

Specified Diagnosis of Breast Cancer on the basis of Immunohistochemical Images Analysis

Oleh Berezsky^a, Oleh Pitsun^a, Tamara Datsko^b, Bohdan Derysh^a, Ivan Tsmots^c, Vasyl Teslyuk^c

^a West Ukrainian National University, 11 Lvivska st., Ternopil, 46001, Ukraine

^b I.Ya. Horbachevsky Ternopil State Medical University, m.Voli, 1, Ternopil, 46001, Ukraine

^c Lviv Polytechnic National University, S. Bandera st., 12, Lviv, 79000, Ukraine

Abstract

The article shows that cancer is a global problem of the 21st century. The work provides the relevance of scientific research on the diagnosis automation. The object of research is breast cancer. The authors analyzed modern methods of automated diagnosis. It is shown that the main methods for making an automated diagnosis are convolutional neural networks, Naive Bayes, SVM, Decision Tree. The researchers analyzed the immunohistochemical method for clarifying the diagnosis based on histological analysis. Approach has been developed to a more specified diagnosis based on the analysis of immunohistochemical images. The module for automated diagnosis is implemented in software. Computer experiments have been carried out to clarify the diagnosis of breast cancer subtypes.

Keywords 1

Breast cancer, automated diagnosis, CNN, immunohistochemical analysis.

Introduction

Cancer is one of the global medical problems of the 21st century. This disease ranks second after cardiovascular disease and AIDS. According to world statistics obtained from the website 'Worldometers', about six million people died from cancer in the world this year alone. Therefore, this problem is extremely important. Early diagnosis is an effective way to prolong life.

Breast cancer ranks first in the world. In Ukraine, the incidence of breast cancer is 45% higher than in other countries of the world. Pathological diagnostics remains the main research method.

Histological diagnosis continues to be the main determining method of research and remains a prerequisite for adequate justified treatment. However, currently, histological examination is insufficient for the treatment and prognosis of relapses. Therefore, it becomes necessary to conduct an immunohistochemical study.

1. Literature review

To automate the diagnosis, an analysis of the literature was conducted.

Researcher A. Albu used decision trees to diagnose hepatitis infection and predict the patient's health [1].

IDDM'2020: 3rd International Conference on Informatics & Data-Driven Medicine, November 19–21, 2020, Växjö, Sweden
EMAIL: ob@tneu.edu.ua (O. Berezsky); o.pitsun@tneu.edu.ua (O. Pitsun); datsko_t@tdmu.edu.ua (T. Datsko); dbb@tneu.edu.ua (L. Dubchak), tsmots@tneu.edu.ua (I. Tsmots), vasyi.teslyuk@gmail.com (V. Teslyuk)
ORCID: 0000-0001-9931-4154 (O. Berezsky); 0000-0003-0280-8786 (O. Pitsun); 0000-0001-9283-2629 (T. Datsko), 0000-0002-7215-9032 (L. Dubchak), 0000-0002-4033-8618 (I. Tsmots), 0000-0002-5974-9310 (V. Teslyuk)



© 2020 Copyright for this paper by its authors.
Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
CEUR Workshop Proceedings (CEUR-WS.org)

S. Sayed, S. Ahmed and R. Poonia analyzed various factors influencing breast cancer. To do this, scientists have developed a model using hologlotropy of the decision tree, which determines the criteria for this disease [2].

The authors of [3] compared the methods of the decision tree and neural networks. The researchers used neural networks to predict the disease of the urinary system.

H. S. Laxmisagar and M. C. Hanumantharaju analyzed DSS systems for the detection and analysis of breast cancer. The article discusses various classification methods for the WBCD dataset based on the machine-learning algorithm [4].

Article [5] used texture features and decision tree classifier for diagnosis of tumor in brain.

T. He et al. developed a deep learning analytical model for breast cancer treatment. This model calculates the probabilities of biopsy recommendations [6].

Maheshwar and G. used various classifiers to classify breast cancer: decision tree, Naive Bayes, KNN, SVM. The authors experimented on a dataset with the UCI Machine Repository. The results showed that the decision tree classifiers are accurate among all classifiers that have been used to predict breast cancer. As a development of their work, scientists propose the use of classifiers together with evolutionary algorithms [7].

The article developed the IDSS, in which the inference engine automated the process of identifying suspicious regions and classified the regions into benign and malignant based on mammography data. The inference engine used ANN for classification [8].

Researchers C. G. Tams and N. R. Euliano in the article [9] provided substantial recommendations for the creation and organization of DSS.

