

DEEP NETWORK-BASED METHOD AND SOFTWARE FOR SMALL SAMPLE BIOMEDICAL IMAGE GENERATION AND CLASSIFICATION

Berezsky O. M. – Dr. Sc., Professor, Professor of the Department of Computer Engineering, West Ukrainian National University, Ternopil, Ukraine.

Liashchynskyi P. B. – Post-graduate student of the Department of Computer Engineering, West Ukrainian National University, Ternopil, Ukraine.

Pitsun O. Y. – PhD, Associate Professor, Associate Professor of the Department of Computer Engineering, West Ukrainian National University, Ternopil, Ukraine.

Melnyk G. M. – PhD, Associate Professor of the Department of Computer Engineering, West Ukrainian National University, Ternopil, Ukraine.

ABSTRACT

Context. The authors of the article investigated the problem of generating and classifying breast cancer histological images. The widespread incidence of breast cancer explains the problem's relevance. The automated diagnosing procedure saves time and eliminates the subjective aspect. The study's findings can be applied to cancer CAD systems.

Objective. The purpose of the study is to develop a deep neural network-based method and software tool for generating and classifying histological images in order to increase classification accuracy.

Method. The method of histological image generation and classification was developed in the research study. This method employs CNN and GAN. To improve the classification accuracy, the initial image sample was expanded using GAN.

Results. The computer research of the developed method of image generation and classification was conducted on the basis of the dataset located on the Zenodo platform. Light microscopy served as the basis for obtaining the image. The dataset contained three classes of G1, G2, and G3 breast cancer histological images. Based on the developed method, the accuracy of image classification was 96%. This is a higher classification accuracy compared to existing models such as AlexNet, LeNet5, and VGG16. The software module can be integrated into CAD.

Conclusions. The developed method of generating and classifying images is the basis of the software module. The software module can be integrated into CAD.

KEYWORDS: computer-aided diagnosis system, breast cancer, deep neural networks, generative competitive networks, convolutional neural networks.

ABBREVIATIONS

GAN is a generative adversarial network;
CNN is a convolutional neural network;
DNN is a deep neural network;
CAD is a computer-aided diagnosis;
AWS is a Amazon Web Services;
Zenodo is a general-purpose open repository developed under the European OpenAIRE program and operated by CERN;
ReLU is a rectified linear unit;
Batch Norm is batch normalization;
ROC is a receiver operating characteristic;
CLI is a command line interface;
IS metric is a metric based on the Google Inception V3 image classification model;
FID is Fréchet inception distance;
AlexNet is a name of a convolutional neural network architecture designed by Alex Krizhevsky;
LeNet5 is a name of a convolutional neural network structure proposed by LeCun;
VGG16 is a name of a Visual Geometry Group convolutional neural network;
ResNet50 is a name of a 50-layer Residual Network;
DenseNet201 is a name of 201-layer Densely Connected Network;

CSAResnet is a channel and spatial attention embedded Resnet network;
DAMCNN is a dual attention multiscale convolutional neural network;
DSS is a decision support system;
DeepGrade is a histological grade model;
DenseNet is a Densely Connected Network;
CA-BreastNet is a Coordinated Attention Breast Network;
DHE-Mit-Classifier is a Deep Heterogeneous Ensemble mitotic Classifier;
SVM is a Support Vector Machine;
SSDHO is a Shuffled Shepherd Deer Hunting Optimization;
LR is a Logistic Regression;
MLP is a Multilayer Perceptron;
U-Net is a Network with U-shaped structure;
PyTorch is an open source machine learning framework Python Torch;
AWS SageMaker is an Amazon Web Services SageMaker machine learning service;
AWS S3 is an Amazon Web Services Simple Storage Service;
URL is a Uniform Resource Locator;
CrossEntropyLoss is a cross entropy loss;

RAM is a Random Access Memory;
vCPU is a virtual Central processing unit;
GPU is a graphics processing unit;
TFLOPS is a Tera FLoating-point OPERations per Second;

Adam optimizer is an adaptive moment stochastic gradient descent method.

NOMENCLATURE

I_{imp} is a set of input images;
 I_l is a training set of images;
 I_t is a test set of images;
 L is a set of CNN layers;
 A_{CNN} is a CNN architecture;
 L_{DEV} is a set of developed CNN layers;
 O_{DEV} is a set of developed CNN operations;
 O_{CNN} is a set of CNN operations;
 P_{CNN} is a set of CNN parameters;
 G_{max} is a pooling function with maximum element definition in the scan window;

G_{avg} is a pooling function with average element definition in the scan window;

G_{ad_avg} is a pooling function with adaptive definition of the average element in the scan window;

i is a layer index;

n is a number of CNN layers;

TP is a true positive;

TN is a true negative;

FP is a false positive;

FN is a false negative;

AC_{ALEX} is a accuracy of AlexNet original image classification;

AC_{LE} is a LeNet5 original image classification accuracy;

AC_{VGG} is a VGG16 original image classification accuracy;

I_{imp}^g is a set of images on the basis of GAN;

I_l^g is a training images on the basis of GAN;

I_t^g is a test images on the basis of GAN;

AC_{ALEX}^g is a AlexNet classification accuracy based on the extended sample;

AC_{LE}^g is a LeNet5 classification accuracy based on the extended sample;

AC_{VGG}^g is a VGG16 classification accuracy based on the extended sample;

AC_{DEV}^g is a classification accuracy for developed CNN architecture;

$G1$ is a Nottingham grading for breast cancer type number 1;

$G2$ is a Nottingham grading for breast cancer type number 2;

$G3$ is a Nottingham grading for breast cancer type number 3;

$\langle C \rangle$ is a set of convolution functions;

$\langle G \rangle$ is a set of pooling functions;

$\langle K \rangle$ is a set of activation functions;

V is a batch normalization operation;

$\langle B \rangle$ is a set of functions of a fully connected network;

c_i is an activation function in layer i ;

v_{c_i} is a parameters of activation function in layer i .

INTRODUCTION

Breast cancer in women is still a major medical and social problem that demands immediate attention. According to recent statistics, breast cancer continues to be the most common kind of malignant neoplasm in women. In 2020, the incidence rate of breast cancer in European Union nations was 13.3% of all new cases [1]. According to the American Cancer Society, breast cancer is also the most common among American women as of 2021 [2].

When analyzing morbidity data in Ukraine for the years 2021–2022, it is important to take into account that both periods were characterized by specific conditions in the country: the long course of the COVID-19 pandemic and military operations, which affected the work of both medical institutions and the cancer registration system. The report's data for 2022 cannot accurately represent the country's real onco-epidemiological process [3].

In 2021, women's oncological incidence was dominated by breast cancer, skin cancer, and neoplasms of the body and cervix, accounting for 54.5% of identified diseases. Deaths from breast cancer, colon cancer, esophageal cancer, and ovarian cancer accounted for the majority of the overall structure of mortality among women (48.8%) [3].

Biomedical images are widely used to diagnose diseases in oncology. Let us define biomedical images.

A biomedical image is a structural and functional image of human and animal organs, used to diagnose diseases and study the anatomy and physiology of the human body [4].

Cytological, histological, and immunohistochemical images are used to diagnose [4] oncological diseases.

Accurate cancer diagnosis involves histological analysis of materials. Histopathology is the microscopic examination of thin slices of damaged tissue. Histopathologists examine tissues and offer diagnostic information based on their findings.

Histological stains are frequently used to improve the capacity to visualize or differentiate microscopic structures. Chemical fixatives are used to protect tissues against destruction while also preserving the structure of cells and subcellular components.

Cytopathological, histopathological, and immunohistochemical examinations are used to learn about the features of the tumor, its degree of dissemination, and the best treatment option [5].

The diagnostic process normally begins with cytopathological investigation, which might reveal the existence of abnormalities in the cells. Following that, a histological analysis of the resected tumor is undertaken for a more in-depth investigation. Immunohistochemical tests can supplement these findings by giving further information about the tumor's biological characteristics.

Cytopathological investigations involve the examination of cells obtained during a tumor biopsy or puncture. They enable the assessment of anomalies in cellular structure and the identification of a suspected cancer process. Histopathological tests involve a detailed analysis of the removed tumor and adjacent tissues under a microscope. This helps determine the type of cancer, its aggressiveness, and its penetration into adjacent tissues.

Immunohistochemical investigations employ antibodies to identify specific proteins in tissues. They make it possible to more precisely detect the subtype of cancer and examine the presence of particular chemicals that might suggest prognosis and treatment alternatives. The molecular genetic subtype of the tumor is evaluated by immunohistochemistry expression of estrogen, progesterone, and oncoprotein HER-2/neu receptors, as well as detection of tumor cell proliferation using Ki-67.

The subject of research is the process of histological image generation and classification.

The object of research is deep neural networks used for image synthesis and classification.

The purpose of the research is to develop a method and software tool for breast cancer automatic diagnosis based on histological image analysis.

1 PROBLEM STATEMENT

Let the given set of original images is I_{inp} . Let us divide this set into two subsets: I_l and I_b , and $I_{inp} = I_l \cup I_b$.

In addition, the architecture of A_{CNN} CNN is given. The CNN architecture can be represented through multiple layers:

$$A_{CNN} = \{L_i, i = \overline{1, n}\}.$$

The classification accuracy is determined by the accuracy measure:

$$AC = \frac{TP + TN}{TP + TN + FP + FN}.$$

Classification accuracy depends on the number of layers and their parameters. The classification accuracy function is then presented in the following form:

$$AC = f(L, P_{CNN}).$$

For the known AlexNet, LeNet5 and VGG16 architectures, based on the original images, we get the following classification accuracies:

$$AC_{ALEX}, AC_{LE}, AC_{VGG}.$$

Then, we perform affine distortions on the input original images and generate I_{inp}^g on the basis of GAN. Let us divide these images into I_l^g and I_b^g ones, that is:

$$I_{inp}^g = I_l^g \cup I_b^g.$$

Based on the extended sample, we obtain the following classification accuracies for the known architectures AlexNet, LeNet5 and VGG16, respectively:

$$AC_{ALEX}^g, AC_{LE}^g, AC_{VGG}^g.$$

For the developed CNN architecture, we have the following classification accuracy: AC_{DEV}^g .

