Applications of neural networks in the diagnosis of lymphoproliferative diseases

V G Nikitaev\(^1\), A N Pronichev\(^1\), E V Polyakov\(^1\), V V Dmitriev\(^1\), S A Kobelev\(^1\), A V Mozhenkova\(^2\), N N Tupitsin\(^2\), M A Frenkel\(^2\), V Y Selchuk\(^1,3\), O P Grebennikova\(^3\), G V Titova\(^3\)

\(^1\)National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe shosse 31, 115409, Moscow, Russia.
\(^2\)Federal State Budgetary Scientific Institution «N.N. Blokhin Russian Cancer Research Center», Kashirskoe shosse 24, 115478, Moscow, Russia.
\(^3\)A.I. Evdokimov Moscow State University of Medicine and Dentistry, Delegatskaya ulica, 20/1, 127473, Moscow, Russia

E-mail: VGNikitayev@mephi.ru

Abstract. The work is devoted to one of the promising directions in the field of artificial intelligence - recognition of blood cells and bone marrow to differentiate leukocytes by type in the diagnosis of lymphoproliferative diseases using computer microscopy. The article presents a study using neural networks to classify cells by type.

1. Introduction

Modern diagnostics of hematological diseases based on the combined data of morphological research with immunophenotypical results. The study of the morphology of blood cells in blood and bone marrow smears is one of the main methods of diagnosis of acute leukemia. Important sign of the development or exacerbation of oncohematological diseases is the presence of lymphocytes with atypical or pathological morphology in the blood, even in small quantities.

Morphological classification of leukocytes is difficult to formalize, and the same cells by some researchers can be classified as normal, and others as atypical. The use of immunophenotyping to determine the differentiation of surface antigens of cells allows only to determine the total number of an immunophenotype without the possibility of assessing the "thin" morphological features of cells [1-5].

The development and application of new methods and means of image processing using computer microscopy to identify structural differences between normal and pathological cell elements for the diagnosis of acute leukemia is relevant.

Computer microscopy with mathematical apparatus of pattern recognition in comparison with traditional (visual analysis of microscopic images) research has immeasurably wider possibilities for the study of the structure of chromatin filaments of cell nuclei and allows to objectify the data obtained in the form of numerical indices [6-7]. Modern directions of research of application of computer microscopy are:

* Obtaining additional independent objective criteria for the differential diagnosis of acute leukemia;
Establishment of correlation between the structure of blasts and their immunophenotypic status;

Recently, numerous attempts have been made to automate microscopic studies to improve the accuracy of detection of young cell forms. One of these areas is the use of neural networks for the recognition of normal and atypical blood cells and bone marrow.

In a number of scientific publications it is noted that the use of cell type recognition methods for normal blood makes it possible to achieve the frequency of correct classification above 95%, while for pathological cells - myeloid and lymphoid series, this figure does not exceed 88% [8-12].

The article [13] analyzes the possibilities of classification of bone marrow and blood cells using the methods of "K-nearest neighbors", "Linear vector quantization", "Multilayer neural network", "support vector machine". The classification accuracy of 81%, 83%, 90% and 91% was obtained on the test sample for the presented classifiers. The database contained 108 images with 258 cells.

The aim of this work is to improve the accuracy of recognition of blood cells and bone marrow in the diagnosis of lymphoproliferative diseases using a neural classifier.

2. Materials and methods.

The sample was formed for research from 6674 images of them 1029 - lymphocytes of blood donors, 5419 - lymphoblasts bone marrow, of which 2415 patients with T- acute lymphoblastic leukaemia (T-ALL), 3004 cells with B- acute lymphoblastic leukaemia(B-ALL), 1495 – leukemic lymphoid cells of the bone marrow of patients with follicular lymphoma. Four classes of blood cells were formed from the obtained images and assigned numbers ("B-ALL" - 1," follicular lymphoma " - 2," Norm "-3," T - ALL " - 4). For these classes of images the characteristics of color components of halftone images of color models RGB, XYZ, HSL, Lab, Luv, HLC, HLS, HSV, YUV, YIQ, YCbCr, CMY texture, wavelet, statistical and Fourier characteristics were calculated [14-18].

