The blood smear image processing for the acute leukemia diagnostics

V. G. Nikitaev, O. V. Nagornov, A. N. Pronichev, E. V. Polyakov, V. V. Dmitrieva

***Abstract*—**One of the signs of acute leukemia is the blast cells emergence in the peripheral blood. If you suspect acute leukemia it is mandatory to analyze a blood smear under the microscope. In this paper we proposed a method of automated blood smears image analysis for detection of blast cells. An experiment on the estimation of classification errors of blast cells in blood smears is described. The analysis of the experimental results confirms the novelty of the proposed method for solving the problem of automation of diagnostics of acute leukemia.

***Keywords*—**blood cells classification, automation of acute leukemias diagnostics, computer microscopy, image processing.

I. Introduction

Acute leukemia is neoplastic disease of the blood system in which hematopoiesis process disruption takes place. In the past this disease was accompanied by inevitable death. Today there are effective treatments, but a delay in starting of medical procedures significantly complicates the process of treatment and reduces the patient's chances for recovery. Therefore, timely and accurate diagnosis is required for successful healing.

It should be borne in mind that in the initial stage of development of diseases specific symptoms of acute leukemia are absent. Manifestations of acute leukemia is largely similar to the manifestation of anemia - weakness, shortness of breath, pallor, prolonged viral infection, etc.

The situation may be clarified by clinical analysis of blood. Leukocyte formula in acute leukemia usually abnormal. In addition, blast cells can be detected in the peripheral blood along with the usual types of cells [1]. The detection of blast cells in the peripheral blood is one of the signs that determine the presence of acute leukemia. This diagnosis is confirmed by microscopic examination of samples of bone marrow. Therefore, the first step in the diagnosis of acute leukemia is a peripheral blood analysis for the purpose of blast cells detection.

The most-used hematologic analyzers in clinic-diagnostic laboratories do not allow for identifying surely blast cells in peripheral blood. That is why the microscope study of cells of peripheral blood is needed when deviations of the blood clinical analysis results from normal ones take place.

Usually doctor carries out such investigation in microscope by visual observation of images of the blood smear on glass. The main complexity of the analysis is due to classification of atypical, reactive and immature forms of the white blood cells. There are activated lymphocyte, myelocyte, blast cells among them.

Fulfillment of such analysis is difficult and requires the high qualification of doctor while the resulting estimation of the blood cell type is heavily reliant on subjective factors.

The computer microscopic complex with specific software presents a way for solution of problem of the objectivity enhancement and accuracy of the results of the blood smear analysis.

Information about the percent content of different type cells allows us to make conclusion about either norm or pathology, and take decision about further investigations.

II. Problem Formulation

Currently unresolved problem is automatic detection of blast cells among other cell types in peripheral blood.

A. Common characteristics of the blood cells types under microscopic study

The following types of leucocytes are normal in peripheral blood: lymphocytes, monocytes, banded and segmented neutrophils, eosinophiles and basophiles.

The young, immature, reactive and atypical forms of leukocytes can be identified in peripheral blood under hematologic diseases. For example, such cells are blasts (monoblasts, lymphoblasts, erythroblasts, myeloblasts), prolymphocytes, promyelotics, myelocytes, hairy cell lymphocytes, lymphoid cells, monocytoid cells, atypical mononuclear cells etc. Different types of leucocytes fulfill various functions. Therefore, determination of ratio of different types of leucocytes, presence of young forms and pathological cell forms, changes of cell morphology presents

V. G. Nikitaev is with the National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, Russia (e-mail: v.nikitaev@bk.ru).

O. V. Nagornov is with the National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, Russia (e-mail: nagornov@yandex.ru).

A. N. Pronichev is with the National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, Russia (e-mail: a.n.pronichev@inbox.ru).

E. V. Polyakov is with the National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, Russia (e-mail: voterstreit@inbox.ru).