Classification accuracy of the developed architecture depends on the following parameters:

$$AC_{DEV}^g = f(L_{DEV}^g, P_{DEV}^g)$$

and

$$AC_{DEV}^g > \max(AC_{ALEX}^g, AC_{LE}^g, AC_{VGG}^g).$$

Therefore, it is necessary to find a CNN architecture with a certain number of layers and parameters to satisfy the following condition:

$$A_{CNN} = \arg \max_{L_{DEV}, O_{DEV}} AC_{DEV}^g(P_{DEV}^g, L_{DEV}, I_l^g).$$

2 REVIEW OF THE LITERATURE

The use of deep neural networks for cancer detection based on histological image processing has been widely addressed in research studies. For example, in the article [6], an ensemble of CNNs was investigated for malignancy identification using histological and cytological images. In [7], a multi-scale deep learning model was developed for breast cancer classification. In [8], the authors explored multi CNNs (VGG16, ResNet50, and DenseNet201) to detect mitotic cells.

Some researchers worked on the development of a deep learning-based approach for breast cancer image classification [9]. Others focused on the comparison of different CNN architectures for breast cancer classification [10]. Besides, segmentation and classification of cell nuclei in histology was discussed in [11]. The authors of [12] proposed a modified residual neural network-based method for breast cancer detection based on histological images. And the authors of [13] used an ensemble of deep multiscale CNN networks, namely CSAResnet and DAMCNN.

Additionally, a DSS for diagnosis of oncopathologies using histological images was developed by researchers in [14]. A CNN-based classifier for the automatic classification of breast cancer histological images was developed by the authors of [15]. A new DeepGrade histological class model was developed in [16]. The authors employed deep learning to analyze histological images. In [17], researchers combined two CNN architectures with the use of fractal geometry to improve histological image classification accuracy. In [18], the authors reviewed histological image classification methods for diagnosing breast cancer.

Thus, many researchers focused on the classification of breast cancer histological images. For example, the article [19] is devoted to the detection of breast cancer based on the CNN ensemble using histological images. In [20], the authors developed a method based on the combination of convolutional and recurrent deep neural networks for the classification of breast cancer histological images. In [21], the authors improved the DenseNet network and synthesized the CA-BreastNet model for the classification of breast cancer histological images. In [22], the authors analyzed modern algorithms for the automatic classification of breast cancer based on histological images. In [23], a heterogeneous ensemble based on CNN “DHE-Mit-Classifer” was developed. This ensemble was used to analyze mitotic nuclei in breast cancer histological images.

Besides, in [24], the authors analyzed the color and texture features of histological image slides. These features were used to count breast cancer mitosis. Article [25] was also devoted to the detection of mitosis in breast cancer. SVM, Naive Bayes, and Random Forest classifiers were used in the research study. The authors of the study [26] developed a DNN neural network based on the SSDHO optimization method to classify six classes of breast cancer images. In [27], the authors analyzed two methods of machine learning SVM and LR. The study also considered combinations of CNN + LR and CNN + SVM.

In the article [28], a review of machine learning methods for breast cancer diagnosis was carried out. In [29], the authors also used CNN to classify histological images of breast cancer. In the research study [30], the authors developed a 3-tier CNN model, which was used to classify breast cancer histological images. And researchers in [31] analyzed VGG16, VGG19, and ResNet50 networks for the classification of breast cancer histological images. A comparison of MLP and CNN networks was made in [32]. These networks have been used in breast cancer detection. The authors of [33] used a cascade deep learning network with U-Net architecture for segmentation and a ResNet network for breast cancer classification.

The authors of the article [34] developed a method of manual feature selection and applied a DNN to classify breast cancer. The authors of the article [35] analyzed the use of artificial intelligence methods in DSS for diagnos-

ing breast cancer. In the article [36], a method and a software tool for diagnosing skin diseases were developed.

These authors have experience in the development of methods, algorithms, and software tools for diagnosing oncological diseases based on the analysis of cytological, histological, and immunohistochemical images. For example, in works [36–42] technology and software systems for the analysis of biomedical images for diagnosis were developed. A number of publications [42–45] are devoted to the development of methods and algorithms for the analysis of biomedical images.

3 MATERIALS AND METHODS

The developed method of generating and classifying histological images consists of the following steps:

1. Formation of the initial dataset of histological images of three classes: G1, G2, and G3 based on affine distortions.
2. Experimental study of known architectures of neural networks for histological image classification and evaluation of classification accuracy on a given sample.
3. Expanding the histological image sample based on GAN networks.
4. Determining the improved neural network accuracy on the extended sample.
5. Designing new neural network architecture.
6. Comparison of accuracy of neural network architectures for histological images.

This section details steps 1, 3, and 5. Steps 2, 4, and 6 are described in section 5.

Let's create an initial data set. Diagnosticians determine the type of breast cancer using a histological examination. The Nottingham scale assesses the degree of difference in study findings. The following types of breast cancer are distinguished by the Nottingham grading: G1, G2, and G3. Breast tumor differentiation is determined by the degree of differentiation between pathological and normal cells, as well as the tumor cell growth rate. The authors of this article used cytological and histological images of breast cancer [46] from a private image database on the Zenodo platform. All images are anonymized, which complies with European standards [47].

The original sample contains images by class: G1 – 9 images, G2 – 100 images, and G3 – 76 images. This sample was expanded to 100 images in each class by applying affine distortions [48]

An example of histological images of different cancer types is shown in Fig. 1.

Based on the initial data set, we will form an extended data set using GAN.

The generator and discriminator architectures are based on the ResNet Block, borrowed from the ResNet [49].

The generator takes a noise vector with a Gaussian distribution of dimensions 1×100 as input and produces a $64 \times 64 \times 3$ image as output. The generator's architecture may be generally divided into three levels, as shown in Fig. 2.

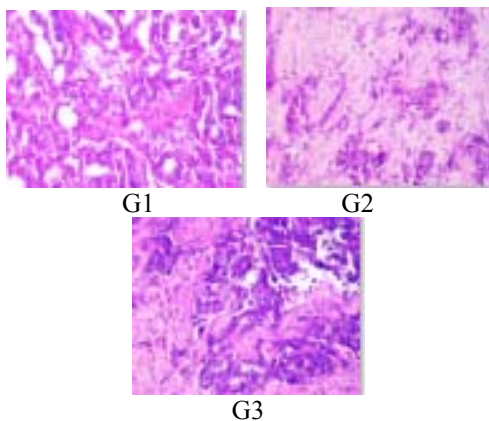


Figure 1 – Examples of histological images of type G1, G2, G3

The first layer is the pre-processing layer. In this layer, a one-dimensional input noise vector is read and sent to the linear layer for further conversion into a three-dimensional array.

The second layer is the main computing layer. It consists of four ResNet units. After each block, nearest-neighbor interpolation was used to enlarge the image by two times. In addition, following the third ResNet block, the Self-Attention block was used [50]. All convolutional layers use step 1. ReLU is also applied as the activation function.

The last layer is the output layer. Batch normalization, activation, another convolution layer, and the final Tanh activation function are all used here.

The discriminator is a convolutional neural network that takes an image of 64x64x4 pixels as input. The discriminator consists of five ResNet blocks. A Self-Attention mechanism is applied after the first block.

To reduce image dimensionality, the Average Pooling operation with a kernel size of 2x2 and a step of 2 is used

in each ResNet block. However, dimensionality reduction is not used in the final block. Convolution layers use step 1.

ReLU is also applied as an activation function.

The output of the discriminator is two linear layers. The first has 1280 neurons, whereas the second has only one. The architecture of the discriminator is shown in Fig. 3.

IS and FID metrics were used to evaluate the quality of synthesized images [51, 52].

The next step is to design the new CNN architecture.

The CNN architecture will be presented in set of layers:

$$A_{CNN} = \{L_i, i = \overline{1, n}\}.$$

Let us also define a set of CNN operations:

$$O_{CNN} = \{C, P, K, V, B\}.$$

Each CNN layer is a separate operation with parameters. For example, we detail a set of convolution operations:

$$C = \{c_1 \langle v_{c_1} \rangle, \dots, c_i \langle v_{c_i} \rangle, \dots, c_n \langle v_{c_n} \rangle\}.$$

Other conversion operations are determined similarly.

The architecture of the developed CNN is presented in Figure 4.

The CNN architecture consists of nine convolutional layers, four pooling layers, and one fully connected (linear) layer.

