Proceedings of the International Conference on Applied Artificial Intelligence and Computing (ICAAIC 2022)
IEEE Xplore Part Number: CFP22BC3-ART; ISBN: 978-1-6654-9710-7

Detection of Leukemia using Machine Learning

[1]M. Iswarya, Shivakami. A[2], Aravindha Mira. R[3], Karpagam. A[4]
[1]Assistant Professor, Sri Sai Ram Engineering College,
[2,3,4]Student, Sri Sai Ram Engineering College

2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC) | 978-1-6654-9710-7/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICAAIC53929.2022.9792725

ABSTRACT

Leukemia is a type of cancer which damages the blood-forming tissues that produces rbc's, wbc's and blood platelets, which is present inside bone marrow. Leukemia is the cancer that mainly occurs within blood forming tissues in the human body. It may slowly affect all the organs inside the body without any external indication leading to several other disorders. The diagnosis of leukemia in early stages of the disease is difficult as the current technologies are more time consuming. Studies suggest that India ranks 3rd highest in reported cases of Blood Cancer, after USA and China. If a person is already diagnosed with any one among other types of cancer, then the chances of developing leukemia is nearly 30% .The symptoms of leukemia are tiredness,flu,fatigue and weight loss which makes it very difficult to predict manually by clinical methods. Since it is malignant and fatal it is necessary that it must be predicted earlier in order to undergo sufficient treatment and thereby saving millions of lives.As compared to all other types cancers, leukemia cases have been increasing dramatically in the entire world.Since the symptoms of leukemia are being very common, people fail to detect it in the early stage.The main objective of our project is to diagnose leukemia in the primary stage by analysing peripheral blood smear images and thus treatment could be given in specific and successful way.Thus our project focusses on using the best machine learning algorithms for detecting leukemia in a very early stage. This research study collects peripheral blood smear images(PBS), which are microscopic images of the blood samples. The images collected are preprocessed and then segmentation is performed where the images are segmented based on pixels and enlarged portion of affected area of the blood sample image is used for feature extraction process. Finally classification is done using CNN in order to detect leukemia.

Keywords*:-Leukemia,Bone marrow, Peripheral Blood Smear,Segmentation, Convolutional Neural Networks*

I.INTRODUCTION

Blood has three types of cells namely white blood cells are the ones that fight with the infection, red blood cells usually carry oxygen, and platelets helps in blood clotting.Our bone marrow usually makes millions of fresh blood cells, and 93% of them are erythrocytes. When a person is diagnosed with leukemia, the body produces more white blood cells than it actually needs and these leukemic cells could not fight with the infection in the body in the way normal white blood cells fight.They slowly start to affect the working of our internal and external organs. After some time, it will result in lack of red blood cells to supply oxygen, enough platelets to clot your blood, or enough normal white blood cells to fight infection.Thus the severity of the disease increases drastically.Thus proper diagnosis and treatment has to be given in the early stage itself.Leuklemia can be broadly classified into acute leukemia and chronic leukemia.Acute leukemia requires fast treatment as soon as it is detected. In chronic leukemia, some parts in our body will produce too many blood cells. Some of the types of chronic leukemia initially shows no symptoms and can be undiagnosed for several time period.By using medical science bone marrow biopsy and spinal tap tests are done for leukemia detection. Bone marrow biopsy is a test which is done for detecting leukemia in the human body.Yet another approach is usage of Spinal tap where fluid from your spinal cord is taken and diagnosed with which spread of leukemia can be detected. These clinical diagnosis may take more time to know the exact results and are more expensive and painful techniques.Therefore by using machine learning algorithms the diagnosis becomes faster and easier.The

Proceedings of the International Conference on Applied Artificial Intelligence and Computing (ICAAIC 2022)
IEEE Xplore Part Number: CFP22BC3-ART; ISBN: 978-1-6654-9710-7

blood smear images collected will be the input given to the machine learning code.After the process of segmentation and extraction the final results of whether a person is affected with leukemia or not will be identified.

