"Innovations in Science and Technologies" scientific electronic journal
www.innoist.uz
ISSN : 3030-3451
Volume 1
№ 1
February, 2024

DIAGNOSIS OF MICROSCOPIC BLOOD SAMPLES FOR EARLY DETECTION OF WHITE BLOOD CELL DISEASES

Eshmurodov D.E[1], Tulaganova F.K[2].
[1]TATU professor, [2]TATUdoctoral student.
e-mail:tulaganovafotima043@gmail.com

Annotation: The immune system is one of the most important human systems, it fights against all diseases and protects the body from viruses, bacteria, etc. White blood cells (WBC) play an important role in the immune system. To diagnose blood disorders, doctors analyze blood samples to describe the characteristics of WBCs. The characteristics of WBC are determined based on the chromatic, geometric and textural characteristics of the WBC nucleus. Manual diagnosis is subject to many errors and different opinions of experts and takes a long time; however, artificial intelligence techniques can help solve all these problems. Automatic diagnosis of WBC type helps hematologists to identify different types of blood disorders. This work aims to overcome manual diagnosis by developing automated systems for classification of microscopic blood sample datasets for early detection of WBC diseases. Several proposed systems have been used

Key words: X-ray films, Machine Learning, diagnosis, performance, algorithms, differentia, dental panoramic.

Blood contains a great amount of information that can be used to evaluate and analyze a person's health. It consists of 45% red blood cells (RBCs) and 55% plasma, in addition to less than 1% WBCs (WBC) and platelets [1]. Plasma transports nutrients, proteins, hormones, and minerals to all body parts through blood vessels and removes harmful elements in the form of waste. The RBC rate ranges between four to six million per microliter. Hemoglobin is the most important component of RBCs, hauling oxygen to all regions of the body [2]. Platelets are responsible for blood clotting and range from 150,000 to 450,000 per microliter in a normal person [3]. The WBC rate ranges from 4500 to 11,000 per microliter in a normal person. WBCs are the main component of immune cells that is responsible for providing the body immunity to fight viruses and resist diseases, as well as protecting the body from infections such as fungi, bacteria, and viruses [4]. WBCs are created by the bone marrow, lymphoid tissue, and some important glands. There

are five types of WBCs: eosinophils, lymphocytes, monocytes, neutrophils, and basophils. An increase or decrease in the count of these WBCs causes various chronic and fatal diseases [5]. Additionally, many diseases, such as bacteria, leukemia, immunodeficiency syndrome, and infections, appear due to an increase or decrease in four types of WBC: eosinophils, lymphocytes, monocytes, and neutrophils. The number of images available for basophils is few due to their low incidence (0–3%). Due to this lack of images and their low importance, we will focus on the other four types of WBCs in this work. Increased neutrophil cells in the blood can occur due to bacteria, endotoxins, exotoxins, and fungi [6,7]. Lymphocyte cells increase due to diseases such as hepatitis, whooping cough, viruses, bordetella, leukemia, and brucellosis; other diseases, such as HIV, chickenpox, rubeola, and tuberculosis, reduce lymphocyte cells. Malaria, listeriosis, and viral and bacterial infections cause an increase in monocyte cells [8]. Eosinophil cells increase due to diseases such as allergies, parasites, and atopic diseases [9]. The WBC type and number are diagnosed with a blood test (hemogram) on a peripheral smear. The examination is based on spreading blood on a microscope slide and then evaluating the WBCs under the microscope [10]. For ease in the diagnosis of blood samples under a microscope, WBCs are divided into polynuclear (eosinophils, neutrophils, and basophils) and granulocytic (lymphocytes and monocytes) cells. Common procedures to calculate the count of WBC types include determining the type and location of micrographs and the shape and color of each cell. Thus, the process requires great effort, takes a long time, and is prone to errors caused by blood experts due to their differing diagnostic opinions. Computer-aided automated diagnostic systems can help to solve these challenges, reduce manual errors, and obtain reliable diagnostic accuracy

The dataset contains all slide images containing noise caused when the blood samples were mixed with components such as Wright's stain or a methylene blue or eosin (red) dye mixture. The dataset also contains all slide images containing noise resulting from the diversity of microscope devices, from their accuracy and optical reflections, or from the methods of storing the dataset. These noises pose a challenge for obtaining an ultra-accurate diagnosis of the input images. Therefore, the first step in biomedical image processing is pretreatment; in this study, the noise was removed, and WBC edge contrast was increased with two filters: median and Laplacian.