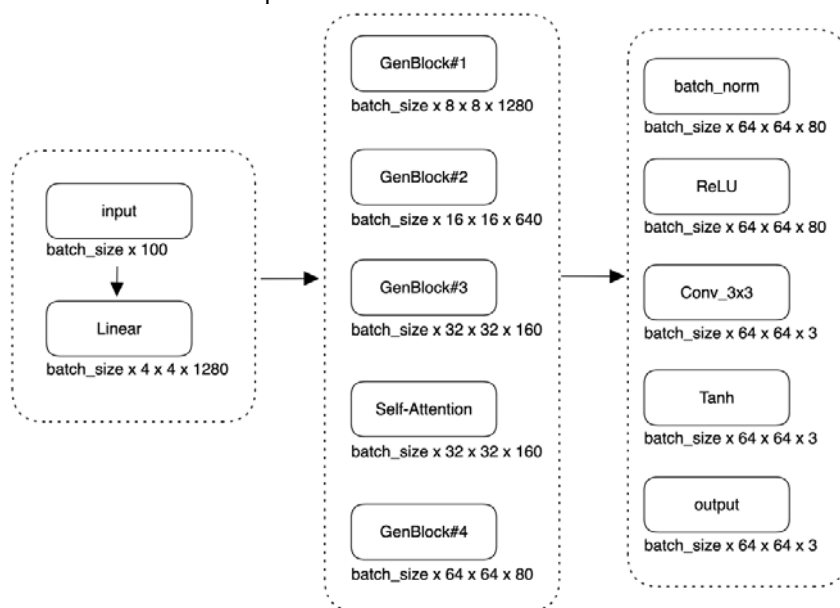


Figure 2 – Generator

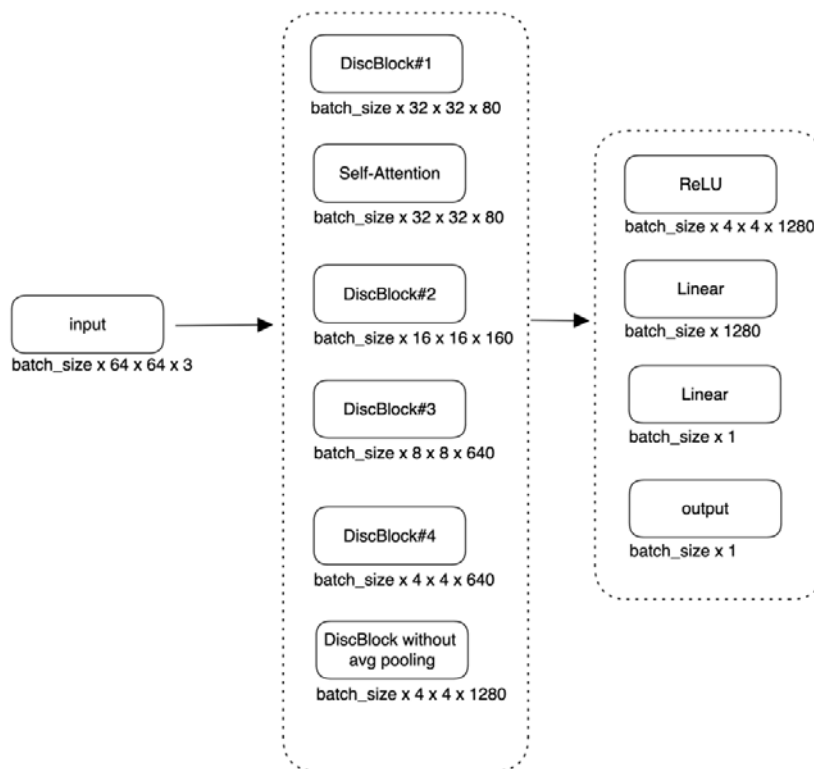


Figure 1 – Discriminator

All convolutional layers use step 1. Maximum pooling layers with a 3x3 kernel and step 2 are used to minimize the image's dimensionality.

An image with a size of 64x64x3 pixels is sent to the network input. The image is then passed to the first convolutional layer, which uses 64 feature maps with a 3x3 kernel.

The next two layers are convolutional blocks, which consist of successive layers of 3x3 kernel convolution, ReLU activation, and batch normalization. The first block employs 64 feature maps, whereas the second employs 128. A maximum pooling layer is then applied after these blocks. Accordingly, now the image size is 32x32x128.

Then there are two convolutional blocks with a maximum pooling layer at the end. However, here the convo-

lution layers use a 1x1 kernel and the same number of feature maps – 128. The size of the image after these layers is 16x16x128.

The next two convolution blocks are identical to the first two and also use the same number of feature maps – 128. At the end, a maximum pooling layer is applied. The image size is 8x8x128.

Next, there are two convolution blocks, identical to the previous ones, but the pooling layer is no longer applied after them. The size of the image remains the same – 8x8x128.

The output layer consists of sequential batch normalization layers, an adaptive average pooling layer with the number of output nodes 1, and a linear layer with 3 nodes.

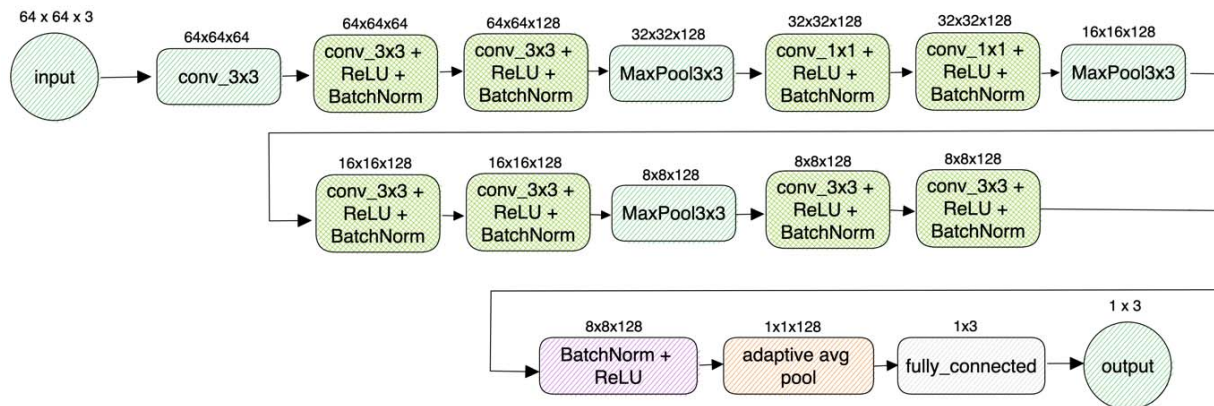


Figure 4 – Developed CNN Architecture

The CNN parameters are shown in Table 1. The formalized description of the developed CNN is as follows:

$$\begin{aligned}
 A_{CNN} = & \langle I_{inp_1} \rangle \langle 64 \times 64 \times 3 \rangle; \\
 & \langle c_2 \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle c_3 \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle k_3 \rangle \langle \text{ReLU} \rangle; \langle v_3 \rangle \langle \text{Batch Norm} \rangle; \\
 & \langle c_4 \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle k_4 \rangle \langle \text{ReLU} \rangle; \langle v_4 \rangle \langle \text{Batch Norm} \rangle; \\
 & \langle g_{max_5} \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 2 \rangle; \\
 & \langle c_6 \rangle \langle \text{kernel size} = 1 \times 1, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle k_6 \rangle \langle \text{ReLU} \rangle; \langle v_6 \rangle \langle \text{Batch Norm} \rangle; \\
 & \langle c_7 \rangle \langle \text{kernel size} = 1 \times 1, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle k_7 \rangle \langle \text{ReLU} \rangle; \langle v_7 \rangle \langle \text{Batch Norm} \rangle; \\
 & \langle g_{max_8} \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 2, \text{padding} = 1 \rangle; \\
 & \langle c_9 \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle k_9 \rangle \langle \text{ReLU} \rangle; \langle v_9 \rangle \langle \text{Batch Norm} \rangle; \\
 & \langle c_{10} \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle k_{10} \rangle \langle \text{ReLU} \rangle; \langle v_{10} \rangle \langle \text{Batch Norm} \rangle; \\
 & \langle g_{max_{11}} \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 2, \text{padding} = 1 \rangle; \\
 & \langle c_{12} \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle k_{12} \rangle \langle \text{ReLU} \rangle; \langle v_{12} \rangle \langle \text{Batch Norm} \rangle; \\
 & \langle c_{13} \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle k_{13} \rangle \langle \text{ReLU} \rangle; \langle v_{13} \rangle \langle \text{Batch Norm} \rangle; \\
 & \langle v_{14} \rangle \langle \text{Batch norm} \rangle; \langle k_{14} \rangle \langle \text{ReLU} \rangle; \\
 & \langle g_{max_{15}} \rangle \langle \text{kernel size} = 3 \times 3, \text{stride} = 1, \text{padding} = 1 \rangle; \\
 & \langle f_{16} \rangle \langle \text{clauses} = 3 \rangle.
 \end{aligned}$$

Table 1 – Developed CNN parameters

Layer number	Layer type	Options
1	Image input	64x64x3 image
2	Convolution	3x3 kernel convolution with stride 1 and same padding
3	Convolution	ReLU followed by 3x3 kernel convolution with stride 1 and padding 1 followed by Batch Normalization
4	Convolution	ReLU followed by 3x3 kernel convolution with stride 1 and padding 1 followed by Batch Normalization
5	MaxPooling	3x3 kernel max pooling with stride 2
6	Convolution	ReLU followed by 1x1 kernel convolution with stride 1 and padding 1 followed by Batch Normalization
7	Convolution	ReLU followed by 1x1 kernel convolution with stride 1 and padding 1 followed by Batch Normalization
8	MaxPooling	3x3 kernel max pooling with stride 2
9	Convolution	ReLU followed by 3x3 kernel convolution with stride 1 and padding 1 followed by Batch Normalization
10	Convolution	ReLU followed by 3x3 kernel convolution with stride 1 and padding 1 followed by Batch Normalization
11	MaxPooling	3x3 kernel max pooling with stride 2
12	Convolution	ReLU followed by 3x3 kernel convolution with stride 1 and padding 1 followed by Batch Normalization
13	Convolution	ReLU followed by 3x3 kernel convolution with stride 1 and padding 1 followed by Batch Normalization
14	BatchNorm	Batch normalization followed by ReLU
15	Adaptive Average Pooling	3x3 kernel adaptive average pooling with output nodes
16	Output	Linear layer with 3 output nodes

4 EXPERIMENTS

For computer experiments, special software has been developed. The software implementation is based on the reliability and scalability of Amazon Web Services cloud infrastructure, allowing efficient use of cloud computing resources. The program focuses on the development and deployment of a PyTorch-based CNN model for histological image classification. The infrastructure of the software is shown in Figure 5.

