http://doi.org/10.35784/iapgos.2951
received: 08.05.2022 | revised: 17.06.2022 | accepted: 17.06.2022 | available online: 30.06.2022

METHODS FOR DETECTING AND SELECTING AREAS ON TEXTURE BIOMEDICAL IMAGES OF BREAST CANCER

Ainur Orazayeva[1], Jamalbek Tussupov[1], Waldemar Wójcik[2], Sergii Pavlov[3], Gulzira Abdikerimova[1], Liudmyla Savytska[3]
[1]L.N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan, [2]Lublin University of Technology, Lublin, Poland, [3]Vinnytsia National Technical University, Vinnytsia, Ukraine

Abstract. *This paper is devoted to topical issues - the development of methods for analyzing texture images of breast cancer. The main problem that is resolved in the article is that the requirements for the results of pre-processing are increasing. As a result of the task, images of magnetic resonance imaging of the breast are considered for image processing using texture image analysis methods. The main goal of the research is the development and implementation of algorithms that allow detecting and isolating a tumor in the breast in women in an image. To solve the problem, textural features, clustering, orthogonal transformations are used. The methods of analysis of texture images of breast cancer, carried out in the article, namely: Hadamard transform, oblique transform, discrete cosine transform, Daubechies transform, Legendre transform, the results of their software implementation on the example of biomedical images of oncological pathologies on the example of breast cancer, it is shown that The most informative for image segmentation is the method based on the Hadamard transform and the method based on the Haar transform. The article presents recommendations for using the results in practice, namely, it is shown that clinically important indicators that make a significant contribution to assessing the degree of pathology and the likelihood of developing diseases, there are other information parameters: diameter, curvature, etc. Therefore, increased requirements for the reliability, accuracy, speed of processing biomedical images.*

Keywords: biomedical image processing, textural features, clustering, orthogonal transformations

METODY WYKRYWANIA I WYRÓŻNIANIA OBSZARÓW W TEKSTUROWANYCH OBRAZACH BIOMEDYCZNYCH RAKA PIERSI

Streszczenie*. Niniejszy artykuł jest poświęcony aktualnemu tematowi - opracowaniu metod analizy obrazów tekstury raka piersi. Główny problem, który został rozwiązany w artykule, polega na tym, że wymagania wobec wyników przetwarzania wstępnego są coraz większe. W wyniku realizacji zadania rozpatrzono obrazy rezonansu magnetycznego piersi przeznaczone do przetwarzania metodami teksturowej analizy obrazu. Głównym celem badań jest opracowanie i wdrożenie algorytmów wykrywania i odróżniania na obrazie guza w piersi u kobiet. Do rozwiązania tego problemu wykorzystuje się cechy tekstury, grupowanie i transformacje ortogonalne. W artykule przedstawiono metody analizy obrazów teksturowych raka piersi, tj. transformatę Adamarda, transformatę skośną, transformatę dyskretno-cosinusową, transformatę Dobeshiego, transformatę Lejandre'a, oraz wyniki ich implementacji programowej na przykładzie obrazów biomedycznych patologii onkologicznej w przypadku raka piersi. Metoda oparta na transformacie Adamarda oraz metoda oparta na transformacie Haara są najbardziej przydatne do segmentacji obrazów. W artykule przedstawiono zalecenia dotyczące wykorzystania wyników w praktyce, a mianowicie wykazano, że inne parametry informacyjne, takie jak średnica, krzywizna itp. są ważnymi klinicznie wskaźnikami, które w istotny sposób przyczyniają się do oceny stopnia patologii i prawdopodobieństwa rozwoju choroby. W związku z tym wzrastają wymagania dotyczące niezawodności, dokładności i szybkości przetwarzania obrazów biomedycznych w urządzeniach diagnostycznych*.

Słowa kluczowe: przetwarzanie obrazów biomedycznych, cechy tekstury, klasteryzacja, transformacje ortogonalne

Introduction

Despite the ubiquitous presence of textures in images, there is no single and formal approach to the description of texture and its strict definition at the moment. Texture analysis methods are usually developed for each individual case.

Recently, there have been many works devoted to the analysis of dynamic textures and color images. Analyzing the textures of color images, additional designations are introduced for their characteristics, based on measuring the intensity levels of each color and distributing them over the image field. When analyzing dynamic structures that change in time, space-time is introduced, which is the third dimension, which is connected to two spatial coordinates [7, 18, 19].