Preliminary preprocessing of initial data was carried out, consisting in normalization of values according to the following algorithm: 1. A training sample of 75% of the obtained number of cell images is formed; 2. For each feature value of the object is calculated the mean and standard deviation; 3. The normalization of feature values \( y = (x-M(x))/\sqrt{D(x)} \) in the training and then in the test sample is carried out. The neural network of direct distribution was used as a classifier.
A program in Python with the library "Scikit-learn" was developed to conduct the research. The following network settings were used: {activation function: logistic, alpha: 0.01, hidden_layer_sizes: 200, learning_rate: invscaling, random_state: 0, solver: adam}. The number of neurons in the inner layer is 100. The research model is shown in Figure 1.

3. Results
As a result of testing, the following values of the proportion of correctly recognized cells for the corresponding classes were obtained: ("Norm" - 93%, "follicular lymphoma" -95%, "T – ALL" -88%, "B – ALL" - 89%). The experiment was conducted in the following sample images of the nuclei of white blood cells: normal - 1029, follicular lymphoma - 1495, T-OLL – 2415 and B-OLL – 3004 images. The average value of correctly classified cells was 91% for the test sample.

The matrix of classification results for the test sample, formed by 25% of the sample of all cells, including 7943 images of leukocyte nuclei, is presented in table 1.

|                | B     | Limfoma | Norma | T    |
|----------------|-------|---------|-------|------|
| **B**          | 675   | 2       | 2     | 68   |
| **Limfoma**    | 3     | 359     | 16    | 0    |
| **Norma**      | 3     | 14      | 225   | 3    |
| **T**          | 74    | 1       | 0     | 541  |
4. Conclusion
The paper studies the capabilities of the neural classifier for the recognition of blood cells in lymphoproliferative diseases. As a result of the experiment, the classification accuracy for the test sample under consideration is 91%.
Further development will be aimed at the analysis of classification errors and improvement of cell classification in order to determine the cells of patients with lymphoproliferative diseases.

Acknowledgments
The reported study was funded by RFBR according to the research project № 18-29-09115.

References
[1] Frenkel M A 2007 Laboratory diagnosis of acute leukemia Clinical Hematology edited by Professor M A Volkova (original Russian title: Laboratornaya diagnostika ostykh leykozov Klinicheskaya onkogematologiya pod redaktsiy prof. M.A.Volkovoy) (Moscow Medicine publisher) pp 306-319
[2] Tupitsin N N 2007 Immunodiagnostics of acute leukemia and non-Hodgkin's lymphoma. Clinical Hematology edited by Professor M. A. Volkova (original Russian title: Immunodiagnostika ostykh leykozov i nehodzhkinskih limfom. Klinicheskaya onkogematologiya pod redaktsiy prof. M.A.Volkovoy) (Moscow Medicine publisher) pp 338-370
[3] Frenkel M A 2014 (in Russian) Haematopoiesis immunology 12(1-2) 18
[4] Volkova S A and Borovkov N N 2013 Fundamentals of clinical Hematology (original Russian title: Osnovi klinicheskoy gamatologii) (N. Novgorod: Lower GMA publisher) p 400
[5] Fyias A T 2013 (in Russian) Hematological malignancies: clinical features, diagnosis, treatment (original Russian title: Gemooblastosi: klinika, diagnostika, lechenie) (Grodno: GrGMU publisher) p 172
[6] Nikitaev V G et al. 2018 Journal of Physics: Conference Series 945(1) 012008
[7] Nikitaev V G 2015 Measurement Techniques 58(6) 215
[8] Rawat J et al. 2015 Procedia Computer Science 70 748-756
[9] Alfèrez S et al. 2015 American journal of clinical pathology 143(2) 168-176
[10] Wang Q et al. 2017 Biomedical Optics Express 8(6) 3017-3028
[11] Rawat J et al. 2017 Multimedia Tools and Applications 76(18) 19057-19085
[12] Linder J and Zahniser D J 2012 MLO: Medical Laboratory Observer 44 14-16
[13] Bhende P G 2012 International Journal of Science & Technology 2 59-64.
[14] Pronichev A. N et al. 2017 Journal of Physics: Conference Series 784(1) 012003
[15] Nikitaev V G et al. 2015 Proceedings of the 6th International Conference on Bioscience and Bioinformatics (ICBB’15) Dubai United Emirates February 22-24 p 120
[16] Nikitaev V G et al. 2015 Proceedings of the 15th International Conference on Signal Processing, Computational Geometry and Artificial Vision (ISCGAV’15) Sliema Malta August 17-19 p 49
[17] Nikitaev V G et al. 2016 International journal of biology and biomedical engineering 10 2016
[18] Mozhenkova A V et al. 2017 Journal of Physics: Conference Series 798(1) 012126