IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 10, Issue 11, November 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.165

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011022 |

Automatic Detection of Acute Lymphoblastic Leukemia from Bone Marrow Microscopic Images using Convolutional Neural Network

Aniket Sharma, Arbaz Attar

Undergraduate Student, Department of Computer Science Engineering, MIT World Peace University, Pune, Maharashtra, India

Undergraduate Student, Department of Information Technology, RMD Sinhgad School of Engineering, Savitribai Phule Pune University, Pune, Maharashtra, India

ABSTRACT: The growth of many aberrant cells is brought on by leukemia (blood cancer), which starts in the bone marrow. The primary classifications of leukemia include both acute and chronic. Compared to chronic leukemia, acute leukemia deteriorates quickly. Acute lymphocytic leukemia (ALL), acute myeloid leukemia (AML), chronic lymphocytic leukemia (CLL), and chronic myeloid leukemia (CML) are the four most prevalent kinds of leukemia. Leukemia-related death has been ranked among the top 10 most deadly causes of human death. Therefore, effective clinical decision support for the categorization of acute leukemia type has become essential. The goal of this thesis is to develop a system for detecting acute lymphocytic leukemia (ALL) using images processing techniques and convolution neural network (CNN) classifier, thus automate the detection process. The proposed approach is assessed using ALL-IDB2. The proposed approach performs better with an accuracy rate of 96 %. Additionally, broader use of the approach could offer an indication of the pre-leukemic condition and the residual sickness in the future.

KEYWORDS: Acute Leukemia Classification, Convolution neural network, Image processing.

I. INTRODUCTION

For pathologists and other medical professionals, cancer detection is a constant problem when it comes to diagnosis and therapy. Another area of image processing research is the development of convolution neural networks for the classification of leukemia.

Given that they show a rapid production of lymphoid cells in the bone marrow, a high number of white blood cells and lymphoblasts in the blood might raise doubts about ALL [1] and a bone marrow biopsy, with generally more than 20% of all cells being leukemic lymphoblasts, offers definitive evidence of ALL [2]. One of the research elements that has been employed for leukemia diagnosis in the past several years is image processing. Image processing is commonly used to diagnose blood, perform X-rays, CT scans, MRIs, and many more tests to find diseases. To the author's knowledge, there hasn't been any study on image enhancement for leukemia, even though image processing is frequently employed in medical applications [3].

