A REVIEW ON WHITE BLOOD CELL IMAGE ANALYSIS FOR LEUKEMIA DISEASE CLASSIFICATION USING DEEP LEARNING TECHNIQUES

[1]I.Vinurajan, [2]Dr. K. P. Sanal Kumar, [3]Dr. S. Anu H Nair

Research Scholar, Department of Computer and Information Sciences, Annamalai University, Chidambaram, India[1], Assistant professor PG Department of Computer Science, R.V Government Arts College, Chengalpattu, India[2], Assistant professor Department of CSE, Annamalai University, Chidambaram, India [Deputed to WPT, Chennai[3]

nivinrajan01@gmail.com[1], sanalprabha@yahoo.co.in[2], anu_jul@yahoo.co.in[3]

ABSTRACT

Blood cancer is one of the most common and dangerous cancers. Leukemia is a form of blood cancer caused by uncontrolled and abnormal White Blood Cell (WBC) production in the bone marrow. Early diagnosis of leukemia, gives the chance to cure cancer with the right treatment. Counting the amount of white and red blood cells (RBC) is one approach to detect leukemia. A specific equipment known as a haemocytometer is used to count white blood cells and red blood cells in the traditional method. These tests are time-consuming and extremely complicated, resulting in misclassification.

The detection of cancer cells can also be enhanced by image processing of microscopic images, which is inexpensive due to the circuitry's simplicity. Several researchers have discovered cancer cells, but they failed to examine all criteria at the same time, such as image enhancement, noise reduction, image identification, and so on, resulting in a reduction in accuracy. The purpose and aim of this study is to develop a new system that takes into account all of the above factors. To detect the existence of cancer and cancer stages, the process includes different stages including image acquisition, image pre-processing, image segmentation, edge detection, and feature extraction. The paper helps in measuring the number of white blood cells and red blood cells, as well as their average cell sizes (regular/irregular), which can be used to detect the existence of cancer.

Keywords: blood cell count, disease diagnosis, leukemia disease, cancer detection and white blood cell.

INTRODUCTION

Blood is necessary for life, and many body organs rely on it to function properly. Analyzing the blood constituents can be used to determine the healthiness of blood (i.e., cells). In most cases, blood contains cells as well as plasma, a liquid component [1]. Blood cells make up around 45% of the total amount of blood, while plasma makes up the other 55%. The red blood cells (erythrocytes), white blood cells (leukocytes), and platelets (thrombocytes) are the three types of blood cells [2]. The red blood cells (RBC) make up 40–45% of the blood, while the white blood cells (WBC) make up roughly 1% of the blood. The three types of blood cells each serve a different purpose for the body's organs. White blood cells are formed in the bone marrow and are an essential component of the blood. White blood cells are in charge of the body's immune system, which acts as a defense against foreign elements in the body, particularly disease-causing elements [3]. As shown in Figure 1, Neutrophils, eosinophils, lymphocytes, monocytes, and basophils are among the five types of white blood cells. Similarly, Granulocytes and agranulocytes (nongranulocytes) are two types of blood cells that can be further split; see Figure 2.

