – helicopter.

Description of the architecture of the implemented neural network.

Input layer. The CNN input layer is fed an image of MO. Type: image; dimensionality, $768 \times 768 \times 3$; RGB format; JPEG type. Each image is divided into 3 channels: red, green, blue. Thus, three feature maps with a dimensionality of 768×768 pixels are obtained, which are fed to the wavelet layer.

The topology of connections between neurons of the network on the example of the first channel (red line in Fig. 4) is shown in Fig. 5.

Wavelet layer. The input of the layer is fed three feature maps with a dimensionality of 768×768 pixels, which are decomposed according to the formula of fast discrete wavelet transform (DWT) [11]:

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \cdot \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} P(x, y) \cdot \varphi_{j_0, m, n}(x, y), \quad (12)$$

where x and y are the values of the pixels of the image;

$\varphi_{j_0, m, n}$ are the values of the coefficients of a wavelet function.

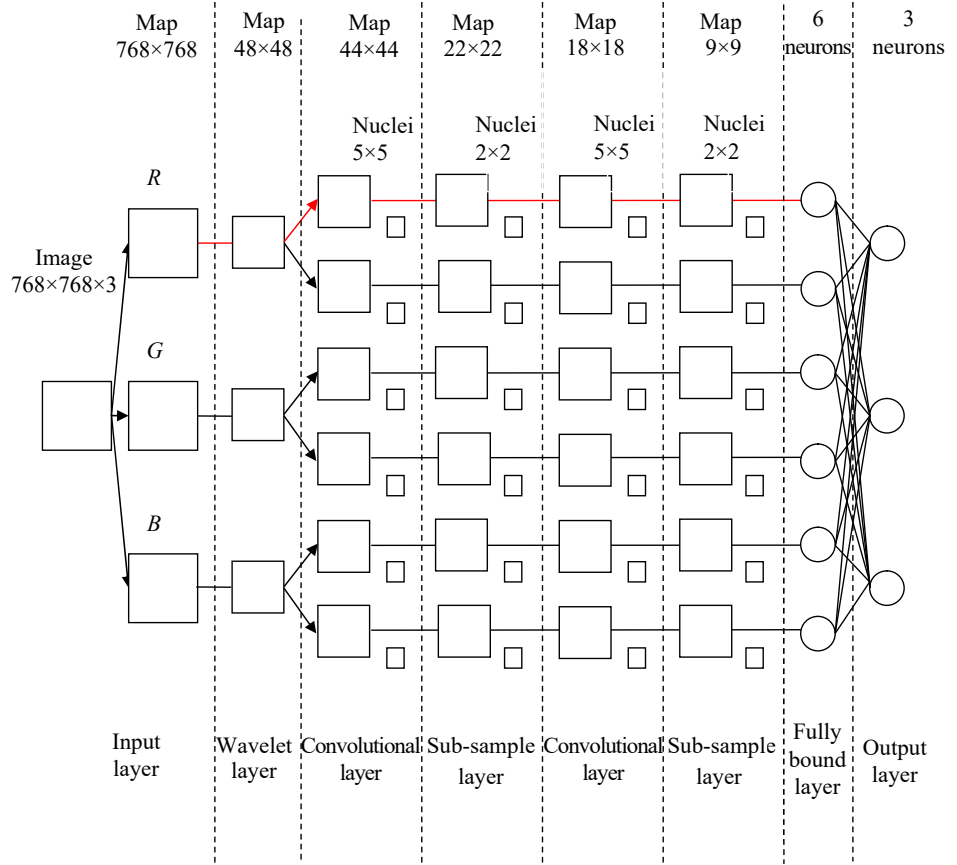


Fig. 4. Convolutional neural network diagram

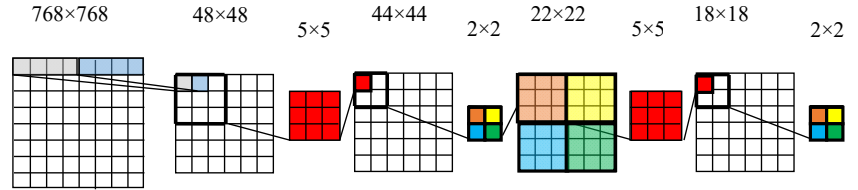


Fig. 5. Elements of a convolutional neural network

As a result, three feature maps with a dimensionality of 48×48 pixels were obtained, which are normalized to each pixel value in the range from 0 to 1, according to the following formula [18]:

$$S(p, \min, \max) = \frac{p - \min}{\max}, \quad (13)$$

where S is the normalization function;

p is the value of a specific pixel color from 0 to 255;

\min – the minimum pixel value – 0;

\max – the maximum pixel value – 255.

The wavelet layer is used to reduce the dimensionality and approximation of the original image. The original image after a fast discrete wavelet transform is reduced by 16 times. This, in turn, would affect the CNN performance. As a basic function, Haar wavelets are used, which have proven effective in practical tasks of digital image processing [9, 16, 20–22].

The size of the three output feature maps of a given layer is 48×48 pixels.

The first convolutional layer. Three 48×48-pixel feature maps are fed to the first convolutional layer, to which a 5×5 convolution is applied (core size, 5×5). A given layer is a map of features (matrix), the number of maps is 6. Each map has a synoptic kernel. The size of the kernel is selected in the range from 3×3 to 7×7. The small size of the kernel does not make it possible to distinguish attributes, and a large one increases the number of connections between neurons. The size of the kernel is selected so that the size of the maps in the convolutional layer is even, which makes it possible not to lose information when reducing the dimensionality in the sub-sample layer. For the proposed CNN, the chosen kernel size is 5×5. The sizes of all maps of the first convolutional layer are the same and are calculated from the following formula [18]:

$$(w, h) = (mW - kW + 1, mH - rY + 1), \tag{14}$$

where (w, h) is the calculated size of the convolutional map;
 mW is the width of the preceding map;
 mH is the height of the preceding map;
 kW is the kernel width;
 kH is the kernel height.

