International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 04 | April - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

CLASSIFICATION OF WBC USING CONVOLUTION NEURAL NETWORKS

[1]Prof.Prathima L, [2]Ms.Ashwini P Patel, [3]Ms.Bhavana V, [4]Ms.Keerthi P, [5]Ms.Prajna H R
Dept. of Information Science and Engineering, JNNCE, Shimoga, India

ABSTRACT

Leukocytes, developed within the cartilage of the bone, account for barely 1% of the overall blood cell counts. Erratic flourishing of leukocytes induces an outbreak of blood cancer. Amongst three of the diverse sorts of cancer in blood, the suggested ponder provides a vigorous instrument for the sorting of subtypes of leukemia and multiple myeloma, utilizing the related dataset. White blood cells with leukemia are not normal that grow throughout cells present in the red blood. WBCs, and platelets and affect the blood and bone marrow. Whereas, multiple myeloma is a different type of cancer that affects plasma cells. It develops in the bone marrow instead of the blood stream. The suggested method uses a deep learning technology called as convolutional neural networks to lessen the likelihood of errors occurring during the human method. The model first extracts the leading highlights from the cell imaging by pre-processing it. Next, the model will be prepared using CNN, and lastly, the cancer type can be predicted. Furthermore, the model's accuracy of 97.33% is higher than Yolov8's and Naive Bayes.

Keywords: Acute Lymphoblastic Leukemia (ALL), Acute Myelogenous Leukemia (AML), Chronic Lymphocytic Leukemia (CLL), Chronic Myelocytic Leukemia (CML), Multiple Myeloma (MM), Deep Learning, CNN

I. INTRODUCTION

Platelets, leukocytes, and haemoglobin are the three primary types of cell found in the blood of humans. They penetrate the bloodstream after their manufacture in the bone marrow. The rapid expansion of erroneous cells that impede ordinary blood cell formation is an important factor in the development of blood malignancies. Leukocytes, often referred to as white blood cells, constitute vital aspects of the body's immune response and defend the human organism from sensitivities and ailments. Nevertheless, malignancy may originate from an overabundance of white blood cells.

Blood malignancies can be identified as three primary varieties. Such a kind of malignancy that strikes the blood along with bone marrow is called ALL. The immune system produces an abundance of immature white blood cells as a consequence. ALL is a conceivably fatal condition that is very hazardous. It strikes adolescents alongside youngsters and accounts for around 25% of pediatric cancer cases. There are two varieties exists in chronic leukemia: CLL and CML.

An alternative form of cancer that is generated within the marrow is called multiple myeloma, which affects the plasma cells. Older people are normally affected. The disease known as lymphoma affects the body's lymphatic system. An overabundance of leucocytes results in an increase in their growth.

A proposed model suggests using computers to identify white blood cell cancer employing related data and computer programs. Identifying several varieties of leucocytes is really important for finding out if someone has blood or immune system problems or certain types of cancer. The way blood cancer is treated depends on how old the patient is, what kind of cancer they have, how fast it is spreading, and where it is in their body [1].

The research method uses both computerized and hands-on methods to count and group white blood particles. Although manual operations require competent employees and take considerable amounts of time, they are exceedingly precise. Despite the fact that computerized approaches are more rapid, they are not error-proof. CNN-based methods are used by deep learning to extract features from raw information. This helps to address crises in the monetary, medical, and numerous other spheres. The model analyzes the input and extracts information by combining both deep learning and image processing. To find important places, the first method is CMYK-moment localization. Following that, a technique known as CNN-based feature fusion is applied. By classifying items using some deep learning approaches, this strategy reduces errors that come from human methods. The model's functionality is evaluated using a range of measures. It uses special computer programs to learn and determine which kind of malignancy the cells contain. With just a modest quantity of data, a new model can predict some cancer types. To further enhance the estimations, data augmentation is applied

Ultimately, the most recent automated method for identifying leukocyte cancers is less costly and more quick to classify. It eliminates the need for professionals and the challenges associated with performing it manually. By increasing the speed and accuracy of cancer detection, deep learning may lead to better patient outcomes and services.

