

Identification of Leukemia Using Convolution Neural Networks

Nitish Panchpor¹, Shilpa Jahagirdar², Manisha Sagade³, Lokesh Patil⁴, Utkarsh Nikam⁵

Department of E&TC, SKNCOE, SPPU, Pune, India

¹nitishpanchpor@gmail.com

²ssjahagirdar.skncoe@sinhgad.edu

³manisha.sagade_skncoe@sinhgad.edu

⁴lokeshpatil193@gmail.com

⁵nikamutkarsh142@gmail.com

Abstract

Leukemia is a type of malignant blood cell cancer that can affect both children and adults. In order to get the most effective treatment, the patient needs early diagnosis. Therefore, a support system is needed for early diagnosis to guide treatment for patients with Leukemia as soon as possible. This paper is focused on with the purpose of fast and early detection of Leukemia cell to prevent spreading of the cancer in patient's body and to the treatment for patient. The proposed system uses convolution network that inputs a blood cell image and output whether cell is infected with cancer or not. The common algorithm for this approach consists of several rigid steps. The basic idea to develop a system which detect the Leukemia from blood cell images with more precision.

Keyword: Convolutional Neural Network, Image Processing, Machine Learning, Leukemia.

I. INTRODUCTION

Leukemia is a blood cells cancer caused by family history of leukemia, radiation exposure and exposure to certain chemicals. Generally, leukemia was classified based on the progression speed and the type of cells. On the basis of leukemia progresses, the leukemia classification is split into two types acute leukemia and chronic leukemia. In acute leukemia, the abnormal blood cells which cannot carry out their proper functions as multiply speedily. The main objective of implemented work is to make a system that has ability detect cancer automatically from the blood cell images. Developed system uses a convolution network where input is taken as a blood cell images and outputs whether the cell is infected with cancer or not. The presence of cancer in blood cell images is often vague, it may overlap with other diagnoses, and can mimic many other benign abnormalities. These variations cause considerable variability among medical personnel within the diagnosis of cancer. Automatic detection of cancer from blood cell images at the level of expert medical personnel would not only have tremendous benefit in clinical settings, it would also be useful in delivery of health care to populations where medical expertise are limited.

II. LITERATURE SURVEY

B. Ramya, V. Uma Rani proposed in [1] about ant-colony optimization technique that utilizes deep learning methods, specifically Convolutional Neural Networks (CACO-CNN), for detection of Acute Myelogenous Leukemia. Dataset consist of 410 microscopic blood images representing varying magnitudes of infection having 184×138 sized are used for the evaluation. To classify the AML, this research paper considered (I) feature. K-fold cross validation, Swarm intelligence for data mining tool and dataset MATLAB is used. The results indicated that the K-fold classifier is having better classification accuracy than the existing classifiers. Further in [2], T. T. P. Thanh, Caleb Vununu,

Sukhrob Atoev, Suk-Hwan Lee, and Ki-Ryong Kwon used a network containing 7 layers, the first 5 layers perform feature extraction and the other 2 layers (fully connected and softmax) Experiments were conducted on MATLAB with 1188 images, 70% (831 images) of them for training and the remaining 30% (357 images) for testing model. Author has augmented the original dataset to 1188 pictures by applying transformations such as blurring, histogram equalization, reflection, translation, rotation and shearing. For data mining tool and dataset MATLAB is used and Data augmentation using Gaussian filter. Proposed CNN model achieved 96.6% accuracy. Authors R. D. A. U. Pallegama, B. G. D. A Madhusanka, H. D. N. S. Priyankara have done classification of ALL generally consist of several stages these include image acquisition, image segmentation, feature extraction, feature selection and classification in [3]. For segmentation and extraction of the ALL cells, author has used Hue, Saturation, Value Colour Space (HSV colour space) method 17 epochs were used get the above training accuracy. Dataset of 1000 images is used and it took 30 second for identify the cells. The proposed system is implemented with Support Vector Machines (SVMs) algorithms for feature extraction and classification. The evaluation result value of the ALL-cancer detection accuracy using above mentioned CNN is 98.53%

