IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Machine Learning-Based Blood Cell Categorization In Smear Images

[1]Rohit Sonawane, [2]Manisha Darak, [3]Abdul Shaikh, [4] Sai Yandralwar,

[1]B.E 4th year, [2]Professor, [3]B.E 4th year, [4]B.E 4th year,

Department of Computer Engineering,

RMDSSOE, Pune-411058, India

Abstract: This research focuses on utilizing machine learning techniques to automate the classification of blood cells into white blood cells (WBCs) and red blood cells (RBCs) from blood smear images. The process involves data preprocessing, feature extraction or deep learning, model training, and evaluation. Various models are explored, including traditional classifiers and convolutional neural networks (CNNs).The model's performance is assessed using standard binary classification metrics. Successful models can be deployed in healthcare for efficient blood cell classification, contributing to streamlined diagnostics. Ethical considerations regarding data privacy are integral to this process. This research aims to provide a valuable tool for healthcare professionals, enhancing the speed and accuracy of blood cell analysis. Keywords – Blood Cell Categorization, Machine Learning , Smear Images, CNN , Image Processing.

***Index Terms* – Blood cell categorization, Medical imaging, Image processing, Smear Images, Machine learning**

I. Introduction

Proposed a machine learning model for classifying blood cells in smears that includes disease detection. Achieved high accuracy in cell type identification and disease diagnosis, improving clinical diagnosis [1]. Blood is a constantly circulating fluid that is made up of formed cells. The process involves the circulation of RBCs, WBCs, and Platelets. The process known as Hematopoiesis takes place in the bone marrow as shown. WBCs are the largest cells and are present in small numbers. The white blood cells have a nucleus surrounded by cytoplasm and their size can range from 8-20µm in diameter, depending on the specific type of WBC. The normal RBC is biconcave discshaped cell with central pallor. They do not have a nucleus and appear pink in color. It measures about 7-8µm in diameter and more in number. Any deviation in size, volume, or shape of red cells represents abnormal RBCs. Platelets are cytoplasm fragments of megakaryocytes, normally 2-3µm in diameter, bluish and second numerous A blood smear is often used as a follow up test to abnormal results on a CBC[2]. Features are data descriptors used to describe data elements such as classification and clustering. A comprehensive understanding of WBC features is critical for differentiating between various types and subtypes of leukemia. Current methods used for WBC detection, segmentation, and classification face several challenges, although they are performed using automatic and manual approaches. Manual detection of WBCs is conducted by pathologists and is typically subject to human error and produces inaccurate results. This process is tedious, time-consuming, and subject to inter and intra-class variations among pathologists. Only 76.6% of the cases showed agreement between pathologists during leukemia diagnosis. [3].

CNN is a well-developed and widely used computer vision technique that can be effectively used for various image data processing tasks. Many medical imaging applications have successfully implemented pre trained neural networks such as ResNet, VGGnet, and Inception. CNN also leveraged transfer learning, which involves training large generic datasets and then training them on a specific classification for a smaller dataset. This is a common problem with medical datasets. Various techniques and algorithms have been proposed to detect the classifications of leukemia. While there are still limitations in this domain, the challenges presented by the current work inspire this study. [4]. Using a generic model, the proposed DCNN predicts the type of cancer using a small dataset. To generalize, data augmentation is used. A comparative analysis shows that the proposed DCNN outperforms the current stateof-the-art CNN models on the dataset thus retrieved and prepared. With less computational time and trainable parameters the proposed model outperforms existing machine learning and predefined deep learning models in predicting the type of cancer [5]. White blood cells (WBCs), also known as leukocytes, are crucial for the human immune system and are produced in bone marrow. They can be classified into lymphocytes, monocytes, and neutrophils. Classifying WBCs is crucial for medical diagnosis and early detection of diseases like cancer. The proposed method uses WBC images to classify different types. A new framework uses CNN, transfer learning, and R-CNN mask techniques to automatically calculate and classify blood cells in blood smear images. It aims to find, predict boundaries, and count red and white blood cells, using R-CNN mask for object detection and segmentation. [6].