II. RELATED WORK

Numerous investigations have gone through the identification of leukemia having acute lymphoblastic syndrome utilizing deep learning and machine learning methods. These studies use pre-trained CNNs and different feature-highlighting approaches, and the results show improvements in the correctness of classifying benign and malignant cases [2]. Leukemia detection by using image analytics and classification algorithms uses pre-processed cell images, along with some methods like K-means clustering for segmentation purposes and neural network for classification of leukemia, which display improvements in automatic digital systems for better detection [4]. For leukemia recognition, an automated method using deep CNNs is proposed [5]. It includes pre-processing and data augmentation, which improve generalization, as well as a weighted ensemble of models based on metrics of performance like the F1-score and AUC.

CNN models consist of three layers, i.e., convolutional layers, pooling layers, and dense connected layers employed for feature extraction, spatial reduction, and classification, respectively. These layers are used in the leukemia classification study to examine both acute and chronic leukemia. Performance is evaluated using metrics such as accuracy and precision [6]. An image-based classifier includes both the KNN and SVM classifiers with structural traits from wavelet and curve-let curves that were been relied on to boost leukemia detection. It leverages machine learning tactics for image processing and demonstrates encouraging accuracy for prompt diagnosis [7]. In order to conquer the issues of poor

generalization and slow convergence in previous models, a lightweight, robust system relied on EfficientNet-B3 has been used for ALL detection [8]. This demonstrates the manner in which methods from deep learning are applied in order to boost leukemia diagnostic accuracy.

A hybrid approach integrating deep neural networks of convolution layer, adaptive multipurpose CAT algorithms, and methods that process the images is currently pursued to improve bone marrow cancer diagnosis, with a focus on Acute Myeloid Leukemia (AML) and Multiple Myeloma (MM). The intension is to mitigate the manual process and enhance the ability to designate the cancer [9]. For the purpose of detecting acute lymphoblastic leukemia (ALL) using blood smear images, a customized deep learning encoder was constructed. This process involves preprocessing, model architecture design based on CNNs, and evaluation using standard metrics, highlighting efforts towards accurate diagnosis through advanced computational techniques [10].

III. PROPOSED METHODOLOGY

Utilizing sophisticated deep learning strategies, the suggested technique appears to be a multi-step preparation for identifying and classifying leukemia images. A summary of each step is provided below.

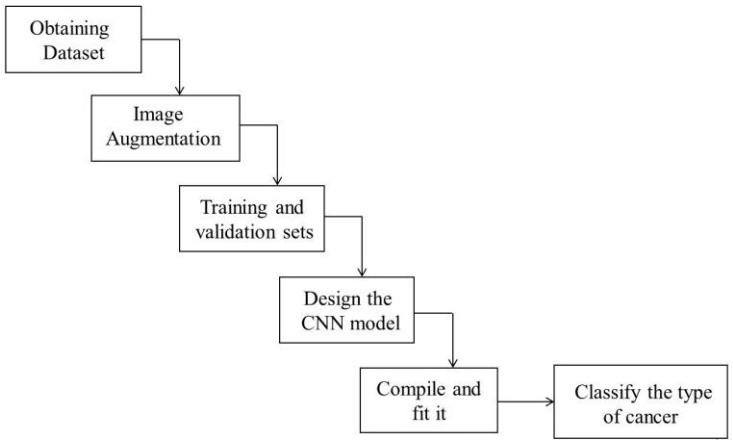

Figure 3.1: Proposed Methodology

A. Description of the cell images of the dataset:

A dataset collection's two distinct subsets are where the dataset for the proposed study was obtained. Images of leukemia patients i.e., those with ALL, AML, CLL, and CML make up the first segment of the dataset. In contrast, images of patients with multiple myeloma make up the second portion of the collection.