Tatdow Pansombut et al implemented a CNN classifier to explore the feasibility of deep learning approach to identify lymphocytes and ALL subtypes, and with support vector machines (SVMs) in [4]. Author Dataset have been rescaled to equal size of 256×256 pixels, and each contains a single cell. This system implemented a CNN, Con-VNet, performance of our deep learning model is evaluated against a dominant approach of SVM classifier, namely, SVM-GA, and two traditional machine learning approaches including MLP (Multilayer perceptron) and random forest. SVM classifier delivered good results above 98.53% accuracy. This is followed by Nizar Ahmed, Altug Yigit, Zerrin Isik and Adil Alpkocak who employed data augmentation using image transformations and finally applying CNN to classify the four Leukemia subtypes in [5]. In this paper Dataset is acquired from 2 sources: The American Society of Hematology (ASH) Image Bank and ALL-IDB dataset provided annotated microscopic blood cell images designed for evaluation of segmentation and classification purposes. Total number of samples reached 2478 and 3843. Python programming language with two open-source image processing libraries, OpenCV and KERAS were used to perform image transformations. Trained and tested on Windows 10. The NVIDIA Cuda Developer Toolkit was used with a GTX 1070 GPU along with TensorFlow-GPU library. Optimizers used were SGD, ADAM. 88% accuracy for the binary classification of one leukemia type and an 81% accuracy for classifying all leukemia subtypes has obtained. In [6], Amjad Rehman, et al proposed deep learning technique using Alexnet model with CNN for the classification of ALL into its subtypes and normal condition. The data sets for this system consist of all the images of ALL subtypes, which include 100 images of L1, 100 images of L2, 30 images of L3 due to rare nature of L3, and 100 images of normal bone marrow. The proposed architecture is implemented in MATLAB 2017a with computer vision toolbox and Alexnet model on GPU. Segmentation using Gram Schmid and k-means clustering algorithm, classifier used k-fold validation where $k=10$. This research work has presented a complete architecture based on deep learning techniques for the classification of ALL and achieved 97.78% accuracy with efficient processing time. Further, Deepika Kumar and author co-authors about a deep learning technique using CNN for the classification of ALL into its subtypes and normal condition in [7]. Dataset was acquired from two different subsets of a dataset i.e., B-Lineage Acute. Lymphoblastic Leukemia having 90 images in total and Multiple Myeloma, i.e., MM having 100 images. The proposed architecture is implemented in MATLAB 2017a with computer vision toolbox and Alexnet model on GPU, classifier used k-fold validation using chi-square test for optimization DCNN. This research work has presented a complete architecture based on deep learning techniques for the classification of ALL and achieved 97.78% accuracy with efficient processing time. Moreover, Ahmed T. Sahlol, Philip Kollmannsberger

& Ahmed A. Ewees first extracted features from WBC images using VGGNet, (CNN architecture) pre-trained on ImageNet in [8]. The extracted features are then filtered using a statistically enhanced Salp Swarm Algorithm (SESSA). The SESSA optimization selected only 1 K out of 25 K features extracted with VGGNet, while improving accuracy at the same time. The proposed system was implemented in Python 3 on Windows 10 64 bit using a Core i5 CPU and 8 GB RAM. The training was performed on Nvidia Tesla P100 GPU nodes (16 GB GPU memory, Ubuntu Linux 16.04) of the high-performance computing cloud *Julia*. This model delivers good results above 93.24% accuracy.

