A Survey on Feature Extraction for Leukemia Detection in Blood Microscopic Images

P. Aiswariya[1]; S. Manimekalai[2]

[1]Research Scholar, PG & Research Department of Computer Science, Theivanai Ammal Collage of Women, Thiruvalluvar University, Tamil Nadu, India.

[1]aispadman@gmail.com

[2] Head & Assistant Professor, PG & Research Department of Computer Science, Theivanai Ammal Collage of Women, Tamil Nadu, India.

[2]mamekaroshini@gmail.com

Abstract

Leukemia is a kind of blood cancer that can harm a person's organs severely. Numerous types of blood cells, such as WBC, RBC, Plasma, Platelets, can be seen in microscopic photographs of blood. WBC can be classified into various types and subtypes. The three main types of WBC are monocytes, lymphocytes, and granulocytes. Monocytes are differentiated into dendritic cells and macrophages, among other types of cells. Different strategies for detecting WBC have been proposed in this research. Segmenting, detecting groups of leucocytes, and extracting their characteristics are some of the approaches used.

**Key-words:** Leukemia, WBC, Feature Extraction.

1. Introduction

There are different approaches to detect WBC have been proposed in this research. Segmenting, detecting groups of leucocytes, and extracting their characteristics are some of these strategies [3]. WBCs, also known as leukocytes, are required for the diagnosis of a variety of disorders. They can be obtained in a variety of ways. They assist in the treatment and prevention of a variety of illnesses. It's a type of malignancy that affects the blood and bone marrow. White blood cells can be produced by aberrant cells. Acute leukemia is always lethal, whereas chronic leukemia progresses slowly. The term "acute" refers to the disease's quick progression [5].

To enable earlier diagnosis and reduce the need for manual screening, an automated blood smear approach for detecting leukemia is required. It can be successfully treated if caught early enough. The purpose of this paper is to explore how to build image segmentation algorithms for acute leukemia. Many efforts have gone into developing technologies to aid in the categorization and picture segmentation of acute leukemia. The process of extracting features from a huge amount of data is known as feature extraction. This method is employed to classify data and enhance image processing [7]. The literature on image segmentation and characteristic identification for the Acute Myelogenous Leukemia (AML) will be reviewed in this paper.

Leukemia is caused by a decrease in the amount of blood cells produced by the bone marrow. This can happen when the bone marrow's supplies are depleted. This symptom could be indicative of a viral or bacterial infection. Fever, a higher risk of bacterial infection, coughing, shortness of breath, chest pain, vomiting, and indications of anemia are all signs of infection [11]. It takes longer and is more accurate depending on the operator's ability. When it comes to cancer treatment, automatic identification of leukemia is really useful.

Figure 1 - A Blood Sample Suffering from Leukemia
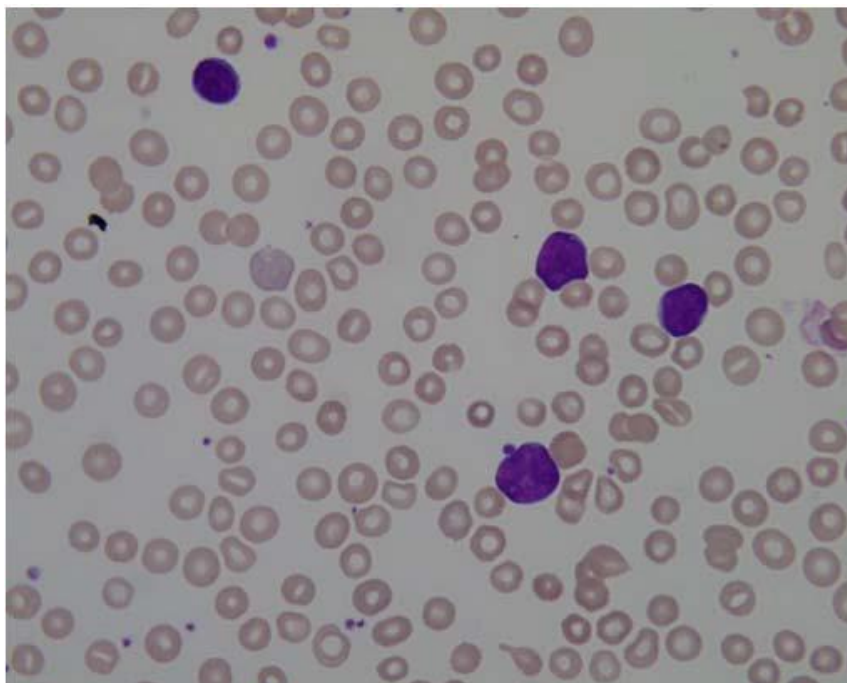
The prime goal of this study is to classify the characters of Leukemia centred on its type and its classification technique. In this field, various techniques are used to identify the cancer cells that

are most likely to cause cancer. The paper reviews the various advantages and disadvantages of these techniques.