First, the images were optimized with an average filter of 4 × 4 pixels and passed sequentially to process all image pixels. The average filter smooths the image by eliminating disparities between nearby pixels and replacing each center (target)

pixel with an average of 15 neighboring pixels. The process is carried out continuously until the whole image is processed. Equation (1) shows how the average filter works.

$$Average\,(M) = \frac{1}{L}\sum_{i=0}^{M-1} y\,(M-1) \qquad (1)$$

where *Average* (*M*) denotes the optimized image (output), $y(M-1)$ denotes the previous input, and *M* denotes the image's pixel count.Because of the blurred edges between WBCs and other cells, the Laplacian filter was used, which detects the edges of WBCs, shows them clearly, and distinguishes them from other blood cells. The Laplacian filter's action mechanism on the region of interest (WBCs) is described in Equation (2)

$$\nabla^2 f = \frac{\partial^2 f}{\partial^2 x} + \frac{\partial^2 f}{\partial^2 y} \qquad (2)$$

where $\nabla^2 f$ refers to a second order differential equation, and *x* and *y* refer to the coordinates in 2D matrices.

Lastly, to obtain an enhanced and clear image, the image produced using the Laplacian filter is subtracted from the image produced using the averaging filter, as described in Equation (3)

Imege optimized $=Average(M)-\nabla 2\ f$

$$LBP\,(x_c, y_c)_{R,P} = \sum_{P=0}^{P-1} s\left((g_p - g_c)\cdot 2^p\right) \qquad (4)$$

where g_c denotes the central pixel, g_p denotes the neighboring pixel, *R* denotes the radius around the central pixel, and *P* denotes the number of neighbors. The binary threshold function *x* is defined according to Equation (5)

$$s\,(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \qquad (5)$$

Secondly, the GLCM algorithm is an algorithm used to extract texture features from the ROI (WBCs). The algorithm distinguishes between a smooth texture and a rough texture using spatial information; the texture is smooth when the pixels are of similar values. In contrast, the texture is coarse when the pixels are of different

values. The pairwise correlations between pixels are determined by the distance d and directions θ between 1 pixel, where θ represents four directions: 0°, 45°, 90°, and 135°. The relationship between distance and directions is when d = 1, then θ between pixels is θ = 0 or θ = 90; when d = √2, then θ between pixels is θ = 45 or θ = 135. This algorithm extracted 13 representative features.

