Classification and Recognition of Medical Images Based on the SGTM Neuroparadigm

Viktor Khavalko[1[0000-0002-9585-3078]], Ivan Tsmots[1[0000-0002-4033-8618]],

Anastasija Kostyniuk[2[0000-0002-8906-4531]] and Christine Strauss[3 [0000-0003-0276-3610]]

[1] Lviv Polytechnic National University, S. Bandery Str., 12, Lviv 79013, Ukraine
{khavalkov, ivan.tsmots}@gmail.com
[2] Stepan Gzhytskyi National University of Veterinary Medicine and Biotechnologies, Pekarska Str., 50, Lviv, 79000, Ukraine
anastasija.kostunyk@gmail.com
[3] University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Vienna, Austria
christine.strauss@univie.ac.at

Abstract. The paper discusses methods and algorithms for medical images pre-processing, their classification and recognition, which are oriented to use in machine vision systems. The structure and description of a number of software subsystems of image processing have been developed. The paper considers and analyzes the effectiveness of using methods for improving the visual quality of images as a stage of images pre-processing before classification. It is shown that image pre-processing is an effective and has significant impact on the accuracy of the images classification. The simulation of methods for improving the images' quality showed the correspondence of the practical results with the theoretical results, confirming to the reliability of the proposed approaches and full working capacity of the developed software product. For implementation of the subsystem of medical images classification, a neuroparadigm of successive geometric transformations model is adapted.

Keywords: Medical Image Processing; Image Classification; Successive Geometric Transformations Model; Neural Network.

1 Introduction

Automatic processing of visual information is one of the most important directions in the field of artificial intelligence. Interest in the problems of computer processing is determined by the expansion of the capabilities of both the computer systems themselves and the development of new technologies for the processing, analysis and identification of various types of images. In order to create effective technologies, methods and algorithms that are developed must meet a number of requirements for speed and accuracy. Usually, each algorithm, having certain characteristics, "specializes" in its type of image. Therefore, in machine vision systems (MVS)[1, 2], a combination of several methods, which solve the same problem in different ways, while providing

2019 IDDM Workshops.

the necessary parameters for the speed and authenticity of identification, is required. In turn, for the efficient functioning of MVS, it is necessary to regular replenish the arsenal of methods and means of pre-processing, compression of images and constructing classifiers, which necessitates the openness of these systems, as well as the need for tools for their design [1-3].

MVS has a long history of development and effective application in many high-tech areas of production. The use of machine vision is wide enough, it covers a wide range of activities, including, for example, such as: large industrial production; accelerated production of unique products; safety systems at work; preliminary control of finished products (for example, quality control, investigation of mistakes made); visual control and control systems (accounting, barcode reading); automated vehicle control; food quality control, etc. AI, machine learning and machine vision (MV) are absolutely essential to medical science in today's world. Making sense of scans and other kinds of medical imaging would be near impossible for a human specifically due to the sheer volume of images that a typical imaging procedure may produce.

In the last few years, AI has become so sophisticated in this field that it's oftentimes no longer necessary to code a machine to search for specific images. Through deep learning and pattern recognition processes, the software can find these images on its own when presented with similar images, thus reducing the time a medical professional is spent "training" the machine. This is extremely useful for radiologists and oncologists with a large caseload.

The AI can then go on to detect subtleties in these images that can reveal patterns humans wouldn't ordinarily be able to figure out. As a result of breakthroughs in deep learning, AI can correlate the subtle features of medical images hidden to human scrutiny with patient diagnoses [3-5]. As a result, medical professionals discover patterns that can influence which images they choose to focus when the AI selects it for them out of thousands. In other words, medical professionals learn from AI which images are cause for concern. This of course then translates to earlier diagnoses for patients, which can greatly increase their chances of survival and cure.

Machine vision is a critical technology for combining the contradictory demands of high quality performance standards in medical technology and reduced costs.

2 Formulation of the problem

The problem of perception and processing of images in real-time systems requires the development of new information bases with minimal redundancy of information, new principles for the construction algorithms with the ability to change the parameters to adapt them to the requirements of a specific task, new dynamic models and mechanisms for quick search of objects and follow them, new architectures for parallel processes for image processing.