2. Related Work

According to a literature review, statistical methodologies can be utilized to predict a problem. The goal of this study is to advance a technique for detecting leukemia cells utilizing an existing image processing technique. The approach is mostly used to classify diseases based on their characteristics.

Dr. Sajith V., et al. [9] discussed Automatic leukemia identification in human blood samples using machine learning and reached an accuracy of (96.67%), which is comparable to accuracies obtained in previous studies such as deep convolutional networks (97.5 percent). Strong Segmentation and Measurements Methods of White Cells in Blood Microscope Images was discussed by Fabio Scotti, et al [7], and superior findings were obtained. Chandranand presented a new method for feature extraction based on the GLCM technique. They employed the Herlev Dataset and got a 94 percent accuracy [14]. For automated leukemia detection, Samanta S., et al presented Fuzzy based blood image segmentation A [10]. Sinha, et al. [18] proposed automated differential blood count and obtained an accuracy of 95%. (96 percent). P. Aiswariya and S. Manimekalai, et al. suggested a Global and Local Entropy Based Segmentation Model for Detecting Leukemia in Blood Images with an accuracy of (84.64%) [6].

3. Pre-processing

Some of the images that were taken may have blurriness and/or noise. In this section, we are trying to minimize these issues so that the images are appropriate for the further footstep in the processing. Various pre-processing techniques are using to enhancing the contrast, remove noise, and isolate regions.

4. Feature Extraction

Feature extraction is a process that takes large amounts of data and turns it into a reduced representation. It works by taking advantage of the features of the existing data.

A. Fractal Dimension

Fractals were utilized in different scientific domains to examine the features of space prior to the emergence of statistical measurements. The Hausdorff dimension is a fractal dimension theory feature. It is split into three sections: Renyi, HD, and packing dimension.

- The color of each nucleus is transformed to a binary picture.
- Using the canny approach, the edge boundary of the generated image is retrieved.
- Over the edges, an R-square grid is placed, and the occupied squares are counted.
- In step 3, the number of squares that can be increased.
- As a result, the Hausdorff Dimension HD can be defined as follows:

$$HD = \frac{log(R)}{(\log R(g))}$$

The number of boxes or squares in the overlay grid is denoted by R. R(s) denotes the number of occupied squares.

B. Local Binary Patterns (LBP)

LBP is a crucial component of computer vision. It's used to categorize textures. The LBP feature vector is made in the following way:

- There are 16x16 cells in the window being studied.
- Divide each pixel's neighbors by 8 for each pixel in a cell.
- Pixels are arranged in a circle.
- Write "1" if the values of the central pixel are greater than the neighbors' values, else "0."
- Use the histogram to visualize the frequency of each number.
- Normalize the histogram.

C. Texture Features

A texture is a part of an image that describes various details. There are various methods for extracting texture features from an image. The first approach is to transform an image into a frequency domain, and then extract its features.