II.SCOPE OF THE PROJECT

1.The aim of our proposed system is to detect the presence of leukemic cells in the blood in our body by using advanced machine learning techniques.

2.The peripheral blood smear images are used and thus detection of the disease is done in a very early stage by using an efficient algorithm thus helping the doctors in treating the patients.

3.Using CNN algorithm the process of identifying the leukemic cells is performed which provides most accurate diagnosis.

III.EXISTING SYSTEM

In general,the diagnosis of leukemia is done by spinal tap,imaging tests and bone marrow biopsy which are more complicated and painful. The existing system makes use of traditional machine learning algorithms like SVM and ANN for diagnosing the cancer of blood forming tissues. Though these algorithms have been used for implementation,the performance of SVM is low when compared to CNN in classifying types of leukemia. And feature extraction in ANN algorithm is not much accurate as compared to CNN algorithm.

IV.LITERATURE SURVEY

The classification and detection of WBC cancer and some of its subtypes are done by various researchers and scholars by using different algorithms.

[1]IoMT-Based Automated Detection and Classification of Leukemia Using Deep Learning

This paper is purely based on Internet of Medical Things which involves cloud computing techniques that enables intercommunication between diagnosis machines for leukemia detection that helps in real time coordination for testing and diagnosis.

[2] Machine Learning in Detection and Classification of Leukemia Using Smear Blood Images.

This project makes use of machine learning technique namely random forest model to detect the leukemic cells.It had an accuracy of 94.3%.This classifier performs test on 65 sample images which seemed to be less accurate.

[3]Automated detection of acute lymphoblastic leukemia from microscopic images based on human visual perception.

It proposes an approach which is purely based on the conventional digital image processing techniques and machine learning algorithms which helps to automatically detect acute lymphoblastic leukemia.

[4] Automated decision support system for detection of leukemia from peripheral blood smear images.

In the above proposed system,. SVM classifier was used for classification of the white blood cells into normal and leukemic cells.The over all accuracy was around 90%.

[5] . Acute myeloid leukemia diagnosis using deep learning.

In this approach algorithm was designed for classification of a subtype of leukemia named acute myeloid which uses principal component analysis (PCA) based on ABC-BPNN scheme.Here the classification of leukemia cells is performed which attained 95% accuracy.

V.PROPOSED SYSTEM

In the proposed system,peripheral blood smear images which are microscopic images of the blood samples are taken as the input datasets.The images collected are preprocessed by using various filters.The process of segmentation is performed by using zack algorithm where wbc detection happens and lymphocytes are grouped together.Feature extraction is followed by segmentation process and various features including cell structure,size,nuclei,cytoplasm are extracted and finally classification is done using CNN with which we detect leukemia and hence the lives of millions of people can be saved by detecting it in a very early stage.We have used CNN for classification purpose because comparing

Proceedings of the International Conference on Applied Artificial Intelligence and Computing (ICAAIC 2022)
IEEE Xplore Part Number: CFP22BC3-ART; ISBN: 978-1-6654-9710-7

to existing models namely SVM classifier,ANN,KNN and so on the performance is efficient with CNN method and accuracy of the project is high.

VI.SYSTEM IMPLEMENTATION

The implemention starts with collecting the peripheral blood smear images..Once the images are collected they are fed as input for classification process to take place.Initial stage of classification is data preprocessing followed by segmentation of the wbc's separately,then feature extraction happens and finally classification is performed.Classification is the stage by which we detect the presence of leukemic cells by undergoing comparision of the input image with the features set.It is performed by using CNN algorithm.The final stage will provide the desired output of whether leukemic cells are present in the blood or not.