Spectral analysis is a powerful tool for analyzing signals and images, because it has long been noted that the spectrum is very sensitive to various changes in the structure of signals and images.

To perform spectral analysis, it is necessary to first decompose the signal or image into frequencies. For this, different sets of basis functions are used. The corresponding algorithms are called transforms: cosine, Hadamard, Haar, oblique, etc. Note that the Haar and Daubechies, Mueller-matrix polarimetry (MMP) [13, 14] for the transforms are the simplest wavelet transforms. These methods, according to the theory of signal processing, can be applied to stationary random processes, but often we do not have this. However, it is possible to select for analysis areas that are considered conditionally stationary (in other words, quasi-stationary) and whose size is sufficient to obtain statistically correct results [14, 15].

1. Method

The program is implemented in the Matlab environment and allows performing spectral transformations of six types: 1) cosine, 2) Hadamard of order 2^n, 3) Hadamard of order n = p + 1, p = 3(mod4) is a prime number, i.e. based on the Legendre symbol, 4) Haar, 5) oblique, 6) Daubechies-4, 7) MMP.

To normalize the Haar function, the following formula is used:

$$\alpha_{mk}(\theta)=\begin{cases}\sqrt{2^m}, & \theta\in\left[\frac{k}{2^m},\frac{k+1/2}{2^m}\right)\\ -\sqrt{2^m}, & \theta\in\left[\frac{k+1/2}{2^m},\frac{k+1}{2^m}\right).\\ 0, & \theta\in\left[\frac{k}{2^m},\frac{k+1}{2^m}\right)\end{cases} \tag{1}$$

The Daubechies-4 transform is given [6, 7] by the following matrix

$$M=\sqrt{2}\begin{bmatrix} h_0 & h_1 & h_2 & h_3 & & & & \\ & & h_0 & h_1 & h_2 & h_3 & & \\ & & & & h_0 & h_1 & h_2 & h_3 \\ h_2 & h_3 & & & & & h_0 & h_1 \\ h_3 & -h_2 & h_1 & -h_0 & & & & \\ & & h_3 & -h_2 & h_1 & -h_0 & & \\ & & & & h_3 & -h_2 & h_1 & -h_0 \\ h_1 & -h_0 & & & & & h_3 & -h_2 \end{bmatrix} \tag{2}$$

Matrix elements are calculated using the formulas below:

$$h_0=(1+\sqrt{3})/8,\qquad h_1=(3+\sqrt{3})/8,$$
$$h_2=(3-\sqrt{3})/8,\qquad h_3=(1-\sqrt{3})/8. \tag{3}$$

In most cases, iterative algorithms are used to select areas in grayscale images. For example, the k-mean algorithm [16, 17] is mainly used in the space of pixel brightness and divides images into regions according to a given number. The main feature of this algorithm is its simplicity and speed of execution [15–17].

This method is implemented in the MatLab environment and is used for image analysis. The image is divided by software into separate parts. When studying textures, you can use various non-standard approaches that use orthogonal transformation. For example, the original image is divided into square windows that do not intersect. Experiments have shown that it is better to take a large window size for its intended purpose: 32×32, 64×64, etc. We carry out an integral transformation of each window. In the two-dimensional case, the frequency spectra have the form of a matrix. We place the elements of the matrix in the form of vectors. For example, we can arrange the rows of a matrix sequentially at the end of each other. As a result, we carry out the procedure of clustering these vectors [1, 2, 4].

2. Experimental results

During the experiment, 30 X-ray images of the mammary glands were taken as initial images. Each of the images has dimensions of 1024×1024.

The images were first classified by an expert physician into three classes:

- Images of the mammary gland without pathology were taken as sample No 1 (Fig. 1a).
- Images of the breast on the right side with different stages of pathologies were considered as sample No 2 (Fig. 1b).
- Images of the breast on the left side with different stages of pathologies were considered as sample No 3 (Fig. 1c).