V. V. Dmitrieva is with the National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Moscow, Russia (e-mail: dmitrievq@yandex.ru).

valuable diagnostic information. For example, appearance of blast cells in peripheral blood can testify to presence of the acute leukemia or blastic crisis under the chronic myeloid leukemia, while appearance of atypical mononuclear characterizes the infectious mononucleosis and a set of other viral diseases. Along with that atypical mononuclear often looks similar to some of blast cells. And it generates diagnostics errors.

Visual analysis of microscopic images of these cells does not allow us to correspond cell to one or the other type uniquely. A doctor who analyzes the cells, has to be highly qualified and experienced in such analyses to prevent error conclusions in the hematologic disease diagnostics.

B. Brief characteristics of differences in the blood cells types under microscopic study

Let us consider some peculiarities of microscopic images of the blood cells at different diseases.

Lymphoid series. The lympohocytes possesses most variability of normal types of blood cells characteristics estimated by microscopic images. Small lymphocytes are standard with point of view of norm. The sizes of their nuclei are comparable with normal erythrocytes. Small lymphocytes are normal and pathological. Pathological small lympohocytes are both tumoral under chronic lymphatic leukemia and reactive, for example, under whooping-cough.

There is no significant increase of prolymphocytes in blood during reactive states. In opposite case it testifies to lymphoma or to chronic lymphocytic and prolimfocytic leukemia.

The atypical mononuclear cells are the most variable in characteristics among pathatological blood cells. They also named as the activated lymphocytes, wide-cytoplasmic lymphocytes, the immune cells, monocytoid lymphocytes, the cells of infectious mononucleosis, appeared in blood due to any infections in organism. The structure of chromatin reminds ones of lymphoblasts. Atypical mononuclear cells are similar to lymphoid cell morphology. The former are close to cells of a tumor of the lymphatic system.

The blast lymphocyte are pathological cells appearing in blood under acute leukemia. The chromatin structure is similar to typical mononuclear cell. Also blast lymphocyte can be interpreted as myeloblasts without graininess. It is possible to differ the acute myeloid leukemia from the acute lymphoblastic leukemia in rare immunological cases.

The lymphoid cells are due to complexity in recognition. These cells differ from lymphocytes, lymphoblasts, atypical mononuclear cells and lymphoblasts. Recognition is carried out by the elimination method or by other cells that attend together with lymphoid cells.

Myeloid series. Myeloblasts are similar to small and big lymphocytes. Myeloblasts have thin structure of the chromatin nuclei that reminds sand and differs from lymphocytes. They can be found in peripheral blood at pathology. Complexity of recognition of monoblasts and myeloblasts is connected with existence both monoblasts and myeloblasts properties in these cells.

Monocytes in peripheral blood of adult consists of 3..8% of leucocytes in norm. An increase of monocytes takes place during inflammatory disease, some chronical infectious (for example tuberculosis), chronic myelocytes leukemia and acute leukemias with a monocytic component. Sometimes it is very complex to differ monocytes from large activated T-lymphocytes under infectious mononucleosis or from the lymphoma cells.

Monocytoid cells under leukemia are at different stages of maturation and it is difficult to distinguish them with confidence from the normal monocytes.

C. Image processing is method for blast cells detecting in blood smear

Complexity of the blast cell identification in peripheral blood under the acute leukemia diagnosis is due to significant texture variability of blast, immature, reactive and atypical cells, weak formalization for description of these cell types. For this reason the non-experienced doctors can identify the blast cells with errors and consider them as the other types of cells

To increase objectivity and accuracy of the blood cell type determination it is necessary to automatize an analysis of such cells. There are the automated systems for calculating the leukocyte formula of peripheral blood in norm. Currently, a lot of research is devoted to the development of methods for automatic detection of blood cells in images of blood smears [2]-[4]. But there are not image processing systems which reliably detect blast cells.

The objective of this paper is to describe automated method of image processing based on wavelet transformation for detection blast cells in peripheral blood. Experimental investigation of adequateness of the assumed mathematical models is described also. The model for description of the studied objects is key one to carry out the problem recognition.

To produce the objective validation of the proposed models of the blood cell type recognition it is necessary to fulfill the experimental studies by real blood cells images obtained by method of computer microscopy under analysis of the stained blood smears.