The developed software is made up of two different Python files: *train.py* and *predict.py*, each of which serves a specific purpose in the entire workflow.

In the *train.py* file, we describe a convolutional neural network model. We define the architecture by specifying parameters such as the number of convolution layers, filters, activation functions, and loss functions. This file also outlines the model training process.

To train the CNN model, we use AWS SageMaker and a cloud-based machine learning service. This service offers a scalable and controlled environment for efficient GPU training.

After the training process is completed, the model is downloaded and stored in the AWS S3 service. A URL is generated in AWS SageMaker to make the model available for use. The URL allows you to utilize the model in both the AWS cloud architecture and other web apps.

The software *predict.py* allows the use of the trained model to classify new data. This application is designed to be a CLI tool. To obtain classification results for the images, just call the file, specifying the path of the directory containing the images as the first parameter. The model analyses images and produces classification findings that may be further analyzed or integrated into other research procedures. The classification results are output to the *output.txt* file for convenience, with the data separated into two columns – the image file and the class.

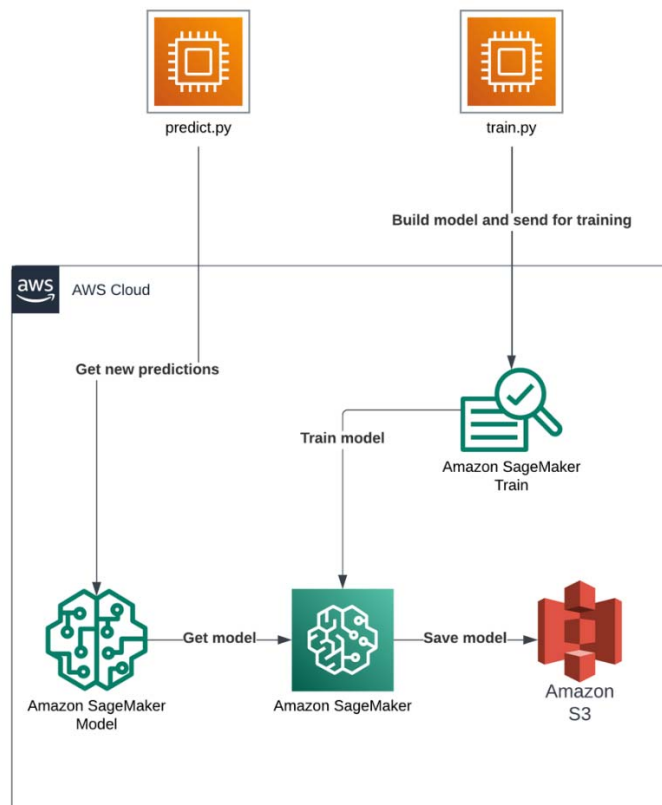


Figure 5 – Software infrastructure

5 RESULTS

Computer experiments were performed on the original and expanded samples for AlexNet, LeNet5, and VGG16. Experiments have also been carried out for the developed CNN architecture on the expanded sample.

AlexNet, LeNet5, and VGG16 original sample experiments.

100 images were used for each class. The training sample was divided into 80/20. The same training parameters were used for all networks: the Adam optimizer

(1e-4), the number of epochs was 100, the batch size was 100, and the loss function was CrossEntropyLoss.

Computer experiments were carried out using the three most prominent architectures: AlexNet, LeNet5, and VGG16. The results of the experiments are as follows: AlexNet classification accuracy was 74%, LeNet5 classification accuracy was 57%, and VGG16 classification accuracy was 70%.

ROC curves for these architectures are shown in Fig. 6–8.

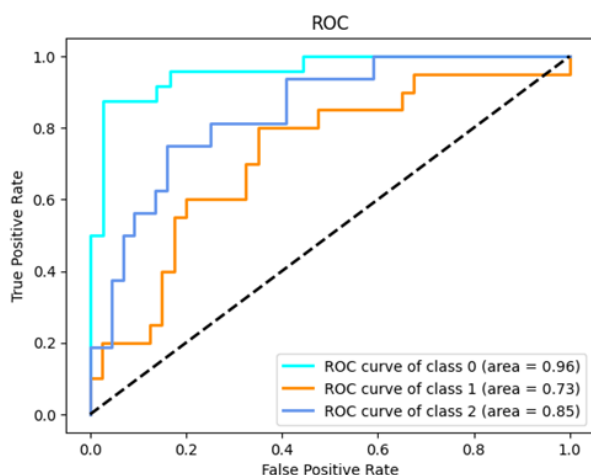


Figure 6 – ROC curve of AlexNet

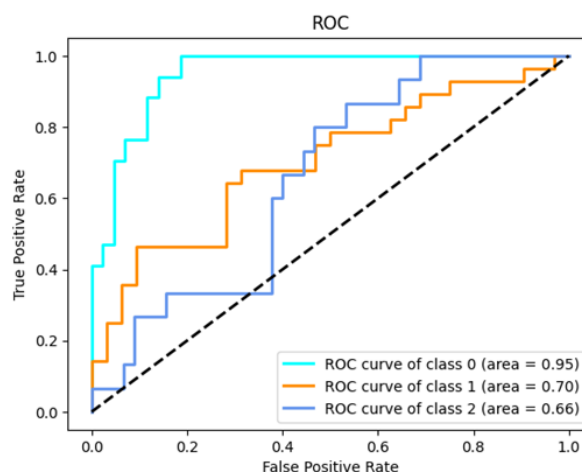


Figure 7 – ROC curve of LeNet5

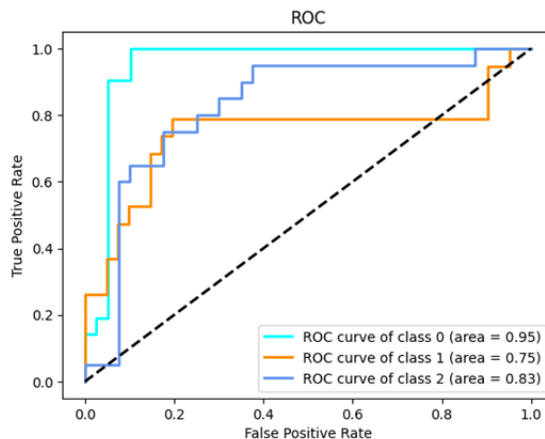


Figure 8 – ROC curve of VGG16

The original sample of images (100 images per class) was used as the training dataset. The Python programming language and the Pytorch framework were used to write the code. The GAN network training parameters were as follows: Adam optimizer, generator learning rate – $1e-4$, discriminator – $4e-4$, number of epochs – 100000, batch size – 96, and loss function – HingeLoss.

A virtual machine with the following configuration was used to run the experiments: 16 GB RAM, 10 vCPU x 2.2 GHz, Nvidia Tesla V100 GPU 16 GB (13.2 TFLOPS). The GAN network was trained for 11 hours. As a result of the experiments, the values of the metrics for the network were as follows: IS – 3.024, FID – 68.

Examples of synthesized images are shown in Fig. 9.

AlexNet, LeNet5, and VGG16 extended sample experiments. The training sample was increased to 3000 images per class using a generative-competitive network. The sample has 9,000 images in total. The training sample was 80/20 divided. Three architectures were tested on computers: AlexNet, LeNet5, and VGG16. The same training parameters were used for all networks – the Adam optimizer ($1e-4$), the number of epochs was 10, and the batch size was 100. The results of the experiments

are as follows: AlexNet classification accuracy was 85%, LeNet5 – 90%, and VGG16 – 91%.

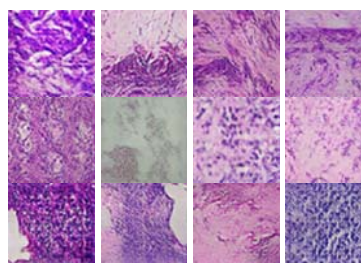


Figure 9 – Examples of synthesized images

The ROC curves for these architectures are shown in Fig. 10–12.

Experiments on extended samples for the developed CNN. The number of images per class and training parameters was similar to those for experiments with classical architectures. As a result of the experiments, the classification accuracy was 96%. The ROC curve is shown in Fig. 13.

A comparison of neural network architectures is shown in Table 2.