Figure 2- Comparison of Texture Features
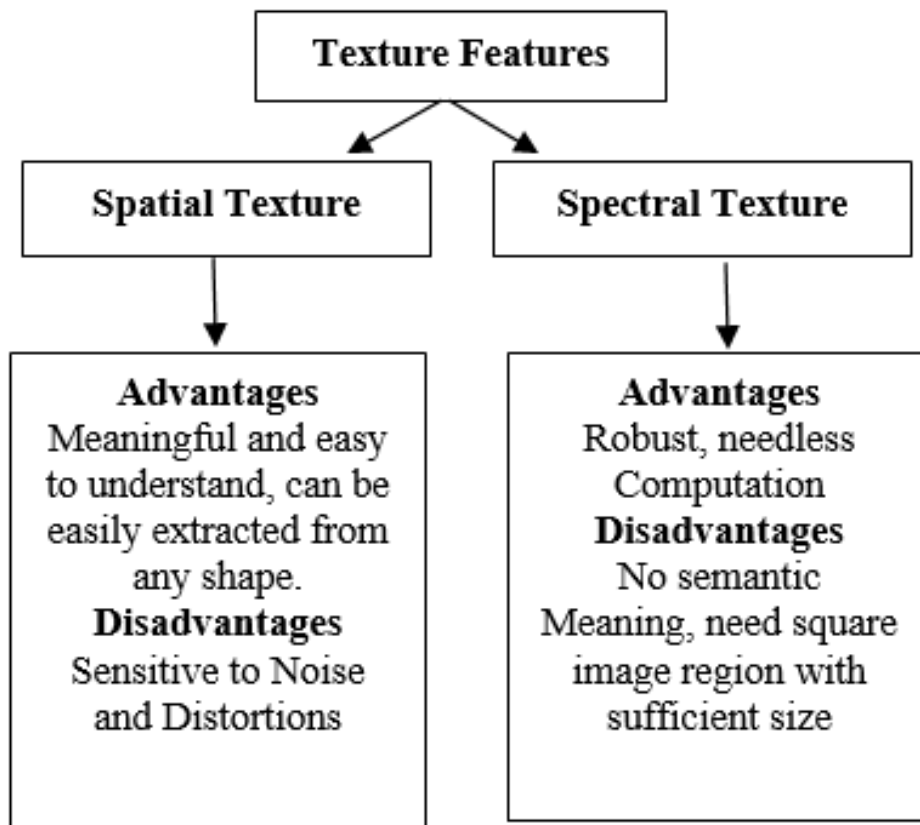
D. Shape Features

Shape descriptors are numbers that describe a given shape. They are subdivided into two groups: contour and region based. We will discuss two techniques that can extract features from regions by shaping them according to their shape. In addition, this topic will also discuss the relative locations of regions and their absolute locations.

E. Color Features

It is possible to define it based on the model and the color space chosen. LUV, RGB, and HMMD are some of the color spaces available. The color histogram, Color Moments, Color coherence vector, and color correlogram are all important color properties. In Table II, many color coding systems are shown. The benefits and drawbacks of each strategy are listed in the tables below. The dominant color descriptor, CSD, and SCD, respectively, are denoted by DCD, CSD, and SCD.

Table 1 - Comparison of Color Descriptors

Color method	Advantages	Disadvantages
Histogram	Easy to use, simple to compute and understand	Sensitive to noise, high dimension, no spatial info
CM	Compact, robust	Not enough to describe all colors, no spatial info
CCV	Spatial info	High dimension, high computation cost
Correlogram	Spatial info	Very high computation cost, sensitive to noise
SCD	Compact on need, scalability	No spatial info, less accurate if compact
CSD	Spatial info	Sensitive to noise, rotation and scale
DCD	Compact, robust, perceptual meaning	Need post-Processing for spatial info

5. Conclusion

This paper discussed various feature extraction methods used for leukemia detection various techniques related to the study of advantages and disadvantages are also presented. Finally, this paper covered various aspects of research, including the various features and guidelines for future work.

References

S.K Mishra, Leukemia diagnosis based on machine Learning Algorithms, *IEEE,* 2020.

Jasmine Minj, Leukemia Disease Detection and classification using Machine Learning Approaches: *A Review, IEEE,* 2020.

Mohamed Loey, Mukdad Naman and Hala Zayed, Deep Transfer Learning in Diagnosing Leukemia in Blood Cells, *MDPI,* Vol. 9, 2020.

Komal Nain Sukhia et al., "Automated acute lymphoblastic leukaemia detection system using microscopic images", *IET Image Process,* vol. 13, no. 13, pp. 2548-2553, 2019.

Pouria Mirmohammadi, *Automatic recognition of acute lymphoblastic leukemia using multi-SVM classifier, 115*(8), 2018.

P. Aishwariya and Manimekalai. S., "Global and Local Entropy Based Segmentation Model for Detecting Leukemia in Blood Images", *Annals of R.S.C.B,* vol. 25, pp. 1914-1926, 2021.

Xiaofu Huang, Ming Chen, Peizhong Liu, Yongzhao Du, Texture Feature-Based Classification on Transrectal Ultrasound Image for Prostatic Cancer Detection, *Computational and Mathematical methods in Medicine,* 2020.

