Acute lymphoblast leukemia and multiple myeloma classification using neural network and multilevel Otsu thresholding

[1]Dishant Totade, [2]Ashish Shrivastava,

[1]Research Scholar, [2]Assistant Professor

Department of Computer Science and Engineering
Lovely Professional University, Punjab, India.

Abstract: Acute Lymphoblast leukemia and multiple Myeloma is a kind of blood cancer that affects the white blood cells. Early diagnosis is very important to prevent the progression of cancer. The main objective of this research paper is to classify the blood cancer from the microscopic image of the patient's blood smear. Microscopic blood image analysis results in the early diagnosis of leukemia and myeloma with lower cost. It is less costly to use image for diagnosis, compare to the equipment and methods used in the field of Hematology. The aim of this study is to identify, whether the cell belong to leukemia class or myeloma class. For this 30 microscopic images from both classes is used to train the Convolution Neural Network model and Multi Otsu algorithm to extract the cell region of interest. A degree of accuracy 97.72% gives high performance of the proposed method.

Keywords: Neural network, multilevel Otsu, image processing, multiple myeloma, acute lymphoblast leukemia.

I. Introduction

Blood cells are of various type, there are white blood cells (fights viruses and bacteria in the body) Red blood cells (carries oxygen throughout the body) Platelets (small cells that prevent clotting) and from these three cells blood is formed. But due to abnormalities in cells, diseases may happen. Acute lymphoblastic leukemia and multiple myeloma are the type of blood cancer associated with white blood cells. It is essential to have a proper guided system to detect these cells from the blood smear microscopic images. Multiple Myeloma is very complex diseases. It is a type of white blood cells located in the bone marrow. These white plasma cells helps to fight bacteria and virus in the body. When cancer affects the blood it causes abnormalities in the white plasma cells. This leads to weaken the immune system of the body. Symptoms cannot be found with myeloma until it expands throughout the body. With the help of hematologist monitoring and active surveillance on patient, it can be improved. In the later stage of blood cancer it leads to primary malignant bone tumor. Common symptoms of myeloma are bone pain, Back pain, possible fractures, weakness and fatigue. Leukemia cancer is also white blood cells diseases, it affect the growth of plasma membrane in the white blood cell. Leukemia also forms inside the bone marrow and as it expands can cause tumor inside the bone. Leukemia stops the evolution of the white blood cells and promotes the functionless cells in the bone marrow. It is very important to get clinical care in the early stage as soon as possible. Leukemia cells goes through bloodstreams to other parts of body organs, brain, liver where they grows and divides.

II. Literature Review

This is the related work of blood cancer detection and classification like Thanh et al. [1] used convolution neural network with histogram analysis and data augmentation to classify normal and abnormal leukemia from ALL-IDBI image dataset which consist of 1188 images that gave the accuracy of 96.6%. Nizar et al. [2] identifying the different type of leukemia subtype from microscopic cell images using multi class convolution neural network with accuracy of 81.74%. Faisal et al. [3] classifying acute myeloid leukemia (AML) and acute lymphoblast leukemia (ALL) using back propagation neural network with HSV color feature, texture feature extraction using GLCM (gray level co-occurrence matric) with accuracy of 86.66%. Adnan et al [4] proposed a methods for the identification of normal blood cells or ALL infected cell using spatial filter and automatic edge detection on ALL-IDB dataset [12] with neural classifier gives the accuracy of 90%. Vyshnav et al. [5] have compared two deep learning techniques i.e. mask-RCNN and Unet neural network for the detection and segmentation of multiple myeloma cancer cells, which gives the accuracy 93.99% and 89.62% for mask-RCNN and Unet respectively. The details have been discussed of mark-RCNN in [11]. Samabia et al. [6] presented a computer aided methodology for multiple myeloma detection and diagnosis. Pre-trained and fine-tuned AlexNet is used for diagnosis of myeloma. Microscopic images are classified into normal and blast using AlexNet as feature extractor and support vector machine for classification. Rohit et al. [7] detects white blood cells cancer diseases (AML) Acute myeloid leukemia and (ALL) Acute lymphoblast leukemia using combination of Gaussian distribution, Otsu Adaptive Threshold, K-mean clustering, GLCM (Gray Level Co-occurrence Matrix) have been used as feature extractor to train CNN for the classification with accuracy of 97.3%. Deepika et al. [8] compared different machine learning algorithms and Dense Convolution neural network for the classification of acute lymphoblast leukemia and multiple myeloma on SN-AM

dataset. In which Support Vector Machine (SVM), Random forest, decision tree, naive bayes have shown the accuracy of 73.02%, 96.83%, 96.77%, and 74.6% respectively. For VGG-16, CNN have shown the accuracy of 90.1%, 97.25% respectively.