The aim of this paper is consideration and analysis the effectiveness of using methods for improving the visual quality of images as a stage of images pre-processing before classification. Beside this it will be showed that image pre-processing are effective tool and has significant impact on the accuracy of the images classification. A new SGTM neuro-paradigm has been applied to accomplish this task and build an effective medical image classification subsystem. The results of computer simulations will allow to evaluate the efficiency, speed and accuracy of the classification of medical images before and after pre-processing.

3 Functional scheme of the machine vision system

The general functional scheme of the MVS is shown in Fig. 1. The image of the object through the camera is captured and transmitted to the image processing system, in which the pre-processing of the received image is performed. Through the visual control unit, supervision and control over the image preprocessing process are performed and, accordingly, those algorithms are selected that are necessary for the accomplishment of the task. In more detail, the structure, methods, algorithms and tasks that are solved by the image processing system are discussed below. If necessary, the processed information about the object is displayed on the visual control device. On the basis of the received information, the communications controller selects control signals that activate actuators that have a targeted effect on the object. In addition, MVS can record the results of image analysis on the media and output to the printer [2, 6,7].

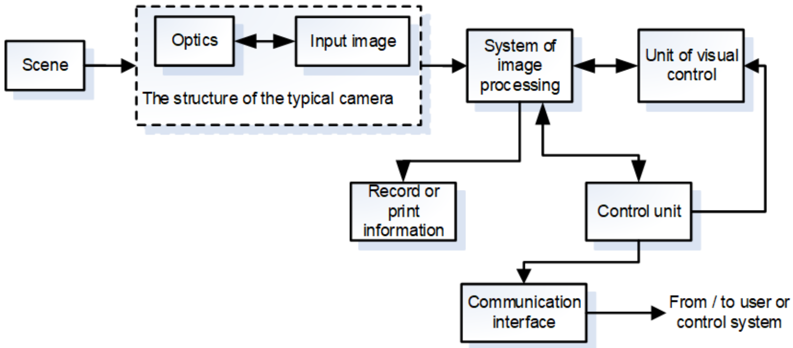


Fig. 1. MVS functional scheme

An important part of the system is the control unit. Its function includes managing the parameters of processing units, as well as synchronization of processes performed in the system.

The system of image processing consists of two units – primary processing unit and unit of secondary processing. The image analysis device (secondary processing) serves to select and recognize the object, determine its coordinates and position. If necessary, the processed information about the object is displayed on the visual control device. On the basis of the information received, the communications controller selects control signals that activate actuators that have a targeted effect on the object.

4 System of image processing and classification

High-quality images are a key requirement for efficient use of machine vision and reduced processing time. To obtain graphic information, the image must be processed and analyzed. Next, this information is compared with the database of known objects, on the basis of which a corresponding decision is made [5-9]. Typically, methods for processing and analyzing graphic details are integrated into a single algorithm. Among such complex methods one can distinguish the following.

4.1 Realization of system of image processing

To simulate image quality algorithms, software designed exclusively to improve image quality is developed, which allows you to explore the quality improvement of images by various methods with the possibility of the task of arbitrary filter masks.

The main functions of the program are [6, 9-11]:

- linear scaling;
- reduction of noise in the image using linear filtering;
- noise reduction in the image using median filtering;
- underscoring the boundaries.

Linear zooming or linear contrasting is performed in accordance with the above theoretical views. Scaling requires two image passes. During the first pass, the minimum and maximum values of brightness in the image are determined, while during the second pass, the calculation of the brightness of each point is performed.

For color images, this algorithm applies separately to each component of the color image. Since these components are three (R is a red component, G is green, B is blue), then six passes of the initial image are required, which requires significant computing resources.