Shafique, S, Tehsin S, Acute lymphoblastic leukemia detection and classification of its subtypes using pre-trained deep convolutional neural networks, *Technol. Cancer Res. Treat,* Vol. 17, pp. 1-7, 2018.

Catarina Barata, M. Emre Celebi, Jorge S. Marques, A Survey of Feature Extraction in Demoscopy Image Analysis of Skin Cancer, *IEEE,* Vol. 14, 2016.

Mohapatra, Saurav, Dipti Patra, Sudhakar Kumar, and Siddhartha Satpathi. "Kernel induced rough c-means clustering for lymphocyte image segmentation." *IEEE,* 1-6, 2012.

Kun Ding hung, "A computer assistant method for leukemia neucleus segmentation and recognition in blood smear images.", *ACM,* Vol. 85, 2012.

Abdul Nasir, A.S., M.Y. Mashor, and H. Rosline. "Unsupervised colour segmentation of white blood cell for acute leukaemia images", *IEEE,* 142-145, 2011.

Raje, Chaitali, and Jyoti Rangole, "Detection of Leukemia in microscopic images using image processing." *IEEE,* 255-259, 2014.

Y. Jia, E. Shelhamer, J. Donahue, and S. Karaye Caffe, Convolutional architecture for fast feature embedding, *In Proc. of MM,* 675, 2016.

F. Xing and L. Yang, Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: a comprehensive review", *IEEE reviews in biomedical engineering,* 9, 234- 263, 2016.

G. Castellano, L. Bonilha, L.M. Li, and F. Cendes, Texture analysis of medical images", *Clinical Radiology,* 59(12), 1061-1069, 2004.

S. Mohapatra, S. Samanta, D. Patra, and S. Satpathi, Blood image segmentation for automated leukemia detection", *in Proc. ICDeCom,* pp. 1-5, 2011.

S. Mohapatra, D. Patra, and S. Satpathi, Acute leukemia detection for Image analysis of blood microscopic images", *in Proc. IECR,* pp. 215-219, 2010.

S. Mohapatra, D. Patra, and S. Satpathi, Automated cell nucleus segmentation and acute leukemia detection in blood microscopic images", *in Proc. ICSMB,* pp. 49-54, 2010.

N. Sinha and A.G. Ramakrishnan, Blood cell segmentation using EM algorithm, *In Proc. 3rd Indian Conf. Comput. Vis,* 445-450, 2002.

Sonali C. Sonar U. An Efficient techniques for white blood cells. *International journals of science and research,* Vol. 6, 2015.

Suma Varsha, Sushma, "Detection of White Blood cells Nuclei using automatic segmentation", *IJREAM,* 2018.

S. Mohapatra, S. Samanta, D. Patra, and S. Satpathi, Fuzzy based blood image segmentation for automated leukemia Detection, 1-5, 2011.

Sos Agaian, Monica Madhukar and Anthony T. Chronopoulos, *Acute Myelogenous Leukemia Detection in Blood Microscopic Images using Automated Screening System,* 2014.

Gokhan Sengual, White Blood cell classification by SURF image matching, *PCA and Dedogram, Biometric Research,* Vol. 6, 2006.

N. Sinha and A.G. Ramakrishnan, Blood cell segmentation using EM algorithm, *in Proc. 3rd Indian Conf,* 445- 450, 2011.

Jha K.K and Dutta H.S. Mutual Information based hybrid model and deep learning for Acute Lymphocytic Leukemia classification in single cell blood smear images, *Computer methods and programs in biomedicine, 179,* 104-987, 2019.

Zhao J, Zhang M, Zhou Z, Chu J and Cao F, Automatic classification and classification of leukocytes using convolutional neural networks, *Medical & biological engineering & computing, 55*(8), 1287-1301, 2017.

Sahlol A.T, Kollmannsberger P and Ewees A.A, Efficient classification of white blood cell leukemia with improved Swarm optimization of deep features, *Scientific Reports, 10*(1), 1-11, 2020.

Putri R.R, Mandyartha E.P and Indrawanti A.S, Automatic Identification of Acute Lymphoblastic Leukemia on Blood Cell An image Using Geometric Features, *IOP Conference Series: Materials Science and Engineering, 462*(1), 2019.

Tuba E, Strumberger I, Bacanin N, Zivkovic D and Tuba M, Acute Lymphoblastic Leukemia Cell classification in Microscopic Digital Images Based on Shape and Texture Features, *International Conference on Swarm Intelligence,* 142-151, 2019.