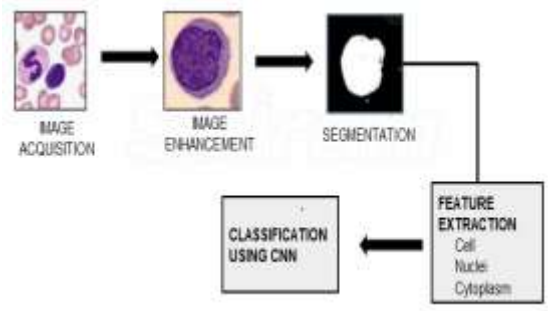

Fig1 System architecture diagram

VII.MODULE DESCRIPTION:

The entire implementation is split up into fourmodules for easier processing.

Module 1: ImagePreprocessing
Module 2: Segmentation
Module 3:Feature extraction
Module 3: Classification of leukemic cells
The output obtained at the end of imagepreprocessingstage will be segmented in the next stage and by using convolutional neural network technique it undergoes classification process successfully.

A.IMAGE PREPROCESSING

Data preprocessing is the first stage of classification process.For this,peripheral blood smear images are taken as input datasets and further they are resized,filteration happens,noises are removed and edges are soothened.Median filter is used for removing noise from the images and weiner filter is used for removing blurriness of the cell clusters.At the end of preprocessing the input data set will be ready for segmentation process.

B.DATASET

We have downloaded datasets which are available publicly from the Kaggle website. The dataset is basically the peripheral blood smear images(PBS).These PBS images are obtained from blood samples which are basically stained using microscopic slides in the laboratory.In our project 1000 samples are used as the input datasets of which 80% is used for training and remaining 20% for testing.

C.SEGMENTATION

Segmentation is a method of isolating and extracting segments from the entire data which is mainly applied for feature extraction and selection.After preprocessing is completed,segmentation will be performed.The main task of segmentation is to reduce the complexity and further to reduce the size of high resolution images.Zack algorithm is used for wbc detection and lymphocytes isolation which is one of the types of wbc's.
The two main steps involved in this algorithm for segmentation purpose are,

- Localization of leukocytes
- Extraction of region of cells

Thus in the segmentation process,leukocytes are separated from the entire cell clusters and particular region of cells are extracted to undergo further process.

D.FEATURE EXTRACTION

In this stage,feature extraction is performed and various features are extracted with the help of which classification is done.Some of the features are nucleus size and cytoplasm features by using lbp algorithm.Here a separate set of data are collected for leukemic cell[cancerous cell]. Once the features are extracted for

Proceedings of the International Conference on Applied Artificial Intelligence and Computing (ICAAIC 2022)
IEEE Xplore Part Number: CFP22BC3-ART; ISBN: 978-1-6654-9710-7

the particular cell which is given as the input for classifying,analysis is done by using CNN algorithm where comparison between normal and leukemic cells happens followed by classification.

E.CLASSIFICATION

Finally,by processing the image datasets which are blood smear images of the patient, we proposed convolutional neural network algorithm in the ANACONDA software which develops a model for leukemia detection so that when an PBS image is inserted as an input, it compares with the leukemic cells as well as normal cells and predict the output..Now this will give the doctors an idea of the exact stage of leukemia in the patient's body.

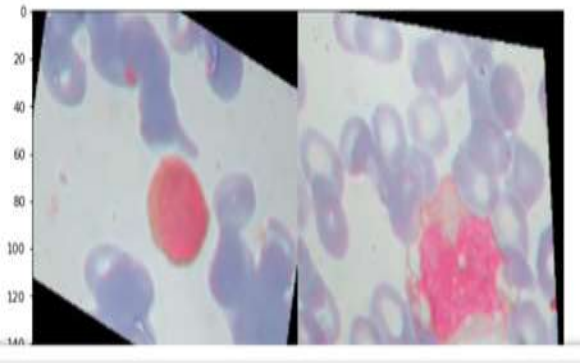

Fig2-Comparison and analysis of normal and leukemic cells

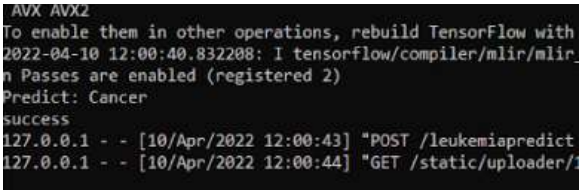

Fig 3- Output of the prediction status of leukemia

F.PERFORMANCE

ACCURACY:

The overall accuracy of the project is calculated by finding accuracy vs epoch graph where the calculation is done for both training and testing data. The accuracy is found to be greater than 90%.