III. METHODOLOGY

We have applied three main stages in this project they are: Image processing, Image Augmentation and Neural Network. The dataset consist of 30 images of Acute Lymphoblast Leukemia and 30 images of Multiple Myeloma. Microscopic images were captured from bone marrow aspirate slides of patient's diagnosis with B-Acute Lymphoblast leukemia (ALL) and Multiple Myeloma (MM). Slides were stained using Jenner - giemse stain. Images were capture 1000x magnification using Nikon Eclips-200 microscope [9]. The detail knowledge of dataset and the cancer image archive (TCIA) is discussed in [10].

3.1 Image Processing

In image processing stage, we applied five steps to extract the white blood cells regions of interest area from the microscopic image of both classes (ALL and MM). Steps are shown in figure (1).

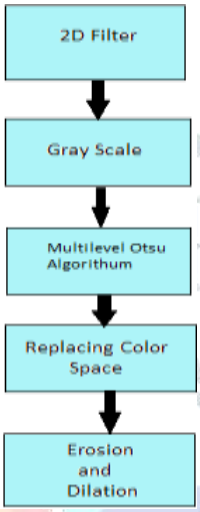

fig. 1.

As, it can be seen in figure (a) and (b) the microscopic image of ALL and MM is very complex. First step is to apply a 2 Dimension filter with kernel M size of matrix (7, 7) on both classes. It has been found that for the leukemia class the following kernel of equation (1) is suitable.

$$M_{7\times 7} = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{bmatrix}$$