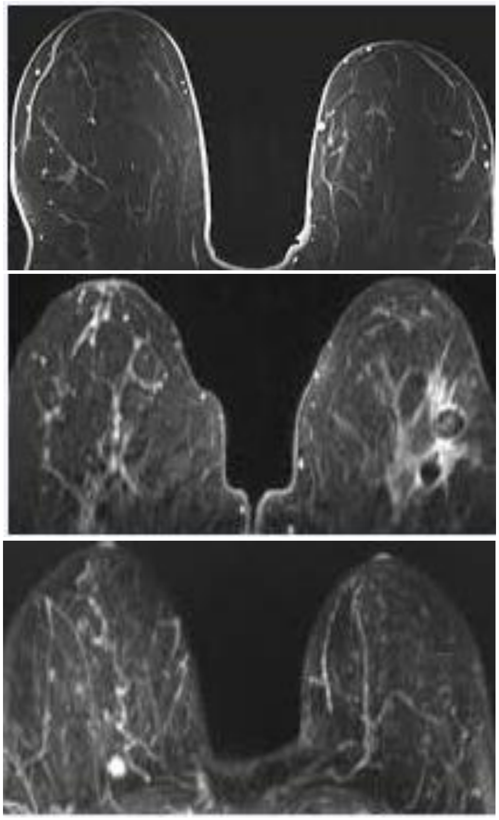

Fig. 1. Sample No 1 of the image of the mammary gland without pathology (a); sample No 2 of the image of the breast on the right side with different stages of pathologies (b); sample No 3 the image of the breast on the left side with different stages of pathologies (c)

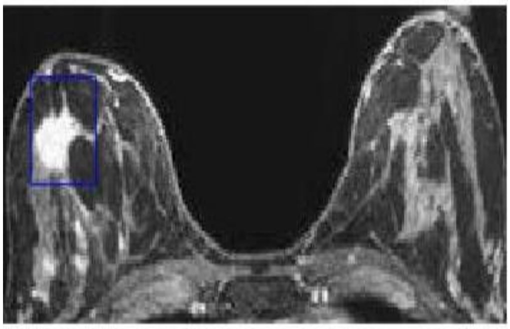

Fig. 2. Selection of a fragment for calculating the orthogonal transformation

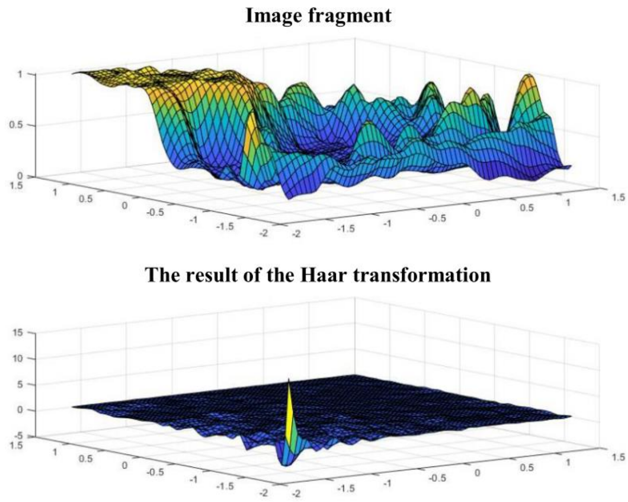

Fig. 3. Graph of the original brightness function in the window and the result of the Haar transformation

The entire code of the main program can be manually inserted into the Matlab workspace, and the results will immediately appear – graphics and 7 text files. Graphs of the results of transformations will be presented each in a separate window, and all at the same time in one window. Matlab allows graphs to be rotated and viewed from different angles. Text files store the original data and the results of the six transformations listed above. All text files are automatically saved in the work directory, inside the Matlab system [3, 10, 12].

It can be concluded that orthogonal transformations are effective for mammographic images; all images were clustered. As an example, after applying an orthogonal transformation to figure 4, as shown in the result of figure 5, 4% of women were found to have a tumor on the left side of the breast [5 9, 11].

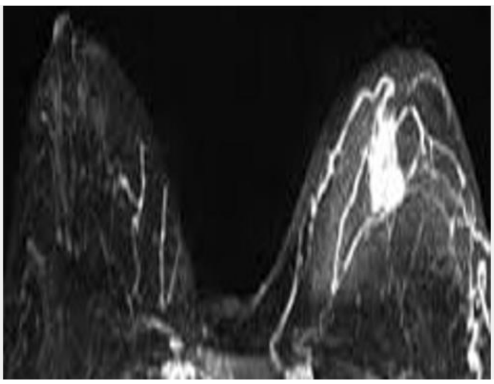

Fig. 4. Original image

a)

b)