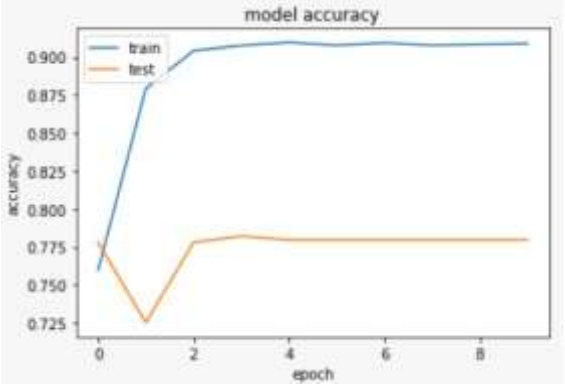

Fig 4Accuracy vs epoch

VIII.CONCLUSION

Leukemia is a major problem nowadays that should be detected and requires proper treatment.Among many types of cancers,leukemia is the deadliest since it affects 23% of the total population Further the symptoms of leukemia are not much exposive.Thus, the novelty of our idea is to provide an algorithm to diagnose leukemia in cost-effective method and with greater evaluation accuracy.The proposed non-invasive method detects the occurrence of leukemia at an early stage. Furthermore, the results can be extended for therapeutic applications.

IX.FUTURE ENHANCEMENT

The future of this project is to provide an algorithm to detect the subtypes of leukemia namely acute lumphocytic,acutemyeloid,chronic lymphocytic and chronic myeloid.On detecting the leukemic type,the treatment could be started at an earlier stage thus millions of lives can be saved.

X.REFERENCES

[1]IoMT-Based Automated Detection and Classification of Leukemia Using Deep Learning,*Ahmad Almogren,2021*

[2] Machine Learning in Detection and Classification of Leukemia Using Smear Blood Images: A Systematic Review,*Mustafa Ghadezadeh,2021*

[3]Automated detection of acute lymphoblastic leukemia from microscopic images based on human visual perception,*J. Zidek,2020*

[4]Automated decision support system for detection of leukemia from peripheral blood smear images. *K. Prasad,2019*

[5] Acute myeloid leukemia diagnosis using deep learning,*F.Hussain,2019*

[6]Identification of Leukemia Subtypes from Microscopic Images Using Convolutional Neural Network,*Nlzar Ahmed,2019*

[7]A decision-tree approach for the differential diagnosis of chronic lymphoid leukemias and peripheral B-cell lymphomas,L.O.Moreas,2019

[8] Detection of Acute Lymphocytic Leukemia Using Blast Cell Morphological Features, Dharani.T ,2019

[9]A decision-tree approach for the differential diagnosis of chronic lymphoid leukemias,*Orfao.A,2019*

[10]A decision support system for Acute Leukaemia classification based on digital microscopic images,*Ahmed,Osama,2018*

[11] Leukemia Blood Cell Image Classification Using. Int. J. Comput. Theory ,*Sukhrob, A.; Suk-Hwan, L,2018*

[12]A decision support system for Acute Leukaemia classification based on digital microscopic images,*Ahmed H,2018*

[13]Detection and Classification of Blood Cancer from Microscopic Cell Images Using SVM KNN, *Rathore.Y,2017*

[14] Microscopic Segmentation Using Fuzzy C Means for Leukemia Diagnosis, *Karthikeyan, T.; Poornima,2017*

[15]Diagnosis and management of AML in adults, *Amadori S,*2017