This field aims to develop a system capable of classifying 10 abnormal red blood cells using the Decision-Tree Algorithm and the Radical Basis Function Network. The system, which includes input image, pre-processing, and feature extraction, achieved an accuracy rate of 83.3% in classifying these cells. [7]. This study uses VGG-16 deep learning to classify blood cells in smear images using bounding boxes. A backbone model and CNN model were proposed, with improvements in accuracy. The model was tested on Kaggle datasets, resulting in 98.08%, 96.34%, and 84.52% improvements.[8].

II. RELATED WORK

This study provides a concise summary of previous research in the area of automated blood cell segmentation and classification. In this study, a new object detection method utilizing the YOLO algorithm was tested to detect and count blood cells. The method achieved an accuracy of 96.1% for red blood cells (RBC) and 86.89% for white blood cells (WBC) using 364 annotated images. The iterative circle method was used to segment WBC and RBC, achieving an average accuracy of 98.4% for WBCs and 95.3% for RBCs in 100 images. Cruz et al. proposed a method for RBC counting using blob analysis and watershed transform in the Hue Saturation value (HSV) component. Achieved a 96% average accuracy with 10 blood samples. Wei and colleagues Developed a technique to distinguish between white blood cells and red blood cells using hue and saturation components. The work's author achieved an accuracy of 92.9% for 100 Wright-Giemsa stained images. The author suggested a method for blood cell counting by utilizing Nearest Neighbor and SVM techniques and manually cropping each cell. With 368 images from the ALL-IDB dataset, this method achieved an average accuracy of 98% for RBCs and 99.2% for WBCs. Please rewrite the following passage in a different way. Thank you. We introduced a deep learning semantic segmentation approach for RBC and WBC segmentation, achieving 94.93% and 91.11% accuracy respectively. Additionally, we presented a Faster R-CNN model for classifying WBC and RBC variants, achieving a 98% accuracy for the BCCD dataset. While deep neural networks and several machine learning algorithms are leveraged in the categorization of blood cells. We have developed a straightforward and reliable approach for differentiating between white blood cells (WBC) and red blood cells (RBC) in images with multiple stains.

2.1 "Classification of blood cells into white blood cells and red blood cells from blood smear images using machine learning techniques"

The automatic classification model improves the hematological procedures, quickens the diagnosis process and enhances the accuracy of the evaluation process. Thus in this paper, we used a semi-automatic method to segment and classify blood cells into White Blood cell (WBC) and Red Blood Cell (RBC). Texture features of a cell are extracted using Gray Level Co-occurrence Matrix (GLCM) and fed to the classifiers like Naive Bayes classifier, K-nearest neighbors, decision tree, K-means clustering, random forest, logistic regression, ANN and SVM. The performance parameters are compared and found that the logistic regression is best suited for the work with the 97% accuracy

2.2 "Feature Extraction of White Blood Cells Using CMYK-Moment Localization and Deep Learning in Acute Myeloid Leukemia Blood Smear Microscopic Images"

Artificial intelligence has revolutionized medical diagnosis, particularly for cancers. Acute myeloid leukemia (AML) diagnosis is a tedious protocol that is prone to human and machine errors. In several instances, it is difficult to make an accurate final decision even after careful examination by an experienced pathologist. However, computer-aided diagnosis (CAD) can help reduce the errors and time associated with AML diagnosis. White Blood Cells (WBC) detection is a critical step in AML diagnosis, and deep learning is considered a state-of-the-art approach for WBC detection. However, the accuracy of WBC detection is strongly associated with the quality of the extracted features used in training the pixel-wise classification models. CAD depends on studying the different patterns of changes associated with WBC counts and features. In this study, a new hybrid feature extraction method was developed using image processing and deep learning methods. The proposed method consists of two steps: 1) a region of interest (ROI) is extracted using the CMYK-moment localization method and 2) deep learning-based features are extracted using a CNN-based feature fusion method. Several classification algorithms are used to evaluate the significance of the extracted features. The proposed feature extraction method was evaluated using an external dataset and benchmarked against other feature extraction methods. The proposed method achieved excellent performance, generalization, and stability using all the classifiers, with overall classification accuracies of 97.57% and 96.41% using the primary and secondary datasets, respectively. This method has opened a new alternative to improve the detection of WBCs, which could lead to a better diagnosis of AML.