$$\text{Kernel of ALL} = [M]\frac{1}{19} \qquad (1)$$

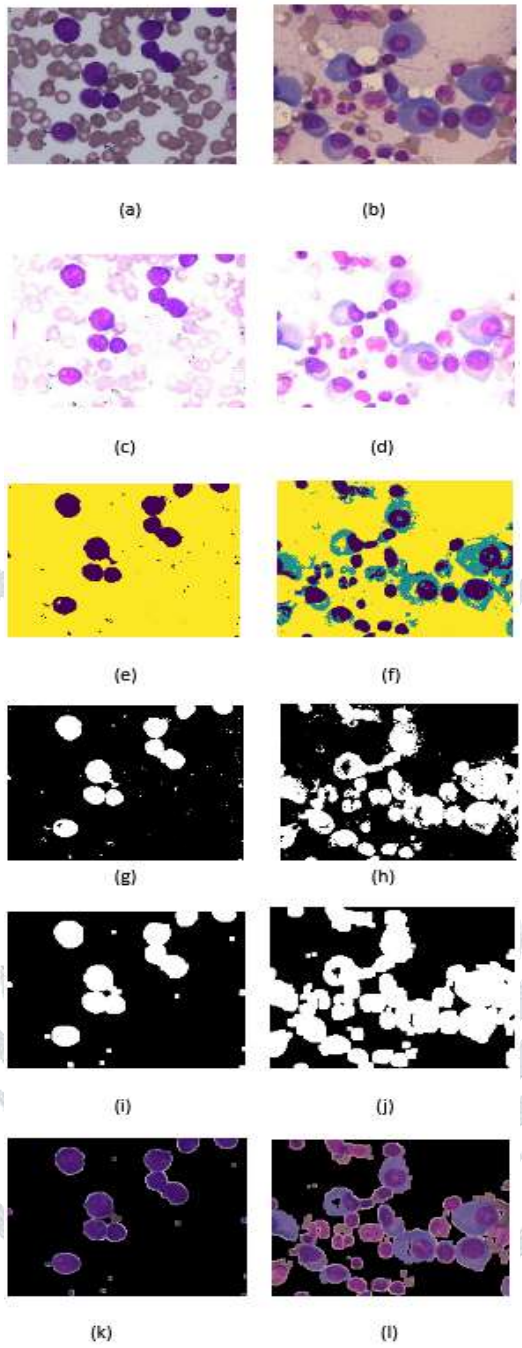

Fig. 2. Image Processing Steps (a) Acute Lymphoblast leukemia, (b) Multiple Myeloma, (c) and (d) 2D filter, (e) and (f) Multi Otsu Output, (g) and (h) Replaced Color, (i) and (j) Erosion and Dilation, (k) and (l) Bitwise AND operation

Similarly for the myeloma class the following kernel of equation (2) is suitable.

$$\text{Kernel of MM} = [\text{M}]\frac{1}{23} \qquad (2)$$

Consider α and β are the original images figure (a) and (b) belong to class ALL and MM respectively and x, y are the dimension of pixels. Then from equations (1) and (2) for the background removal, we apply the following equation.

$$= \sum \quad \sum_{x=1\ y=1}^{x=n\ y=n} \alpha(x,y) \times Kernel\ of\ ALL(x,y) \qquad (3)$$

$$= \sum \quad \sum_{x=1\ y=1}^{x=n\ y=n} \beta(x,y) \times Kernel\ of\ MM(x,y) \qquad (4)$$

After iterating equations (3) and (4) on images the output is shown in figure (c, d) respectively. Then converting the filtered images into grayscale images. Next step is to apply Multilevel Otsu Algorithm [13] for the segmentation of the image. Multi Otsu helps to divide the image pixels data into different classes based on threshold values. We applied Multi Otsu with 3 classes

on multiple myeloma because there are three major parts in this image, background, white blood cells nuclei and cytoplasm figure (d) and its classes are yellow, dark blue, green pixels figure (f). Similarly for the leukemia image multi Otsu with 2 classes is applied because only white blood cells with nuclei figure (c) is present and its classes are yellow and dark blue pixels figure (e). The segmented image have less complexity, now it is easy to extract the cell area. Next step is to replace the color space of segmented images with black and white color space. Yellow is replaced with black color. Green and dark blue is replace with white color. Masked images is created figure (g) and (h) with cells area in white color. But there are unwanted small dark spots in the cell area of interest. To remove this spots, we adopted Erosion [14] and Dilation [14] methods. These methods work on the corners and edges of the object in the images. Erosion removes the pixels (reduce the pixels) from the objects in the image. Dilation adds the pixels to the object. We applied first erosion with 2 iteration and then dilation with 3 iteration both with (17, 17) kernel size result shown in figure (i) and (j). To extract the cell region from the original image using the output of erosion and dilation we used bit wise AND operation in python's opencv library. By taking the reference of masked image bit wise AND operation considers the white space and extracts exact feature map from the original image which results figure (k) and (l) white blood cells area. Later this image processing approach is applied on whole dataset to train the neural network classifier.

3.2 Image Augmentation

After Image processing, it is very necessary to use image augmentation. It is possible to train neural network with limited data with the help of image augmentation technique. It enlarges the data and creates the transformed images. Image augmentation function that we have used on images are brightness, flip vertical, flip horizontal, zoom image, shift width, shift height, rotate image, transformation.

3.3 Neural Networks

We trained binary classifier on the dataset with traditional approach by creating neural network architecture using open source CNN frame work keras. Input layer of network with input size of 300×300. First hidden layer of 32 feature map with 3×3 kernel. second hidden layer of 32 feature map with 3×3 kernel and activation function of rectified linear unit (RELU) followed by dimension reduction function max pooling. third hidden layer of 64 features map with 3×3 kernel size followed by max pooling function. Next, flatten function to convert two dimensions feature map into one dimension feature vector. Next, two fully connected layer with 64 and 128 dimensions feature vector with RELU activation function. To avoid the over fitting of training dataset on model a dropout layer of 0.5 is used followed by the output layer with sigmod activation function to give the label 1 for ALL class and 0 for MM class. For training, one of the traditional loss function binary cross-entropy and network optimizer Adam is used in this study. The model is trained with 30 epoch and loss of training stop decreasing at 0.07. Improving accuracy of training and testing is shown in figure Fig.3. Decreasing loss value of training and testing is shown in Fig.4.