The article developed a method for diagnosing breast cancer based on a decision tree combined with a set of features. The results were validated against a Wisconsin clinic and showed a classification accuracy of 94.3% [10].

The article [11] shows the application of the Bayesian method in DSS for the diagnosis of breast cancer.

L. Hussain, W. Aziz, S. Saeed, S. Rathore and M. Rafique used robust classification methods such as SVM, Bayesian approach and decision tree in the article. These studies were carried out based on mammographic data to detect breast cancer [12].

So, the main groups of algorithms and methods used to solve problems of automatic diagnosis are based on:

1. Convolutional neural networks.
2. Deep learning methods for using Naive Bayes, SVM and Decision Tree.
3. Decision support systems based on the above methods.

2. Problem statement

Analysis of the literature has shown that an actual task is to automate the diagnosis of breast cancer. The traditional method of diagnosis is based on the analysis of histological images. To clarify the diagnosis, an immunohistochemical method is used. Therefore, the objectives of this work are:

1. Analysis of immunohistochemical images.
2. Development of an approach to making a more precise diagnosis of breast cancer.
3. Conducting and analyzing computer experiments..

3. Analysis of immunohistochemical images.

For a more specified diagnosis, an immunohistochemical research method is used. The immunohistochemical method makes it possible to determine the receptor status of the tumor, the proliferative potential, and the molecular genetic type. The prognosis of the course of the disease and the response to treatment depend on the molecular genetic characteristics of the tumor.

Immunohistochemical analysis takes into account the presence or absence of receptors for hormones estrogen (ER) and progesterone (ER), receptors for epidermal growth factor HER2 / neu, and other molecular and genetic markers in tumor cells.

When studying the expression of estrogen and progesterone receptors, also quantitative, qualitative indicators are defined. Used Allred technique, where two criteria are assessed: the number of positive cells the color intensity. The scores obtained are in the total score from 1 to 5 (taking into account percentages from 1 to 100).

For example: estrogen receptor α (DAKO, clone EP1) - a positive reaction in 80% of tumor cells (PS = 5) of significant intensity (IS = 3) TS = 5 + 3 = 8 - a positive result.

Progesterone receptor (DAKO, clone PgR 636) - a positive reaction in 30% of tumor cells (PS = 3) of significant intensity (IS = 3) TS = 3 + 3 = 6 - a positive result. (Fig.1,2).

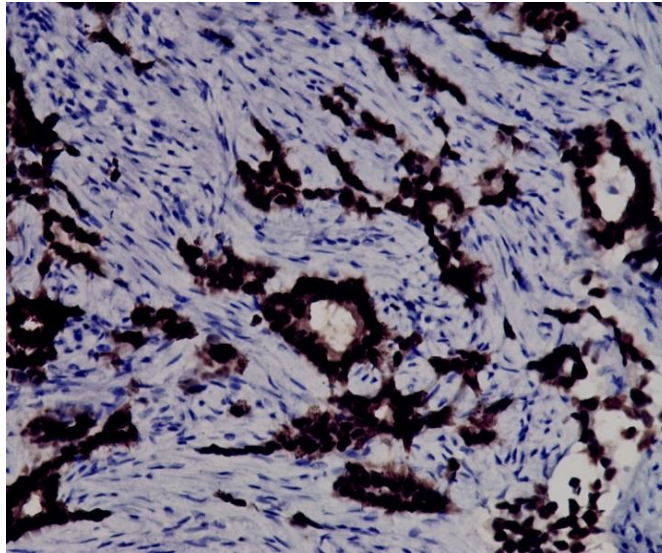


Figure. 1: Estrogen receptor α (DAKO, clone EP1) - a positive reaction in 80% of tumor cells (PS = 5) of significant intensity (IS = 3) TS = 5 + 3 = 8 - a positive result. Incr. x 200