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011022 |

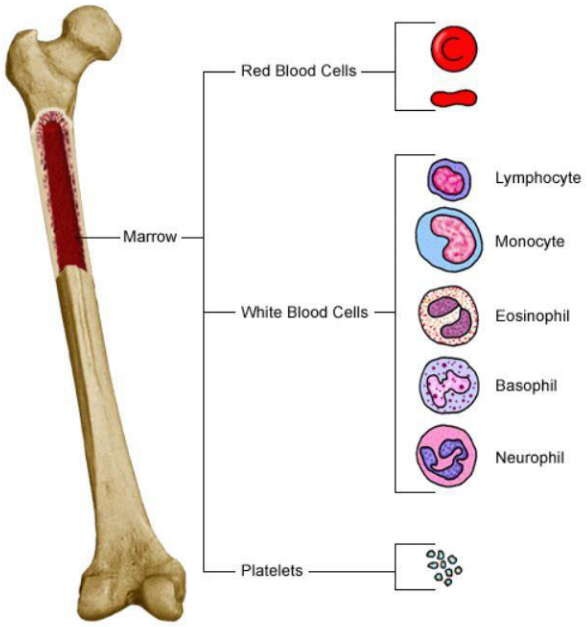


Fig. 1. Bone marrow subparts

In addition, advances in image processing technology, which typically tries to identify an object by extracting important features from an image, can assist in spotting abnormalities in the human body. The feature extraction method, which is applied after the segmentation procedure, is crucial in separating leukemic cells from normal cells. The WBC's nucleus is where the characteristics are taken out. Radius, roundness, standard deviation, major axis, and minor axis of the WBC nucleus are often retrieved parameters [4]. If performed by a trained individual, the manual method for counting blood has a 100% recognition rate but is also time-consuming. However, while automated counting is quicker, there is a greater chance that the count will be off. Thus, each approach has advantages and disadvantages. When analysing a lot of images, Deep Learning is known to perform better than traditional Machine Learning. Convolution Neural Networks (CNNs) combine different multilayer perceptron and, with a little pre-processing, display effective results. As each convolution layer of the network learns a new feature that is present in the images and results in a high activation, CNNs themselves serve as feature extractors [5]. Recognizing the characteristics of normal and diseased cells is a crucial criterion for achieving a robust and successful computerized diagnosis. Only a few studies have looked at how to choose important discriminating characteristics from the segmented areas to successfully aid in ALL diagnosis. In this research, we propose a novel approach to address the problems mentioned above to classify acute leukemia using a convolutional neural network (CNN). To distinguish between normal and abnormal cell images from an ALL-IDB database, this approach performs exceptionally well throughout the classification phase.

II. RELATED WORK

2.1. OVERVIEW OF ACUTE LYMPHOCYTIC LEUKEMIA (ALL)

One of two classes of acute leukemia that arise from early (immature) forms of lymphocyte cells in bone marrow is acute lymphocytic leukemia (ALL). Most deaths occur in adults, which is unsurprising given that children's bodies can frequently withstand rigorous therapy better than adults. ALL is the most prevalent of the four major varieties among children but the least prevalent in adults. Normal white blood cells have an average size that is similar to the red blood cells around them, but leukemia-infected white blood cells often have an average size that is roughly two times the size of the surrounding red blood cells. Most ALL-infected cells have nuclei that take up between 80 and 90 percent of the total cell volume, leaving just 20 to 30 percent of the space for cytoplasm. Each cell has a smooth surface.

2.2 RECENT LEUKEMIA DETECTION

The capacity to diagnose ALL in the early phases of therapy and prevention is crucial because it is frequently detected in children between the ages of 5 and 15 and individuals over 50. After reviewing the previously published research in this area, it became clear that many researchers had tried out various segmentation techniques on leukemia

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011022 |

images to get better results. These techniques included thresholding, region-based methodologies, edge detection, clustering, artificial neural networks, fuzzy processes, watershed algorithms, and others.

In "White Blood Cell Classification and Counting Using Convolutional Neural Network" [6], CNN was used to categorize and count WBCs based on microscopic blood images. Additionally, when the three CNN models (Alexnet, GoogleNet, and ResNet-101) were evaluated, it was found that AlexNet outperformed GoogleNet and ResNet-101 in the task of classifying and counting based on 21 images of microscopic blood samples. Model will be published to the cloud platform so that any medical professionals or facilities that use the service can use method to identify leukemia patients as discussed in "Cloud based Acute Lymphoblastic Leukemia Detection Using Deep Convolutional Neural Networks" [7].

The identification system's neural network underwent training and testing using three different learning strategies. Utilizing various learning strategies aims to find the best neural classifier for the application and the data at hand. According to the findings of the experiments, this method of recognizing infected blood cells can be successfully applied in practical situations [8].

In [9], the method successfully separates the largest nucleus blob in the produced blood cell image to support hematologists in the leukemia detection process. The given test images are classified as benign or malignant by the KNN classifier. With distinct sets of training data, the algorithm matures further because less training results in less efficiency.

Presented a low-cost digital pathology system that can identify malaria and provide a CBC and WBC count from an image of a blood smear. To increase the detection's precision, the situation of partially visible WBCs is treated independently. The system is economical and performs nearly as well as a human pathologist [10]..