IV. Result and Performance Measure

In order to evaluate the performance of the trained model. We used the testing data to predict the binary classes. The testing data consist of 20 images of MM (Multiple myeloma) class and 20 images of ALL (Acute Lymphoblast leukemia). The results is calculated in the following metrics, Precision, Recall, F1 Score, Sensitivity, Specificity and for the Confusion Metrix following term are used True Positive (tp) are the predicted true results for the given condition which are actually true, True Negative (tn) are the predicted false results for the given condition which are actually false, False Positive (fp) are the predicted true results for the given condition which are actually false. False Negative (fn) are the predicted false results for the given condition which are actually true, (N) is total number of sample.

$$\text{Accuracy} = (tp + tn) \div N$$

$$\text{Precision} = tp \div (tp + fp)$$

$$\text{Recall} = tp \div (tp + fn)$$

$$\text{F1 score} = 2tp \div (2tp + fp + fn)$$

$$\text{Sensitivity} = tp \div (tp + fn)$$

$$\text{Specificity} = tn \div (tn + fp)$$

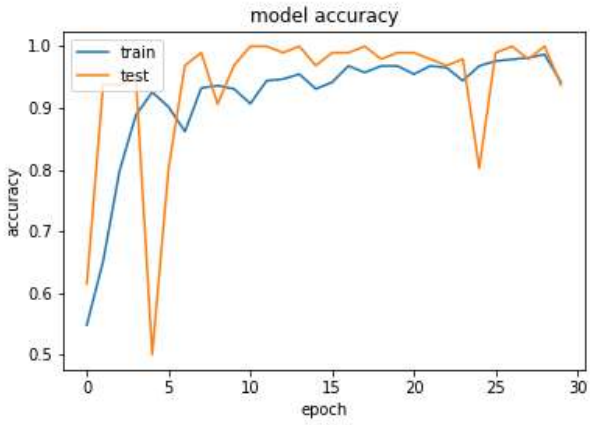

fig.3. model accuracy on testing and training.

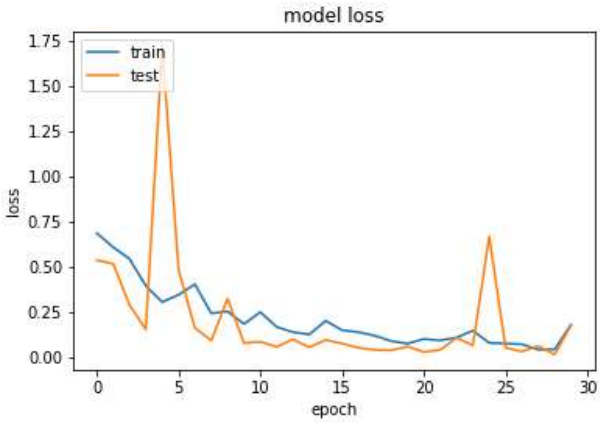

fig.4. model loss on testing and training

Confusion matrix Fig.5. is created on the predicted labels using sklearn python library in which True classes on the left side and predicted classes on the bottom. Model predicted 23 classes correct out of 23 in MM and 20 classes correct out of 21 in ALL classes. From the figure Fig.5, tn = 20, tp = 23, fp=1, fn=0.

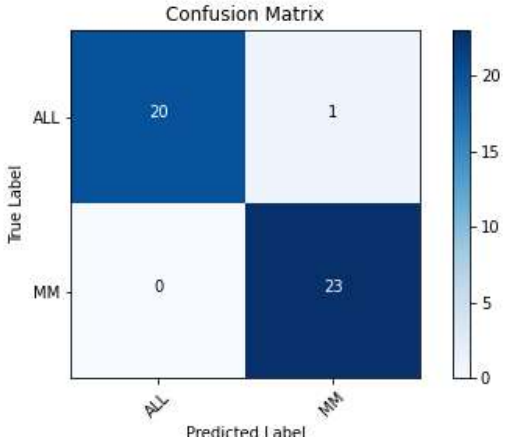

fig.5. confusion matrix of predicted label

Table 1. Predicted Results

S.No		
1.	Accuracy	97.72%
2.	Precision	95.83%
3.	Recall	99.25%
4.	F1 Score	97.87%
5.	Sensitivity	99.98%
6.	Specificity	95.23%