III. Problem Solution

A. The method for automated analysis of blood smears

The visual recognition of cells types is based on the characteristics of size, color, shape, texture and contrast of the objects. The blast cells detection problem solution is associated with a search for informative descriptions of the features used to distinguish blast cells from non-blast cells. In this regard, the authors propose an approach based on wavelet analysis of blood cell images.

The method for automated analysis of blood smears consists of a number of stages:

1. The formation digital images for computer analysis by scanning blood smear under a microscope
2. The searching of leukocytes on a digital image.
3. The division of leukocytes into two groups -normal and abnormal cells.
4. Classification of abnormal leukocytes in blast and non-blast cells

At the first stage a digital image of a blood smear is formed using the method of computer microscope in transmitted light in bright field (Fig.1).

Fig.1. Computer microscope system for automatic scanning of blood smear on glass.

The digital image is an input one obtained in the computer microscopy system from the blood smear painted after the Romanovsky-Gimza method. Computer-controlled motorized specimen stage of microscope provides a scanning of the analyzed region of a blood smear. A digital image is formed by the camera mounted on the microscope. A color of each pixel is coding by three color components RGB (red, green and blue) with the depth encoding of 8 bits or more for each component.

Due to the fact that the main feature that distinguishes blast cells from non-blast is the structure of chromatin in the nucleus of white blood cell, it is necessary in the kernel image to see finer structure. Therefore, we use the highest possible magnification of a light microscope to obtain an image of leukocytes. The projection of a blood smear on the photosensitive matrix of the recording camera is formed by using an immersion lens of microscope with 1.3 aperture at 100 fold magnification. The availability of more magnification, which give more fine detail than that obtained in the mentioned conditions, is limited to the physical limit of the wavelength of electromagnetic radiation in the visible wavelength range. The operator, observing the image in the microscope with the specified lens, sees the image, magnified 1000 fold (using the normal eyepiece of the microscope with a magnification of 10 fold). Automatic scanning of a blood smear and a microscopic images camera photographing provide creating in computer memory digital images of great number fields of view of a blood smear. An example of such image is presented on Fig.2.

Most of the observed cells are red blood cells. The presented example has only two of leukocyte in the field of observation.

The second stage of processing is the separation of leucocytes from erythrocytes. This procedure can be performed on the basis of registering differences in color staining of erythrocytes and leukocytes.

The third stage is separation of cells into two groups – normal and abnormal white blood cells. Neutrophils can be easily distinguished from other cells by the shape of the nucleus. Eosinophils and basophils are distinguishable by color and the presence of granules in the cytoplasm.
Normal lymphocytes and monocytes are distinguishable by size, shape and density of the chromatin of the cell nucleus.

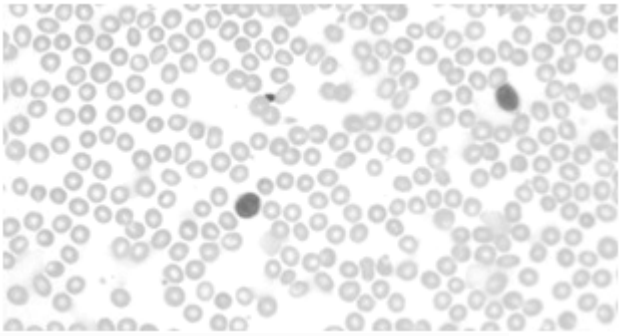

Fig.2. Blood smear image under the microscope

On the fourth stage blast cells are detected among abnormal cells. This stage image processing includes: segmentation of leukocyte image to highlight its nucleus; wavelet transform of cell nucleus image; calculated features that describe the specific features of the abnormal cells; classification of cells in blast and non-blast.

The quality of classification depends on the selection of traits that describe a cell. One of the main characteristics of the blood cell microscopic analysis at acute leukemia diagnosis is a structure of the chromatin nucleus. The structure of the leucocyte nuclei observed on microscopic images of blood smear can be characterized by qualitative criterions as smoothness, heterogeneity, grain size, roughness, tenderness, reticulation. The methods of texture analysis can be implemented for quantitative description of structure of the blood cell chromatin nuclei [5]-[8]. The method of structural elements can be used too [9].