B. Augmentation of cell images data:

Through rotating, rescaling, shearing, reordering, and extracting the highlights, the gathered dataset is enhanced. For the model to perform successfully throughout the assessment (testing) stage, there must

of the collected dataset be an astounding amount of information available, i.e., in varying sizes, points, and postures. By using various techniques to enhance photographs, information can be created from already-existing data. Each of the image related to dataset have been shifted by a factor of 90 since our framework ought to be versatile in identifying characteristics displayed in regardless of direction. In the absence of data augmentation, overfitting occurs, which makes the model difficult to extrapolate the examples that weren't included in the training set.

C. Training, Validation and Testing:

Upon rearranging the data, it is split into groups for training and testing. Training datasets are groups of labeled information that are used to teach deep learning models. They help models understand patterns and features better in order to predict or do tasks. The validation dataset is employed to see if the computational model is appropriately nurtured or not. Testing datasets are distinct sets of information used to see how well the model works. The testing datasets are barely exploited when training the framework; rather they are employed for verification of how accurately the model can predict new data.

D. CNN Architecture:

A deep learning CNN consists of three layers: a convolutional layer, a pooling layer, and a fully connected (FC) layer. From the layer that does the convolutions to the fully connected layer, the CNN gets more complicated. CNN can find more and more detailed parts of an image and gradually figure out what the whole object is.

Convolution Layer: This is the first layer of a CNN where majority of the calculations happen. Another layer of convolution can come following the first layer. The process of convolution uses a filter to scan the image and see if a certain feature is there. The kernel moves across the whole picture many times. After each time, the totality of the input pixels and the filter is calculated. The pattern of dots at the terminus is identified as a feature map or convolved feature. In this layer, the images are finally transformed into numbers. CNN is better able to understand the image and identify key patterns as a result.

Pooling Layer: The pooling layer moves a small region across the input image in a manner akin to convolutional layer. However, the pooling layer diminishes the aggregate of data and can also lead to a loss of information, which is different from the convolutional layer. This intermediate layer makes things simpler and helps CNN work best.

Fully Connected Layer: The FC layer is where image classification happens within the CNN in line with the highlights extracted within the past layers. Here, all the inputs or hubs from one layer are affiliated with each node of another layer. The layers in the CNN are not utterly interconnected because it would make the network too crowded, cause more losses, lower the quality of the output, and cost more to do.

E. Classify the type of cancer:

The model first recognizes the edges of the white blood cancer cells present in the testing image and then yields the output as type of cancer.

IV. COMPARISON ANALYSIS

S.No	Algorithm	Accuracy	Precision	Recall	Loss
1	YOLO	90%	83%	90.9%	10%
2	Naive Bayes	93%	94.66%	93.8%	-
3	CNN	97.33%	97.37%	97.33%	8.04%