By substituting the values $mW=mH=48, kW=kH=5$ into formula (14), the sizes of the maps of the first convolutional layer are obtained:

$$(w, h) = (48 - 5 + 1, 48 - 5 + 1) = (44, 44).$$

In the initial state, the values of each convolutional layer feature map are 0. The values of the kernel weights are set randomly in the region from -0.5 to 0.5. The kernel slides (Fig. 5) over the preceding map and performs a convolution operation according to the following formula [18]:

$$(f \cdot q)[m, n] = \sum_{k, l} p[m - k, n - l] \cdot q[k, l], \tag{15}$$

where p is the original image matrix; q is the convolution kernel.

The first convolutional layer can be described by the following formula [19]:

$$x^l = f(x^{l-1} \cdot k^l + b^l), \tag{16}$$

where x^l is the output of the layer l ; f is the activation function; b^l is the l layer shift coefficient; k^l is the convolution kernel of the layer l .

Due to the boundary effects, the size of the initial matrices decreases according to the following formula [19]:

$$x_j^l = f\left(\sum_{i=1}^N x_i^{l-1} \cdot k_j^l + b_j^l\right), \tag{17}$$

where x_j^l is the map of features j (the output of layer l); b_j^l is the l layer shift coefficient for the feature map j ; k_j^l is the kernel of the convolution j of the map of layer l .

At the output from the layer, we have six output feature maps with a size of 44×44 pixels.

The first sub-sample layer. The first sub-sample layer receives six feature maps of 44×44 pixels. A given layer reduces the dimensionalities of the maps of the first convolutional layer from 44×44 to 22×22 pixels. Each kernel of the sub-sample layer is 2×2 in size, which makes it possible to reduce the preceding maps of the convolutional layer by 2 times, from 44×44 to 22×22. The entire map of features is

divided into cells per 2×2 element, from which the maximum value is selected (Fig. 5). Mathematically, the sub-sample layer can be described by the following formula [19]:

$$x^l = f(a^l \cdot \text{sub}(x^{l-1}) + b^l), \tag{18}$$

where x^l is the output of the layer l ;

a^l is the l layer shift coefficient;

sub is the operation of selecting local maximum values.

The size of the six output feature maps of the first sub-sample layer is 22×22 pixels.

The second convolutional layer. The second convolutional layer receives six feature maps of 22×22 pixels. The second convolutional layer is a map of features (matrix); the number of maps is 6. The kernel size is 5×5.

Substituting the values $mW=mH=22, kW=kH=5$ yields the size of maps in the second convolutional layer:

$$(w, h) = (22 - 5 + 1, 22 - 5 + 1) = (18, 18).$$

In the initial state, the values of each map of the second convolutional layer are 0. The values of the kernel weights are set randomly in the region from -0.5 to 0.5. The kernel slides over the preceding map and performs a convolution operation according to formula (15). The second convolutional layer can be described by formula (16). Due to the boundary effects, the size of the initial matrices decreases according to formula (17). The size of the six output feature maps of a given layer is 18×18 pixels.

The second sub-sample layer. The second sub-sample layer receives six feature maps of 18×18 pixels. The second sub-sample layer reduces the dimensionalities of the maps of the second convolutional layer. Each kernel of the second sub-sample layer is 2×2 in size, which makes it possible to reduce the preceding maps of the second convolutional layer by 2 times, from 18×18 to 9×9. The entire map of features is divided into cells per 2×2 element, from which the maximum value is selected. Mathematically, the second sub-sample layer can be described by formula (18). The size of the six output feature maps of the second sub-sample layer is 9×9 pixels.

Fully connected layer. The fully connected layer receives six 9×9-pixel feature maps. Feature maps are converted into 6 feature vectors (81 pixels each). Each vector is fed to its neuron of a fully connected layer (Fig. 4). The fully connected layer optimizes the nonlinear function, improves the quality of MO recognition, and can be described by the following formula [19]:

$$x_j^l = f\left(\sum_{i=1}^N x_i^{l-1} \cdot w_{i,j}^{l-1} + b_j^{l-1}\right), \tag{19}$$

where x_j^l is the map of features j (the output of layer l);

b^l is the l layer shift coefficient for the feature map j ;

$w_{i,j}^l$ is the matrix of the weight coefficients of layer l .

Output layer. The output layer is associated with all neurons of the fully connected layer (Fig. 4). Responds to CNN. The number of neurons corresponds to the number of recognized classes. The output of the first neuron with a value close to 1 means belonging to class 1. The output of the second neuron with a value close to 1 means belonging to class 2. The output of the third neuron with a value close to 1 means belonging to class 3.

CNN training using EBPA is carried out in several stages:

Step 3. 1. Initial initialization of CNN weights w_{ij} with small random values in the region from -0.5 to 0.5 .

Direct walk along the CNN:

Step 3. 2. The input layer of the CNN is fed a training image of MO whose dimensionality is $768 \times 768 \times 3$. Each image is divided into 3 channels with a dimensionality of 768×768 pixels: red, green, blue.

Step 3. 3. Three images with a dimensionality of 768×768 pixels are decomposed according to formula (12) of the discrete wavelet transform and are normalized according to formula (11).

Step 3. 4. Next, in the following layers, the weighted summation is carried out and, also, a nonlinear transformation is performed using the following activation function [19]:

$$s_{pj}^l = \sum_{i=1}^N w_{ij}^l \cdot x_{pi}^l = \sum_{i=1}^N w_{ij}^l \cdot y_{pi}^{l-1}, \quad (20)$$

where s_{pj} is the weighted sum of the input signals of the j -th neuron of layer l ;

w_{ij}^l is the weight of the connection between the i and j neurons;

y_{pi}^{l-1} is the output of the i -th neuron in the preceding layer ($l-1$) of image p .

The output signal of the j -th neuron of the l -th layer:

$$y_{pj}^l = f_j(s_{pj}^l). \quad (21)$$

In a general case, the recurrent relationship that determines the output of a neuron in an arbitrary layer is written as follows [19]:

$$y_{pj}^l = f_j\left(\sum_{i=1}^N x_i^l \cdot w_{ij}^l + b_j^l\right). \quad (22)$$

Thus, a sequential direct propagation of the input training image P along the neural network is carried out.