2.3 “Lightweight EfficientNetB3 Model Based on Depth wise Separable Convolutions for Enhancing Classification of Leukemia White Blood Cell Images”

Acute lymphoblastic leukemia (ALL) is a form of cancer that occurs when there is an overproduction of immature white blood cells (WBCs) in the bone marrow. Recently, various classification models utilizing machine learning (ML) and deep learning (DL) algorithms have been developed. The research study suggests a new DL-assisted robust model using the EfficientNet-B3 and depth wise separable convolutions to classify acute lymphoblastic leukemia and normal cells in the white blood cell images dataset. The proposed model is lightweight and aims to improve accuracy in the classification task. The new and lighter EfficientNet-B3 utilizes fewer trainable parameters in order to improve the efficiency and performance of leukemia classification. Additionally, we are considering two publicly available datasets to assess how effective and applicable the proposed lightweight EfficientNetB3 is. Additionally, our discovery indicates that the suggested lightweight EfficientNet-B3 model is trustworthy and applicable for supporting clinical research and healthcare professionals in the detection of leukemia.

2.4 “Automatic Detection of White Blood Cancer From Bone Marrow Microscopic Images Using Convolutional Neural Networks”

Leukocytes, which are produced in the bone marrow, comprise approximately one percent of the total blood cell count. The excessive proliferation of these white blood cells results in the development of blood cancer. The proposed study offers a strong method for categorizing Acute Lymphoblastic Leukemia (ALL) and Multiple Myeloma (MM) using the SN-AM dataset, in the field of cancer classification. Acute lymphoblastic leukemia, also known as ALL, occurs when the bone marrow produces an excessive amount of lymphocytes, a type of white blood cell. Conversely, Multiple myeloma (MM) is a distinct type of cancer that results in the accumulation of cancer cells in the bone marrow instead of their release into the bloodstream. By utilizing convolutional neural networks, the proposed model eliminates the potential for errors in the manual process using advanced deep learning techniques. Therefore, the model can effectively be utilized as a tool for identifying the cancer type present in the bone marrow.

2.5 “Blood Cell Classification using Neural Network Models”

Blood cells classification is a crucial aspect in medical diagnosis. Several machine learning models have been proposed under various researches for classification of blood cells in recent years. However, traditional machine learning algorithms are limited in the accurate detection of abnormal cells. In this study, we propose deep learning based approach for blood cell classification and evaluate the efficiency of multi-layer neural network model built for the classification of the various types of White Blood Cells using Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) in combination. The proposed method leverages the strengths of both CNN and RNN and gives better results

2.6 “An Automated Blood Cells Counting and Classification Framework using Mask R-CNN Deep Learning Model Guided”

Bioengineering is the art of applying engineering principles, techniques, and technologies to biology and medicine for general healthcare applications. Analyzing human biological samples such as blood, has become essential for physicians to diagnose and follow diseases evolution. Traditional blood cells counting techniques used in laboratories are time consuming and laborious. They can lead to inaccurate results due to the human intervention in this complicated process. In this paper, we propose an automated blood cells counting framework using convolutional neural network (CNN), instance segmentation, transfer learning, and mask R-CNN techniques. Red and white blood cells are identified, classified, and counted from microscopic blood smear images. The obtained results reveal highly detection rate of different blood cells. In addition, unlike other state-of-theart techniques, our proposed method has the ability to identify overlapped and faded cells.