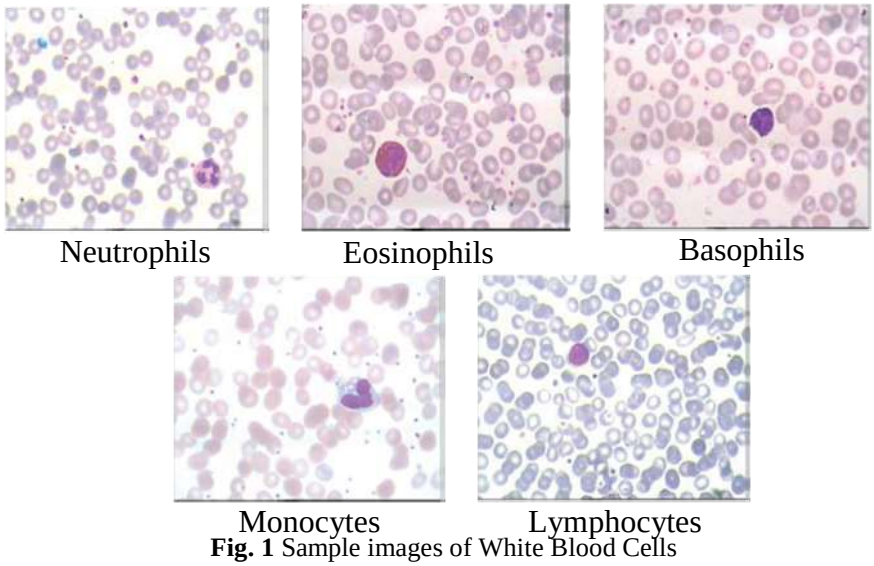

Fig. 1 Sample images of White Blood Cells

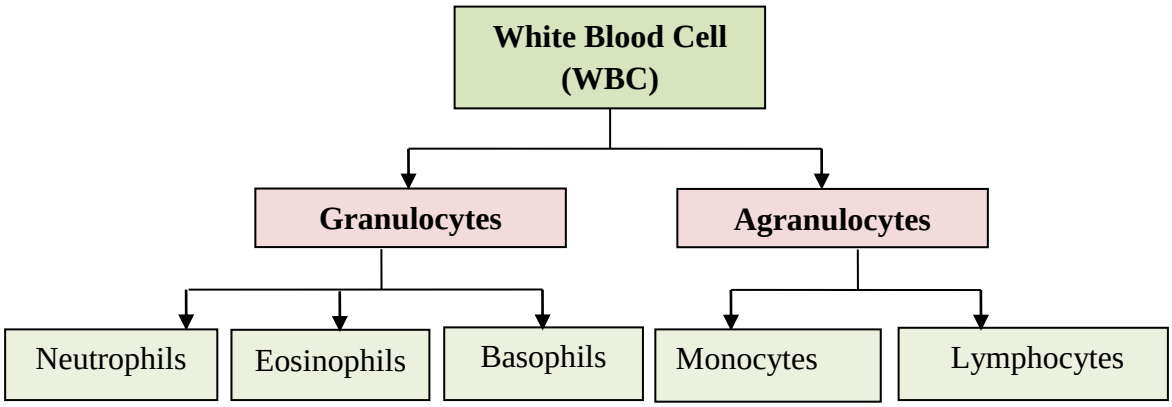

Fig. 2 Types of White Blood Cell

Granulocytes are white blood cells with visible granules, whereas agranulocytes are white blood cells with no visible granules under a microscope [4]. The granulocytes class includes neutrophils, eosinophils, and basophils, while the agranulocytes class includes monocytes and lymphocytes. The percentages of neutrophils, eosinophils, lymphocytes, monocytes, and basophils in the blood are 40–60%, 1–4%, 20–40%, 2–8%, and 0.5–1%, respectively [5].

The five types of white blood cells each have diverse functions and represent different aspects of a patient's health (subjects). Proper identification, in particular, allows for the counting of various white blood cells to assess their presence in the correct or predicted proportions. Different white blood cells can also be separated after being identified for a more thorough investigation for abnormalities [6]. White blood cell analysis, both quantitative and qualitative, provides a lot of information about a patient's health. For example, it is possible to monitor patients for diseases such as leukemia, immune system abnormalities, and malignant cells [7]. Traditionally, identification has been done in a laboratory setting, when images of blood cells are collected, dyed with particular chemicals, and then evaluated under a microscope by a specialist. This technique, however, is sensitive and requires that the human professional make no examination errors. Unfortunately, after several hours of testing, specialists can become exhausted and make incorrect white blood cell identifications.

The remaining sections of the paper are organized as follows: Section 2 is a survey of the research literature in the field of automated detection systems for leukemia identification and classification, as well as related work, with a concentration on image segmentation and classification, that has been done over the years. Section 3 examines dataset collection and uses data augmentation techniques to improve the dataset images. Section 4 focuses on current research in deep learning techniques, while Section 5 focuses on the research's conclusions and future scope.