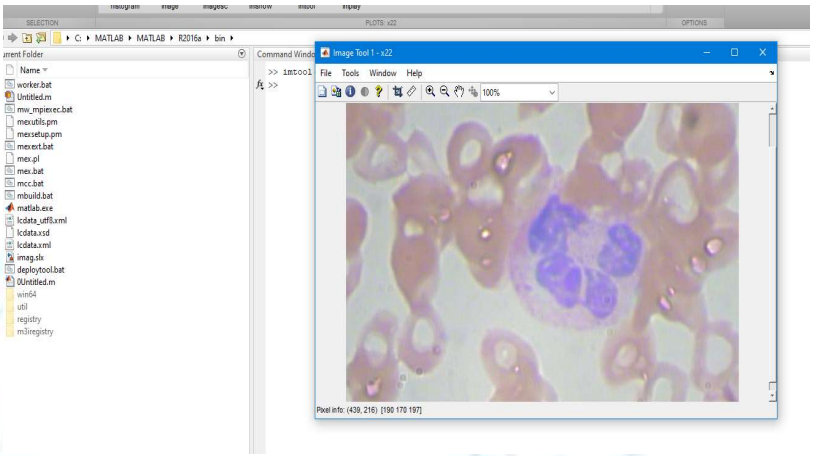

Fig 1: Cancer image output

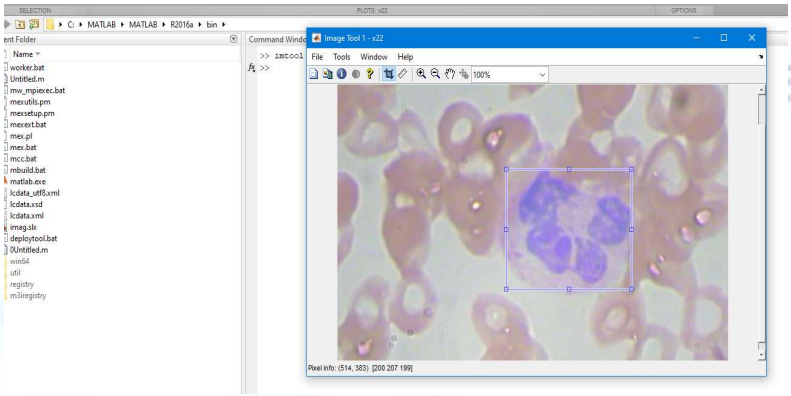

Fig 2: Determining the disease in the image

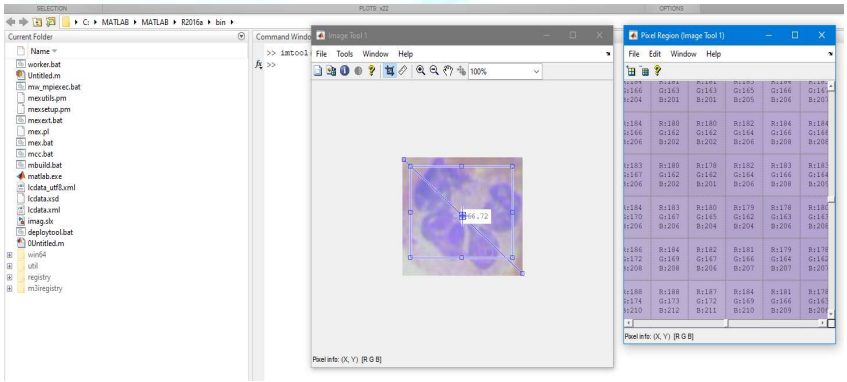

Fig 3. diagnosis of blood cell imaging

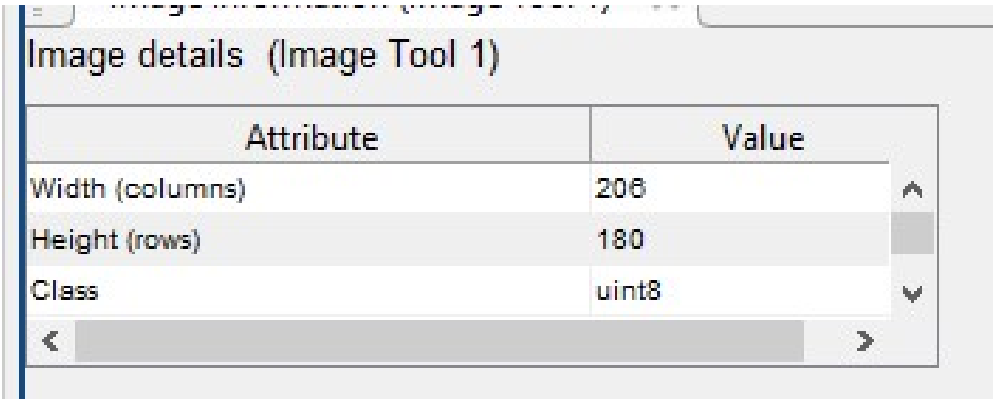

Fig 4. describe the image

$$\text{Accuracy} = \frac{TN + TP}{TN + TP + FN + FP} * 100\% \quad (13)$$

$$\text{Precision} = \frac{TP}{TP + FP} * 100\% \quad (14)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} * 100\% \quad (15)$$

$$\text{Specificity} = \frac{TN}{TN + FP} * 100\% \quad (16)$$

$$AUC = \frac{\text{True Positive Rate}}{\text{False Positive Rate}} = \frac{\text{Sensitivity}}{\text{Specificity}} \quad (17)$$

where the true positive (TP) is the unhealthy WBCs that have been correctly diagnosed, true negative (TN) is the healthy WBCs from correctly diagnosed normal patients, false negative (FN) is the blasted WBCs diagnosed as normal, and false positive (FP) is a normal WBC count diagnosed as blasted WBCs.