CNN return walk.

Step 3. 5. The total RMS error E for all neurons of the output layer of the neural network is determined, which is calculated as the difference between the required (reference) output d and the real (actual) output y of the last l -th layer of the neural network [19]:

$$E = \frac{1}{2} \cdot \sum_{i=1}^N (y_i^l - d_i^l)^2. \quad (23)$$

Step 3. 6. CNN training is based on the adaptive correction of weight coefficients w_{ij}^l in such a way as to minimize the value of the RMS error.

Minimization of the error function $E(w)$ is carried out on the basis of the delta rule, according to which the adjustment of weight coefficients is carried out in line with the following formula [19]:

$$w_{ij}^l(t+1) = w_{ij}^l(t) - \eta \cdot \frac{\partial E}{\partial w_{ij}^l}; \quad (24)$$

$$w_{ij}^l = -\eta \cdot \frac{\partial E}{\partial w_{ij}^l}, \quad (25)$$

where η is the learning rate coefficient that determines the value of the correction step, $0 < \eta < 1$; t is the number of the iteration of the training.

Since the total RMS error E depends on the CNN results, and the CNN output is formed using a nonlinear activation function (22) from the weighted sum of the input signals:

$$f(s_j^l) = f\left(\sum_{i=1}^N w_{ij}^l \cdot x_i^l\right), \quad (26)$$

then the partial derivative of the error function is written as a derivative of the function in the following form [19]:

$$\frac{\partial E}{\partial w_{ij}^l} = \frac{\partial E}{\partial y_j^l} \cdot \frac{\partial y_j^l}{\partial s_j^l} \cdot \frac{\partial s_j^l}{\partial w_{ij}^l}. \quad (27)$$

The weighted sum of the input signals s_j^l is a function of the synaptic weights; the expression of the partial derivative $\frac{\partial s_j^l}{\partial w_{ij}^l}$ is equal to the value of the input signal of the current layer of the multilayer neural network, which is simultaneously the output of the neuron of the preceding layer:

$$\frac{\partial s_j^l}{\partial w_{ij}^l} = x_i^l = y_i^{l-1}. \quad (28)$$

Hence, it follows

$$\frac{\partial E}{\partial w_{ij}^l} = \frac{\partial E}{\partial s_j^l} \cdot x_i^l. \quad (29)$$

Once the designation is introduced

$$\delta_j^l = \frac{\partial E}{\partial s_j^l},$$

the following equality is derived:

$$\frac{\partial E}{\partial w_{ij}^l} = \delta_j^l \cdot x_i^l. \quad (30)$$

$$\delta_j^l = \frac{\partial E}{\partial s_j^l} = -(y_j^l - d_j) \cdot y'(s) = -(y_j^l - d_j) \cdot \frac{\partial y_j^l}{\partial s_j^l}. \quad (31)$$

By substituting expression (29) in (23), and using equality (27), the following expression for the correction of weights is obtained:

$$\Delta w_{ij}^l = -\eta \cdot \delta_j^l \cdot y_i^{l-1}. \quad (32)$$

This formula is used to correct the weight coefficients starting from the output layer of the network and towards the input.

Step 3. 7. Checking the criterion of stopping the learning algorithm. If at least one of the following criteria is met, the training stops:

- the learning error has reached a predetermined value;
- the learning error does not decrease or decrease slightly;
- the generalization error begins to increase, indicating the onset of retraining.

If the criterion for stopping the learning algorithm is not met, then the transition to Step 3. 2. is carried out, and the next iteration of training is performed.

At the end of the EBPA algorithm, the neural network is considered trained and ready for use.

An example of using the EBPA algorithm to correct the coefficients of synaptic connections. A neural network with two inputs, two neurons of a fully connected layer. and two

output neurons (Fig. 6). It is required that the neural network should produce 0.01 and 0.99.

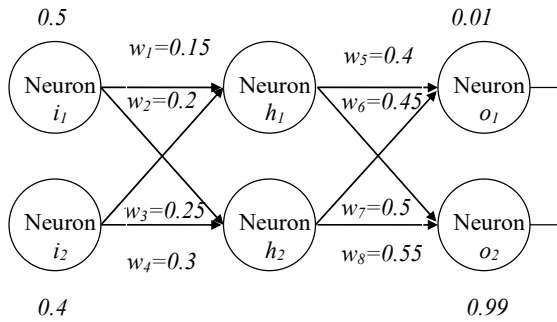


Fig. 6. Neurons in a convolutional neural network

Direct walk along the CNN.

Step 1. Direct distribution of the input example along the CNN is carried out.

Taking into consideration the input values, the weighted sum of the input signals of the 1st neuron is equal to:

$$s_{h1} = \sum_{i=1}^N x_i \cdot w_i = 0.15 \cdot 0.5 + 0.2 \cdot 0.4 = 0.155. \quad (33)$$

As a simplification of the example, the logistic activation function (a simplified version of SoftMax) is used. The output value of the 1st neuron:

$$y_{h1} = f(s) = \frac{1}{1 + e^{0.155}} = 0.4630. \quad (34)$$

For the 2nd neuron, $y_{h2} = 0.4386$ is obtained. Repeating this process for neurons of the output layer, using the output data from the neurons of the preceding layer as inputs, $y_{o1} = 0.4910$, $y_{o2} = 0.389596$ are obtained.

Return walk.

Step 2. Determining the ANN RMS error.

Taking into consideration the input values, the RMS error is:

$$E_{o1} = \frac{1}{2} \cdot (0.01 - 0.4910)^2 = 0.1157. \quad (35)$$

For $E_{o2} = 0.18033$. The general error of the neural network is composed of the following errors $E = E_{o1} + E_{o2} = 0.1157 + 0.1803 = 0.296$.

Since the desired and actual values of the ANN outputs do not coincide, correction (adaptation) of the weight coefficients of the synaptic connections of the neural network is necessary.

