Detection of Leukemia Using Machine Learning Algorithms

K P Jayavikash¹, P Gopinath¹, U Nandheeswaran¹, R Sivabalan¹, T Divya Bharathi¹
¹Department of Biomedical Engineering, KPR Institute of Engineering and Technology, Coimbatore.
17bme022@kpriet.ac.in

Abstract. Detection of disorders of Blood usually include visual inspection through a microscope. This may help to classify various conditions related to human blood. The paper details about evolving a system that could help in diagnosis and detection of Leukemia at an early stage. Only then controlling and prevention can be done. The proposed system focuses on White Blood Cells as they are the affected ones in case of Leukemia. The affected cells differ in shape and color features, which are used as classifier inputs.

1. Introduction

Medical Imaging involves technique and process that are used to generate images of different parts of our body for diagnostic and treatment purposes. Medical Imaging has become an integral part of Medicine over the past decade or so. Advanced equipment has been developed for detecting, analyzing and storing of medical images. The main aim of analyzing through images is to formulate as much detail as possible regarding identification, diagnosis, control, treatment, monitoring and assessment of diseases.

Cancer is among the deadliest diseases. Leukemia is a cancer affecting the blood forming cells. It is a blood cancer that happens when there is a lot of abnormal WBC’s are formed by the Bone marrow. Basic components of Blood are - Plasma, RBC, WBC and also platelets. Blood is made of 45% RBC and 55% plasma. WBC and platelets are present less than 1%. In case of Leukemia WBC’s will be produced more in number and this would lead to disruption in the Blood system. The presence of unusual and extra white blood cells can be exposed when the blood sample is collected and inspected by a hematologist. [1-3] The activity would be hectic and could take a lot of time. And moreover any work that involves human work is vulnerable to errors because of emotional disturbance and the person’s physical capability limit. Also it is really tough to get steady results from the perceivable inspection.

Automated system in image processing can really overcome the errors and problems in visual inspection. The system developed will be mainly on microscopic images to differentiate normal cells from Leukemia cells. Early detection helps providing fast and appropriate treatment for the condition.

With an automated system many images can be processed, scanning time can be reduced, accuracy can be improved. [4-7] Examining and classifying leukemia is based on size, structure and color features of White
Blood cells (WBC).

2. Literature review

A lot of algorithms are available and were used by various people to segment and classify medical images. By using the thresholding process it segments the image accurately from the white and black pixel it can be clearly segmented. There are various algorithms like K means clustering and Watershed methods. Also GFCVR (Gaussian feature convolutional visual recognition) was used [8-11]. Commonly used methods were KNN classifier and SVM classifier.

We have used K-mean clustering algorithm for segmentation of color images of microscopic blood images and VGG Architecture was used to extract shape and color features of the image and then finally classify if the input image is Normal or Leukemia. Figure 1 shows the flowchart.

3. Methodology

Figure 1. Flowchart

i. Image Acquisition:

Images were obtained from ALL-IDB online database. The dataset has 108 images collected in 2005. It has over 39,000 elements, where lymphocytes have been marked by experts. [12-14] These images were taken using a microscope with magnifications ranging from 300 to 500. These images present in the dataset were taken using a microscope attached with a camera. The obtained images were present in JPG and TIF formats. And has a resolution of 2592 x 1944. Figures 2-8 shows the image results. Figure 9 shows the graph.

Preprocessing was not done as the obtained images were already clear and did not need any processing.
4. Image Segmentation

Segmentation - progression of sectionalization of the image into its numerous segments. segmentation was classically used for locating matters and borders in images. We have used K means clustering algorithm. We have used color images for segmentation. The ROI (Region of Interest) is WBC’s as those are the ones affected in case of leukemia. The input image is changed into HSV color space.

Figure 4. Hue (Cancer and Normal cell)

Figure 5. Saturation (Cancer and Normal cell)
5. Feature Extraction and Classification

It is the part of dimensionality reduction procedure, in which, an early set of the raw data is divided and condensed to more controllable groups. So when you want to process it will be easier.