LITERATURE SURVEY

Many researchers have developed machine learning-based classification and diagnosis of leukemia diseases using white blood cell counts images in the last few decades. For example, Hiremath [8] used a variety of segmentation approaches, including k-mean clustering and the Expectation Maximization algorithm. Farnoosh Sadeghian et al. [9] provided a framework for extracting the nucleus and cytoplasm region in white blood cells utilizing active contour, snake algorithm, and zack thresholding. Morphological approaches to detect cancer cells have been proposed by Yujie, Yuhki, and Kohei [10]. Counting and segmenting for distinct blood cells has been proposed by Vinod and Kimbahune, neleh [11]. Red blood cell counting and extraction have been proposed by Poomeokrakl and Neatpisamvanit [12]. Fatin A.Dawood [13] proposed an unsupervised nucleus segmentation of microscopic white blood cells images using a histogram equalization technique.

Similarly, Red Blood Cell Classification by Support Vector Machine was proposed by Navi D Jambhekar [14]. With the Circular Hough transform approach, Nasrul and Muhammad [15] locate red blood cells. Based on the Zack's Method, Heidi Berge [16] proposed segmenting red blood cells in a thin blood smear image. In the urine microscopy, Guitao et al. [17] proposed the hough transform for detecting and collecting red blood cells.

Kendall Preston [18] addressed the current state of the art in blood smear analysis automation as well as other related fields such as multi-resolution microscopy. For segmenting and border detection of cells in images of peripheral blood smear slides, Nicola Ritter [19] introduced Otsu adaptive thresholding and watershed transform. Roy Dimayuga et al. [20] employed histogram thresholding to separate the nuclei of leukocytes or white blood cells from the rest of the cells in the image. On the basis of thresholding, Madhuri and Patil [21] presented the otsu adaptive thresholding in segmenting and extracting the WBC in the blood cell image. Sumeet [22] developed a watershed segmentation algorithm for RBC counting and used the Circular Hough Transform approach to quantify the quantity of red blood cells. Nur Alom [23] employed wavelet transformation to partition different blood cells and then a fuzzy inference system to make a final conclusion on blood cancer based on the number of cells counted. Fauziah Kasmin et al. [24] proposed that leukemia could be detected in a microscopic image. Sharif J.M et al. [25] proposed a method for RBC segmentation based on the making and watershed algorithms for automated RBC counting.

The most common problem found in many surveys and studies is the failure of over segmentation methods and machine learning models to produce adequate results in terms of accuracy. Furthermore, manual leukemia disease diagnosis takes time and may not produce reliable outcomes. To address the challenges, we must design an automatic leukemia diagnosis classification and detection system using a deep learning model.

1. Dataset Collections

The dataset is an essential component of the deep learning algorithm, it is critical to collect properly labeled data. However, WBCs, particularly bone marrow smears, lack a big open dataset. As a result, a dataset was collected for this research. Seoul National University Hospital's Department of Laboratory Medicine prepared bone marrow smear samples. The Wright-Giemsa staining method was used on the bone marrow aspiration. A light microscope was used to image the prepared slides at a magnification of x1000. A total of 200 images with

a pixel resolution of 1080 x 1330 were collected from 30 slides of 10 human subjects, with five to ten nonoverlapping acquisition spots chosen at random for each slide. The image preparation was done anonymously with data that had already been acquired. The institutional review board of Seoul National University Hospital approved exempt status to the study protocol.

The entire images were manually cropped into single-cell patch images with a resolution of 96 x 96 pixels. Two expert hematologists labeled and validated that the single-cell images were placed in the relevant classes after collecting a total of 2,174 cropped images. The dataset included ten WBC classes in various stages of maturation, including four consecutive stages of the erythroid series pronormoblast (C1), basophilic normoblast (C2), polychromatic normoblast (C3), and orthochromatic normoblast (C4) and six consecutive stages of the myeloid series myeloblast (C5), promyelocyte (C6), myelocyte (C7), metamyelocyte (C8), band neutrophil (C9), and segmented neutrophil (C10). As shown in Figure 3, examples of WBCs at various phases of maturity.