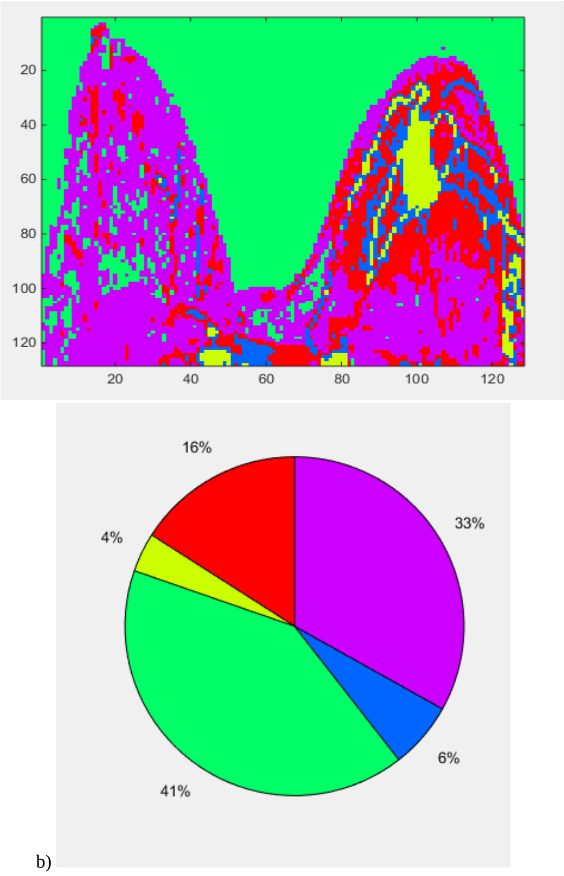

Fig. 5. The result of clustering by the Daubechies method, the window size is 8×8 (a); percentage result of Daubechies clustering result (b)

The percentages of the presence of breast tumors according to the considered figures are presented in the table below (table 1).

Table 1. Percentage of biomedical image processing by 6 methods of orthogonal transformations

Methods	**Images**													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
haara	4	2	2	3	8	3	3	6	2	3	5	3	5	4
dobeshi	5	2	2	4	8	3	3	4	2	2	5	3	5	4
discrete	4	2	2	4	8	3	3	5	1	3	5	3	5	4
naklon	4	2	2	4	8	3	3	6	3	2	5	3	5	4
legandr	4	2	2	4	8	3	3	5	2	3	5	3	5	4
hadamard	4	2	2	4	8	3	3	5	1	3	5	3	5	4

Methods	**Images**													
	15	16	17	18	19	20	21	22	23	24	25	26	27	28
haara	6	10	4	5	3	6	5	5	8	7	3	9	4	9
dobeshi	2	11	3	4	2	1	2	4	8	7	3	6	4	8
discrete	4	11	3	5	2	3	5	4	8	7	3	6	4	9
naklon	4	11	4	5	3	4	5	5	8	7	3	6	4	9
legandr	2	11	4	5	2	4	5	5	8	7	3	6	4	8
hadamard	2	11	3	5	4	4	5	3	8	7	3	6	4	9

From the considered methods, a graphical representation of the Haar method is shown in the following diagram (figure 6).

Of the considered methods, a graphical representation of the Haar method is shown in the following diagram (figure 6).

Checking the results of cluster solutions. To check the correct distribution over clusters, the following formula (4) is used. Percentage of correct class definitions:

$$P = \frac{w(1)' * 100}{w(1)} \tag{4}$$

where

- $w(1)'$ is the number of correct objects in the cluster;
- $w(1)$ number of all considered images.

The selected methods determine the pathology of breast cancer in images by an average of 98% (table 2).

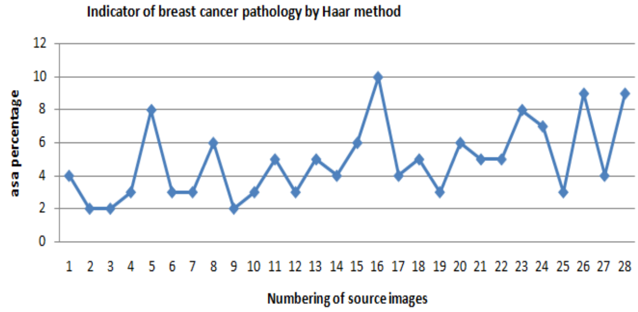

Fig. 6. Indicator of the pathology of breast cancer by the Haar method

Table 2. Clustering result

Methods used	The result of determining the pathology in percent (*P*)	
	Without pathology	With pathology
Haara	97%	97%
Dobeshi	98%	97%
Discrete	98%	97%
Naklon	98%	97%
Legandr	98%	97%
Hadamard	98%	97%
Mueller-matrix polarimetry	98%	97%

3. Conclusions

This investigation is devoted to the study of texture images. The source is mammography images. The main result is the creation of software tools and experiments on image processing. The program is implemented in the Matlab environment, which allows performing spectral transformations of six types: 1) cosine, 2) Hadamard of order 2^n, 3) Hadamard of order n = p + 1, p = 3(mod4) is a prime number, i.e. based on the Legendre symbol, 4) Haar, 5) oblique, 6) Daubechies-4, 7) Mueller-matrix polarimetry.