V. RESULT AND ANALYSIS

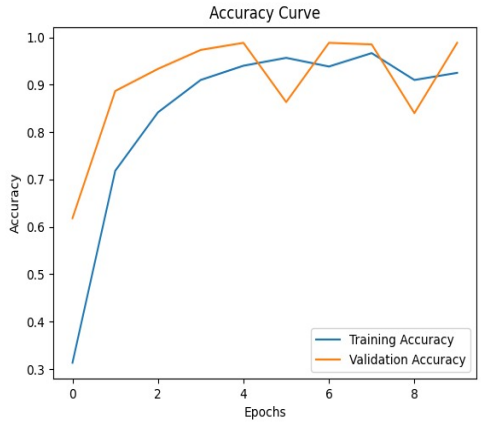

Figure 5.1: Accuracy Curve

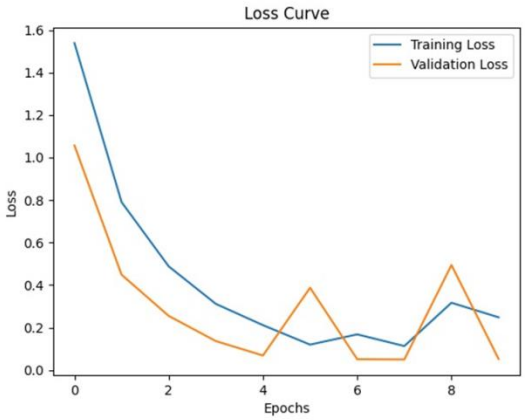

Figure 5.2: Loss Curve

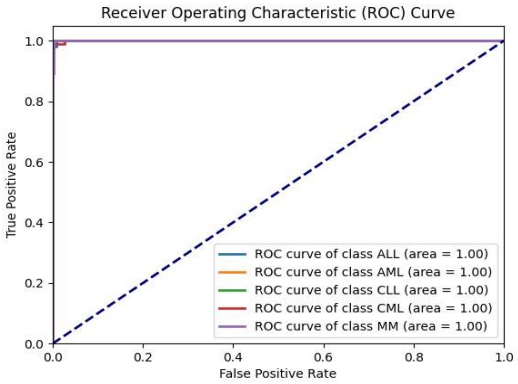

Figure 5.3: ROC Curve

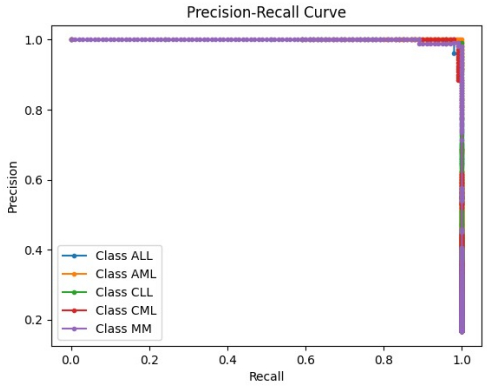

Figure 5.4: Precision-Recall Curve

Accuracy curves in CNNs track validation and training accuracies over epochs, highlighting generalization and model learning; discrepancies indicate overfitting and encourage regularization and dropout as mitigating measures for better diagnostic reliability. CNN's effectiveness is evaluated through the quantification of prediction discrepancies, which is how the loss curve measures model performance. Accuracy curves for training and validation measure generalization and learning, directing model development for dependable diagnostic results. CNN's potential is to distinguish between different cancer

kinds is demonstrated by ROC curves, and the overall discriminative performance is measured by AUC values, which help physicians prioritize the accuracy of leucocyte diagnosis and guide treatment plans. PR curves guide model optimization for precise recognition and organization by evaluating CNN performance across leukemia kinds by balancing accuracy and recall. The repercussions of a threshold on precision and recall is shown by each point on the curve, which is important for comprehending the pluses and liabilities of the framework for leukemia detection.

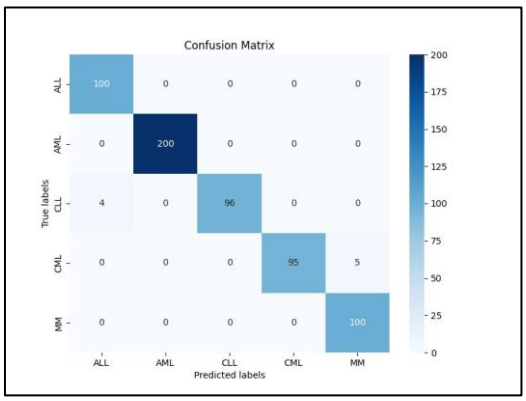

Figure 5.5: Confusion Matrix

The Confusion Matrix provides information relating to its efficacy of the framework for various categories of leukemia by classifying model predictions and emphasizing values comprising True positives, True negatives, False positives and False negatives. Nuanced evaluation is made possible by constraints such as precision, recall, and F1 score that are obtained from the Confusion matrix. These metrics help optimize CNNs for precise evaluation and therapy tactic in the healthcare industry.