To describe the texture we use the wavelet analysis to isolate the regularities in the structure of the blood cell chromatin nuclei. A wavelet transform is used to divide spatial function into wavelets [10], [11].

The following model of the digital image processing is applied to obtain quantitative criterions of the leukocyte chromatin nuclei.

Firstly HSL formats of images is obtained from RGB format (H is color tone, S is saturation, L is level), where each component is coding by 8-bit number. The HSL components are chosen to correspond the psychophysical perception of color by people. Thus image representations in RGB and HSL formats are considered as 6 independent half-tone images with gray brightness gradation in range from 0 to 255.

Next the wavelet transformation is applied for these six images with application of the Haar, Daubechies with three scales of decomposition 2, 4, 8, Cohen-Daubechies-Feauveau 9/7 wavelet-functions[12]. These set of wavelet functions is chosen as the base one to analyze the texture images.

As a result of the first level wavelet transform of an half-tone images with gray brightness we get the four frequency sub-band representation of the wavelet coefficients for each of wavelet-functions (Fig.3).

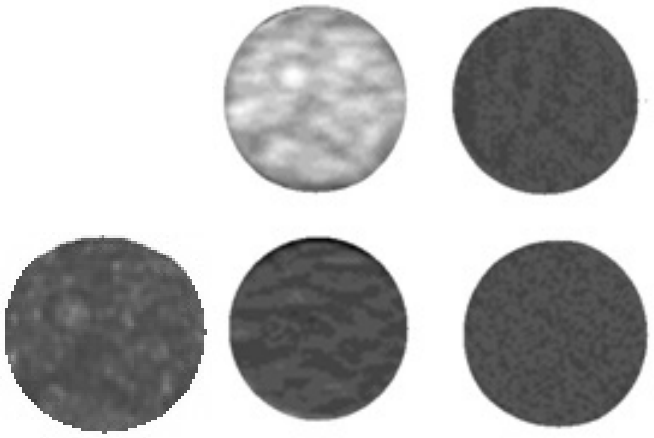

Fig.3 The image of the cell nucleus and images of four sub-band representation of first level wavelet-transformation result for G-color component and Haar wavelet function

For each wavelet transformation of image for six components (R,G,B,H,S,L) and for each frequency sub-band representation we calculate the aggregated characteristics by equations:

the mean value of the wavelet coefficients

$$Mean = X_{mean} \sum_{k=1}^{n} X_k / n$$

the maximum value of the wavelet coefficients

$$Max = \max_k (X_k),$$

the range of variation of the wavelet coefficients

$$Range = \max_k (X_k) - \min_k (X_k),$$

the energy of the wavelet coefficients

$$Energy = \sum_{k=1}^{n} (X_k^2),$$

the entropy of the wavelet coefficients

$$Entropy = -\sum_{k=1}^{n} (X_k LnX_k),$$

the variance of the wavelet coefficients

$$Disp = \sum_{k=1}^{n} (X_k - X_{mean})^2 / n \cdot$$

Here n is the number of wavelet coefficients resulting from wavelet transform of image of the nucleus of white blood cell in some of color component, X_k – value of k-th wavelet coefficient, k - the index that identifies the number of the wavelet coefficient and varying from 1 to n. In addition to the above mentioned parameters the return values of the specified signs are used: ReMean=Mean^{-1}, ReMax=Max^{-1}, ReRange=Range^{-1}, ReEnergy=Energy^{-1}, ReEntropy= Entropy^{-1}, ReDisp=Disp^{-1}.

B. Formation of representative sampling of cells for experiment

To test the effectiveness of the proposed method of descriptions blast cells an experiment was conducted by processing real images of blast and non-blast cells.

There was created the representative sampling of two classes for execution of our experiments. The total number of cells was 1289.

The first class is formed by the blast cells and includes lymphoblasts (141), myeloblasts (326) and monoblasts (69).