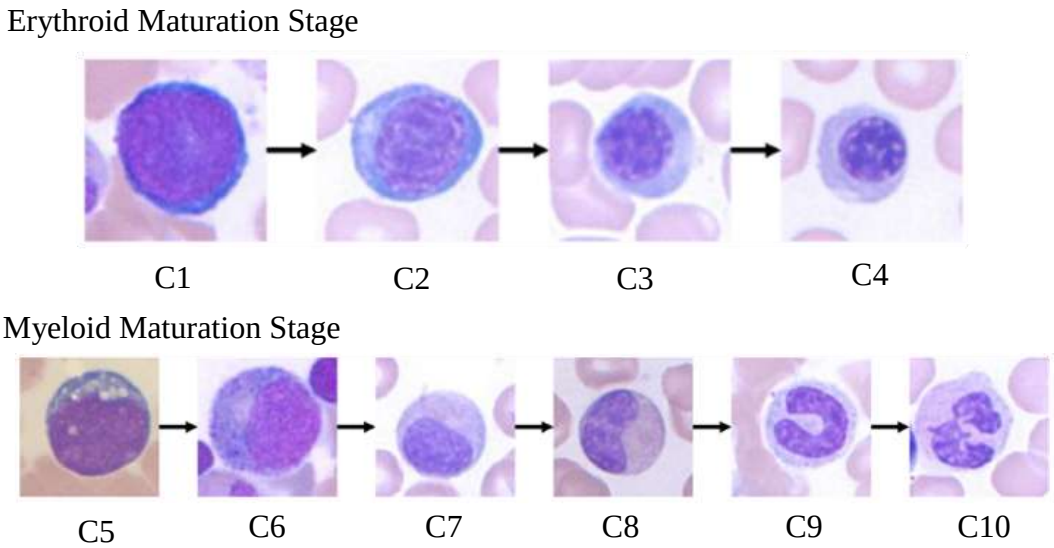

Fig. 3 Various maturation stages of WBCs

The distribution of classes in the collected dataset was unbalanced. Because the natural distribution of white blood cells is uneven, as demonstrated in White blood cell differential count, this problem was unavoidable while collecting the dataset. However, because only a few features are learned from classes with a small number of data, this imbalanced dataset can be an issue while training the network.

Oversampling was used for classes with relatively few data items in an attempt to resolve the problem of imbalanced data caused by the heterogeneous distribution of white blood cells. We manually cropped multiple images of the same cell at slightly different centers during data preparation for classes with little data. This led to a greater number of data points being collected, as well as a greater variety of data.

2. Current Research in Deep Learning Models

Deep learning is a machine learning technique that is used to build a model that learns from data such as images, videos, audio, and text and performs tasks such as classification and detection. Scene classification is used to classify scenes using various scene classification models such as Convolutional Neural Network (CNN) [26] and Deep CNN models [27]. Deep learning-based approaches include CNN, transfer learning (VGG-16 and VGG-19), Recurrent Neural Network (RNN), Auto Encoder, and Generative Adversarial Network. For image classification, CNN and transfer learning models are commonly used. CNNs are extensively used for image segmentation and classification. Despite the fact that CNNs were first created in 1989, their outstanding success in the ImageNet Competition in 2012 garnered them even more attention. The computational complexity of

CNN design is increasing as the number of layers, neurons with millions of weights, and connections between different neurons increases. The basic block diagram of CNN is shown in Figure 4, which includes convolutional, pooling, activation function, and fully linked layers, each of which serves a different purpose.