2.7 “Image Classification of Abnormal Red Blood Cells Using Decision Tree Algorithm”
The study aims to increase the number of abnormal red blood cells that can be detected using image processing. This study used Decision-Tree Algorithm as a machine learning algorithm in classifying. As a result, the system detected and classified a total of ten abnormal red blood cells. Images used in the system came from hospitals’ past patients. In addition, a camera is used to capture the slides. The image was then inserted into the program. The system processed and classified the image. the results show the name of the abnormal red blood cells detected in the image within the system including the soft copy of the list.

2.8 “Detection of RBCs, WBCs, Platelets Count in Blood Sample by using Deep Learning”
A whole blood cell count is an essential check in scientific diagnosis to assess common fitness conditions. Blood cells are traditionally counted manually using a cell counting chamber (hemocytometer) in conjunction with various laboratory compounds and solvent (chemical) compounds, which is timeconsuming and very dull to experiment with. In this work, a machine learning system mastering the method of automated computerized counting of blood cells is proposed. By using machine learning and deep learning techniques, the blood cells and their counts can be identified with the best accuracy compared to the other existing techniques. The CNN is used for the image classification. One of the best techniques for achieving the best accuracy in the least amount of time for the blood cells dataset is VGG-16 technique. The proposed system is a combination of the CNN and VGG-16 methods. The learnt models are generalized for the training and testing of the different datasets. In general, this computer-aided system of detecting the blood cells is more useful for practical applications. According to the results of this research, the accuracy of counting blood cells is in the range of 90 to 95 percent.

2.9 “A Comparative Evaluation of Deep Learning Methods in Automated Classification of White Blood Cell Images”
The identification and characterization of a patient's blood sample are required for the diagnosis of blood-related disorders. As a result, the medical implications of automated methods for identifying and categorizing various kinds of blood cells are considerable. However, deep convolutional neural networks (CNN) and standard machine learning algorithms have performed well in the categorization of blood cell pictures. Red, White, and Platelets are all types of blood cells. Leucocyte, commonly known as the immune cell, is a type of blood cell that plays a vital part in human immune function. Depending on shape info and granulated data in leukocytes, white blood cells are usually split by hematologists into two different categories: non granular cells (lymphocytes and monocytes) and granular cells. The CNN portion receives the pre-trained weight parameters from the image dataset using the transfer learning approach. Also, We have used two different scenarios, the first scenario of using CNN directly gave us pictures. used SVM in the second scenario. Then we compare the best category results. The classification results showed that the accuracy of CNN is 98.4%, while the accuracy of Support Vector Machine (SVM) is 90.6%. Other classifications can be added to the proposed system to improve its performance.

2.10 “Classification of white blood cells from microscopic images using CNN”
White Blood Cells also known as leukocytes plays an important role in the human body by increasing the immunity by fighting against infectious diseases. The classification of White Blood Cells, plays an important role in detection of a disease in an individual. The classification can also assist with the identification of diseases like infections, allergies, anemia, leukemia, cancer, Acquired Immune Deficiency Syndrome (AIDS), etc. that are caused due to anomalies in the immune system. This classification will assist the hematologist distinguish the type of White Blood Cells present in human body and find the root cause of diseases. Currently there are a large amount of research going on in this field. Considering a huge potential in the significance of classification of WBCs, a deep learning technique named Convolution Neural Networks (CNN) will be used which can classify the images of WBCs into its subtypes namely, Neutrophil, Eosinophil,

Lymphocyte and Monocyte. The results of various experiments executed on the Blood Cell Classification and Detection (BCCD) dataset using CNN are reported in this project.