III. PROPOSED ALGORITHM

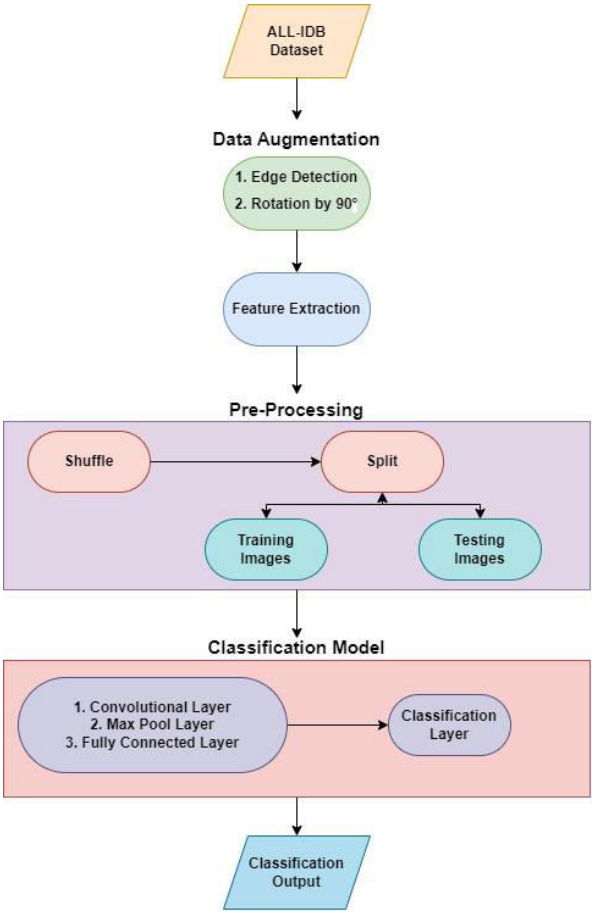


Fig. 2. Proposed Algorithm

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

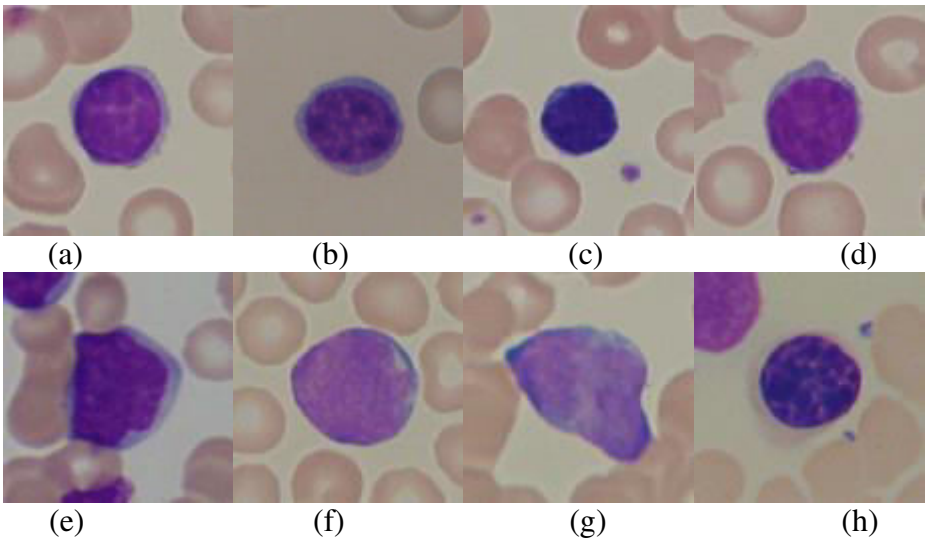
| DOI: 10.15680/IJIRCCE.2022.1011022 |

A. *Image Collection:*

Image dataset is collected from Department of Computer Science - Università degli Studi di Milano. Dataset contains images which has been captured with an optical laboratory microscope coupled with a Canon PowerShot G5 camera. All images are in JPG format with 24-bit color depth, resolution 2592 x 1944.

B. *Image Pre-processing:*

It is a method used to enhance the quality of the image. The obtained images underwent pre-processing and enhancement. The handling of null values, one-hot encoding, normalization, multi-collinearity, scaling the data, shuffling and dividing the data, etc. are all examples of pre-processing. In the suggested study, the transformed data generated through feature selection is first normalized, and then following randomization, they are split into training and testing sets. The dataset is split into 20% for model testing and 80% for training.



C. *Data Augmentation:*

The picture is rotated, and edges are extracted, and then the SN-AM dataset is enhanced. Training and testing sets are created from the scrambled photos. For the model to generalize well during the evaluation (testing) phase, there should be a sizable amount of data available because the object of interest must be present in various sizes, poses, and lighting conditions. By using a variety of picture manipulation techniques, data may be created from already-existing data. Without data augmentation, overfitting occurs and the model has a hard time generalizing to new examples that weren't in the training set.