Step 3. Correction of coefficients of synaptic connections in the direction opposite to the direct propagation of input signals.

Taking into consideration the input values, the partial derivatives of the error function are equal to:

$$\frac{\partial E}{\partial w_5} = \frac{\partial E}{\partial y_{o1}} \cdot \frac{\partial y_{o1}}{\partial s_{o1}} \cdot \frac{\partial s_{o1}}{\partial w_5}. \quad (36)$$

$$\frac{\partial E}{\partial y_{o1}} = -(0.01 - 0.4910) = 0.4810. \quad (37)$$

$$\frac{\partial y_{o1}}{\partial s_{o1}} = 0.4910(1 - 0.4910) = 0.2499. \quad (38)$$

$$\frac{\partial s_{o1}}{\partial w_5} = 0.4386 \cdot 1 + 0 = 0.4386. \quad (39)$$

$$\frac{\partial E}{\partial w_5} = 0.4810 \cdot 0.2499 \cdot 0.4386 = 0.00527. \quad (40)$$

To reduce the error, this value, multiplied by the learning rate η (for the proposed method, the value 0.3 was experimentally selected), is subtracted from the current weight:

$$w_5^+ = w_5 - \eta \cdot \frac{\partial E}{\partial w_5} = 0.4 - 0.3 \cdot 0.00527 = 0.39842. \quad (41)$$

The remaining coefficients of synaptic connections are calculated in a similar way.

5. The results of studying the effectiveness of recognition of monitored objects by a convolutional neural network using a discrete wavelet transform

5.1. Investigating the indicators of quality of image recognition of monitored objects

The coefficient of recognition of the shape of images acquired by a UAV optical-electronic means depends on a group of factors. The main ones are:

- the scale of the image;
- brightness coefficient;
- lighting conditions;
- mutual position of the UAV and the monitored object;
- the mode of processing of shooting materials.

All these factors are very variable, so the coefficient of recognition of the shape of even the same type of object could vary significantly. That causes individual objects to stand out on the same aerial image with different probabilities due to changes in the value of the shape recognition coefficient. Whereas the ratio of spatial resolution to maximum geometric size for these objects can be unchanged. Fig. 7 shows plots that characterize the dependence of the probability of correct recognition P_p on the value of the shape recognition coefficient B at different values of spatial resolution R for the same object the size of L , at $R_3 > R_2 > R_1$. The dependence plots were built in the Microsoft Excel 2016 programming environment.

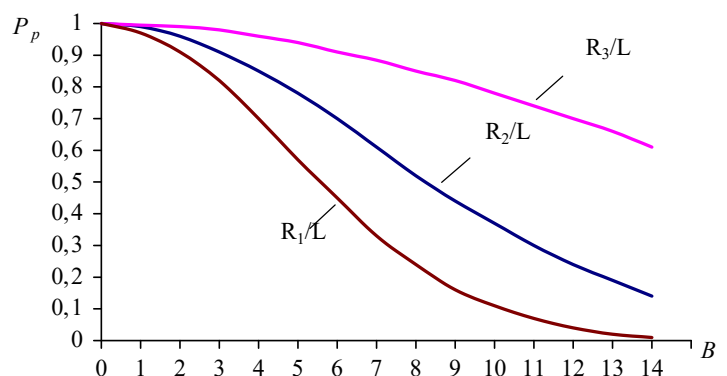


Fig. 7. Plots of the dependence of the recognition probability P_p on the value of the shape recognition coefficient B at different values of spatial resolution R

The plots in Fig. 7 demonstrate that to increase the probability of correct recognition, it is necessary to reduce the value of the spatial resolution R or the value of the shape recognition coefficient B . A value of the spatial resolution is a characteristic of the optical-electronic observation system. It determines the potential probability of correct recognition of objects in aerial photographs. Thus, in practice, the value of the resolution that is necessary for the recognition of objects with a probability of $P_p \geq 0.8$ is determined based on the Johnson criterion (Table 1).

Table 1
Value of spatial resolution for image recognition according to Johnson's criterion [18]

Recognition level	Task	Number of spatial resolution R values per minimum L_{\min} size of the object
Class recognition	The operator recognizes the image to the class of the object (for example, an airplane, a helicopter, a tank)	7.6...9.6
Type recognition	The operator recognizes the image to the type of object (for example, the type of tank)	10...16

L_{\min} is understood as the minimum size of the projection of the object onto the plane (the projection of the object is perpendicular to the line of vision of the optical device of a UAV). In accordance with Johnson's criterion, the condition under which the recognition of the MO image on an aerial photograph can be carried out with a probability of not less than 0.8 can be represented as follows:

$$L_{\min} / R \geq 10 \dots 16. \tag{42}$$

Meeting condition (42) is necessary but not sufficient to conduct recognition of the image of MO with the predefined probability. Therefore, in cases where, when condition (42) is met, the value of the probability of recognition does not exceed the specified level, it is necessary to reduce the value of the shape recognition coefficient. Such an event occurs when the shape of the image of the object is distorted (overlapped) by its own or falling shadow.

5. 2. Assessing the effectiveness of the method for recognizing monitored objects by a convolutional neural net with the use of discrete wavelet transform

Information processing was carried out at the ground control point. For aerial photography, a UAV was used, which is equipped with a Sony ILCE-7R camera. A given camera has the following characteristics and functions:

- matrix type - 35-mm full-frame CMOS-matrix Exmor™ (35.9x24 mm);
- recording format (photo) - RAW (sony arw 2.3 format), JPEG support (DCF ver. 2.0, Exif Ver.2.3, compatible with MPF Baseline);
- image size (pixels): 4,800x3,200 (15M); 3,680x2,456 (9.0M); 7,360x4,912 (36M).

An example of an aerial photograph taken by the digital camera Sony ILCE-7R from a height of 1,400 meters from a UAV is shown in Fig. 8.