2.11 "Classification of White Blood Cells: A Comprehensive Study Using Transfer Learning Based on Convolutional Neural Networks"
White blood cells (WBCs) in the human immune system defend against infection and protect the body from external hazardous objects. They are comprised of neutrophils, eosinophils, basophils, monocytes, and lymphocytes, whereby each accounts for a distinct percentage and performs specific functions. Traditionally, the clinical laboratory procedure for quantifying the specific types of white blood cells is an integral part of a complete blood count (CBC) test, which aids in monitoring the health of people. With the advancements in deep learning, blood film images can be classified in less time and with high accuracy using various algorithms. This paper exploits a number of state-of-the-art deep learning models and their variations based on CNN architecture. A comparative study on model performance based on accuracy, F1-score, recall, precision, number of parameters, and time was conducted, and DenseNet161 was found to demonstrate a superior performance among its counterparts. In addition, advanced optimization techniques such as normalization, mixed-up augmentation, and label smoothing were also employed on Dense Net to further refine its performance.

III. Conclusion

The availability of high-quality, diverse blood cell images is fundamental to the success of the system. Access to a well-annotated dataset is crucial for training and validation. The necessary technical infrastructure, including computational resources, software libraries, and expertise in computer vision and image processing, must be in place to implement the algorithm effectively. Careful selection of image processing and machine learning algorithms, along with model development and optimization, are essential to achieving accurate blood cell detection. Adhering to ethical considerations and healthcare regulations, such as patient data privacy and safety, is imperative when dealing with medical image analysis. The availability of GPUs or other hardware accelerators can significantly improve the efficiency of the algorithm, making it practical for real-world applications.

IV. Reference

[1] Navya K.T, Keerthana Prasad, Brij Mohan Kumar Singh. " Classification of blood cells into white blood cells and red blood cells from blood smear images using machine learning techniques."

[2] Tusneem Ahmed Elhassan, Mohd Shafry Mohd Rahim, Tan Tian Swee, Siti Zaiton Mohd Hashim and, Mahmoud Ahjurf. "Feature Extraction of White Blood Cells Using CMYKMoment Localization and Deep Learning in Acute Myeloid Leukemia Blood Smear Microscopic Images"

[3] Amreen Batool and Yung Cheol Byun, "Lightweight EfficientNetB3 Model Based on Depth wise Separable Convolutions for Enhancing Classification of Leukemia White Blood Cell Images"

[4] Deepika Kumar, Nikita Jain, Aayush Khurana, Sweta Mittal, Roman Sekerik and Suresh Chandra Satapathy " Automatic Detection of White Blood Cancer From Bone Marrow Microscopic Images Using Convolutional Neural Networks"

[5] Jagrit Mitra, Kartik Vijayran, Kartikeya Verma & Anurag Goel." Blood Cell Classification using Neural Network Models".

[6] Najmeddine Dhieb, Hazim Ghazzai, Hichem Besbes, Yehia Massoud Wei Qi Yan "An Automated Blood Cells Counting and Classification Framework using Mask R-CNN Deep Learning Model Guided.'

[7] Virgnia Mari E. Batitis, Merwin Jhan G. Caballes, Abigail A. Ciudad, Miccaela D. Diaz, Russel D. Flores and Roselito E. Tolentino." Image Classification of Abnormal Red Blood Cells Using Decision Tree Algorithm".

[8] Allaparthi HemaSri, Mopuru Devi Sreenidhi , Valluru Venkata Krishna, Gamidi Vasanth, V.Murali Mohan and T.Satish "Detection of RBCs, WBCs, Platelets Count in Blood Sample by using Deep Learning".

[9] Hersh Abdulrahman Muhamad, Shahab Wahhab Kareem, and Amin Salih Mohammed, "A Comparative Evaluation of Deep Learning Methods in Automated Classification of White Blood Cell Images".

[10] J.Setu Sai Sowmya Kumari, K,Chandra Sekhara Chari, M.Leelavathi, K.Pavan Kumar and, Mrs.N.Rajeswari "Classification of white blood cells from microscopic images using CNN".

[11] Thinam Tamang, Sushish Baral, May Phu Paing. "Classification of White Blood Cells: A Comprehensive Study Using Transfer Learning Based on Convolutional Neural Networks"