In [9], the author explored two architectures namely, VGG16 and MobileNet for our hybrid model. For experiments, 70% of the images of each class are assigned to the training set, 20% to the validation set, and the remaining 10% to the test set. Different optimizers used were adaptive moment estimation (Adam), stochastic gradient descent (SGD), root mean square propagation (RMS Prop). For implementation, Python using the Keras package with Tensorflow as the deep learning framework backend is used and run on Nvidia GeForce GTX 1080 Ti GPU with 11GB RAM. Proposed model delivered high accuracy (96.17%) on dataset with Adam optimizer. Prayag Tiwari, Jia Qian, Qiuchi Li, Benyou Wang, Deepak Gupta, Ashish Khanna, Joel J.P.C. Rodrigues, C. Victor Hugo, used two convolutional layers and one pooling layer, and it is followed by the fully-connected structure with one hidden layer and output layer in [8]. In proposed algorithm, forward pass and backward propagation algorithm has been implemented. The experiment dataset contains around 13k augmented pictures of blood cells with having the blood cell type labels. And was carried out on Intel(R) Core (TM) i7-7700K CPU @ 4.20 GHz RAM 16.0 GB Operating System Ubuntu 18.04 with Nvidia GeForce GTX 1060 3 GB graphics card. In this paper, Author proposed a CNN-based model which gives 94% accuracy.

III. METHODOLOGY

Machine learning technique become popular choices for medical image analysis process. For example, works in convolutional neural networks have used as a methodology in microscopic analysis. and also used deep learning method based on a convolutional neural network to detect the cytoplasm region in cancer cell segmentation. In this technique, a convolutional neural network applied for an image classification method which directly uses RGB values of the cell images for the learning procedure that extracts image features through a multilayer architecture automatically. For such approach, using image processing techniques feature values are extracted from the object information of the image. Many techniques have been developed by researchers to detect leukemia. One of the most used technique is Convolution Neural Network which is based on computer vision in some recent years. Algorithm for this method have various steps like image pre-processing, clustering, filtering, segmentation, feature selection or extraction, classification, and evaluation.

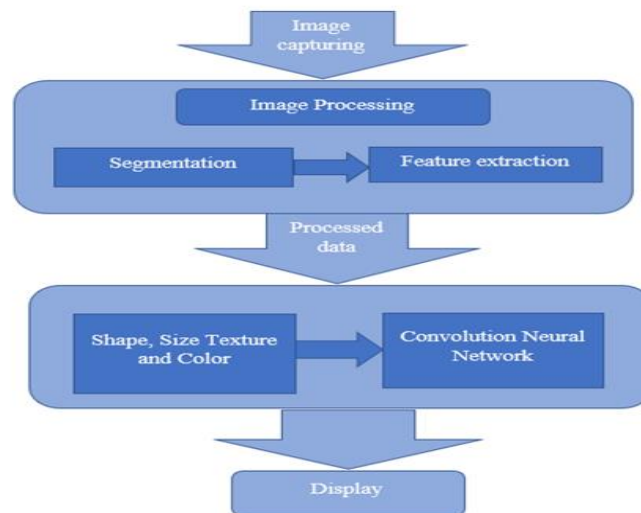


Fig 1. Block diagram of system

The CNN consist of layers of neurons and it is made for two-dimensional pattern recognition. CNN has three main types of layer that are pooling layer, convolutional layer, and fully connected layer. The first 6 layers of the convolution networks are convolution layer. convolutional layer sometimes also called feature extractor layer. The nonlinear transformation sublayer provides the ReLU activation function. In poolig layer, max pooling sublayer applies to a 2*2 filter to the image that results to reduce the image size to its half. The eighth layer is the flatten layer. The flatten layer is used to transforms a multidimensional array into one-dimensional array by simply concatenating the entries of the multidimensional array together. There are various algorithms before CNN become popular that people used for image classification. People try to create features from images and then given those features into some classification algorithm. Yann le cun in 1998 presented this concept first time. he used a single convolution layer for digit classification. That was later popular by Alex net in 2012 that is used multiple convolution layers.