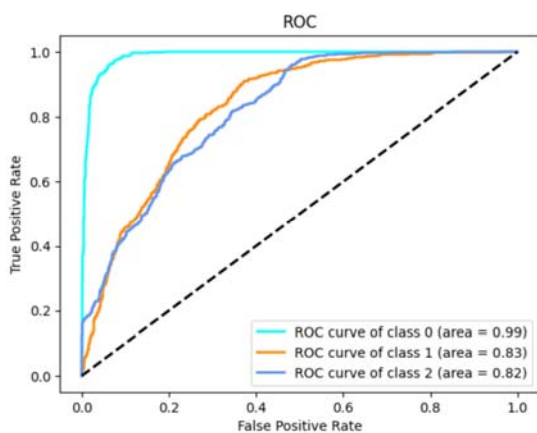


Figure 10 – ROC curve of AlexNet

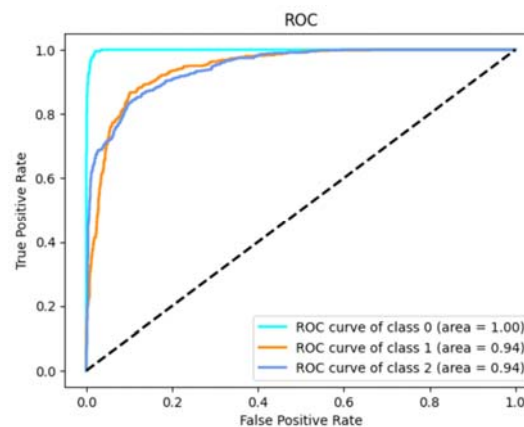


Figure 11 – ROC curve of LeNet5

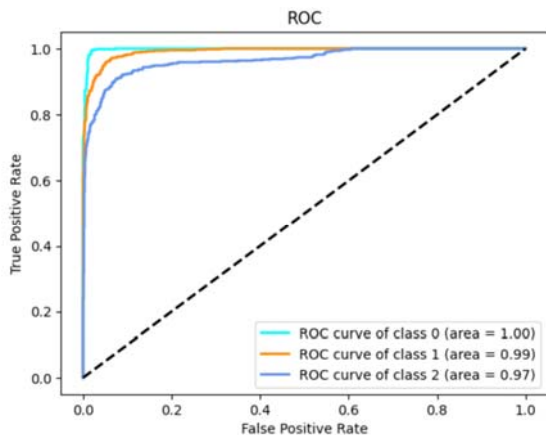


Figure 12 – ROC curve of VGG16

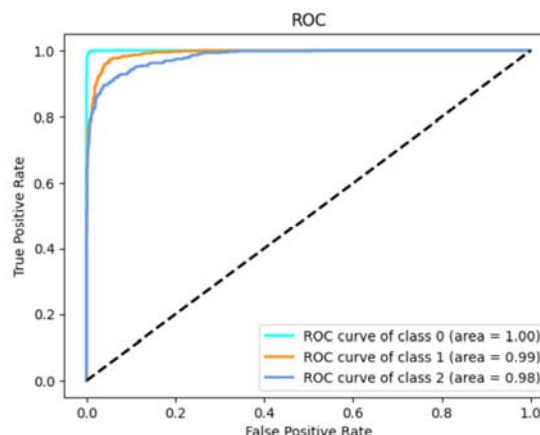


Figure 13 – ROC curve of the developed CNN

Table 2 – ANN comparison

Article	ANN	Accuracy %
[53]	AlexNet	85
[54]	LeNet5	90
[55]	VGG16	91
Proposed CNN	Custom CNN	96

6 DISCUSSION

The limitation of using histological images to diagnose different forms of breast cancer is the limited initial sample size.

According to studies of CNN applications, every CNN model performs better on larger datasets. Large datasets offer a more diversified set of samples to train on, allowing the model to generalize the data more effectively. CNNs learn a broader range of characteristics and patterns when trained on a huge dataset, making them more resistant to variations in the data. Overtraining is more likely with smaller datasets. Simultaneously, the model learns to recall rather than generalize training data. Because of the higher variety in the data, large datasets make retraining a model challenging. The model has the capacity to handle large datasets. In large datasets, the model has the ability to learn complex and hierarchical features. This is especially important for deep CNNs like VGG16, which have many layers. Large datasets allow these models to extract meaningful features at different levels of abstraction. With a large dataset, the optimization process (e.g., gradient descent) usually works more efficiently.

GAN was used to expand the initial sample.

The initial original sample contained the following number of images per class: G1 – 9 images, G2 – 100 images, and G3 – 76 images. Based on affine distortions, the sample is expanded to 100 images in each class. 7200 artificial images were generated using GAN. The GAN network has been trained for 11 hours. At the same time, it was possible to obtain the value of IS metrics – 3.024, and FID – 42.552.

For the three architectures of AlexNet, LeNet5, and VGG16, the following results were obtained for the initial samples: AlexNet classification accuracy was 74%, LeNet5 classification accuracy was 57%, and VGG16 classification accuracy was 70%. On the expanded sample, the

results of the experiments are as follows: AlexNet classification accuracy was 85%, LeNet5 – 90%, and VGG16 – 91%.

The developed CNN architecture showed an accuracy of 96% on extended samples.

The developed program can be used in CAD. This will allow accurate histological image classification when making a diagnosis.

CONCLUSIONS

The problem of histological image classification to identify different breast cancer types was examined in this article. The authors used CNN in the research study. Low classification accuracy was revealed by analysis of known CNN based on original samples. The initial GAN-based image sample was enlarged for this purpose. On extended samples, AlexNet, LeNet5, and VGG16 showed a significant increase in classification accuracy.

In comparison to the well-known AlexNet, LeNet5, and VGG16 architectures, the newly developed CNN architecture demonstrated higher classification accuracy.

Scalability, dependability, and cost-effectiveness were provided for model training and classification using AWS SageMaker and AWS S3. Separating the training and prediction steps into two files (train.py and predict.py) ensured modularity and ease of maintenance.

With a user-friendly and simple interface, the CLI tool (predict.py) simplified the prediction process.

The cloud architecture ensures that the trained model is stored in a safe and accessible place and can be used for prediction in any application.

The architecture of the software system allows efficient training of models and classification of new data in a cloud environment, making it suitable for scalable machine learning applications.

The scientific novelty of the article is the development of a method of small initial sample image generation and classification.

The practical value of the article is the development of software for image generation and classification, which can be used as a separate module in CAD.

Prospects for further research are the investigation of methods of automatic design of convolutional networks and generative-competitive networks and the development of CADs for automatic diagnosis in oncology.

REFERENCES

1. Breast cancer burden in EU-27. [Electronic resource]. Access mode: https://ecis.jrc.ec.europa.eu/pdf/Breast_cancer_factsheet-Oct_2020.pdf
2. Cancer Facts for Women. American Cancer Society. [Electronic resource]. Access mode: <https://www.cancer.org/cancer/risk-prevention/understanding-cancer-risk/cancer-facts/cancer-facts-for-women.html>
3. Cancer in Ukraine 2021–2022: Incidence, mortality, prevalence and other relevant statistics. Bulletin of the National Cancer Registry of Ukraine № 24, 2021–2022. [Electronic resource]. Access mode: http://www.ncru.inf.ua/publications/BULL_24/PDF_E/bull_eng_24.pdf
4. Berezsky O., Pitsun O., Verbovy S. et al. Computer diagnostic tools based on biomedical image analysis 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), 21–25 Feb. 2017. Lviv, Ukraine, 2017, pp. 388–391.
5. Aswathy M. A., Jagannath M. Detection of breast cancer on digital histopathology images: Present status and future possibilities, *Informatics in Medicine Unlocked*, 2017, Vol. 8, pp. 74–79. DOI: 10.1016/j.imu.2016.11.001
6. Dey S., Mitra S., Chakraborty S. et al. GC-EnC: A Copula based ensemble of CNNs for malignancy identification in breast histopathology and cytology images, *Computers in Biology and Medicine*, 2023, Vol. 152, P. 106329. DOI: 10.1016/j.combiomed.2022.106329
7. Chattopadhyay S., Dey A., Singh P. K. et al. MTRRE-Net: A deep learning model for detection of breast cancer from histopathological images, *Computers in Biology and Medicine*, 2022, Vol. 150, P. 106155. DOI: 10.1016/j.combiomed.2022.106155
8. Shihabuddin A. R., Sabeena K. S. Multi CNN based automatic detection of mitotic nuclei in breast histopathological images, *Computers in Biology and Medicine*, 2023, Vol. 158, P. 106815. DOI: 10.1016/j.combiomed.2023.106815
9. Vo D. M., Nguyen N.-Q., Lee S.-W. Classification of breast cancer histology images using incremental boosting convolution networks, *Information Sciences*, 2019, Vol. 482, pp. 123–138. DOI: 10.1016/j.ins.2018.12.089
10. Saikia A. R., Bora K., Mahanta L. B. et al. Comparative assessment of CNN architectures for classification of breast FNAC images, *Tissue Cell*, 2019, Vol. 57, pp. 8–14. DOI: 10.1016/j.tice.2019.02.001
11. Dogar G. M., Shahzad M., Fraz M. M. Attention augmented distance regression and classification network for nuclei instance segmentation and type classification in histology images, *Biomedical Signal Processing and Control*, 2023, Vol. 79, P. 104199. DOI: 10.1016/j.bspc.2022.104199
12. Gupta V., Vasudev M., Doegar A. et al. Breast cancer detection from histopathology images using modified residual neural networks, *Biocybernetics and biomedical engineering*, 2021, Vol. 41, pp. 1272–1287. DOI: 10.1016/j.bbe.2021.08.011
13. Karthik R., Menaka R., Siddharth M. V. Classification of breast cancer from histopathology images using an ensemble of deep multiscale networks, *Biocybernetics and biomedical engineering*, 2022, Vol. 42, pp. 963–976. DOI: 10.1016/j.bbe.2022.07.006
14. Dovbysh A., Shelehov I., Romaniuk A. et al. Decision-making support system for diagnosis of oncopathologies by histological images, *Journal of Pathology Informatics*, 2023, Vol. 14, P. 100193. DOI: 10.1016/j.jpi.2023.100193
15. Roy K., Banik D., Bhattacharjee D. et al. Patch-based system for Classification of Breast Histology images using deep learning, *Computerized Medical Imaging and Graphics*, 2019, Vol. 71, pp. 90–103. DOI: 10.1016/j.compmedimag.2018.11.003
16. Wang Y., Acs B., Robertson S. et al. Improved breast cancer histological grading using deep learning, *Annals of oncology*, 2021, Vol. 33, № 1, pp. 89–98. DOI: 10.1016/j.annonc.2021.09.007
17. Roberto G. F., Lumini A., Neves L. A. et al. Fractal Neural Network: A new ensemble of fractal geometry and convolutional neural networks for the classification of histology images, *Expert Systems With Applications*, 2021, Vol. 166, P. 114103. DOI: 10.1016/j.eswa.2020.114103
18. Hussein E. H., Emama M. M., Ali A. A. Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review, *Expert Systems with Applications*, 2021, Vol. 167, P. 114161. DOI: 10.1016/j.eswa.2020.114161
19. Majumdar S., Pramanik P., Ram S. Gamma function based ensemble of CNN models for breast cancer detection in histopathology images, *Expert Systems With Applications*, 2023, Vol. 213, P. 119022. DOI: 10.1016/j.eswa.2022.119022
20. Yana R., Ren F., Wangb Z. et al. Breast cancer histopathological image classification using a hybrid deep neural network, *Methods*, 2020, Vol. 173, pp. 52–60. DOI: 10.1016/j.ymeth.2019.06.014
21. Dianzhi Y., Jianwu L., Tengbao C. et al. SECS: An effective CNN joint construction strategy for breast cancer histopathological image classification, *Journal of King Saud University – Computer and Information Sciences*, 2023, Vol. 35, pp. 810–820. DOI: 10.1016/j.jksuci.2023.01.017
22. Aresta G., Araújo T., Kwok S. et al. BACH: Grand challenge on breast cancer histology images, *Medical Image Analysis*, 2019, Vol. 56, pp. 122–139. DOI: 10.1016/j.media.2019.05.010
23. Sohail A., Khan A., Nisar H. et al. Mitotic nuclei analysis in breast cancer histopathology images using deep ensemble classifier, *Medical Image Analysis*, 2021, Vol. 72, P. 102121. DOI: 10.1016/j.media.2021.102121
24. Wahab N., Khan A. Multifaceted fused-CNN based scoring of breast cancer whole-slide histopathology images, *Applied Soft Computing Journal*, 2020, Vol. 97, P. 106808. DOI: 10.1016/j.asoc.2020.106808
25. Rehman M. U., Akhtar S., Zakwan M. et al. Novel architecture with selected feature vector for effective classification of mitotic and non-mitotic cells in breast cancer histology