D. *Feature Extraction:*

In Deep Learning, feature selection is critical and has a significant impact on the model's performance. A model's performance suffers when it has a lot of features. Thus, a process known as feature selection selects the features that have an impact on the result. The main benefits of this method are the following: Less over-fitting since the likelihood of predictions based on noise is eliminated by less redundant data, a shorter training period since there are fewer training data available, Accuracy increases as false data and outliers are eliminated

E. *Classification:*

A Convolutional Neural Network (ConvNet/CNN)[5] is a Deep Learning method that can take in an input picture, give various elements and objects in the image importance (learnable weights and biases), and be able to distinguish between them. Comparatively speaking, a ConvNet requires substantially less pre-processing than other classification techniques. ConvNets have the capacity to learn these filters and properties, whereas in basic techniques filters are hand-engineered.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011022 |

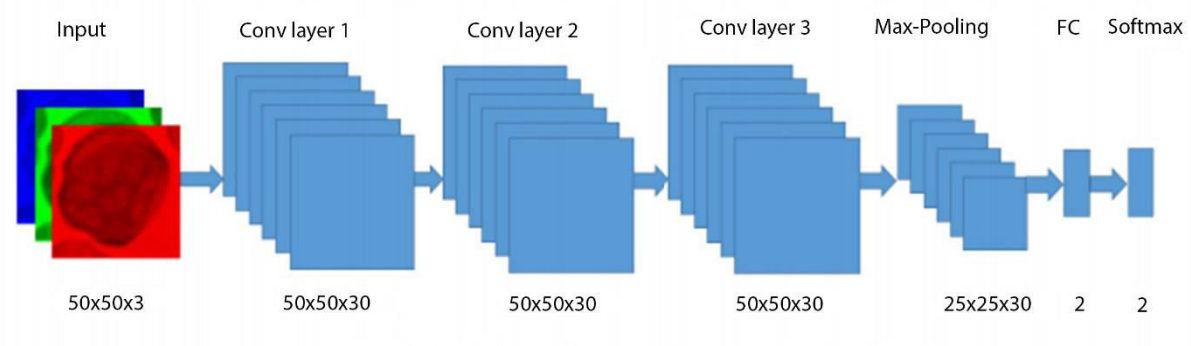


Fig. 3. Architecture of Network

- Convolution: In order to find a line, a filter image (such as a line detector) looks at every area of the input image. The filter is turned on if it notices a line. Until it reaches the end of the input picture, it will advance one unit to the right. The feature map is an array that stores every location's information. Locations with lines will have high values, while those without will have zero values.
- Max Pooling: A particular style of pooling layer that scales down or reduces the resolution of the incoming input layer. By doing this, overfitting is prevented and the cost of computation is significantly reduced.
- Fully Connected Layers: This layer is found at the network's outermost point. It links up all of the active areas in the layers that came before it. Its output is an N-dimensional vector with N the number of classes the network has been trained to distinguish between; there are 5 classes in our study. The probability that the object belongs to the class is contained in each element of the vector. The classification outcome is the component with the highest probability.

IV. RESULTS AND DISCUSSION

In this paper, we presented a four-layer network. The first three layers are used for feature detection, and the next two layers (Fully linked and Softmax) are used for feature classification. The supplied picture is [50x50x3] in size. The filter size (or receptive field) is 5x5. We shift the filters one pixel at a time with a stride of 1. Zero padding is two. We will be able to regulate the spatial scale of the final image (we will use it to exactly preserve the spatial size of the input volume so the input and output width and height are the same). We discovered throughout the experiment that, in our situation, changing the original image's size during convolution resulted in a 40% reduction in accuracy. As a result, convolution layer 1's output picture has the same size as the input image.

The structure of convolution layers 2 and 3 is identical to that of convolution layer 1. The stride is 1, the zero-padding is 2, and the filter size is 5x5. In our situation, there are 30 feature maps (the channel or the depth). The accuracy will drop by 50% depending on whether there are fewer or more feature maps than 30. Through experimentation, we discovered that removing Convolution layers 2 and 3 also causes accuracy to drop by 50%.