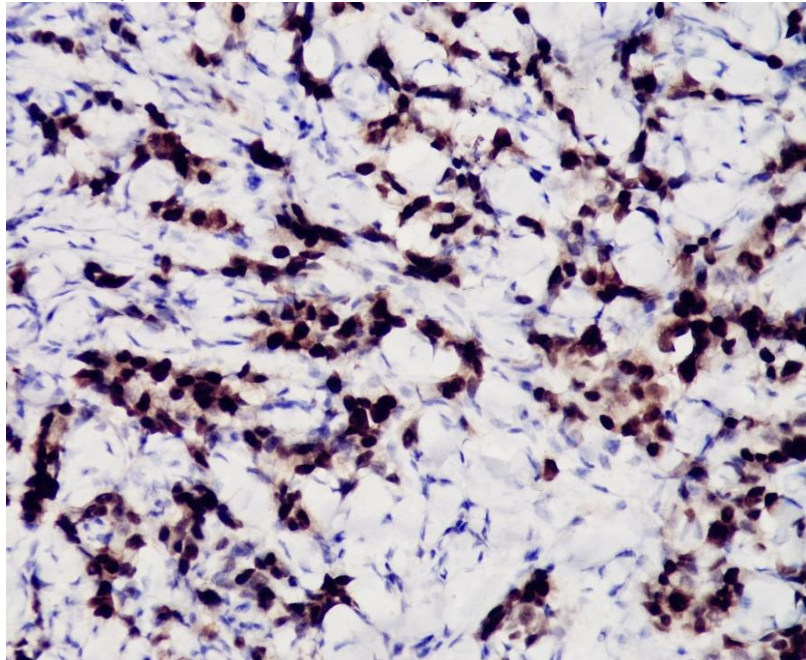


Figure: 2: Progesterone receptor (DAKO, clone PgR 636) - a positive reaction of 80% of tumor cells (PS = 5) of significant intensity (IS = 3) TS = 5 + 3 = 8 - a positive result. Incr. x 200.

For determination of quantitative indicators, it is necessary to segment positive cells and determine their number and intensity.

4. Approach to the specified diagnosis on the basis of immunohistochemical images

Stages of the specified diagnosis on the basis of immunohistochemical images are given in figure 3.

The first step is to load the input image into the computer's memory. After that, the input image parameters are determined for further automatic pre-processing and segmentation. The pre-processing algorithm includes a filtering and histogram alignment step. The segmentation stage is implemented on the basis of the k-mens algorithm and thresholding. Based on the rules of fuzzy logic, the image is pre-processed and segmented automatically [13,14].

To form the rules of diagnosis on the basis of immunohistochemical images, the relative area of cell nuclei to the area of the whole image is calculated. Another parameter that is calculated for diagnosis is the color intensity of the cell nuclei. The calculation of the intensity of cell nuclei occurs after the stage of calculating the area and is based on the selected pixels after the segmentation stage.

In the second stage, the degree of cancer differentiation by Nottingham gradation is assessed on the basis of histological images.