The second class non-blast cells consists of atypical mononuclears (57), lymphoid cells (78), large granular lymphocytes (97), hairy cell leukemia lymphocytes (116), myelocytes (82), monocytes(26), monocytoid cells(33), prolymphocytes (82) and promyelocytes (182). First of all under formation of cells of the second class turn we consider non-blast cells that have properties similar to blast cells in morphological picture of image of the cell nuclei. Therefore, we do not present here typical cells for norm as small lymphocytes, neutrophils, basophils, eosinophils because as a rule it is not complex to differ them from blast cells.

Type of cell is determined by method of expert estimation. High-experienced doctors of clinical laboratory diagnostics were experts.

An example of microscopic images of atypical mononuclear and lymphoblasts is shown in Fig. 4. It illustrates similarity of cells responsible for very different deceases that require different treatments, respectively.

C. Methodology of experiment

The methodology of experimental research of the developed method for the detection of blast cells with the use of wavelet analysis was based on two main stages.

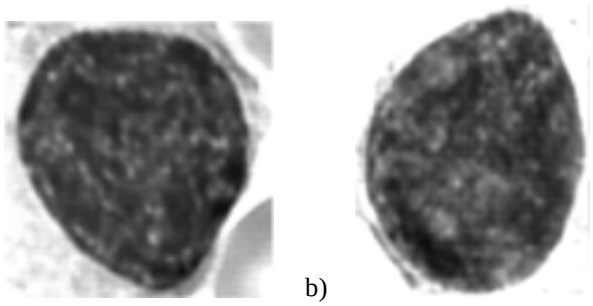

Fig.4 The microscopic images of atypical mononuclears (a) and lymphoblast (b) in blood smear after Romanovsky dyeing with Giemsa solution.

The aim of the first phase of the experiment was experimental evaluation of the accuracy of the proposed method in identifying types of cells – "blast"- "neoblast".

In the second stage of the experiment we evaluated the accuracy of determining the blasts type of among blasts cells (differentiation of blasts cells by type).

In practice the full set of parameters mentioned above is excess. In this study there was stated a problem to find a pair parameters that provide minimal error of classification of studied cells by the blast and non-blast types. In addition the classification is carried out in space of parameters by the standard method in the Manhattan metrics.

To find out the unknown pair of parameters we turn over possible pairs parameters and estimate the classification error in determination of blast and non-blast cells for the sampling. The pair that provides the minimal error of classification was chosen as the optimal for recognition.

D. Analysis of experiment results

It was determined the optimal pair of parameters that provides total error of classification of cells by blast and non-blast types equal to 18%. Distribution of the cell imagines of the sampling in the parameter space and boundary of their separation by classes is shown in Fig.5.

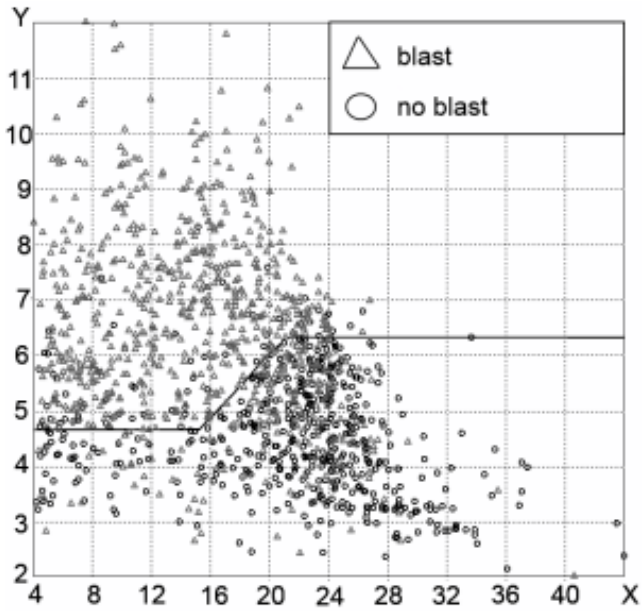

Fig. 5. Distribution of two classes of blast and non-blast cells in 2D space of parameters. Axises correspond to parameters: the first parameter Y=H-Daubechies_2-ReDisp*10^5 is along the vertical axis, the second parameter X=L-Haar-ReEntropy*10^5 is along the horizontal axis.