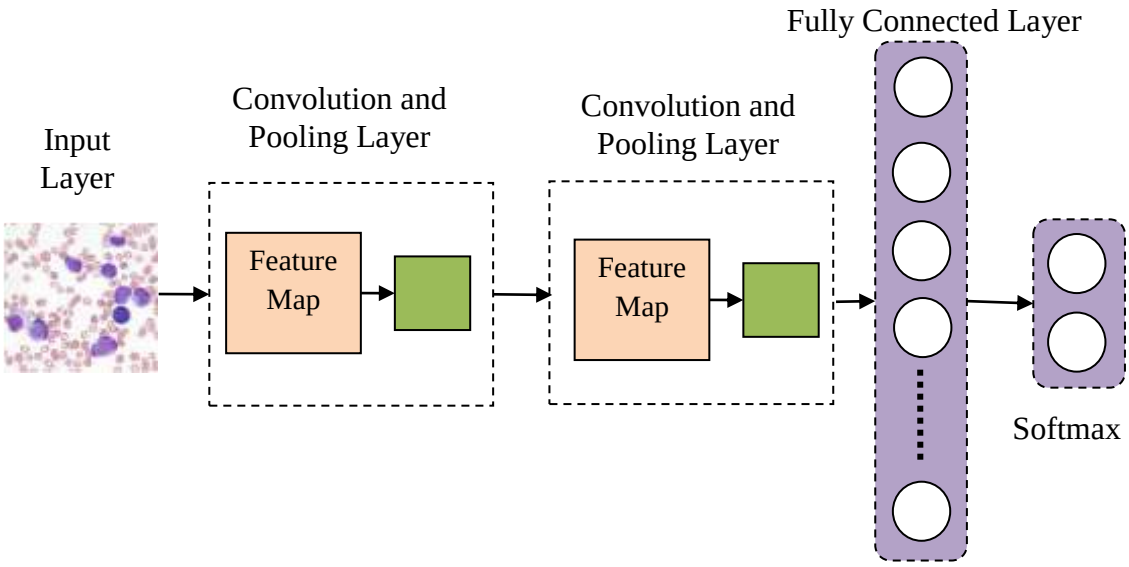

Fig. 4 Architecture of Convolutional Neural Network Model

By convolving input images across the kernel, the convolutional layer generates feature maps. As the value is transferred to the next layer, the results from the preceding convolutional layers are downsampled using the maximum or average of the defined neighborhood in the pooling layer. The loss function is coupled with the remaining layers of CNN model output to give a forecast of the input data. Finally, network parameters are determined by minimizing the loss function between prediction and ground truth labels while maintaining regularization requirements. In addition, backpropagation is used to change the network weights at each iteration until convergence.

CONCLUSION

In this literature review, we have addressed to survey various traditional and machine learning based approaches used for counting the white blood cell production in the bone marrow. One of the methods for detecting leukemia is to count the number of white and red blood cells. In the traditional technique, white blood cells and red blood cells are counted using specialized equipment called as a haemocytometer. These tests are time-consuming and difficult, which leads to misclassification.

There is a great need for further study into developing automated methods to assist pathologists in identifying and classifying leukemia disease, according to the literature. Since, we've noticed a remarkable increase in the number of studies applying automated ways for assessing microscopical smear images for detecting leukemia disease. Deep learning techniques, particularly Convolutional Neural Network and Deep Convolutional Neural Network-based approaches for image classification will be in high demand in recently.

REFERENCES

1. W. King, K. Toler, and J. Woodell-May, “Role of white blood cells in blood-and bone marrow-based autologous therapies,” BioMed Research International, Vol. 2(8), Article ID 6510842, pp. 1-8, 2018.
2. S. Fathima and K. Farhath, “Blood cells and leukocyte culture—a short review,” Blood Research and Transfusion Journal, Vol. 1(2), Article ID 555559, pp. 12-20, 2017.