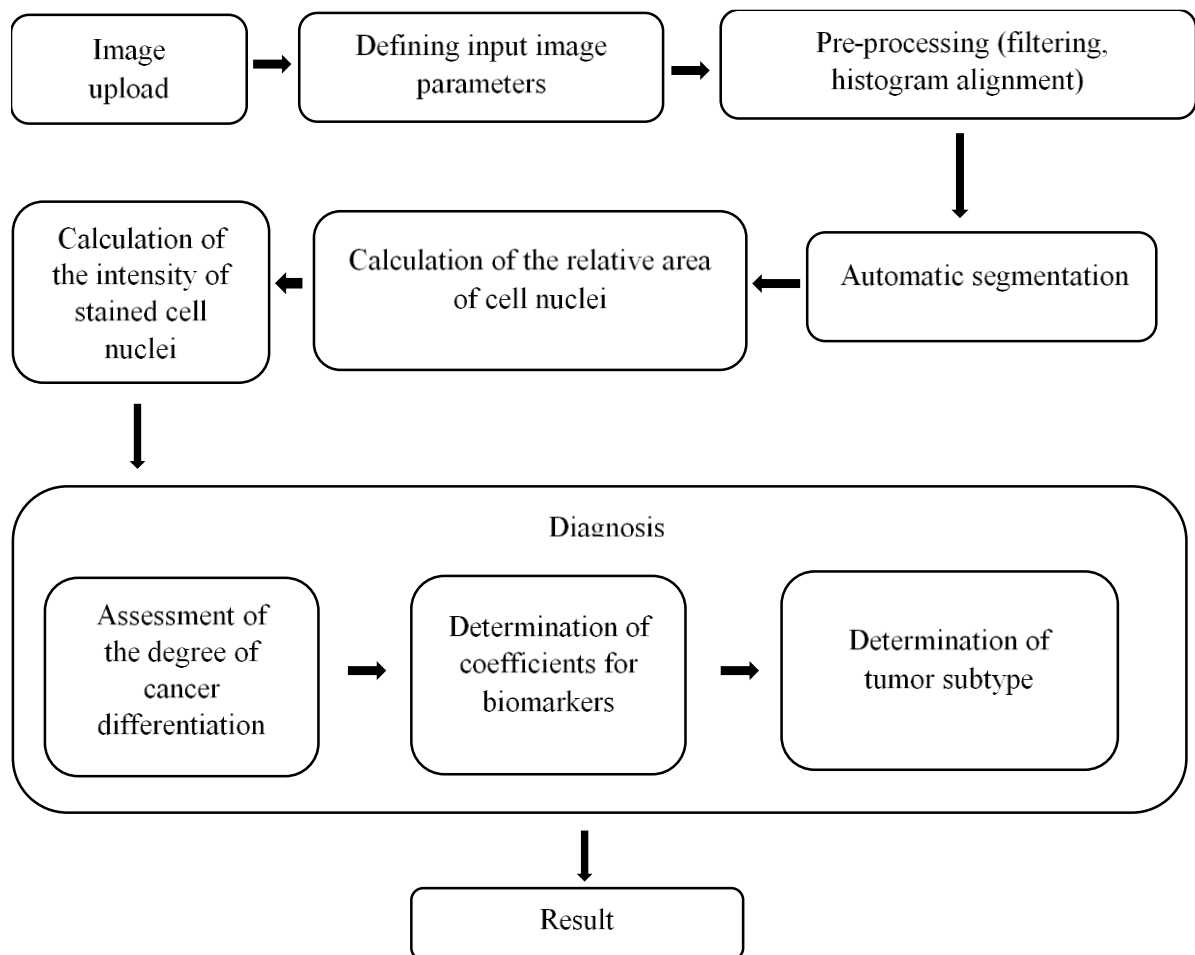


Figure 3: Stages of the specified diagnosis on the basis of immunohistochemical images

The third stage involves determining the coefficients for biomarkers ER, PR, HER2 / neu, KI-67 based on the analysis of stained nuclei and their intensity. The parameters of segmentation algorithms for immunohistochemical and histological images are different and require a separate approach;

At the fourth stage, the tumor subtype is determined.

5. Description of the computer program

It has been developed a number of intelligent automated microscopy systems [15].

As part of the intelligent system HIAMS [16] a separate software package based on Java and the OpenCV has been developed. UML diagram is shown in Figure 4.