Note that this is the best result among possible pairs of studied parameters. The first parameter is calculated by equation of the inverse dispersion from a set of wavelet coefficients under trasformation of image with color component of H and wavelet function Daubechies with the second scale of decomposition (H- Daubechies_2-ReDisp) and the second parameter is calculated by the inverse entropy from the wavelet transformation coefficients for image of color component L with usage of the Haar wavelet functions (L-Haar -ReEntropy).

Results of recognition of the blast and non-blast cell types separately for the blast and non-blast cell groups from the selection size are in Table 1.

The results of experiment for recognition of type of blast cell among three types of blasts are presented in Table 2. Here we apply the same parameters as in problem of recognition of blast and non-blast cells.

Table 1

	Blast (%)	no blast (%)
Blast	88	12
No blast	25	75

Values in Table 2 are percent quantity of cells of the first column divided by cells of the first line. Note that sum of values in each line is equal to 100%.

Analysis of data in Tables 2 shows that difference between the blast types is significant less than one between blasts and non-blasts for applied classificator in the sampling. Cells distribution in 2D space of parameters for three types of blast is shown on Fig.6.

Table 2

	Lympho Blast (%)	Myelo Blast (%)	Mono Blast (%)
lymphoblast	54	17	29
myeloblast	17	63	20
monoblast	10	32	58

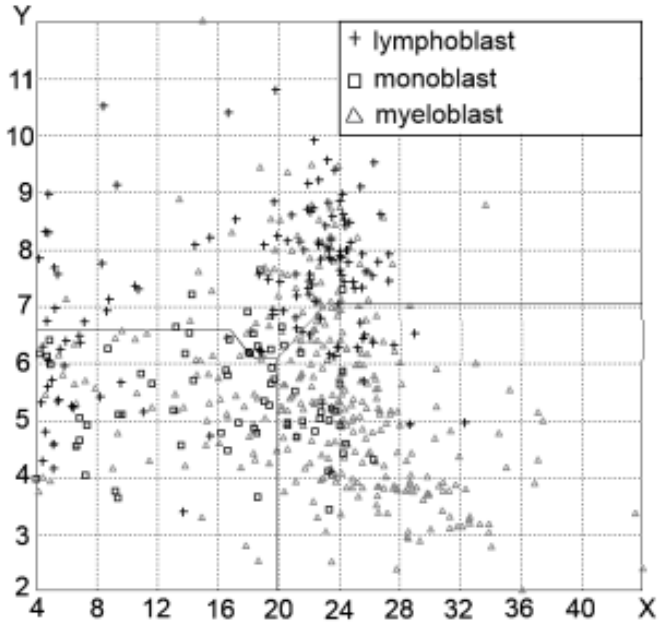

Fig.6. Cells distribution in 2D space of parameters for three types of blast. Yellow- lymphoblast; purple - myeloblast; blue – monoblast. Axises correspond to parameters: the first parameter Y=H-Daubechies_2-ReDisp*10^5 is along the vertical axis, the second parameter X=L-Haar-ReEntropy*10^5 is along the horizontal axis.

IV. Conclusion

The recognition of blast cells is one of the key problems of image processing of blood smears in the diagnosis of acute leukemia.

The analysis of the experiments shows that the assumed mathematical models for description of the blood cells by the wavelet analysis are the effective means for the automated blast cell classification problem solution.

In spite of a part of cells (18%) from the sampling was classified wrong, the obtained result characterizes the method positively. The sampling of the non-blast cells was formed from cells visually similar to the blasts. That is why the result of 82% of right recognitions is satisfactory in these conditions.