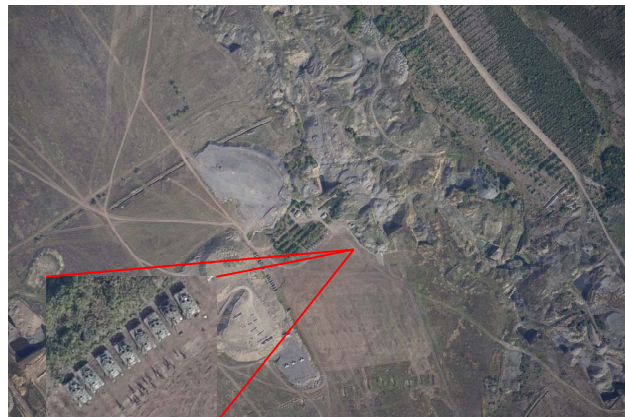


Fig. 8. Aerial photograph acquired from Sony ILCE-7R camera

Characteristics of the digital aerial photograph: resolution, 7,360x4,912; color depth, 24 bit/pixel; file size, 8,792,288 bytes; focal length, 55 mm. The number of images that were used to prepare images of MO for the training and test samples is 100. One aerial photograph can contain several tens of MOs (Fig. 8). Images for the training and test samples were prepared using the ABBYY Screenshot Reader software.

As a training sample, 100 images were prepared for each class, a total of 300 images of MOs were used for training. Sample type: image; dimensionality, 768x768; JPEG format.

As a test sample, 50 images were prepared for each class, a total of 150 MO images were used for testing. Test sample type: image; dimensionality, 768x768; JPEG format.

The main parameter for assessing the effectiveness of the proposed method is the time of MO recognition. Based on this, we measured the time spent on this procedure. Testing was conducted on a Dell computer that is equipped with an Intel(R)Kernel(TM)2 Q9400 processor with a clock frequency of 2.67 GHz and 8 GB of RAM. The time over which 300 MOs were recognized by class using the proposed method was 0.42 s. For comparison, the following neural networks were taken: ConvNets, ResNet. We studied the effectiveness of ANN for the recognition of monitored objects in the computer environment of mathematical modeling MATLAB R2017a.

The results of comparing the MO recognition time for different ANNs are shown in Fig. 9. Dependence plots were built in the Microsoft Excel 2016 programming environment.

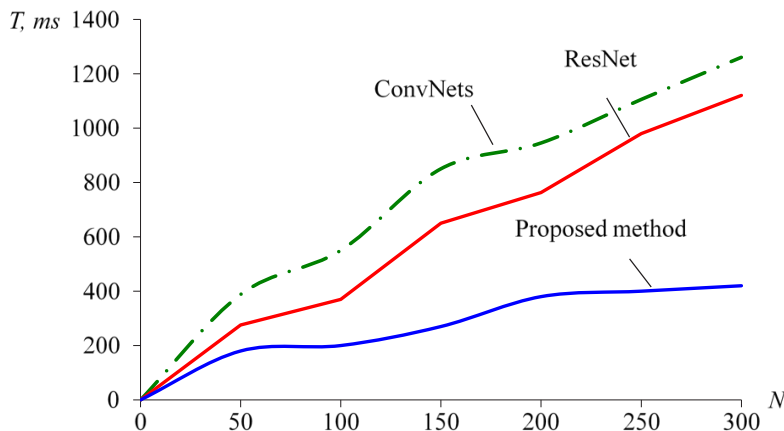


Fig. 9. Comparing the time for recognizing monitored objects - T , ms, depending on the number of images - N

The time over which 300 MOs were recognized by classes by ResNet and ConvNets ANNs was, respectively, 1.12 and 1.26 s. The time for making a decision on MO recognition based on the proposed method decreased from 0.7 to 0.84 s in comparison with the neural networks ResNet and ConvNets. The main gain in time is achieved due to the smaller size of the MO image after DWT.

Convergence. The convergence of ANN shows whether the ANN architecture and the learning algorithm (coefficients of synaptic connections, learning rate) are correctly chosen in accordance with the task set. If the error decreases with each epoch of learning, then the ANN converges. If the error changes upwards or downwards (several times), in this case, the ANN does not converge. To ensure convergence, the learning algorithm (coefficients of synaptic connections, learning speed) is replaced. If convergence is not achieved even in this case, it is necessary to change the architecture of the neural network. Fig. 10 illustrates the assessed convergence of ANN during training.

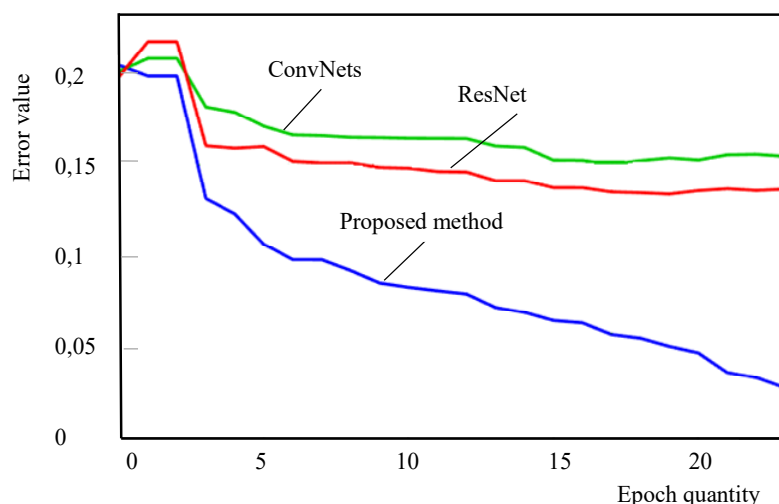


Fig. 10. Estimation of convergence of artificial neural networks

Fig. 10 shows that the proposed ANN has good convergence. After 3 epochs, the learning error decreases.

Adequacy. ANN is adequate if the learning outcomes converge to very close values (or one) – a necessary condition that there is some law (dependence) between the output and input data, which is implemented by the neural network.