Conclusions

When hematologists perform a blood test, if the results of the tests are related to problems with the immune system, the doctor orders a peripheral smear test, which contains essential information about the health of the immune system and related diseases. When tests are analyzed by a doctor, all subtypes of WBCs are also examined. These procedures are performed manually and are, therefore, prone to errors and are time-consuming. Artificial intelligence techniques can help to solve these challenges. This study presented several analytical systems to diagnose microscopic blood sample datasets with high efficiency. Because the microscopic

blood samples contain noise, the average and Laplacian filters were applied to increase the contrast of WBCs and remove the noise. The proposed methods are divided into three types, and each type contains more than one algorithm. The initial method proposed includes neural network algorithms (ANN and FFNN) based on the hybrid features extracted using the LBP and GLCM algorithms. These two algorithms achieved promising results for dataset diagnosis. In the second proposed method, the dataset was diagnosed using four pre-trained CNN models: AlexNet, ResNet-50, GoogLeNet, and ResNet-18. The dataset was diagnosed using the transfer learning method to extract deep features. All models achieved superior results for the early detection of WBC diseases. In the third proposed method, the dataset was diagnosed using hybrid CNN models and SVM techniques, including AlexNet with SVM, ResNet-50 with SVM, GoogLeNet with SVM, and ResNet-18 with SVM. These techniques consist of two blocks, the first of which uses CNN models to extract feature maps. SVM is used to classify deep features in the deep and second block. For the detection of specific WBC diseases, all hybrid technologies produced superior results.

In future work, this dataset will be classified with another dataset using hybrid systems based on fused CNN features. Handcrafted features will be combined with the CNN features to achieve superior accuracy when classifying WBC species.

References

1. Bhuvan, C., Bansal, S., Gupta, R., & Bhan, A. (2020, February). Computer based diagnosis of malaria in thin blood smears using thresholding based approach. In *2020 7th International conference on signal processing and integrated networks (SPIN)* (pp. 1132-1135). IEEE.

2. Yue, G., Zhang, S., Li, Y., Zhou, X., Zhou, T., & Zhou, W. (2023, October). Subjective quality assessment of enhanced retinal images. In *2023 IEEE International Conference on Image Processing (ICIP)* (pp. 3005-3009). IEEE.

3. Kandhari, R., Bhan, A., Bhatnagar, P., & Goyal, A. (2021, February). Computer based diagnosis of Leukemia in blood smear images. In *2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)* (pp. 1462-1466). IEEE.

4. Guan, Y., & Wang, Z. (2022, August). Blood Cell Image Recognition Algorithm based on EfficientNet. In *2022 IEEE International Conference on Mechatronics and Automation (ICMA)* (pp. 1640-1645). IEEE.

5. Hafeez, M. A., Imran, A., Khan, M. I., Khan, A. H., Nawaz, A., & Ahmed, S. (2022, May). Diagnosis of Liver Disease Induced by Hepatitis Virus Using Machine

Learning Methods. In *2022 8th International Conference on Information Technology Trends (ITT)* (pp. 154-159). IEEE.

6. Elhassan, T. A. M., Rahim, M. S. M., Swee, T. T., Hashim, S. Z. M., & Aljurf, M. (2022). Feature extraction of white blood cells using CMYK-moment localization and deep learning in acute myeloid leukemia blood smear microscopic images. *IEEE Access*, *10*, 16577-16591.

7. Eshmurodov D. Iskanderova S. Tulaganova F . (January 29-31, 2024) Algorithms for detection of cells in blood images