The task of reducing noise in an image using linear filtering and underscoring the boundaries in terms of software implementation is one and the same task - the difference is only in the coefficients of the mask. It consists of direct multiplication of the mask coefficients by the pixel value of the image and finding its sum, which is written to the output file. After that, the mask shifts (1) to one pixel and the operation is repeated. Mask size 3x3, as a result, each pixel in the original image is replaced by the sum of pixels in the neighborhood of 3x3, including the central pixel [6]:

$$g(i,j) = \frac{1}{k}\sum_{m=-1}^{1}\sum_{n=-1}^{1} a(1-m,1-n)\, f(i-m,j-n), \tag{1}$$

where $f(i-m,j-n)$ - the value of the input image pixels; $a(1-m,1-n)$ - value of the coefficients of the mask; $g(i,j)$ - the value of the pixels of the original image, k - the coefficient, set by the user.

Often, the value of the coefficient k (2) is the sum of the mask elements:

$$k = \sum_{m=-1}^{1}\sum_{n=-1}^{1} a(1-m,1-n). \tag{2}$$

The main task of this coefficient is to provide in the output image, after calculations, values that are within one byte per component of the color image. But there may be other meanings if it improves the quality of the image and the program provides the output after computing to write to an array of float followed by bringing the result using linear scaling to the standard dynamic range for each component RGB - 0-255.

For color images, this algorithm applies separately to each component of the color image. Simplified algorithm's graph diagram is shown in Fig. 2.

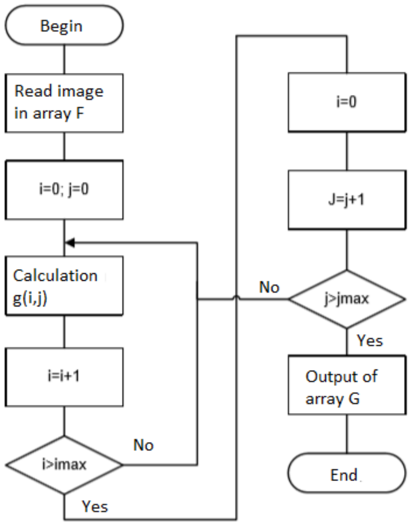


Fig. 2. Flow chart of the algorithm for linear filtering

4.2 Main approaches to the image classification

After the images preprocessing and bringing them to the necessary form the next important step is their classification.

At the present stage of the development of artificial intelligence it is often accepted to apply methods of machine learning for the implementation of classification and image recognition tasks [10-12]. Systems based on machine learning methods are used in systems of machine vision, for identification of objects in images, analysis of human language and texts, etc. [10-16].

Traditionally, pattern recognition (or their classification) was carried out on the basis of information signs. Consequently, the construction of patterns recognition systems (patterns-recognition) or systems based on the methods of machine learning, needed expert knowledge to develop methods and rules for the allocation of features (feature extraction). The selection of attributes is the conversion of the initial "raw" data (such as the pixel value to the image) into a suitable representation (vector of attributes) from which the learning system (classifier) can identify and classify the images submitted to the input. Such methods of machine learning are limited in the ability to process natural data in its original form [4].

Machine learning methods are divided into two main categories: supervised learning and unstructured learning. Learning methods with the teacher share the input data into a set of pre-defined classes. For training this classifier, you need a training sample that contains labeled samples of different classes. Learning methods without a teacher do not require training data, but they do not match a certain class with the input data, but only study the patterns in the input data and share the input data on similar groups (clusters) [5]. Since the task of the work is to analyze the methods of classification, then we will only consider the methods of teaching with the teacher.

Existing types of classifiers are systematized according to different criteria and their short characteristics are given in Table. 1 [10-17].

Table 1. Variety of approaches to classification depending on the criteria

Criterion	Type	Short Description
Use or non-use of training data	Classification with teacher	According to approaches to the classification with the teacher, the input data is shared using a set of samples as learning data
	Classification without teacher	Classification approaches without teacher, known as clustering, do not take into account educational data labels for the classification of input data
	Semi-automatic training	Learning in semi-automatic approaches takes place using data with and without labels
Taking into account or not taking into account any assumption of output data distribution	Parametric classifiers	Parametric classifiers are based on the assumption that the probability density function for each class is known
	Nonparametric classifiers	Nonparametric classifiers are not limited to any assumptions about the distribution of input data
Consideration of one classifier or ensemble	One classifier	A single classifier is used to assign a label to an object
	Ensemble of classifier	When determining the label for an object, the results of several (ensemble) classifiers are taken into account
The use or non-use of hard fission technology, where each object belongs to only one cluster	Hard classifier	Technologies of hard classification do not take into account further changes of different classes
	Soft (fuzzy) classifier	Fuzzy classifiers model gradual marginal changes, providing an assessment of the degree of similarity of all classes
Formation by classifier of the probability distribution of belonging to all classes	Probabilistic classifier	The classifier is capable of estimating the probability distribution for a given set of classes for a given sample
	Nonprobabilistic classifier	The approach defines only the most suitable class for the input image