The most effective way to check the ANN model for adequacy is to compare the results with a known solution to the problem (if such a solution is known). The results of MO recognition experiments on the test sample are given in Table 2.

Table 2

Results of recognizing the monitored objects by class

Neural net	Image quantity, pc.	Time for class recognition, s	Recognition percentage, %
ConvNets	150	0.63	83
ResNet	150	0.56	88
Proposed	150	0.21	94

Table 2 shows that the percentage of recognition of the monitored objects by class for the test sample by the proposed CNN is 94 % (ConvNets – 83 %, ResNet – 88 %).

A sufficient condition for the adequacy of CNN is the predictability of results in the range of the entire set of data, including those that did not participate in the training.

We assessed the adequacy of the proposed method for different MO orientations in an image (Fig. 11). As a test sample, 20 images were prepared for each group of orientations, for each class, a total of 240 MO images were used for testing (these images were not used for training). Test sample type: 768×768 image; JPEG format.

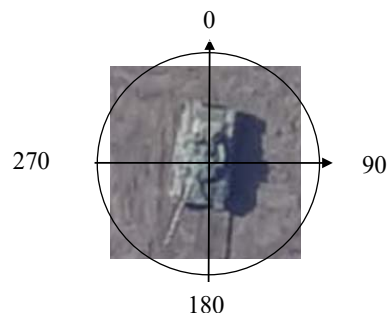


Fig. 11. Orientation of the monitored object, the class of tank (180°, 270°)

Our results (Table 3) show that for different orientations of MOs in an image, the recognition accuracy indicators change insignificantly (they are predictable). The proposed method has demonstrated a gain in the accuracy of MO recognition, compared to ConvNets ANN, from 2 to 7 %, compared with ResNet ANN, from 8 to 9 %.

The convergence of the test results (Tables 2, 3), as well as the comparison of the results with known ANNs (ResNet, ConvNets), shows the adequacy of the proposed method.

Sensitivity. To date, the established approach to assessing the sensitivity of ANN has not been adopted. Some authors [18, 19], when assessing ANNs, use approaches that are accepted in medicine (*a first approach*), others assess sensitivity to noise (data distortion) (*a second approach*).

Table 3

Indicators of accuracy in recognizing monitored objects for groups of different orientations, %

Orientation	Accuracy of monitored object recognition, %		
	ResNet	ConvNets	Proposed method
(0°, 90°)	85±2	86±4	93±2
(90°, 180°)	82±3	89±3	91±3
(180°, 270°)	85±5	88±5	94±2
(270°, 360°)	83±4	87±3	91±3

For *a first approach*, sensitivity is defined as:

$$P = \frac{a}{a+d} \cdot 100\%, \tag{43}$$

where *a* is the number of correctly made decisions on the classification of MO; *d* is the number of wrong decisions made.

The results of experiments to assess sensitivity on test samples (a first approach) are given in Table 4.

Table 4

Results of sensitivity assessment on the test sample

Neural net	Image quantity, pc.	Sensitivity, %
ConvNets	150	85
ResNet	150	89
Proposed	150	96

Table 4 shows that for the test sample of the proposed CNN, the sensitivity is 96 % (ConvNets – 85 %, ResNet – 89 %).

For a second approach, sensitivity is defined as the accuracy of MO recognition by class depending on the noise level. For the experiment, an additional test sample was prepared by adding pulsed noise (imnoise function) to the original images in the mathematical modeling computer environment MATLAB R2017a.

An example of a noisy image is shown in Fig. 12 (from left to right: class image – tank with the noise of 0, 10, 20, and 30 %, respectively).

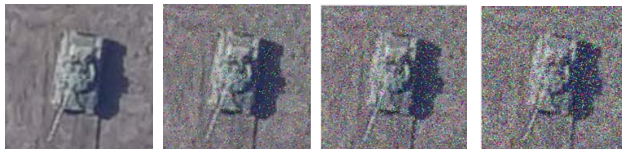


Fig. 12. Examples of test images of the class of tank with added noise

During the testing, a sample of 80 images of each class was used, a total of 240 MO images were used for testing. Test sample type: 768×768 image; JPEG format.

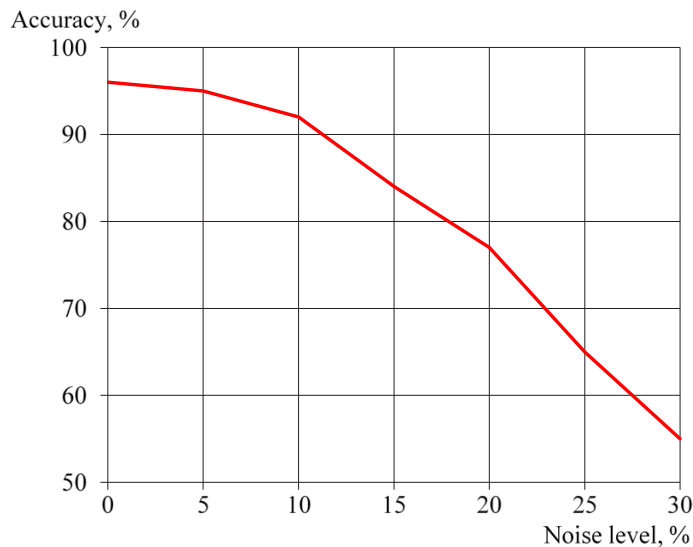


Fig. 13. Image recognition results based on the noise level

Our experiment shows (Fig. 13) that the proposed method makes it possible to recognize MO objects by classes in the presence of noise in the image. At the same time, the accuracy of MO recognition is significantly reduced at a noise level exceeding 20 %.

Reliability. The reliability of study results is the confirmation that the findings (patterns, recognition accuracy) are identical for a certain class of MO under the selected experimental conditions, limitations, and assumptions.

The main way to confirm the study’s reliability is its verification. For verification, CNN is tested on various test samples. The results are checked with each other and, in the case of their repetition (coincidence, proximity), a decision is made on the reliability of the experiments.