The algorithms that were considered in this paper allowed us to effectively isolate areas on the analyzed images that are characterized by different stages of breast cancer. More precisely, doctors are interested in early diagnosis of breast pathology in women.

The images used in this article were taken from the Department of Computed Tomography of the Republican Diagnostic Center, 28 images were taken from there, including 26 images with pathology, 2 images without pathology. Based on these data, an experiment was conducted on image processing, and when using spectral transformations of six types, the program shows a 2% error [8, 20].

Looking at figure 6 we can say that when processing images with 6 methods of orthogonal transformations, the percentage of clustering result of 2% and below is 98% of images without pathology, and above 2% – images with pathology. This corresponds to the result of determining the pathology by 98%.

In our further studies, the parameters of indicators can be associated with different stages of breast cancer and other characteristics. The software system can be trained by examples using algorithms based on brain-computer or other approaches commonly used in machine learning. After training, the system will be able to predict the values of the parameters.

References

[1] Abdikerimova G. B. et al.: Algorithms and software for the analysis of disordering the structure of cellular walls. Bulletin of the Novosibirsk Computing Center – Series „Computer science" 40, 2016, 1–14.

[2] Abdikerimova G. B. et al.: Software tools for cell walls segmentation in microphotography. Journal of Theoretical and Applied Information Technology 96, 2018, 4783–4793.

[3] Ablyazov T., Baizakov N.: Theory and Practice of Territories Spatial Development Based on the Smart City Concept. In: Rudskoi A., Akaev A., Devezas T. (eds): Digital Transformation and the World Economy. Studies on Entrepreneurship, Structural Change and Industrial Dynamics. Springer, Cham 2022 [http://doi.org/10.1007/978-3-030-89832-8_9].

[4] Arazayeva A. R. et al.: Efficiency of Biomedical Breast Cancer Image Processing Using Filters. The National Academy of Sciences of the Republic of Kazakhstan ¬– Physics and Information Technology Series 1(341), 2022, 69–76 [http://doi.org/10.32014/2022.2518-1726.118].
[5] Bayzakov S., Forrest J. Y.-L., Baizakov N. A.: Modeling the management of the economies of developing countries. Advances in Systems Science and Applications 19 (2), 2019, 101–119 [http://doi.org/10.25728/assa.2019.19.2.673].
[6] Fu K. S., Mu J. K.: A survey on image segmentation. Pattern Recognition 13(1), 1981, 3–16.
[7] Grady L.: Random walks for image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 28(11), 2006, 1768–1783.
[8] Karpinski M., Ziubina R., Azatov A., Shaikhanova A., Teliushchenko V., Falat P.,: Information Security Software Using Quality and Reliability Criteria. IEEE 5th International Symposium on Smart and Wireless Systems within the Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS), 2020, 1–5 [http://doi.org/10.1109/IDAACS-SWS50031.2020.9297069].
[9] Pavlov S. et al.: System of three-dimensional human face images formation for plastic and reconstructive medicine. In: Arras P., Luengo D. (Eds.): Teaching and subjects on bio-medical engineering Approaches and experiences from the BIOART-project. Acco cv, Leuven 2021, 187–203.
[10] Rani R.: Performance analysis of different orthogonal transform for image processing application. International Journal of Applied Research 1(12), 2015, 844–847.
[11] Serik M. , Yerlanova G., Karelkhan N. , Temirbekov N.: The Use of the High-Performance Computing in the Learning Process. International Journal of Emerging Technologies in Learning (iJET) 16(17), 2021, 240–254 [http://doi.org/10.3991/ijet.v16i17.22889].
[12] Sidorova V. S.: Hierarchical Cluster Algorithm for Remote Sensing Data of Earth. Pattern Recognition and Image Analysis 22(2), 2012, 373–379.
[13] Stashkevich A. T. et al.: Differential Mueller-matrix tomography of the polycrystalline structure of biological tissues with different damage durations. Proceedings of SPIE 12040, 2021, 120400G.
[14] Stashkevich A. T. et al.: Multiparameter polarization-phase microscopy of optically anisotropic networks of biological crystals. Proceedings of SPIE 12040, 2021, 120400F.
[15] Stashkevich A. T. et al.: Spectral polarimetry of laser images of biological fluid layers in the differentiation of necrotic conditions. Proceedings of SPIE 12040, 2021, 120400C.
[16] Wang J. Z., Du Y.: Scalable integrated region-based image retrieval using IRM and statistical clustering. Proceedings of the 1st ACM/IEEE-CS joint conference on Digital libraries 2001, 268–277.
[17] Wang J. Z., Li J., Wiederhold G.: SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries. IEEE Transactions on Pattern Analysis and Machine Intelligence 23(9), 2001, 947–963.
[18] Wójcik W., Pavlov S., Kalimoldayev M.:. Information Technology in Medical Diagnostics II. Taylor & Francis Group, CRC Press, Balkema book, London 2019 [http://doi.org/10.1201/ 9780429057618].
[19] Wójcik W., Smolarz A. (ed.): Information Technology in Medical Diagnostics. CRC Press 2017.
[20] Zhao F., Nagarathnam M.: 5418 Parallel Computer Architecture and Programming Project Report: Implementation and Comparison of Parallel LZ77 and LZ78 Algorithms. 2021.