Classifiers can be divided into parametric and non-parametric ones. The parametric ones include, for example, the maximum likelihood method, since it works on the assumption that the probability density function for each class is given by the gaussian distribution [10]. Nonparametric classifiers are not based on any assumptions about the distribution of input data. Given the fact that in most cases the distribution function is unknown, nonparametric classifiers have become much more widely used.

An important property of the classifiers is the possibility not only of the input data to a certain class (the classifier's output), but also to determine the probability of belonging to each of the classes, on the basis of which it is easy to choose the most reliable class. Such a feature is, for example, logistic regression. The resulting probabilities for each class can be used for post processing of classification results, for example, to combine outputs from different models in the ensemble, filtering noise, etc.

The most common approaches for classification problems are artificial neural networks [16-19], logistic regression [8], support vector machine (SVM) method [14,15] and random forest [18].

4.3 Application of Neural Networks in the Recognition Problem

Neural networks are successfully used in solving many problems of pattern recognition [4, 10, 12]: symbol recognition, pattern recognition, and many others. At present, the most common problems in image recognition and identification are the use of classical neural network architectures (multilayer perceptron, networks with radial basis function, convolutional neural networks, etc.). However, as the analysis of these works shows, the application of classical neural network architectures to the classification problem is not always is effective:

- usually, a neural network ensemble (2-3 neural networks trained with different initial values of synaptic coefficients and the order of the images input), which negatively affects the computational complexity of the problem and, accordingly, the time of execution;
- as a rule, classical neural network architectures are used in conjunction with auxiliary methods for selecting the plot part of the image (color segmentation, contour allocation, etc.) that require high-quality and punitive pre-processing of learning and working data that is not effective;
- neural network architectures are extremely sensitive to the effects of various external factors (changing the shooting conditions, the presence of individual features in the image, changing the orientation).

In addition, there are difficulties in the application of traditional neural networks to the real tasks of recognition and classification of images.

4.4 Architecture of ANN on the base of SGTM and modelling

To solve the classification problem, an approach of using neural networks based on the successive geometric transformations model (SGTM) is proposed [20-22]. Architecture of neural network based on the SGTM is shoved on Fig 3. Each input vector in the SGTM is considered as a point in N dimensional space, where N is the number of vector components.

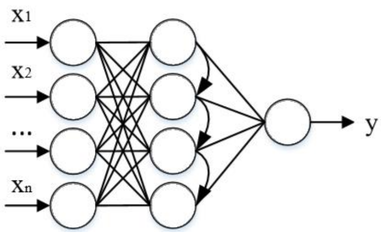
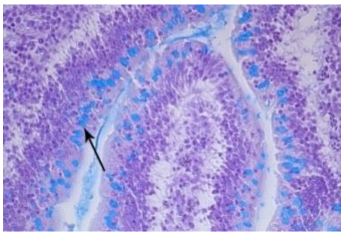


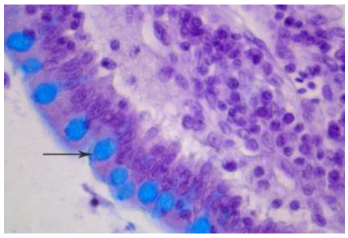
Fig. 3. Architecture of neural network based on the SGTM

Since each of the processed images had its own dimensions and characterized by the presence of the background, it required a pre-processing using the above described algorithms in order to obtain the input of such images whose classification gave it the most accurate result. Before starting to directly solve the problem of image classification using neural networks based on SGTM, the subsystem of received medical images pre-processing was applied. This is primarily due to several different reasons:

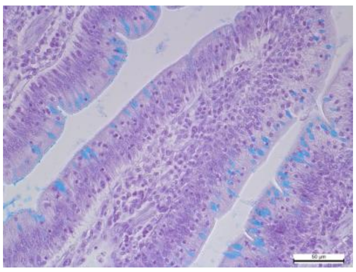
- Resize the original images as follows, to keep the image sizes multiple of the frame size (two image sizes used - 600x420 and 640x480).
- Improving the image quality by increasing the contrast in order to highlight of medical images features (shown by the arrows in Fig. 4a and Fig. 4b).



a)



b)



c)

Fig. 4. Test images: a) test image; b) image 1 for classification; c) image 2 for classification

The input of the neural network sequentially presents data about a single image or an image that corresponds to video stream frames that are separated by frames [23] of different size. In order to investigate the classification accuracy and determine the optimal frame size, it was decided to use different frame sizes, namely 3x3, 4x4, 5x5, 6x6 and 8x8.

That is, if the original image is 600x420 pixels in size, it is divided into 5x5 frames with a total number of pixels 25, then a matrix is fed to the input of the neural network, the number of columns of which is equal to the number of elements in the frame, namely 25, and the number of rows - the number of frames 120x84.

The obtained values of errors are shown in Table 2, and the graphs of these errors are shown in Fig. 6.

Fig. 4 shows three images: a training image (Fig. 4 a) and two images for which the classification procedure was conducted (Fig. 4 b and 4 c). Fig. 4 b is characterized by a higher level of detail and higher scale, whereas the image in Fig 4 c is more general and contains more information. The classification results are positive in both cases, but the classification accuracy is different (Figs. 5, 6 and Table 2).