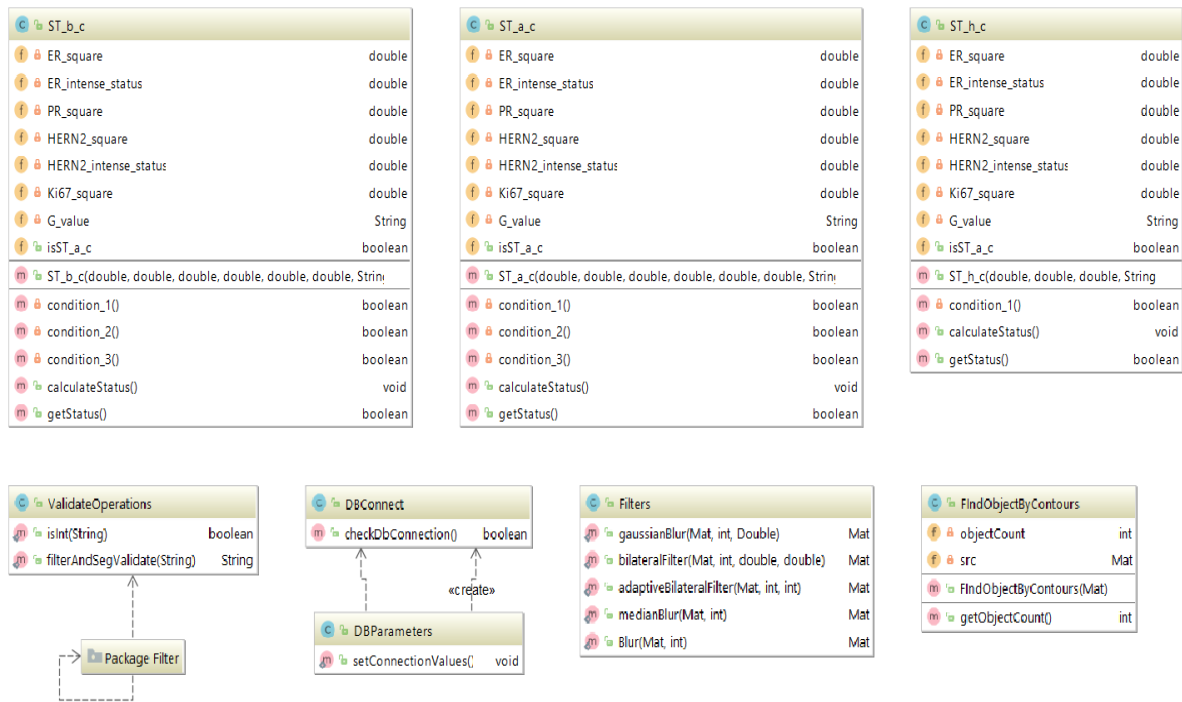


Figure 4: UML class diagram of the software module for diagnosis

The StartImageParams class stores values of RGB parameters. Class "ST_z_c" is responsible for determining the basal-like subtype. This class provides three methods that return "true" if the condition is true. The input parameters are:

biomarker (estrogen), biomarker (progesterone), biomarker (oncoprotein) and color intensity ratio, total ratio, positive cell ratio.

Figure 5 shows the result of the program to determine the subtype of breast cancer.

```

Run: Main x
condition 1: ER square = 0.71, ER intense_status = 2.0
true
condition 2: PR square = 0.23
true
condition 3: Ki67 square = 0.16
true
  
```

Figure 5: Program result to determine the subtype of breast cancer.

This figure shows result of calculating three conditions that characterize the presence or absence of cancer subtype "Luminal A". The parameters "ER square", "PR square", "Ki-67" correspond to the value of the relative area for the biomarker estrogen, progesterone and cell proliferation. The parameter "ER intense_status" corresponds to the value of the intensity of cell nuclei for the biomarker estrogen.

Figure 6 shows a comparative analysis of finding the accuracy of the assessment of conditions 1, 2 and 3 for the cancer subtype "Luminal A". Each expression returns a Boolean value (true or false), which characterizes the truth or falsity of the expression.

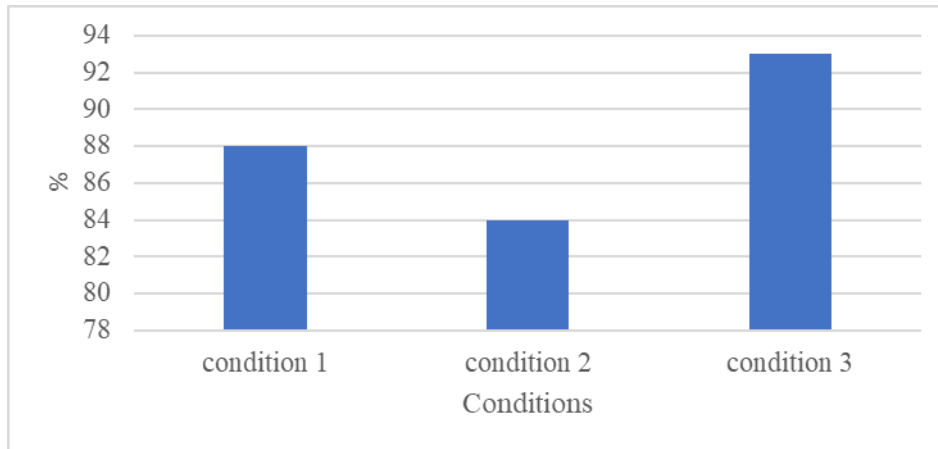


Figure 6: Comparative analysis of finding the accuracy of estimating the presence of expression 1, 2 and 3 for the cancer subtype "Luminal A"

6. Conclusions

1. The analysis of algorithms and methods for automated diagnosis has shown that the main tools for making a diagnosis are convolutional neural networks, deep learning methods using Naive Bayes, SVM and Decision Tree.

2. The analysis of immunohistochemical images to clarify the diagnosis of breast cancer.

3. An approach has been developed for making a more precise diagnosis and its main stages have been described.

4. On the basis of the approach for setting the refined analysis, a software component was developed using the Java and the OpenCV.

5. Conducted computer experiments have shown that the accuracy of making an accurate diagnosis of breast cancer is in the range from 88-93%.