- Thus, the reliability of the proposed method is confirmed:
- by the validity of the choice of initial data, basic assumptions, and limitations;
 - by its verification on various training and test samples, with different orientations of MOs.

6. Discussion of results of studying the method of recognition of monitored objects by an artificial neural network

The operation of any image recognition system consists of several stages. First, the picture highlights the signs of recognition, which are later compared with pre-known sets of features of reference images of classes (types) of objects. Based on the established criterion, the degree of proximity between the features of real and reference images is evaluated. The final stage is the decision to assign the real images to one of the established classes.

A priori information is a component of reference one and should include direct, indirect, and integrated signs of MOs. A priori information is represented in the form of primitives: reference images, characteristic details of MO of a certain type.

In the proposed method, CNN was chosen as the ANN. We trained a given network according to the error backpropagation algorithm. The peculiarity of a given network is that its training is reduced to minimizing the error function, by adjusting the weight coefficients of synaptic connections between neurons.

- CNN has the following advantages [18, 19]:
- one of the best ANNs for image recognition and classification;
 - compared to a fully connected neural network, a much smaller number of adjustable weights;
 - convenient parallelization of calculations, and, consequently, the possibility of implementing algorithms for working and training the network on graphics processors;
 - relative resistance to rotation and shift of the recognized image;
 - training involving the classical error backpropagation algorithm.

In the improved method, approximated wavelet coefficients of the MO images are used as standards. Unlike known methods, a given method takes into consideration:

- the orientation of MO in the image;
- the shape of the image of the object;
- the tone of the MO image;
- the size of the image of MO.

Our study of the proposed method showed that a given method provides higher efficiency of MO recognition. In this case, the time to make a decision by the proposed method decreased on average from 0.7 to 0.84 s compared with the neural networks ResNet and ConvNets for the same training

and test samples. The gain is achieved by using DWT to reduce the size of the MO image.

Our results (Table 3) showed that for different orientations of MO in the image, the recognition accuracy indicators change insignificantly. The proposed method has demonstrated a gain in recognition accuracy, compared to ConvNets ANN, from 3 to 7 %, compared with ResNet ANN, from 8 to 9 %.

The limitation of the proposed method is that it is adapted to recognize MOs in three classes. The CNN was trained on high-contrast digital MO images acquired from aerial photographs by UAV. The shooting was carried out in the daytime, the time of year – summer. Thus, high values of MO recognition accuracy were obtained. For other types of images, the accuracy of recognition by class may vary, which requires additional research.

In the future, work will continue filling the base of reference MOs. It is also planned:

- to assess the sensitivity of the improved method of MO recognition under various conditions of aerial photography;
- to devise a method to search for MO images in aerial photographs;
- to clarify the options for applying the proposed CNN to automate the process of information processing;
- to train the developed CNN for other conditions for acquiring digital images of MOs;
- to improve the structure of CNN and methods of its training;
- to apply CNN to recognize different types of MOs.

7. Conclusions

1. Our study of the quality indicators of MO image recognition, acquired from the optical system of UAV has shown:

- the operator spends 70 percent of the time on detailed processing of an aerial photograph;
- the operator assigns the image to the class of the object (for example, an airplane, helicopter, tank) in the range

of values from 7.6 to 9.6 for the ratio of resolution to the minimum size of the object;

- the operator assigns the image to the type of object (for example, the type of tank) in the range of values from 10 to 16 for the ratio of resolution to the minimum size of the object;
- to improve the efficiency of the entire processing process, it is important to reduce the time for recognition of MO images;
- to increase the probability of correct recognition, it is necessary to reduce the value of the spatial resolution or the value of the shape recognition coefficient;
- even for the same MOs, the values of the recognition features could change, which increases the time of recognition of MO images;
- the reference information that is used during processing does not always make it possible to determine the values of the signs of recognition of a particular MO image;
- there is an insufficient level of automation of the process of recognition of MOs on aerial photographs;
- the low efficiency of recognition of MO images and the entire processing process in general.

2. Evaluation of the effectiveness of the method for recognizing objects of monitoring by a convolutional neural network using DWT showed that the proposed method makes it possible:

- to reduce the size of the image of MO;
- to take into consideration the values of recognition features for each specific MO;
- to use reference images for training CNN;
- to recognize monitored objects by classes: tank, plane, helicopter;
- to reduce the time for MO recognition on average from 0.7 to 0.84 s compared to the ANNs ConvNets, ResNet;
- to improve the accuracy of MO recognition, in comparison with the ConvNets ANN, from 2 to 7 %, compared with the ResNet ANN, from 8 to 9 %.

Improving the performance of an artificial neural network has been obtained by decomposing and approximating the digital image of MO by discrete wavelet transform.

References

1. Pospelov, B., Andronov, V., Rybka, E., Krainiukov, O., Maksymenko, N., Meleshchenko, R. et. al. (2020). Mathematical model of determining a risk to the human health along with the detection of hazardous states of urban atmosphere pollution based on measuring the current concentrations of pollutants. *Eastern-European Journal of Enterprise Technologies*, 4 (10 (106)), 37–44. doi: <https://doi.org/10.15587/1729-4061.2020.210059>
2. Vambol, S., Vambol, V., Kondratenko, O., Suchikova, Y., Hurenko, O. (2017). Assessment of improvement of ecological safety of power plants by arranging the system of pollutant neutralization. *Eastern-European Journal of Enterprise Technologies*, 3 (10 (87)), 63–73. doi: <https://doi.org/10.15587/1729-4061.2017.102314>
3. Vambol, S., Vambol, V., Kondratenko, O., Koloskov, V., Suchikova, Y. (2018). Substantiation of expedience of application of high-temperature utilization of used tires for liquefied methane production. *Journal of Achievements in Materials and Manufacturing Engineering*, 2 (87), 77–84. doi: <https://doi.org/10.5604/01.3001.0012.2830>
4. Semko, A., Rusanova, O., Kazak, O., Beskrovnaya, M., Vinogradov, S., Gricina, I. (2015). The use of pulsed high-speed liquid jet for putting out gas blow-out. *The International Journal of Multiphysics*, 9 (1), 9–20. doi: <https://doi.org/10.1260/1750-9548.9.1.9>
5. Pospelov, B., Andronov, V., Rybka, E., Skliarov, S. (2017). Design of fire detectors capable of self-adjusting by ignition. *Eastern-European Journal of Enterprise Technologies*, 4 (9 (88)), 53–59. doi: <https://doi.org/10.15587/1729-4061.2017.108448>
6. Kustov, M. V., Kalugin, V. D., Tutunik, V. V., Tarakhno, E. V. (2019). Physicochemical principles of the technology of modified pyrotechnic compositions to reduce the chemical pollution of the atmosphere. *Voprosy Khimii i Khimicheskoi Tekhnologii*, 1, 92–99. doi: <https://doi.org/10.32434/0321-4095-2019-122-1-92-99>