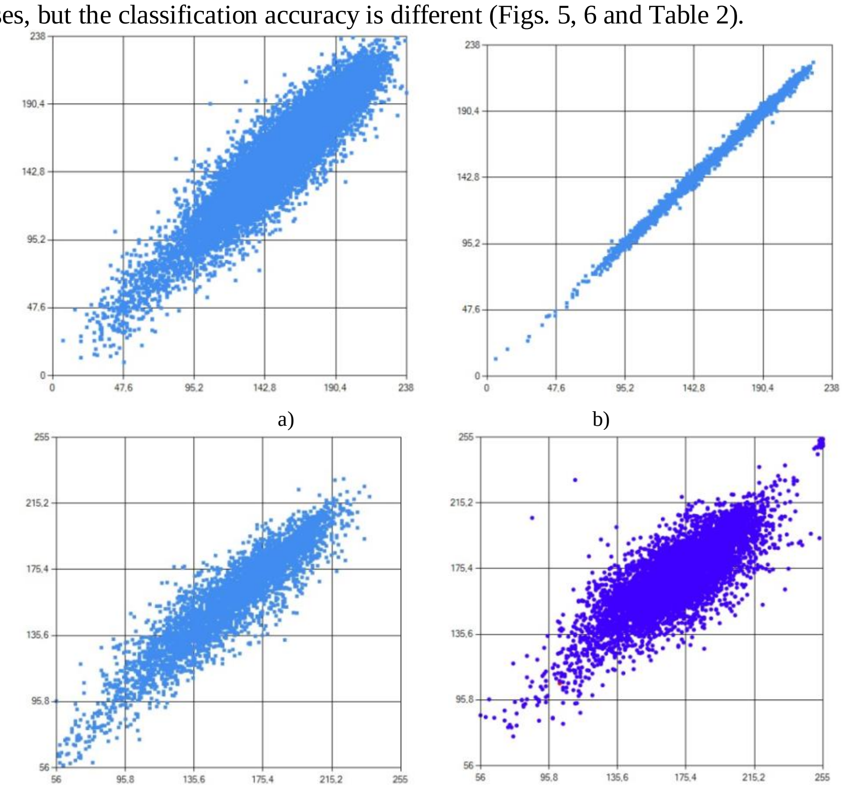


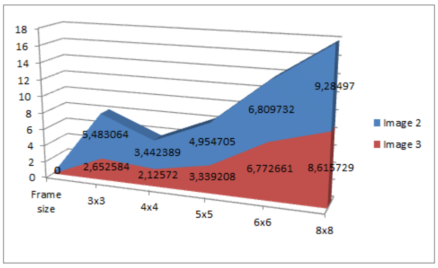
Fig. 5. The result of the classification in the form of scattered plots for: a) for image 1 (Fig. 4 b) frame size 3x3 1; b) for image 2 (Fig. 4 c) frame size 3x3; c) for image 1 (Fig. 4 b) frame size 6x6 1; d) for image 2 (Fig. 4 c) frame size 6x6;

Fig. 5 shows the results of the classifications for each of the images. For frame with size 3x3 we can say that for image 3 (Fig. 4 c) almost all points on Fig. 5 b that the neural network has given after classification, densely placed along the line, which means that the image on which the network trained and the test image belonged to one class and the error is small. What you cannot say about the results obtained for test image 2 (Fig. 4 b). The results of the given neural network are scattered and the classification errors are greater.

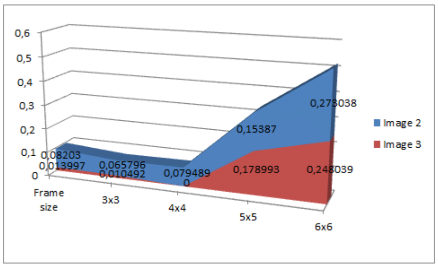
Table 2. Quality indicators of the obtained results

Frame size/ Indicator	Sample	MAPE, %	RMSE	RMSE_M, %
3x3	Image 1	5,483064	0,082030	0.034466
	Image 2	2,652584	0,013997	0,005881
4x4	Image 1	3,442389	0,065796	0,026746
	Image 2	2,125720	0,010492	0,003789
5x5	Image 1	4,954705	0,079489	0,048903
	Image 2	3,339208	0.090919	0,038041
6x6	Image 1	6,809732	0,153870	0,060341
	Image 2	6,772661	0,178993	0,070193
8x8	Image 1	9,284970	0,273038	0,097045
	Image 2	8,615729	0,248039	0,083970

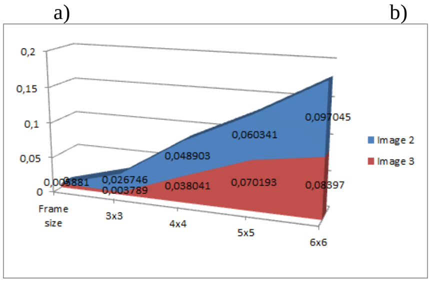
Figure 6 shows graphs of dependance the classification errors of medical images on the frame size.



a)



b)



c)

Fig. 6. Graphs of dependance the classification errors of medical images on the frame size: a) MAPE,%; b) RMSE; c) RMSE_M, %

For frame with size 6x6 we can see (Fig 5,c and Fig.5,d) that classification accuracy is more worse (Table. 2). Table 2 shows the results of errors for the obtained classification results and Figure 6 shows the graphs of the magnitude of the classification error on the frame size.

The graphs in Fig. 6 show that the classification errors increases with the size of the frame. Having analyzed the obtained results, it is safe to say that the most optimal frame size is 4x4, since in this case all three considered errors (MAPE, RMSE and RMSE_M) are minimal and provide the best result in terms of efficiency both in the pre-processing of medical images and their classification.

5 Conclusion

1. The structure and description of a number of software subsystems for MVS was developed, which made it possible to highlight the most important components and to analyze the basic approaches of their realisation.
2. The methods and algorithms of image pre-processing, their classification and recognition, which are oriented for use in MVS, are considered and analyzed.
3. In general, the use of the proposed methods in the development of MVS allows several times to reduce the amount of computational operations with increasing the probability of recognizing objects in images.
4. The SGTM paradigm has been applied for the implementation of neural network medical image classification.
5. The choice of the optimal frame size is substantiated and its influence on the errors value during the medical images classification using neural networks based on the SGTM paradigm is investigated.