The 25x25 Max-Pooling layer has stride is two and filter size is two. There are 2 neuronal in the fully connected layer. Finally, the categorization is done using the Softmax layer.

Writing program that complies the model, then fits the model. The validation_data, _(X_test, y_test)_, and the number of validation_steps are passed to the Keras model.fit function. This means that in addition to showing training loss, accuracy, precision and recall the program will show validation loss, accuracy, precision and recall in its output. Fig. 4. shows the result after training the dataset for 100 epochs and fig. 5. shows the graph of validation accuracy and validation loss.

Loss	Accuracy	Precision	Recall	AUC
0.083 (~0.84)	0.961 (~96%)	0.961 (~0.96)	0.961 (~0.96)	0.997 (~1.0)

Fig. 4. Result after training for 100 epochs

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011022 |

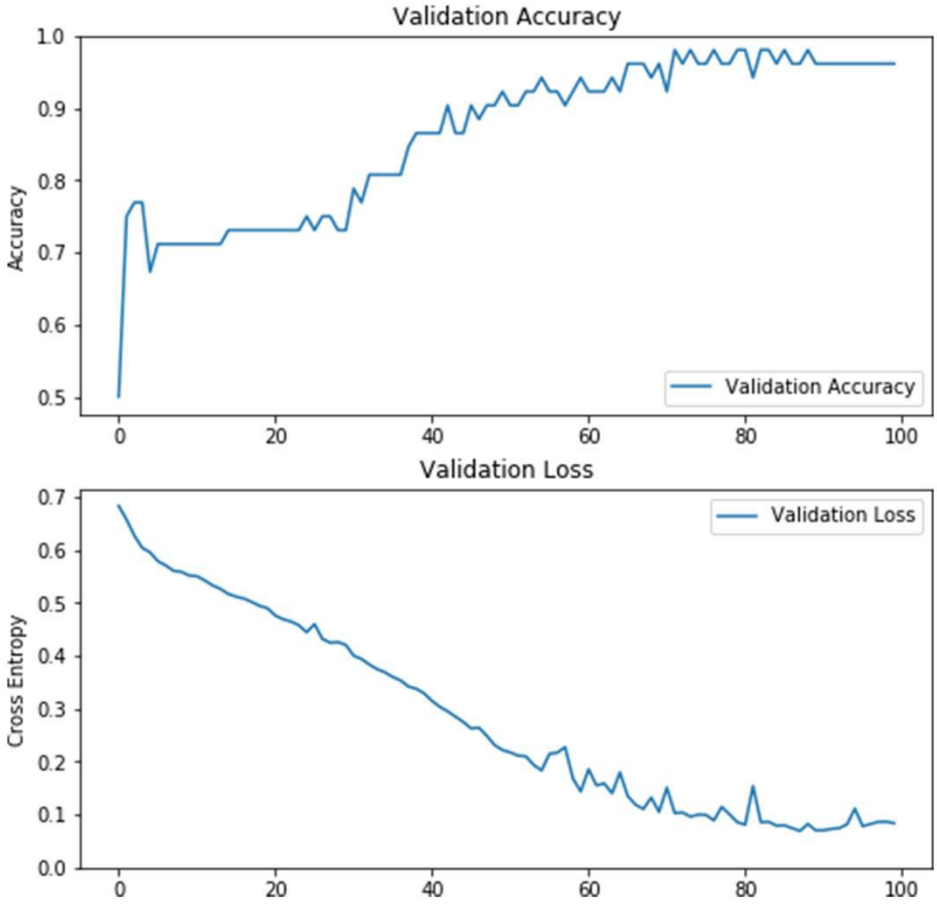


Fig. 5. Graph of validation accuracy and validation loss

V. CONCLUSION AND FUTURE WORK

The original ALL-IDB2 image database, which contains 108 cell images (59 normal cell images and 49 abnormal cell images), was utilized in this experiment. The experiment was run on MATLAB, and our suggested CNN model correctly identified leukemia 96.10% of the time.