- images, *Biomedical Signal Processing and Control*, 2022, Vol. 71, P. 103212. DOI:10.1016/j.bspc.2021.103212
26. Bhausaeheb D. P., Kashyap K. L. Shuffled shepherd deer hunting optimization based deep neural network for breast cancer classification using breast histopathology images, *Biomedical Signal Processing and Control*, 2023, Vol. 83, P. 104570. DOI: 10.1016/j.bspc.2023.104570
27. Gupta K., Chawla N. Analysis of Histopathological Images for Prediction of Breast Cancer Using Traditional Classifiers with Pre-Trained CNN, *Procedia Computer Science*, 2020, №167, pp. 878–889. DOI: 10.1016/j.procs.2020.03.427
28. Yadav R. K., Singh P., Kashtriya P. Diagnosis of Breast Cancer using Machine Learning Techniques – A Survey, *Procedia Computer Science*, 2023, Vol. 218, pp. 1434–1443. DOI: 10.1016/j.procs.2023.01.122
29. Dabeer S., Khan M. M., Islam S. Cancer diagnosis in histopathological image: CNN based approach, *Informatics in Medicine Unlocked*, 2019, Vol. 16, P. 100231. DOI: 10.1016/j.imu.2019.100231
30. Kate V., Shukla P. A 3 Tier CNN model with deep discriminative feature extraction for discovering malignant growth in multi-scale histopathology images, *Informatics in Medicine Unlocked*, 2021, Vol. 24, P. 100616. DOI: 10.1016/j.imu.2021.100616
31. Shalla S., Mehra R. Breast cancer histology images classification: Training from scratch or transfer learning?, *ICT Express*, 2018, Vol. 4, pp. 247–254. DOI: 10.1016/j.ict.2018.10.007
32. Desai M., Shah M. An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN), *Clinical eHealth*, 2021, Vol. 4, pp. 1–11. DOI: 10.1016/j.ceh.2020.11.002
33. Asadi B., Memon Q. Efficient breast cancer detection via cascade deep learning network, *International Journal of Intelligent Networks*, 2023, Vol. 4, pp. 46–52. DOI: 10.1016/j.ijin.2023.02.001
34. Joseph A. A., Abdullahi M., Junaidu S. B. et al. Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer), *Intelligent Systems with Applications*, 2022, Vol. 14, P. 200066. DOI: 10.1016/j.iswa.2022.200066
35. Chaudhury S., Sau K. A BERT encoding with Recurrent Neural Network and Long-Short Term Memory for breast cancer image classification, *Decision Analytics Journal*, 2023, Vol. 6, P. 100177. DOI: 10.1016/j.dajour.2023.100177
36. Lovkin V. M., Subbotin S. A., Oliinyk A. O. et al. Method and software component model for skin disease diagnosis, *Radio Electronics, Computer Science, Control*, 2023, № 1, P. 40. DOI: 10.15588/1607-3274-2023-1-4
37. Berezsky O. M., Berezka K. M., Melnyk G. M. et al. Design of computer systems for biomedical image analysis, *Proceedings of the X th International Conference «The Experience of Designing and Application of CAD Systems in Microelectronics» CADSM 2009, 24–28 February 2009, Lviv-Polyana, Ukraine*. Lviv, Publishing House Vezha&CoC, P. 186–191.
38. Berezsky O., Datsko T., Verbovy S. The intelligent system for diagnosing breast cancers based on image analysis, *Proceedings of Information Technologies in Innovation Business (ITIB)*. Kharkiv, Ukraine, 7–9 October, 2015, pp. 27–30. DOI: 10.1109/ITIB.2015.7355067.
39. Berezsky O., Verbovy S., Dubchak L. et al. Fuzzy system diagnosing of precancerous and cancerous conditions of the breast cancer, *Proceedings of the XIth International Scientific and Technical Conference Computer Sciences and Information Technologies (CSIT'2016)*. Lviv, Ukraine, 6–10 September, 2016, pp. 200–203. DOI: 10.1109/STC-CSIT.2016.7589906.
40. Berezsky O., Dubchak L., Pitsun O. Access Distribution in Automated Microscopy System / O. Berezsky, // *Proceedings of the 14 th International Conference «The Experience of Designing and Application of CAD Systems in Microelectronics» CADSM 2017, 21–25 February 2017, Lviv, Ukraine*. Lviv, 2017, pp. 241–243. DOI: 10.1109/CADSM.2017.7916125
41. Berezsky O., Verbovy S., Pitsun O. Hybrid Intelligent information technology for biomedical image processing, *Proceedings of the IEEE International Conference “Computer Science and Information Technologies” CSIT'2018, Lviv, Ukraine, 11–14 September, 2018*. Lviv, 2018, pp. 420–423. DOI: 10.1109/STC-CSIT.2018.8526711.
42. Berezsky O., Pitsun O., Batryn N., et al. Modern automated microscopy systems in oncology, *Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine, Lviv, Ukraine, 28–30 november 2018*. Lviv, 2018, pp. 311–325.
43. Berezsky O. M., Pitsun O. Y. Evaluation methods of image segmentation quality, *Radio Electronics, Computer Science, Control*, 2018, Vol. 1, pp. 119–128. DOI:10.15588/1607-3274-2018-1-14
44. Berezsky O., Dubchak L., Batryn N., et al. Fuzzy System For Breast Disease Diagnosing Based On Image Analysis, *Proceedings of the II International Workshop Informatics & Data-Driven Medicine (IDDM 2019)*. Lviv, Ukraine, 11–13 November, 2019. Lviv, 2019, pp. 69–83.
45. Berezsky O., Pitsun O., Melnyk G. et al. An Approach toward Automatic Specifics Diagnosis of Breast Cancer Based on an Immunohistochemical Image, *Journal of Imaging*, 2023, Vol. 9, № 1, P. 12. DOI: 10.3390/jimaging9010012
46. Berezsky O., Datsko T., Melnyk G. Cytological and histological images of breast cancer, [Electronic resource]. Access mode: <https://zenodo.org/records/7890874>. DOI: 10.5281/zenodo.7890874
47. Intresoft Consulting. General Data Protection Regulation (GDPR). Recital 26. [Electronic resource]. Access mode: <https://gdpr-info.eu/recitals/no-26/>
48. Liashchynskiy P. Rudi. Lightweight image converter and dataset augmentor, [Electronic resource]. Access mode: <https://github.com/liashchynskiy/rudi>
49. He K., Zhang X., Ren S. et al. Deep Residual Learning for Image Recognition, 2015. [Electronic resource]. Access mode: <https://arxiv.org/pdf/1512.03385.pdf>
50. Lim J. H., Ye J. C. Geometric GAN, 2017. [Electronic resource]. Access mode: <https://arxiv.org/pdf/1705.02894v2.pdf>
51. Salimans T., Goodfellow I., Zaremba W. et al. Improved techniques for training GANs, *Advances in neural information processing systems: Annual Conference on Neural Information Processing Systems 2016, December 5–10, 2016, Barcelona*. Spain, Barcelona, 2016, pp. 2226–2234 DOI: 10.48550/arXiv.1606.03498
52. Heusel M., Ramsauer H., Unterthiner T. et al. GANs trained by a two time-scale update rule converge to a local nash equilibrium, *Proceedings of the 31st International Conference on Neural Information Processing Systems*. NY,

- United States, Curran Associates Inc., 2017, pp. 6629–6640. Access mode: <https://dl.acm.org/doi/10.5555/3295222.3295408>
53. Nawaz W., Ahmed S., Tahir A. et al. Classification Of Breast Cancer Histology Images Using ALEXNET. In: Campilho, A., Karray, F., ter Haar Romeny, B. (eds) Image Analysis and Recognition. ICIAR 2018, *Lecture Notes in Computer Science*. Springer, Cham, 2018, Vol. 10882, pp. 869–876. DOI: 10.1007/978–3-319-93000-8_99
54. Verdhan V. Image Classification Using LeNet. In: Computer Vision Using Deep Learning. Apress, Berkeley, CA., 2021, pp. 67–101. DOI: 10.1007/978-1-4842-6616-8_3
55. Tammina S. Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images, *International Journal of Scientific and Research Publications*, 2019, Vol. 9, Issue 10, pp. 143–150. DOI: 10.29322/IJSRP.9.10.2019.p9420

Received 08.10.2023.