References

1. Forsyth, D., Ponce, J.: Computer Vision: A Modern Approach. Pearson Education, New Jersey (2011).
2. Davies, E.R.: Computer and Machine Vision: Theory, Algorithms and Practicalities, Academic Press, Orlando (2012)
3. Robb, R.A.: Biomedical Imaging, Visualization and Analysis, John Wiley & Sons Inc, New York (2000).
4. Bankman, I.: Handbook of Medical Image Processing and Analysis, 2nd Edition, Academic Press, New York (2008).
5. Haykin, S.: Neural networks and learning machines. 3rd ed. Prentice Hall, New York (2009).
6. Gonzalez, R., Woods, R.: Digital image processing. Tekhnosfera, Moscow (2005).
7. Kryvenchuk, Y., Shakhovska, N., Melnykova, N., Holoshchuk, R.: Smart Integrated Robotics System for SMEs Controlled by Internet of Things Based on Dynamic Manufacturing Processes. In Conference on Computer Science and Information Technologies. pp. 535-549 (2018).

8. Kryvenchuk, Y., Shakhovska, N., Shvorob, I., Montenegro, S., & Nechepurenko, M.: The Smart House based System for the Collection and Analysis of Medical Data. CEUR, 2255. pp. 215-228 (2018).
9. Fukushima, K.: Cognitron: A self-organizing multilayered neural net-work. In Biological cybernetics, vol. 20, no 3-4.121–136 (1975).
10. Bishop, C.: Pattern Recognition and Machine Learning. Springer, New York (2006).
11. Ghamisi, P., Plaza, J., Chen, Y.: Advanced Spectral Classifiers for Hyperspectral Images: A review. In IEEE Geoscience and Remote Sensing Magazine, vol. 5, no 1, 8–32 (2017).
12. Fukunaga, K.: Introduction to statistical pattern recognition, Academic press, New York (2013).
13. Schmidhuber, J.: Deep learning in neural networks: An overview. In Neural networks, vol. 61, 85–117 (2015).
14. LeCun, Y., Yoshua, B., Geoffrey, H. Deep learning. In Nature, vol. 521, no 7553, 436–444 (2015).
15. Liu, P., Cho,o K., Wang, L., Huang, F.: SVM or deep learning? A comparative study on remote sensing image classification. In Soft Computing, vol. 21, no 23, 7053–7065 (2017).
16. Huang, F., LeCun, Y.: Large-scale learning with SVM and convolutional nets for generic object categorization. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 284–291 (2006).
17. Pirotti, F., Sunar, F., Piragnolo, M.: Benchmark of machine learning methods for classification of a Sentinel-2 image. In International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, vol. 41, 335–340 (2016).
18. Krizhevsky, A., Sutskever, I., Hinton G.: ImageNet classification with deep convolutional neural networks. In Advances in neural information processing systems, 1097–1105 (2012).
19. Fukushima, K.: Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. In Biol. Cybern, vol.36, 193–202 (1980).
20. Tkachenko, R., Yurchak, I. and Polishchuk, U.: Neurolike networks on the basis of Geometrical Transformation Machine. In International Conference on Perspective Technologies and Methods in MEMS Design, pp. 77-80 (2008).
21. Tkachenko, R., Izonin, I.: Model and Principles for the Implementation of Neural-Like Structures Based on Geometric Data Transformations. In: Hu Z., Petoukhov S., Dychka I., He M. (eds) Advances in Computer Science for Engineering and Education. ICCSEEA 2018. Advances in Intelligent Systems and Computing, vol 754, pp 578-587 (2019).
22. Tsymbal, Y., Tkachenko, R.: A digital watermarking scheme based on autoassociative neural networks of the geometric transformations model. In IEEE First International Conference on Data Stream Mining & Processing (DSMP), pp. 231-234 (2016).
23. Khavalko, V., Tsmots, I.: Image classification and recognition on the base of autoassociative neural network usage. In IEEE 2nd Ukraine Conference on Electrical and Computer Engineering (UKRCON-2019), pp. 1118-1121 (2019).