V. Conclusion

The present research tried to classify the blood cancer type Acute Lymphoblast Leukemia and Multiple Myeloma. The dataset consists of microscopic images of both types. Multilevel Otsu image processing algorithm is used to extract the abnormal white blood cancer cells from the images. Convolution Neural Networks is used to build the binary classifier. Simulated results show that the proposed system is reliable, efficient and less time consuming for the detection of cancer in the early stages of diagnosis. The proposed methods shows the overall accuracy of 97.72%.

References

[1] Thanh, T. T. P., Vununu, C., Atoev, S., Lee, S., & Kwon, K. (2018). Leukemia Blood Cell Image Classification Using Convolutional Neural Network, (April). https://doi.org/10.7763/IJCTE.2018.V10.1198

[2] Images, M., Convolutional, U., & Network, N. (2019). Identification of Leukemia Subtypes from Microscopic Images Using Convolutional Neural Network.

[3] Asadi, Faisal., Putra, F. M., Sakinatunnisa, M. I., Syafria, F., & Marzuki, I. (2017). Implementation of Backpropagation Neural Network and Blood Cells Imagery Extraction for Acute Leukemia Classification. 2017 5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), (November), 106–110.

[4] Khashman, Adnan., & Abbas, H. H. (2013). Blood Smear Images and a Neural Classifier, 80–87.

[5] Vyshnav, M. T., Sowmya, V., Gopalakrishnan, E. A., V, S. V. V, Menon, V. K., & Soman, K. P. (2020). Deep Learning Based Approach For Multiple Myeloma Detection.

[6] Tehsin, Samabia, Zameer, S., & Saif, S. (n.d.). Myeloma Cell Detection in Bone Marrow Aspiration Using Microscopic Images. 2019 11th International Conference on Knowledge and Smart Technology (KST), 57–61.

[7] Agrawal, Rohit. (2019). Detection of White Blood Cell Cancer using Image Processing. 2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN), (2018), 1–6.

[8] Deepika K., Jain, N., Khurana, A., Mittal, S., Satapathy, C., & Senkerik, R. (2020). Automatic Detection of White Blood Cancer from Bone Marrow Microscopic Images Using Convolutional Neural, x. https://doi.org/10.1109/ACCESS.2020.3012292

[9] Gupta, A., & Gupta, R. (2019). SN-AM Dataset: White Blood Cancer Dataset of B-ALL and MM for Stain Normalization [Data set]. The Cancer Imaging Archive. https://doi.org/10.7937/tcia.2019.of2w8lxr

[10] Clark K, Vendt B, Smith K, Freymann J, Kirby J, Koppel P, Moore S, Phillips S, Maffitt D, Pringle M, Tarbox L, Prior F. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository, Journal of Digital Imaging, Volume 26, Number 6, December, 2013, pp 1045-1057. DOI: 10.1007/s10278-013-9622-7

[11] Dishant Totade,Ashish Shrivastava, "IMAGE ANALYSIS USING NEURAL NETWORKS: A REVIEW", International Journal of Creative Research Thoughts (IJCRT), ISSN:2320-2882, Volume.8, Issue 11, pp.975-980, November 2020, Available at :http://www.ijcrt.org/papers/IJCRT2011116.pdf

[12] 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.

[13] Liao, P. S., Chen, T. S., & Chung, P. C. (2001). A fast algorithm for multilevel thresholding. Journal of Information Science and Engineering, 17(5), 713–727. https://doi.org/10.6688/JISE.2001.17.5.1

[14] Kimmel, R. (2013). Efficient Dilation, Erosion, Opening and Closing Algorithms Efficient Dilation, Erosion, Opening and Closing Algorithms Department of Computer Science, 24(August), 1606–1617.