Accepted 26.11.2023.

УДК 004.932

МЕТОД І ПРОГРАМНИЙ ЗАСІБ ГЕНЕРУВАННЯ І КЛАСИФІКАЦІЇ БІОМЕДИЧНИХ ЗОБРАЖЕНЬ НА ОСНОВІ ГЛИБОКИХ МЕРЕЖ ІЗ МАЛОЮ ВИБІРКОЮ

Березький О. М. – д-р техн. наук, професор, професор кафедри комп’ютерної інженерії Західноукраїнського національного університету, Тернопіль, Україна.

Лящинський П. Б. – аспірант кафедри комп’ютерної інженерії Західноукраїнського національного університету, Тернопіль, Україна.

Піцун О. Й. – канд. техн. наук, доцент, доцент кафедри комп’ютерної інженерії Західноукраїнського національного університету, Тернопіль, Україна.

Мельник Г. М. – канд. техн. наук, доцент кафедри комп’ютерної інженерії Західноукраїнського національного університету, Тернопіль, Україна.

АНОТАЦІЯ

Актуальність. У статті досліджено проблему класифікації гістологічних зображень раку молочної залози. Актуальність проблеми пояснюється широкою розповсюдженістю хвороби – раку молочної залози. Автоматизація процесу постановки діагнозу дає можливість зменшити час і виключити суб’єктивний фактор. Результати дослідження можуть бути використані в CAD в онкології.

Мета роботи – розробка методу та програмного засобу генерування і класифікації гістологічних зображень на основі глибоких нейронних мереж для підвищення точності класифікації.

Метод. У роботі розроблено метод генерування і класифікації гістологічних зображень. Цей метод базується на основі використання CNN і GAN. Для підвищення точності класифікації початкову вибірку зображень розширено за допомогою GAN.

Результати. Комп’ютерне дослідження розроблено методу генерування і класифікації зображень проводилося на основі dataset, який знаходиться на платформі Zenodo. Зображення отримано на основі світлової мікроскопії. Dataset містить три класи G1, G2, G3 гістологічних зображень раку молочної залози. На основі розробленого методу отримано точність класифікації зображень 96%. Це краща точність класифікації порівняно з існуючим моделями типу AlexNet, LeNet5 і VGG16. Програмний модуль може бути інтегрований у CAD.

Висновки. Розроблений метод генерування і класифікації зображень є основою програмного модуля. Програмний модуль може бути інтегрований у CAD.

КЛЮЧОВІ СЛОВА: система автоматизованої діагностики, рак молочної залози, глибокі нейронні мережі, генеративні змагальні мережі, згорткові нейронні мережі.

ЛІТЕРАТУРА

1. Breast cancer burden in EU-27. [Electronic resource]. – Access mode: https://ecis.jrc.ec.europa.eu/pdf/Breast_cancer_factsheet-Oct_2020.pdf
2. Cancer Facts for Women. American Cancer Society. [Electronic resource]. – Access mode: <https://www.cancer.org/cancer/risk-prevention/understanding-cancer-risk/cancer-facts/cancer-facts-for-women.html>
3. Cancer in Ukraine 2021–2022: Incidence, mortality, prevalence and other relevant statistics. Bulletin of the National Cancer Registry of Ukraine № 24, 2021–2022. [Electronic resource]. – Access mode: http://www.ncru.inf.ua/publications/BULL_24/PDF_E/bull_eng_24.pdf
4. Computer diagnostic tools based on biomedical image analysis / [O. Berezsky, O. Pitsun, S. Verbovy et al.] // 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM), 21–25 Feb. 2017 – Lviv, Ukraine, 2017. – P. 388–391.
5. Aswathy M. A. Detection of breast cancer on digital histopathology images: Present status and future possibilities / M. A. Aswathy, M. Jagannath // Informatics in Medicine Unlocked. – 2017. – Vol. 8. – P. 74–79. DOI: 10.1016/j.imu.2016.11.001
6. GC-EnC: A Copula based ensemble of CNNs for malignancy identification in breast histopathology and cytology images / [S. Dey, S. Mitra, S. Chakraborty et al.] // Computers in Biology and Medicine. – 2023. – Vol. 152. – P. 106329. DOI: 10.1016/j.combiomed.2022.106329
7. MTRRE-Net: A deep learning model for detection of breast cancer from histopathological images / [S. Chattopadhyay, A. Dey, P. K. Singh et al.] // Computers in Biology and Medicine. – 2022. – Vol. 150. – P. 106155. DOI: 10.1016/j.combiomed.2022.106155

8. Shihabuddin A. R. Multi CNN based automatic detection of mitotic nuclei in breast histopathological images / A. R. Shihabuddin, K. S. Sabeena // *Computers in Biology and Medicine*. – 2023. – Vol. 158. – P. 106815. DOI: 10.1016/j.combiomed.2023.106815
9. Vo D. M. Classification of breast cancer histology images using incremental boosting convolution networks / D. M. Vo, N.-Q. Nguyen, S.-W. Lee // *Information Sciences*. – 2019. – Vol. 482. – P. 123–138. DOI: 10.1016/j.ins.2018.12.089
10. Comparative assessment of CNN architectures for classification of breast FNAC images / [A. R. Saikia, K. Bora, L. B. Mahanta et al.] // *Tissue Cell*. – 2019. – Vol. 57. – P. 8–14. DOI: 10.1016/j.tice.2019.02.001
11. Dogar G. M. Attention augmented distance regression and classification network for nuclei instance segmentation and type classification in histology images / G. M. Dogar, M. Shahzad, M. M. Fraz // *Biomedical Signal Processing and Control*. – 2023. – Vol. 79. – P. 104199. DOI: 10.1016/j.bspc.2022.104199
12. Breast cancer detection from histopathology images using modified residual neural networks/ [V. Gupta, M. Vasudev, A. Doegar et al.] // *Biocybernetics and biomedical engineering*. – 2021. – Vol. 41. – P. 1272–1287. DOI: 10.1016/j.bbe.2021.08.011
13. Karthik R. Classification of breast cancer from histopathology images using an ensemble of deep multiscale networks / R. Karthik, R. Menaka, M. V. Siddharth // *Biocybernetics and biomedical engineering*. – 2022. – Vol. 42. – P. 963–976. DOI: 10.1016/j.bbe.2022.07.006
14. Decision-making support system for diagnosis of oncopathologies by histological images. / [A. Dovbysh, I. Shelehov, A. Romaniuk et al.] // *Journal of Pathology Informatics*. – 2023. – Vol. 14. – P. 100193. DOI: 10.1016/j.jpi.2023.100193
15. Patch-based system for Classification of Breast Histology images using deep learning / [K. Roy, D. Banik, D. Bhattacharjee et al.] // *Computerized Medical Imaging and Graphics*. – 2019. – Vol. 71. – P. 90–103. DOI: 10.1016/j.compmedimag.2018.11.003
16. Improved breast cancer histological grading using deep learning / [Y. Wang, B. Acs, S. Robertson et al.] // *Annals of oncology*. – 2021. – Vol. 33, № 1. – P. 89–98. DOI: 10.1016/j.annonc.2021.09.007
17. Fractal Neural Network: A new ensemble of fractal geometry and convolutional neural networks for the classification of histology images / [Guilherme Freire Roberto, Alessandra Lumini, Leandro Alves Neves et al.] // *Expert Systems With Applications*. – 2021. – Vol. 166. – P. 114103. DOI: 10.1016/j.eswa.2020.114103
18. Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review / [Essam H. Houssein, Marwa M. Emama, Abdelmgeid A. Ali] // *Expert Systems with Applications*. – 2021. – Vol. 167. – P. 114161. DOI: 10.1016/j.eswa.2020.114161
19. Majumdar S. Gamma function based ensemble of CNN models for breast cancer detection in histopathology images / Samridha Majumdar, Payel Pramanik, Ram Sarkar // *Expert Systems With Applications*. – 2023. – Vol. 213. – P. 119022. DOI: 10.1016/j.eswa.2022.119022
20. Breast cancer histopathological image classification using a hybrid deep neural network / [Rui Yana, Fei Ren, Zihao Wangb et al.] // *Methods*. – 2020. – Vol. 173. – P. 52–60. DOI: 10.1016/j.ymeth.2019.06.014
21. SECS: An effective CNN joint construction strategy for breast cancer histopathological image classification / [Dianzhi Yu, Jianwu Lin, Tengbao Cao et al.] // *Journal of King Saud University – Computer and Information Sciences*. – 2023. – Vol. 35. – P. 810–820. DOI: 10.1016/j.jksuci.2023.01.017
22. BACH: Grand challenge on breast cancer histology images / [G. Aresta, T. Araújo, S. Kwok et al.] // *Medical Image Analysis*. – 2019. – Vol. 56. – P. 122–139. DOI: 10.1016/j.media.2019.05.010
23. Mitotic nuclei analysis in breast cancer histopathology images using deep ensemble classifier / [A. Sohail, A. Khan, H. Nisar et al.] // *Medical Image Analysis*. – 2021. – Vol. 72. – P. 102121. DOI: 10.1016/j.media.2021.102121
24. Wahab N. Multifaceted fused-CNN based scoring of breast cancer whole-slide histopathology images / Noorul Wahab, Asifullah Khan // *Applied Soft Computing Journal*. – 2020. – Vol. 97. – P. 106808. DOI: 10.1016/j.asoc.2020.106808
25. Novel architecture with selected feature vector for effective classification of mitotic and non-mitotic cells in breast cancer histology images/ [M. U. Rehman, S. Akhtar, M. Zakwan et al.] // *Biomedical Signal Processing and Control*. – 2022. – Vol. 71. – P. 103212. DOI: 10.1016/j.bspc.2021.103212
26. Bhausaheb D. P. Shuffled shepherd deer hunting optimization based deep neural network for breast cancer classification using breast histopathology images / D. P. Bhausaheb, K. L. Kashyap // *Biomedical Signal Processing and Control*. – 2023. – Vol. 83. – P. 104570. DOI: 10.1016/j.bspc.2023.104570
27. Gupta K. Analysis of Histopathological Images for Prediction of Breast Cancer Using Traditional Classifiers with Pre-Trained CNN / K. Gupta, N. Chawla // *Procedia Computer Science*. – 2020. – №167. – P. 878–889. DOI: 10.1016/j.procs.2020.03.427
28. Yadav R. K. Diagnosis of Breast Cancer using Machine Learning Techniques – A Survey / R. K. Yadav, P. Singh, P. Kashtriyia // *Procedia Computer Science*. – 2023. – Vol. 218. – P. 1434–1443. DOI: 10.1016/j.procs.2023.01.122
29. Dabeer S. Cancer diagnosis in histopathological image: CNN based approach / S. Dabeer, M. M. Khan, S. Islam // *Informatics in Medicine Unlocked*. – 2019. – Vol. 16. – P. 100231. DOI: 10.1016/j.imu.2019.100231
30. Kate V. A 3 Tier CNN model with deep discriminative feature extraction for discovering malignant growth in multi-scale histopathology images / V. Kate, P. Shukla // *Informatics in Medicine Unlocked*. – 2021. – Vol. 24. – P. 100616. DOI: 10.1016/j.imu.2021.100616
31. Shallu S. Breast cancer histology images classification: Training from scratch or transfer learning? / S. Shallu, R. Mehra // *ICT Express*. – 2018. – Vol. 4. – P. 247–254. DOI: 10.1016/j.icte.2018.10.007
32. Desai M. An anatomization on breast cancer detection and diagnosis employing multi-layer perceptron neural network (MLP) and Convolutional neural network (CNN) / Meha Desai, Manan Shah // *Clinical eHealth*. – 2021. – Vol. 4. – P. 1–11. DOI: 10.1016/j.ceh.2020.11.002
33. Asadi B. Efficient breast cancer detection via cascade deep learning network / B. Asadi, Q. Memon // *International Journal of Intelligent Networks*. – 2023. – Vol. 4. – P. 46–52. DOI: 10.1016/j.ijin.2023.02.001

34. Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer) / [A. A. Joseph, M. Abdullahi, S. B. Junaidu et al.] // *Intelligent Systems with Applications*. – 2022. – Vol. 14. – P. 200066. DOI: 10.1016/j.iswa.2022.200066
35. Chaudhury S. A BERT encoding with Recurrent Neural Network and Long-Short Term Memory for breast cancer image classification / S. Chaudhury, K. Sau // *Decision Analytics Journal*. – 2023. – Vol. 6. – P.100177. DOI: 10.1016/j.dajour.2023.100177
36. Method and software component model for skin disease diagnosis / [V. M. Lovkin, S. A. Subbotin, A. O. Oliinyk et al.] // *Radio Electronics, Computer Science, Control*. – 2023. – № 1. – P. 40. DOI: 10.15588/1607-3274-2023-1-4
37. Berezsky O. M. Design of computer systems for biomedical image analysis / [O. M. Berezsky, K. M. Berezska, G. M. Melnyk et al.] // *Proceedings of the X th International Conference «The Experience of Designing and Application of CAD Systems in Microelectronics» CADSM 2009*, 24–28 February 2009, Lviv-Polyana, Ukraine. – Lviv: Publishing House Vezha&CoC. – P.186–191.
38. Berezsky O. The intelligent system for diagnosing breast cancers based on image analysis / O. Berezsky, T. Datsko, S. Verbovyi // *Proceedings of Information Technologies in Innovation Business (ITIB)*. Kharkiv: Ukraine. 7–9 October, 2015. p. 27–30. DOI: 10.1109/ITIB.2015.7355067.
39. Fuzzy system diagnosing of precancerous and cancerous conditions of the breast cancer / [O. Berezsky, S. Verbovyi, L. Dubchak et al.] // *Proceedings of the XIth International Scientific and Technical Conference Computer Sciences and Information Technologies (CSIT'2016)*. Lviv: Ukraine. 6–10 September, 2016. – P. 200–203. DOI: 10.1109/STC-CSIT.2016.7589906.
40. Berezsky O. Access Distribution in Automated Microscopy System / O. Berezsky, L. Dubchak, O. Pitsun // *Proceedings of the 14 th International Conference «The Experience of Designing and Application of CAD Systems in Microelectronics» CADSM 2017*, 21–25 February 2017, Lviv, Ukraine. – Lviv, 2017. – P. 241–243. DOI: 10.1109/CADSM.2017.7916125
41. Berezsky O. Hybrid Intelligent information technology for biomedical image processing / O. Berezsky, S. Verbovyi, O. Pitsun // *Proceedings of the IEEE International Conference “Computer Science and Information Technologies” CSIT'2018*. Lviv: Ukraine. 11–14 September, 2018. – Lviv, 2018. – P. 420–423. DOI: 10.1109/STC-CSIT.2018.8526711.
42. Modern automated microscopy systems in oncology / [O. Berezsky, O. Pitsun, N. Batryn, et al.] // *Proceedings of the 1st International Workshop on Informatics & Data-Driven Medicine*, Lviv, Ukraine, 28–30 november 2018. – Lviv, 2018. – P. 311–325.
43. Berezsky O. M. Evaluation methods of image segmentation quality. / O. M. Berezsky, O. Y. Pitsun // *Radio Electronics, Computer Science, Control*. – 2018. – Vol. 1. – P. 119–128. DOI:10.15588/1607-3274-2018-1-14
44. Fuzzy System For Breast Disease Diagnosing Based On Image Analysis / [O. Berezsky, L. Dubchak, N. Batryn et al.] // *Proceedings of the II International Workshop Informatics & Data-Driven Medicine (IDDM 2019)*. Lviv, Ukraine 11–13 November, 2019. – Lviv, 2019. – P. 69–83.
45. An Approach toward Automatic Specifics Diagnosis of Breast Cancer Based on an Immunohistochemical Image. / [O. Berezsky, O. Pitsun, G. Melnyk et al.] // *Journal of Imaging*. – 2023. – Vol. 9, № 1. – P. 12. DOI: 10.3390/jimaging9010012
46. Berezsky O. Cytological and histological images of breast cancer. / Oleh Berezsky, Tamara Datsko, Grygoriy Melnyk. [Electronic resource] – Access mode: <https://zenodo.org/records/7890874>. DOI: 10.5281/zenodo.7890874
47. Introsoft Consulting. General Data Protection Regulation (GDPR). Recital 26. [Electronic resource]. – Access mode: <https://gdpr-info.eu/recitals/no-26/>
48. Liashchynskiy P. Rudi. Lightweight image converter and dataset augmentor / Petro Liashchynskiy [Electronic resource]. – Access mode: <https://github.com/liashchynskiy/rudi>
49. Deep Residual Learning for Image Recognition / [Kaiming He, Xiangyu Zhang, Shaoqing Ren et al.] – 2015. [Electronic resource]. – Access mode: <https://arxiv.org/pdf/1512.03385.pdf>
50. Lim J.H. Geometric GAN / Jae Hyun Lim, Jong Chul Ye. – 2017. [Electronic resource]. – Access mode: <https://arxiv.org/pdf/1705.02894v2.pdf>
51. Improved techniques for training GANs. / [T. Salimans, I. Goodfellow, W. Zaremba et al.] // *Advances in neural information processing systems: Annual Conference on Neural Information Processing Systems 2016*, December 5–10, 2016, Barcelona, Spain – Barcelona, 2016. – P: 2226–2234 DOI: 10.48550/arXiv.1606.03498
52. GANs trained by a two time-scale update rule converge to a local nash equilibrium / M. Heusel, H. Ramsauer, T. Unterthiner et al.] // *Proceedings of the 31st International Conference on Neural Information Processing Systems – NY, United States: Curran Associates Inc.*, 2017 – P. 6629–6640. Access mode: <https://dl.acm.org/doi/10.5555/3295222.3295408>
53. Classification Of Breast Cancer Histology Images Using ALEXNET. In: Campilho, A., Karray, F., ter Haar Romeny, B. (eds) *Image Analysis and Recognition. ICIAR 2018*. / [W. Nawaz, S. Ahmed, A. Tahir et al.] // *Lecture Notes in Computer Science*. – 2018. – Vol. 10882. – P. 869–876. Springer, Cham. DOI: 10.1007/978-3-319-93000-8_99
54. Verdhan V. Image Classification Using LeNet. In: *Computer Vision Using Deep Learning*. / V. Verdhan. – Apress, Berkeley, CA., 2021. – P. 67–101. DOI: 10.1007/978-1-4842-6616-8_3
55. Tammina S. Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images / Srikanth Tammina // *International Journal of Scientific and Research Publications*. – 2019. – Vol. 9, Issue 10. – P. 143–150. DOI: 10.29322/IJSRP.9.10.2019.p9420