A Convolution neural network is a multi-layered neural network and it has special architecture to detect complex features. Pre-processing deal with all the transformations on the raw data before it is given to the learning algorithm. The segments of an image as input to a CNN, which labels the pixels. The diagnosis and classification is consist of steps like this include image acquisition, image segmentation, feature extraction, feature selection and classification. The first stage of the work is image acquisition, which is an important step for the diagnosis of leukemia. A prerequisite to efficiently diagnose Leukemia is to set up a standard methodical procedure under which a large collection of good quality, clear and well contrasted blood sample images could be captured. The second stage of the system is image segmentation. The primary purpose of the image segmentation stage is to separate the blast cells from the other surrounding blood components such as red blood cells, platelets and blood plasma. Feature extraction stage is the next stage to extract several features or measurements from the blast cell and its components such as shape, texture and colour Then, these measurements learning to the leukemic identification system. Firstly, training of convolution network using the data in training set to find appropriated filter weights in the three convolutional sublayers and the weights that yield minimum error in the two fully connected layers are used. Then, convolution network is evaluated using the data in the validation set to obtain validation error and cross-entropy loss. The training of convolution network is repeated if required. In the last, the performance of convolution network is evaluated using data in the test set.

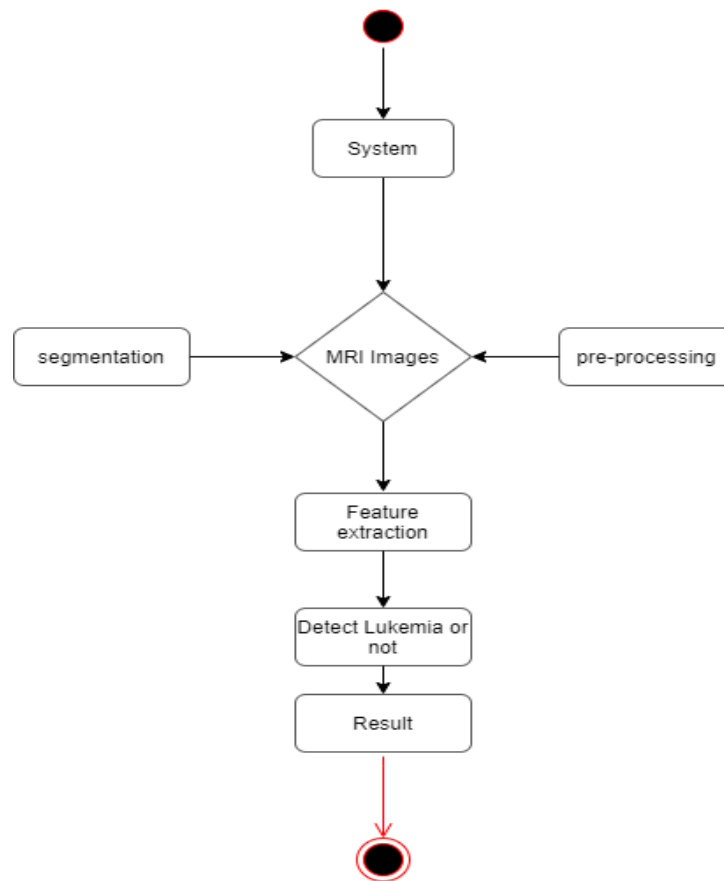


Fig 2. Flowchart

IV. EXPERIMENTATION AND RESULTS

While performing experimentation, authors have converted some of the images in dataset from RGB format to grayscale image. This is a simple method but it has applicability to merely creating more data. Results obtained during experimentations, for diagnosis of leukemia in normal patient and patient infected with leukemia are given in figure. 3 and figure 4.