7. References

- [1] A. Albu, "From logical inference to decision trees in medical diagnosis," 2017 E-Health and Bioengineering Conference (EHB), Sinaia, 2017, pp. 65-68, doi: 10.1109/EHB.2017.7995362.
- [2] S. Sayed, S. Ahmed and R. Poonia, "Holo entropy enabled decision tree classifier for breast cancer diagnosis using wisconsin (prognostic) data set," 2017 7th Intern. Conf. on Communication Systems and Network Technologies (CSNT), Nagpur, 2017, pp. 172-176, doi: 10.1109/CSNT.2017.8418532.
- [3] M. H. Kadhem and A. M. Zeki, "Prediction of Urinary System Disease Diagnosis: A Comparative Study of Three Decision Tree Algorithms," 2014 Intern. Conf. on Computer Assisted System in Health, Kuala Lumpur, 2014, pp. 58-61, doi: 10.1109/CASH.2014.25.
- [4] H. S. Laxmisagar and M. C. Hanumantharaju, "A Survey on Automated Detection of Breast Cancer based Histopathology Images," 2020 2nd Intern. Conf. on Innovative Mechanisms for Industry Applications (ICIMIA), Bangalore, India, 2020, pp. 19-24, doi: 10.1109/ICIMIA48430.2020.9074915.
- [5] G. K. Sundararaj and V. Balamurugan, "An expert system based on texture features and decision tree classifier for diagnosis of tumor in brain MR images," 2014 Intern. Conf. on Contemporary Computing and Informatics (IC3I), Mysore, 2014, pp. 1340-1344, doi: 10.1109/IC3I.2014.7019690.

- [6] T. He et al., "Deep learning analytics for diagnostic support of breast cancer disease management," 2017 IEEE EMBS Intern. Conf. on Biomedical & Health Informatics (BHI), Orlando, FL, 2017, pp. 365-368, doi: 10.1109/BHI.2017.7897281.
- [7] Maheshwar and G. Kumar, "Breast Cancer Detection Using Decision Tree, Naïve Bayes, KNN and SVM Classifiers: A Comparative Study," 2019 Intern. Conf. on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2019, pp. 683-686, doi: 10.1109/ICSSIT46314.2019.8987778.
- [8] H. AlSalman and N. Almutairi, "IDSS: An Intelligent Decision Support System for Breast Cancer Diagnosis," 2019 2nd Intern. Conf. on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1-6, doi: 10.1109/CAIS.2019.8769579.
- [9] G. Tams and N. R. Euliano, "Creating clinical decision support systems for respiratory medicine," 2015 37th Annual Intern. Conf. of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, 2015, pp. 5335-5338, doi: 10.1109/EMBC.2015.7319596.
- [10] L. Yi and W. Yi, "Decision Tree Model in the Diagnosis of Breast Cancer," 2017 Intern. Conf. on Computer Technology, Electronics and Communication (ICCTEC), Dalian, China, 2017, pp. 176-179, doi: 10.1109/ICCTEC.2017.00046.
- [11] J. Gadewadikar, O. Kuljaca, K. Agyepong, E. Sarigul, Yufeng Zheng and Ping Zhang, "Exploring Bayesian networks for automated breast cancer detection," IEEE Southeastcon 2009, Atlanta, GA, 2009, pp. 153-157, doi: 10.1109/SECON.2009.5174067.
- [12] L. Hussain, W. Aziz, S. Saeed, S. Rathore and M. Rafique, "Automated Breast Cancer Detection Using Machine Learning Techniques by Extracting Different Feature Extracting Strategies," 2018 17th IEEE Intern. Conf. On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), New York, NY, 2018, pp. 327-331, doi: 10.1109/TrustCom/BigDataSE.2018.00057.
- [13] O. Berezsky et al. Segmentation of Cytological and Histological Images of Breast Cancer Cells, 8th Intern. Conf. on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS'2015), 24-26 September 2015, Warsaw, Poland, V.1, pp. 287-292, DOI: 10.1109/IDAACS.2015.7340745.
- [14] O. Berezsky et al. Fuzzy system diagnosing of precancerous and cancerous conditions of the breasts, XIth Intern. Scientific and Technical Conf. Computer Sciences and Information Technologies (CSIT'2016), Lviv, 6-10 September, 2016, pp. 200–203, DOI: 10.1109/STC-CSIT.2016.7589906.
- [15] O. Berezsky et al. An Intelligent System for Cytological and Histological Image Analysis, 13 th Intern. Conf. «The Experience of Designing and Application of CAD Systems in Microelectronics» CADSM 2015, 24-27 February 2015, Polyana-Svalyava (Zakarpattia), Ukraine, pp. 28-31, DOI: 10.1109/CADSM.2015.7230787.
- [16] O. Berezsky et al. Modern automated microscopy systems in oncology, International Workshop on Informatics & Data-Driven Medicine, Lviv, Ukraine, 28-30 november 2018, pp. 311-325.