Classification is a process relating to categorization, the process in which objects and ideas are recognized, distinguished.

Feature Extraction and classification are done using VGG Architecture. The shape and color features of the image were obtained from processed input data and then features obtained were fed as the input of the classifier system.
The accuracy of developed system is 93%. This system could be of real help as the time taken is reduced compared to visual inspection and the possible errors due to humans could be avoided.

![Accuracy plot of the Model](image)

**Figure 9** Accuracy plot of the Model

6. **Conclusion**

In this project we have developed a model to detect Leukemia from microscopic images using python. In this method, image processing techniques like segmentation, Dilation and erosion are used to obtain the affected White blood cells. Then the segmented image is used and features like shape and color features are extracted using the VGG16 architecture and the images are classified based on the acquired features. The output of the model will show if the input image is cancerous or not.
References

[1] A. Haldorai and A. Ramu, Security and channel noise management in cognitive radio networks, Computers & Electrical Engineering, vol. 87, p. 106784, Oct. 2020. doi:10.1016/j.compeleceng.2020.106784

[2] A. Haldorai and A. Ramu, Canonical Correlation Analysis Based Hyper Basis Feedforward Neural Network Classification for Urban Sustainability, Neural Processing Letters, Aug. 2020. doi:10.1007/s11063-020-10327-3

[3] D.M.U., Sabino, L.D.F., Costa, L.D.F., E.G., Rizzatti, M.A., Zago, A Texture Approach to Leukocyte Recognition, Real Time Imaging, 10, 2004, pp. 205-206.

[4] M.C., Colunga, O.S., Siordia, S.J., Maybank, Leukocyte Recognition Using EM-Algorithm, MICAI 2009, LNAI 5845, Springer Verlag Berlin Heidelberg, 2009, pp. 545-555.

[5] R. Donida Labati, V. Piuri, F. Scotti, ALL-IDB: the acute lymphoblastic leukemia image database for image processing, in Proc. of the 2011 IEEE Int. Conf. on Image Processing (ICIP 2011), Brussels.

[6] F. Scotti, Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images, in Proc. of the 2005 IEEE Int. Conf. on Computational Intelligence for Measurement Systems and Applications (CIMSA 2005), Giardini Naxos - Taormina, Italy, pp. 96-101, July 20-22, 2005

[7] V. Piuri, F. Scotti, Morphological classification of blood leucocytes by microscope images, in Proc. of the 2004 IEEE Int. Conf. on Computational Intelligence for Measurement Systems and Applications (CIMSA 2004), Boston, MA, USA, pp. 103-108, July 12-14, 2004.

[8] Ruggero Donida Labati, Vincenzo Piuri, Fabio Scotti, The Acute Lymphoblastic Leukemia Image Database for Image Processing, Universita Degli Studi Di Milano.

[9] Behrooz Ghandeharian, Hadi Sadoghi Yazdi and Faranak Homayouni, Modified Adaptive Centre Eighted Median Filter for Uppressing Impulsive Noise in Images, IJRRAS, Vol.1, Issue.3, December 2009.

[10] MedlinePlus, The U.S. National Library of Medicine [Online].http://www.nlm.nih.gov/medlineplus/ency/article/003643.htm

[11] G. Naxos and F. Scotti, CIMSA 2005-IEEE International Conference on Computational Intelligence for Measurement Systems and Applications Automatic Morphological Analysis for Acute Leukemia Identification in Peripheral Blood Microscope Images, no. July, pp. 20–22, 2005.

[12] F. Scotti, Robust segmentation and measurements techniques of white cells in blood microscope images, April, pp. 43–48, 2006.

[13] T. Bergen, D. Steckhan, T. Wittenberg, and T. Zeerfuss, Segmentation of leukocytes and erythrocytes in blood smear images, Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc EMBS’08 - Personalized Healthc. through Technol. pp. 3075–3078, 2008