By applying a thorough learning technique, namely convolutional neural networks, the suggested model completely eliminates the possibility of mistakes in the human process. The model initially pre-forms the images and isolates their best features before being prepared using a modified Convolutional neural network structure.

One of the deadliest diseases for people is leukemia. This study offered a brand-new way for categorizing different types of acute leukemia, employing CNN as one of the building blocks for the clinical decision support system for this form of leukemia. In the upcoming study, we'll think about how to get a better outcome and attempt to categorize four different forms of leukemia.

REFERENCES

1. T. Dharani and S. Hariprasath, "Diagnosis of Leukemia and its types Using Digital Image Processing Techniques," 2018 3rd International Conference on Communication and Electronics Systems (ICCES), 2018, pp. 275-279Hong-ryeol Gil1, Joon Yoo1 and Jong-won Lee2,'An On-demand Energy-efficient Routing Algorithm for Wireless Ad hoc Networks', Proceedings of the 2nd International Conference on Human. Society and Internet HSI'03, pp. 302-311, 2003.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.165

|| Volume 10, Issue 11, November 2022 ||

| DOI: 10.15680/IJIRCCE.2022.1011022 |

2. H. Parvaresh, H. Sajedi and S. A. Rahimi, "Leukemia Diagnosis Using Image Processing and Computational Intelligence," 2018 IEEE 22nd International Conference on Intelligent Engineering Systems (INES), 2018, pp. 000305-000310
3. N. H. Harun, J. A. Bakar, Z. A. Wahab, M. K. Osman and H. Harun, "Color Image Enhancement of Acute Leukemia Cells in Blood Microscopic Image for Leukemia Detection Sample," 2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics (ISCAIE), 2020, pp. 24-29
4. Himali P. Vaghela, Hardik Modi, Manoj Pandya and M.B. Potdar Leukemia Detection using Digital Image Processing Techniques", November 2015 - International Journal of Applied Information Systems (IJAIS).
5. D. Kumar et al., "Automatic Detection of White Blood Cancer from Bone Marrow Microscopic Images Using Convolutional Neural Networks," in IEEE Access, vol. 8, pp. 142521-142531, 2020.
6. M. J. Macawile, V. V. Quiñones, A. Ballado, J. D. Cruz and M. V. Caya, "White blood cell classification and counting using convolutional neural network," 2018 3rd International Conference on Control and Robotics Engineering (ICCRE), 2018, pp. 259-263.
7. S. Nayaki K., J. Denny, M. M. Rubeena and J. K. Denny, "Cloud based Acute Lymphoblastic Leukemia Detection Using Deep Convolutional Neural Networks," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), 2020, pp. 530-536.
8. Khashman, A., Abbas, H.H. (2013). Acute Lymphoblastic Leukemia Identification Using Blood Smear Images and a Neural Classifier. In: Rojas, I., Joya, G., Cabestany, J. (eds) Advances in Computational Intelligence. IWANN 2013. Lecture Notes in Computer Science, vol 7903. Springer, Berlin, Heidelberg
9. D. Umamaheswari and S. Geetha, "Segmentation and Classification of Acute Lymphoblastic Leukemia Cells Tooled with Digital Image Processing and ML Techniques," 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018, pp. 1336-1341
10. A. B. Chowdhury, J. Roberson, A. Hukkoo, S. Bodapati and D. J. Cappelleri, "Automated Complete Blood Cell Count and Malaria Pathogen Detection Using Convolution Neural Network," in IEEE Robotics and Automation Letters, vol. 5, no. 2, pp. 1047-1054.

BIOGRAPHY

Aniket Ranjeet Sharmais an undergraduate student in the Computer Science Engineering Department, MIT World Peace University, Pune. His research interests are Blockchain technology, Machine Learning, Algorithms, Natural Language Processing, etc.

Arbaz Akbar Attar is an undergraduate student, Department of Information Technology, RMD Sinhgad School of Engineering, Savitribai Phule Pune University, Pune. His research interests are Machine Learning, Blockchain, Cyber Security, Cloud Computing, etc.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.165

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details