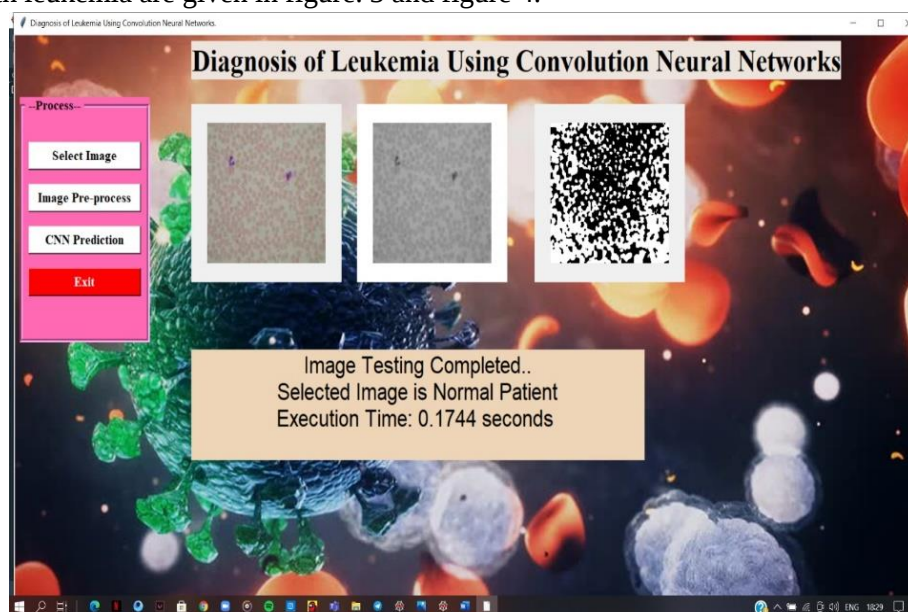


Fig 3. Results obtained for Normal patient

In image processing part authors have successfully converted RGB image into grey scale image using cv library. Define and declared the libraries in python and load the libraries in python. Using in built CNN layers features are extracted for detection. The model is trained to get more precision in the result. After training, module get ready for testing images. When image selected for testing then system shows if the person is infected with leukemia or not.

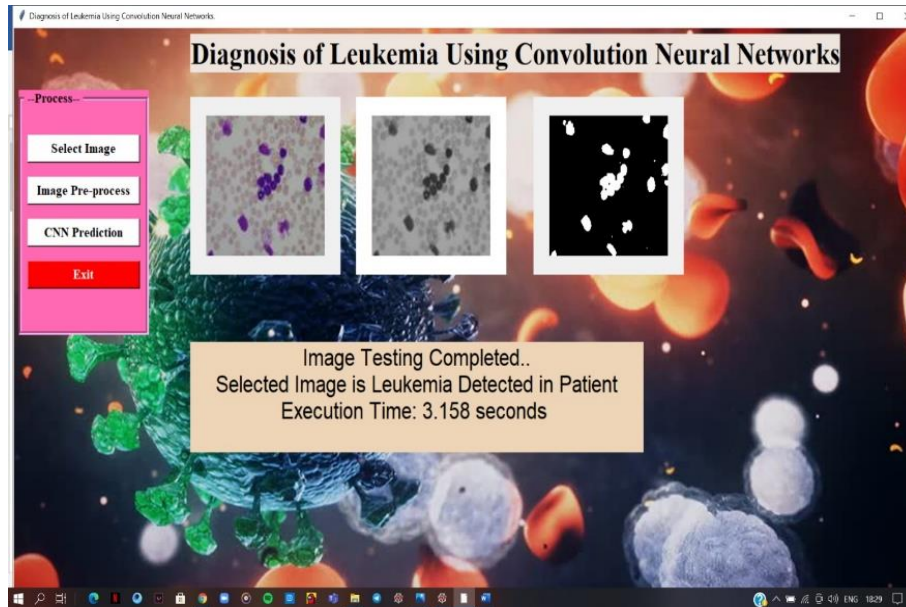


Fig 4. Results obtained for Leukemia detected patient

For the purpose of evaluation, authors have compared method with the previously used methods. While it shows that the accuracy rate after using our CNN model is significantly better than others. The evaluation result value of the leukemia cancer detection training accuracy using above mentioned CNN is 92.88% and that of testing is above 99%. From this it shows that uses of convolution neural network is more effective as compared to previous algorithms.