On the one hand the obtained error in experiment can be explained by significant similarity of cells of different types

against each other in the sampling. On the other hand the used parameters in experiment are not enough sensitive to type of the precognitive cells. Moreover, perhaps it is required a big number of parameters for providing of high quality recognition. Therefore, the further studies have to search high-sensitive parameters, to find out effective classifiers with greater dimensionality of the parameter space, to investigate variability of the blood cells characteristics of different types, and to increase size of the sampling of different cell types.

REFERENCES

[1] Schiffman, FJ. Hematologic pathophysiology / editor, Fred J. Schiffman; with 16 additional contributors. Philadelphia: Lippincott-Raven, 1998. (Lippincott's pathophysiology series). ISBN: 0397515367.

[2] Saraswat, M; Arya, K. "Automated microscopic image analysis for leukocytes identification: A survey". *Micron*. Vol.65, October 2014, pp.20–33.

[3] Alferez, Santiago; Merino, Anna; Bigorra, Laura; Mujica, Luis; Ruiz, Magda; Rodellar, Jose. "Automatic Recognition of Atypical Lymphoid Cells From Peripheral Blood by Digital Image Analysis". *American Journal of Clinical Pathology* Vol. 143, No. 2, 2015, pp. 168-176.

[4] Mohapatra, Subrajeet; Patra, Dipti; Satpathy, Sanghamitra; Jena, Kumar, Rabindra; Sethy, Sudha. "Automated morphometric classification of acute lymphoblastic leukaemia in blood microscopic images using an ensemble of classifiers". *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, No.12, 2014, pp.: 1-14.

[5] Nikitaev V. G., Nagornov O. V., Pronichev A. N., Polyakov E. V., Dmitrieva V. V. "The Use of the wavelet transform for the formation of the quantitative characteristics of the blood cells images for the automation of hematological diagnostics". *WSEAS Transactions on Biology and Biomedicine.* Vol. 12, 2015, pp. 16-19.

[6] Ahmad Chaddad, Camel Tanougast, Abbas Dandache, Ahmed Bouridane, "Extracted Haralick's Texture Features and Morphological Parameters from Segmented Multispectrale Texture Bio-Images for Classification of Colon Cancer Cells", *WSEAS Transaction on Biology and Biomedicine*, Vol. 8, No. 2, 2011, pp. 39-50.

[7] Ampilova, Natalia; Soloviev, Igor. "On Application of Entropy Characteristics to Texture Analysis", *WSEAS Transaction on Biology and Biomedicine*, Vol. 11, 2014, pp. 194-202.

[8] Yao-Ming Yu, "Detecting and Locating of Brain Abnormality in MR Images Using Texture Feature Analysis and Improved Probabilistic Relaxation Methods", *WSEAS Transaction on Biology and Biomedicine* , Vol. 10, No. 2, 2013, pp.57-66.

[9] Nikitaev V. G., Nagornov O. V., Pronichev A. N., Polyakov E. V., Sel`chuk V. Yu., Chistov K. S., Blindar V. N., Dmitrieva V. V., Gordeev V. V. "Model of Description of Leukocytes of Peripheral Blood Based on the Optical Features of the Structure of Nuclei". *Measurement Techniques.*Vol. 57, No. 5, 2014, pp. 560-563.

[10] James E. Whitney, II and Amanda K. Nebel, "Wavelet Analysis and Autoregressive Modeling of the Electromyography Signal". *International journal of biology and biomedical engineering*, 2014, Vol.8, pp.70-78.

[11] Awadhesh Pachauri and Manabendra Bhuyan. "Wavelet Transform Based Arterial Blood Pressure Waveform Delineator". *International journal of biology and biomedical engineering*, 2012, Issue 1, Volume 6, 2012, pp.15-25.

[12] Nikitaev V. G.; Nagornov O. V., Pronichev A. N., Polyakov E. V., Sel`chuk V. Yu., Chistov K. S., Blindar V. N., Dmitrieva V. N., Zaitsev S. M., Gordeev V. V. "Study of the effectiveness of using wavelet analysis in data-acquisition systems for diagnosis of acute leukemias". *Measurement Techniques.* Vol. 57, No. 10, 2014, pp. 1203-1208.