VI. CONCLUSION

In conclusion, the proposed think about presents a novel approach utilizing Convolutional Neural Systems (CNNs) for the sorting of variance of leukemia and multiple myeloma based on cell pictures. By leveraging profound learning strategies, the demonstrate points to robotize the classification handle, in this manner decreasing the potential for blunders related with manual classification strategies. The study's discoveries recommend that the CNN-based show beats existing strategies such as Yolov8 and Gullible Bayes as a matter of exactitude. This shows the adequacy of rigorous grasping in extricating pertinent highlights from cell pictures and precisely foreseeing cancer subtypes. The suggestions of this inquire about are critical, as exact and early determination of cancer subtypes is basic for directing treatment choices and progressing persistent results. Robotizing the classification handle can streamline demonstrative workflows and possibly lead to speedier and more exact analyse, eventually profiting patients and healthcare suppliers alike. Moving forward, encourage approval and refinement of the intended program will be basic to guarantee its unwavering quality and adequacy in clinical settings.

VII. REFERENCES

[1] Jude D.Hemanth and Roman Senkerik, "Automatic Detection of White Blood Cancer from Bone Marrow Microscopic images using Convolutional Neural Networks", IEEE Access, vol. 8, pp. 142521-142531, 2020.

[2] Wahidur Rahman, Mohammad Gazi Golam Faruque and Kaniz Roksana, "Multiclass blood cancer classification using deep CNN with optimized features", ScienceDirect, vol. 18, pp. 1-23, 2023.

[3] Vyshnav M T, Sowmya V, Gopalakrishnan E A, "Deep Learning Based Approach For Multiple Myeloma Detection", Proceeding of 11th International Conference on Computing, communication and networking technology (ICCCNT), 2020, pp. 1-7.

[4] Akshay Belhekar, Kumar Gagare and Ritesh Bedse, "Leukemia Cancer Detection using Image Analytics", Proceeding of 5th International Conference on Computing Communication Control and Automation (ICCUBEA), 2019, pp. 1-6.

[5] Chayan Mondal, Md. Kamrul Hasan and Mohiuddin Ahmad, "Ensemble of Convolution Neural Networks to diagnose Acute Lymphoblastic Leukemia from microscopic Images", ScienceDirect, vol. 27, no. 12, pp. 1-14, 2021.

[6] Astha Ratley, Mrs. Laurine Minj, Mrs. Pooja Pame, "Leukemia Disease Detection and Classification Using Machine Learning Approaches", Proceeding of First International Conference on Power, Control and Computing Technologies (ICPC2T), 2020, pp. 161-165.

[7] Aqsa Jabeen, Sara Jabeen, Syed Adnan Shah, Wakeel Ahmad Rao, "Efficient Features for Effectively Detection of Leukemia Cells", Proceeding of IEEE 23rd International Multitopic Conference (INMIC), 2020, pp. 1-6.

[8] Amreen Batool, Yung-Cheol-Byun, "Lightweight EfficientNetB3 Model Based on Depthwise Separable Convolutions for Enhancing Classification of Leukemia White Blood Cell Images", vol. 11, pp. 37203-37215, 2023.

[9] Manjula Devi Ramasamy, Rajesh Kumar Dhanaraj and Subhendu Kumar Pani, "An improved deep convolutionary neural network for bone marrow cancer detection using image processing", vol. 38, pp. 1-9, 2023.

[10] Niranjana Sampathila, Krishnaraj Chadaga and Neelankit Goswami, "Customized Deep Learning Classifier for Detection of Acute Lymphoblastic Leukemia Using Blood Smear Images", vol. 10, no. 10, pp. 1-16, 2022.

[11] Tusneem Ahmed M.Elhassan, "Feature Extraction of white Blood cells using CMYK Moment Localization and Deep Learning in Acute Myeloid Leukemia Blood Smear Microscopic Images", IEEE Access, vol-10, pp. 16577-16591, 2021.

[12] Fatma M. Talaat and Samah A Gamel, "A2M-LEUK: attention-augmented algorithm for blood cancer detection in children", Springer, vol. 35, pp. 18059-18071, 2023.