7. Popov, O., Iatsyshyn, A., Kovach, V., Artemchuk, V., Taraduda, D., Sobyna, V. et. al. (2019). Analysis of Possible Causes of NPP Emergencies to Minimize Risk of Their Occurrence. *Nuclear and Radiation Safety*, 1 (81), 75–80. doi: [https://doi.org/10.32918/nrs.2019.1\(81\).13](https://doi.org/10.32918/nrs.2019.1(81).13)
8. Yang, X., Lin, D., Zhang, F., Song, T., Jiang, T. (2019). High Accuracy Active Stand-off Target Geolocation Using UAV Platform. 2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP). doi: <https://doi.org/10.1109/icsidp47821.2019.9172919>
9. Thepade, S. D., Dewan, J. H., Erandole, S. S., Jadhav, S. R. (2015). Extended performance comparison of self mutated hybrid wavelet transforms in image compression with hybrid wavelet transforms & orthogonal transforms. 2015 Global Conference on Communication Technologies (GCCT). doi: <https://doi.org/10.1109/gcct.2015.7342675>
10. Zhu, J., Wang, J., Zhu, Q., Liu, P., Li, S. (2018). Reconstruction of Compressed Sensed Images with Multiple-Image Pattern Low-Rank Tensor. 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC). doi: <https://doi.org/10.1109/icivc.2018.8492809>
11. Van, F., Patrick, J. (2019). An introduction to digital images. *Discrete Wavelet Transformations: An Elementary Approach with Applications*. Wiley, 69–123. doi: <https://doi.org/10.1002/9781119555414.ch3>
12. Van, F., Patrick, J. (2019). Biorthogonal wavelet transformations. *Discrete Wavelet Transformations: An Elementary Approach with Applications*. Wiley, 261–320. doi: <https://doi.org/10.1002/9781119555414.ch7>
13. Krishnaswamy, R., NirmalaDevi, S. (2020). Efficient medical image compression based on integer wavelet transform. 2020 Sixth International Conference on Bio Signals, Images, and Instrumentation (ICBSII). doi: <https://doi.org/10.1109/icbsii49132.2020.9167597>
14. Thepade, S. D., Erandole, S. (2013). Effect of tiling in image compression using wavelet transform & hybrid wavelet transform for cosine & kekre transforms. 2013 IEEE International Conference ON Emerging Trends in Computing, Communication and Nanotechnology (ICECCN). doi: <https://doi.org/10.1109/ice-ccn.2013.6528604>
15. Thepade, S. D., Dewan, J. H., Lohar, A. T. (2013). Extended performance comparison of hybrid wavelet transform for image compression with varying proportions of constituent transforms. 2013 15th International Conference on Advanced Computing Technologies (ICACT). doi: <https://doi.org/10.1109/icact.2013.6710497>
16. Paul, A., Khan, T. Z., Podder, P., Ahmed, R., Rahman, M. M., Khan, M. H. (2015). Iris image compression using wavelets transform coding. 2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN). doi: <https://doi.org/10.1109/spin.2015.7095407>
17. Nashat, A. A., Hussain Hassan, N. M. (2016). Image compression based upon Wavelet Transform and a statistical threshold. 2016 International Conference on Optoelectronics and Image Processing (ICOIP). doi: <https://doi.org/10.1109/optip.2016.7528492>
18. Ahanonu, E., Marcellin, M., Bilgin, A. (2020). Lossless Multi-component Image Compression Based on Integer Wavelet Coefficient Prediction using Convolutional Neural Networks. 2020 Data Compression Conference (DCC). doi: <https://doi.org/10.1109/dcc47342.2020.00043>
19. Il'yasov, B. G., Makarova, E. A., Zakieva, E. Sh., Gabdullina, E. R. (2021). Metody iskusstvennogo intellekta v programmnyh prilozheniyah: laboratorniy praktikum po distsiplinam «Metody iskusstvennogo intellekta v upravlenii», «Intellectual'noe upravlenie slozhnymi obektami», «Intellectual'noe upravlenie slozhnymi tekhnicheskimi obektami», «Metody iskusstvennogo intellekta v upravlenii tekhnicheskimi obektami», «Programmnye sistemy i komplekсы v upravlenii kachestvom». Ufa: UGATU. Available at: https://www.ugatu.su/media/uploads/MainSite/Ob%20universitete/Izdateli/El_izd/2021-52.pdf
20. Protsenko, M. M., Pavlun'ko, M. Y., Moroz, D. P., Brzhevs'ka, Z. M. (2019). Procedure of signal filtering based on wavelet transformation. *Modern Information Security*, 1 (37), 64–69. doi: <https://doi.org/10.31673/2409-7292.2019.016469>
21. Protsenko, M., Kurtseitov, T., Pavlunko, M., Brzhevska, Z. (2018). Wavelet transforms application for digital signal analysis. Use of packet wavelet transformation for radio signals processing. *Modern Information Security*, 3 (35), 11–15. doi: <https://doi.org/10.31673/2409-7292.2018.031115>
22. El-Baz, A., Jiang, X., Jasjit, S. (Eds.) (2016). *Biomedical image segmentation*. CRC Press. doi: <https://doi.org/10.4324/9781315372273>