3. S. Liu, Z. Deng, J. Li, J. Wang, N. Huang, Q. Ye, "Measurement of the refractive index of whole blood and its components for a continuous spectral region," Journal of Biomedical Optics, Vol. 24(3), Article ID 035003, pp. 15-21, 2019.
4. P. Zhou, Z. Meng, M. Liu et al., "The associations between leukocyte, erythrocyte or platelet, and metabolic syndrome in different genders of Chinese," Medicine, Vol. 95(44), Article ID e5189, pp. 12-17, 2016.
5. K. E. Badior and J. R. Casey, "Molecular mechanism for the red blood cell senescence clock," IUBMB Life, Vol. 70(1), pp. 32–40, 2018.
6. C. Herron, "Know your WBCs," Nursing Made Incredibly Easy!, Vol. 10(1), pp. 11–15, 2012.
7. A. Mathur, A. S. Tripathi, and M. Kuse, "Scalable system for classification of white blood cells from Leishman stained blood stain images," Journal of Pathology Informatics, Vol. 4(2), pp. 1–15, 2013.
8. P.S.Hiremath, Parashuram Bannigidad, Sai Geeta, "Automated identification and Classification of White Blood Cells (Leukocytes) in Digital Microscopic Images" IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition"Vol. 1, pp. 1-12, 2010.
9. Famosh Sadeghian, Zainina Seman, Abdul Rahman Ramli, Badrul Hiham Abdul Kahar, and M-iqbal aripan "A Framework for white Blood cell Segmentation Microscopic Blood Images using Digital Image Processing", Vol. 5(1), pp. 121-128, 2012.
10. Yujie L, Yuhki Kitazono, Kohei Miyata, Shota Nakashima Shiyuan Yang, enchi erikawa, "An improved Detection Algorithm Baed on Morphology Methods for Blood Cancer Cell Detection", Pediatric Radiology, Vol. 2(5), pp. 643–650, 2014.
11. Vinod V.Kimbahune, Mr. Nelesh j Uke, "Blood Cell Image segmentation and Counting", Int. J. Adv. Res. Sci. Eng. Technol., Vol. 3, pp. 136–142, 2013.
12. J. Poomeokrakl and C. Neatpisamvanit, "Red Blood cell Extraction and Counting", Procedia Comput. Sci. Vol. 5(8), pp. 535–542, 2015.
13. Fatin A.Dawood, "Unsupervised Segmentation of Microscopic White Blood Cell Image Using Histogram Equalization Technique", Journal of Algorithms Comput. Technol, Vol. 1(3), pp. 1-10, 2014.
14. Navi D Jambhekar, "Red Blood Cells Classification using Image Processing", Bioengineering, Vol. 4(1), pp. 10-20, 2013.
15. Nasrul Humaimi Mahmood and Mohammad Araf Manor,"Red Blood cells Estimation using Hough Transform Technique", IPIJ, Vol.3, No.2. April 2012.
16. Heidi Berge, Dale Taylor, riram Krishnan, ad Tania. Dougla. Improved Red Blood Cell Counting in thin Blood smears. Proceedings of IBI, pp.201-207, 2011.
17. Guitao Cao, Cai Zhong, Ling Li and Jung Dong, "Detection of Red Blood Cell in Urine Micrograph". The 3rd International Conference on Bioinformatics and Biomedical Egineering (ICBBE). 2009.
18. KendallPreton Jr., "Image Processing In Medical Micrscopy', IEEE transaction 1986.
19. Nicola Ritter, Jame Cooper "Segmentation and Border Identification of cells in Images of Peripheral Blood Smear Slides", Journal of Neural Network, Vol. 61, pp. 85–117, 2015.
20. Roy A.Dimayuga, Gerwin T. Ong, Rainier Carlo S.Perez, Gefferson O. Siy, Saman C. Sohrabi Langroudi and Miguel O.Gutierrez, "Leukemia Detection Using Digital Image Processing in mat lab". ECE Student from, De La Salle University, Manila. March 26,2010.

21. Madhuri G.Bnamaro: ”Automatic Blood cell Analysis by using digital Image processing”: A preliminary study. IJERT Vol.2. sep.2013.
22. Sumeet Chourasiya: “Automatic Red blood cell counting using watershed segmentation”, IJCSIT, Vol. 5, 2014.
23. .Nur Alom Taalukdar: “Automatic Blood cancer detection using Image processing based on fuzzy system”, IJARCSSE, Vol. 4(2), Aug.2014.
24. Fauziah Kasmin, Anton satria prabuwono, Azizi Abdullah: “Detection of leukemia in human blood sample based on microscopic images: A study”, Journal of Theoretical and applied information technology , Vol. 46(2), pp. 1-10, 2013.
25. J. M. Sharif, M. F. Miswan, M. A. Ngadi, Md ah Hj alam, Muhammad Mahadi bin abdul Jamil, “Red blood cell segmentation uing masking and watershed algorithm”: A preliminary study, ICOBE, penang Malaysia, 2012.
26. Shijie, J.; Ping, W.; Peiyi, J.; Siping, H. Research on data augmentation for image classification based on convolution neural networks. In Proceedings of the Chinese Automation Congress (CAC), Jinan, China, 20–22 October 2017; pp. 4165–4170.
27. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv 2014, arXiv:1409.1556.