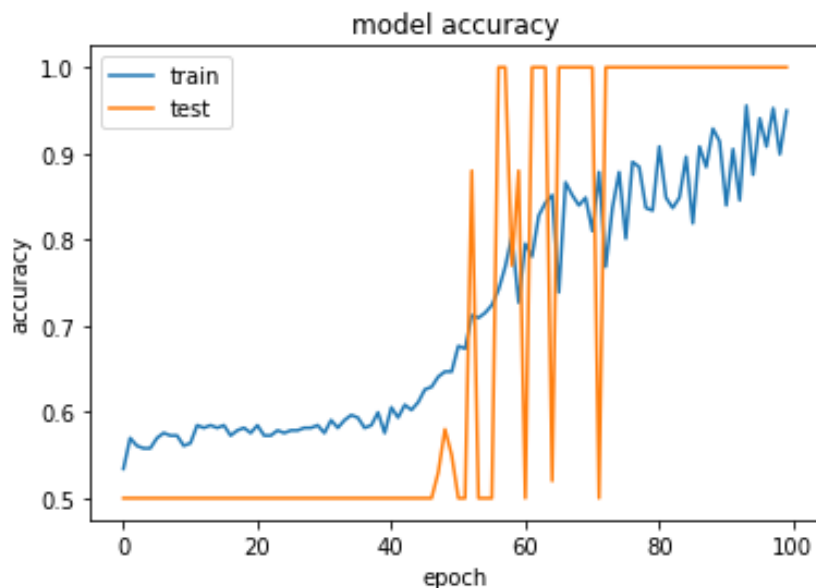


Fig 5. Model accuracy graph.

V. CONCLUSION

This method of classification promises to be used in diagnostic systems for leukemia for early detection of disease in the daily life. Mainly neural networks are used for automatic detection of cancer in blood samples. Neural network is chosen as a classification tool due to its well-known technique as a successful classifier for many real time applications. This technology can help in improving healthcare system.

REFERENCES

- [1] B. Ramya,V. Uma Rani, “*Acute Myelogenous Leukemia Detection using Circumventing Ant Colony Optimization based Convolutional Neural Network*”, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-4, November 2019
- [2] T. T. P. Thanh, Caleb Vununu, Sukhrob Atoev, Suk-Hwan Lee, and Ki-Ryong Kwon, “*Leukemia Blood Cell Image Classification Using Convolutional Neural Network*”, International Journal of Computer Theory and Engineering, Vol. 10, No. 2, April 2018
- [3] R. D. A. U. Pallegama, B. G. D. A Madhusanka, H. D. N. S. Priyankara, “*Acute Lymphoblastic Leukemia Detection using Convolutional Neural Network*”, IJESC, June 2020
- [4] Tatdow Pansombut, Siripen Wikaisuksakul, Kittiya Khongkraphan, Aniruth Phon-on, “*Convolutional Neural Networks for Recognition of Lymphoblast Cell Images*”, Computational Intelligence and Neuroscience, June 2019
- [5] Nizar Ahmed, Altug Yigit, Zerrin Isik and Adil Alpkocak, “*Identification of Leukemia Subtypes from Microscopic Images Using Convolutional Neural Network*”, August 2019
- [6] Amjad Rehman, Naveed Abbas, Tanzila Saba, Syed Ijaz ur Rahman, Zahid Mehmood, Hoshang Kolivand “*Classification of acute lymphoblastic leukemia using deep Learning*”, October 2018
- [7] 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, doi: 10.1109/ACCESS.2020.3012292.
- [8] Sahlol, A.T., Kollmannsberger, P. & Ewees, A.A., “*Efficient Classification of White Blood Cell Leukemia with Improved Swarm Optimization of Deep Features*,” *Sci Rep* **10**, 2536 (2020). <https://doi.org/10.1038/s41598-020-59215-9>
- [9] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider and R. Deters, "A Hybrid Deep Learning Architecture for Leukemic B-lymphoblast Classification," *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, 2019, pp. 271-276, doi: 10.1109/ICTC46691.2019.8939959.
- [10] Prayag Tiwari, Jia Qian, Qiuchi Li, Benyou Wang, Deepak Gupta, Ashish Khanna, Joel J.P.C. Rodrigues, C. Victor Hugo, “*Detection of subtype blood cells using deep learning*,” Cognitive System Research August 2018