M.Sc. Ainur Orazayeva
e-mail: oaris.83@gmail.com

Master, doctoral student of Department of Information Systems, L.N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan.
Scientific direction: biomedical information technologies for the diagnosis of oncological diseases. Deals with the development of biomedical diagnostic systems and methods.

http://orcid.org/0000-0002-2899-9886

D.Sc. Jamalbek Tussupov
e-mail: tussupov@mail.ru

Doctor of Physical and Mathematical Sciences, professor Department of Information Systems, L.N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan.
Scientific direction: mathematics, discrete mathematics, mathematical logic, theory of algorithms, computable algebraic structures, number theory, algebra, graph theory, group theory, computer science (computer science), abstract data types, graphs and algorithms, models of computation, analysis and complexity of algorithms, sequential and parallel algorithms.

http://orcid.org/0000-0002-9179-0428

Prof. Waldemar Wójcik
e-mail: waldemar.wojcik@pollub.pl

He was Director of the Department of Electronics and Information Technology, former long-time dean of the Faculty of Electrical Engineering and Computer Science at Lublin University of Technology. Doctor Honoris Causa of universities in Ukraine and Kazakhstan. He is also a member of European Academy of Science and Arts, International Academy of Applied Radioelectronics, The International Informatization Academy of Kazakhstan and many other scientific organisations of Poland as well as Europe and Asia.
In his research he mainly deals with process control, optoelectronics, digital data analysis and also heat processes or solid state physics.

http://orcid.org/0000-0002-6473-9627

D.Sc. Sergii Pavlov
e-mail: psv@vntu.edu.ua

Academician of International Applied Radioelectronic Science Academy, Professor of Biomedical Engineering and Optic-Electronic Systems Department Vinnitsia National Technical University
Scientific direction: biomedical information optoelectronic and laser technologies for diagnostics and physiotherapy influence.
Deals with issues of improving the distribution of optical radiation theory in biological objects, particularly through the use of electro-optical systems, and the development of intelligent biomedical optoelectronic diagnostic systems and standardized methods for reliably determining the main hemodynamic cardiovascular system of comprehensive into account scattering effects.

http://orcid.org/0000-0002-0051-5560

Ph.D. Gulzira Abdikerimova
e-mail: gulzira1981@mail.ru

Ph.D., Associate professor of Department of Information Systems, L.N. Gumilyov Eurasian National University, Nur-Sultan, Kazakhstan.
Scientific direction: image and signal processing, textural feature analysis, machine learning.

http://orcid.org/0000-0002-4953-0737

Ph.D. Liudmyla Savytska
e-mail: lyudik0304@gmail.com

Docent of Software Department Vinnitsia National Technical University She graduated from VNTU with a degree in "Intelligent Information Processing and Decision Making" with honors and qualified as a software engineer. She successfully defended her dissertation on "Methods of adaptive data compression based on the linear Fibonacci form".
Scientific direction: image and signal processing, machine learning.

http://orcid.org/0000-0003-1130-2621