IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁐ Impact Factor 7.918 ⁐ Vol. 11, Issue 10, October 2022

DOI: 10.17148/IJARCCE.2022.111011

Detection of Leukemia by Convolutional Neural Network

Autade Gayatri Shivdas[1], Bobade Vivek Anil[2], Kanawade Satyam Pravin[3], Sanap Dipali Dinkar[4], Prof. Sharad M. Rokade (Guide)[5]

Sir Visvesvaraya Institute of Technology, Nashik[1-4]

Abstract: Every year, over 900,000 individuals worldwide are diagnosed with Leukemia, i.e., Blood Cancer, but many people are oblivious of the dangers involved with such often incurable diseases. The majority of Blood Cancers are rare, life-threatening illnesses within limited patient populations; together, they account for 7% of all malignancies. Patients may feel abandoned and have difficulty finding the necessary assistance and information due to the complex, often sparse nature of Leukemia. When it comes to Acute Leukemia, if therapy is not started on time, the patient might succumb to the ailment within a few months. It is vital to diagnose Cancer be it of any type, in its early stages to ensure timely treatment and increase chances of survival. Detecting Leukemia manually in labs by medical personnel examining blood samples is a time and resource-consuming procedure. Customarily, the patients suffering do not have the liberty to exhaust their time as they need immediate care. We need systems that can make use of the latest technological developments in artificial intelligence to produce expeditious and more accurate results.

Keywords: Convolutional Neural Network, Support Vector Machine (SVM), Image Processing, neural networks, decision trees.

I. INTRODUCTION

When it comes to Acute Leukemia, if therapy is not started on time, the patient might succumb to the ailment within a few months.

It is vital to diagnose Cancer be it of any type, in its early stages to ensure timely treatment and increase chances of survival.

We need systems that can make use of the latest technological developments in artificial intelligence to produce expeditious and more accurate results.

II. MODELING AND ANALYSIS

Image pre-processing, feature extraction and classification plays major role in detection of white blood cell cancers. The removal of noise is another main factor that yields accuracy to final result. The following flowchart depicts the basic steps involved in automated detection of white blood cancer cells using machine learning and image processing techniques.

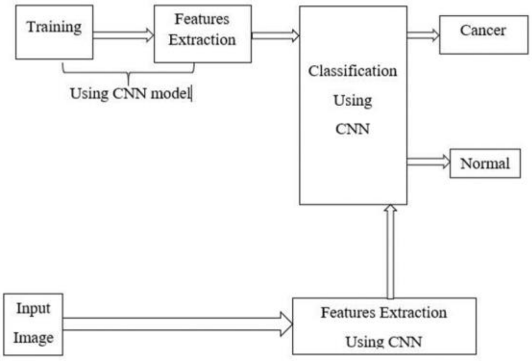

Figure 1: System Architecture

DOI: 10.17148/IJARCCE.2022.111011

Image pre-processing, feature extraction and classification plays major role in detection of white blood cell cancers. The removal of noise is another main factor that yields accuracy to final result. The following flowchart depicts the basic steps involved in automated detection of white blood cancer cells using machine learning and image processing techniques.

INPUT IMAGE This stage is also known as image acquisition. For detecting, blood cancer diseases like leukemia or myeloma, input images from microscope is needed. This informational index or data-set comprises minuscule pictures of blood smear, particularly intended for the assessment and the correlation of algorithms for segmentation and image classification.

IMAGE PRE-PROCESSING The main goal is to de-noise the input image. Preprocessing also includes color relevance wheel RGB images are converted to Grey color space images.

SEGMENTATION OF THE IMAGE It is the process of apportioning an advanced picture into numerous portions. This is one of the most crucial steps which greatly influence on the accuracy of cancer cell detection.

FEATURE EXTRACTION This method used to change the information into set of highlights and is a type of dimensionality decrease. The feature set will separate the critical data from the input information if the highlights are extricated effectively. Some of the features are Haar wavelet features, Hausdorff dimension with and without LBP, Shape features, Gray level co-occurrence features, Haralick's texture features or GLCM, Color feature.

CLASSIFICATION The final stage in the process is classification. It helps to determine the existence of white blood cancer cell in blood image. It includes several algorithms like random forest, support vector machines, nearest neighbour, neural networks, decision trees.

The data set used in this study were taken from the Kaggle Dataset. It consists of 195 images (96 normal and 99 abnormal) and divided into 80% training data and 20% test data.

III. LITERATURE SURVEY

Research conducted by Bibi et. Al proposes an Internet of Medical Things (IoMT)-based architecture for improving and ensuring Leukemia detection. Clinical devices are linked to network resources in the proposed IoMT system using cloud computing. The technology allows patients and healthcare providers to coordinate testing, diagnosis, and treatment of Leukemia in real-time, potentially saving time and effort for both patients and clinicians. The system uses DenseNet-121 and ResNet-34 to identify the different subtypes of Leukemia. [1]

A review paper by Ghadezadeh et. Al examined the present state of all known ML-based Leukemia detection and classification models that handle PBS images in a thorough and systematic manner. The average accuracy of the ML algorithms employed in PBS image analysis to diagnose Leukemia was greater than 97 per cent, indicating that utilizing ML to detect Leukemia from PBS images might provide amazing results. In this survey, Deep learning (DL) outperformed all other machine learning algorithms in terms of precision and sensitivity in recognizing distinct types of Leukemia. [2]

A study by Salah HT et. Al comprised of a compilation of other studies which examine the usage of Machine Learning in diagnosing the different types of Leukemia. Hand-searching of references from similar research and the top Google Scholar results supplemented the automated search. A total of 58 papers were read in their entirety, with 22 studies being included. There were 12, 8, 3, and 1 research that discussed ALL, AML, CLL, and CML, respectively. [3]

A study by M. Akter Hossain et. Al focused on Be Acute Lymphocytic Leukemia (ALL) as it is the most frequent kind of Leukemia. Oncologists are well aware that cancer is considerably easier to treat if discovered early. They suggested a hands-on technique to detect the abnormal blood components prevalent in cancer patients (e.g., neutrophils, eosinophils, basophils, lymphocytes, and monocytes). They selected 14 features to prepare the dataset before determining four essential attributes that are important in determining a Leukemia patient. [4]

A paper by Litjens et. Al proposed "deep learning" as a method for improving the fairness and efficiency of histopathology slide analysis. There is an indication that using prostate cancer identification on biopsy and breast cancer metastasis detection in the sentinel lymph node as examples, this unique technique may lessen the load on pathologists while boosting the objectivity of diagnosis. They discovered that they can exclude 30- 40% of the population. The paper concludes that deep learning has enormous promise for improving prostate cancer detection and classification. [5]

IV. PROPOSED SYSTEM

They would like to use new dataset including four types of Leukemia to evaluate our architecture. In the future work, we would like to use some algorithm tuning in terms of weight initialization, activation function to improve performance of our CNN architecture.

Proposed Workflow The proposed method includes three main steps:

Step 1: Processing The method proposed in this paper used three operations when training the deep learning model. The following operations are:

A . Convert to RGB: In this operation convert all Leukemia images to RGB color model. B. Resize all to 227 × 227: All images have different pixel size because they take from different device, so in this operation change the dimension of all images to fixed 227 × 227 pixel.

C. Augmentation

Step 2: Pre-trained deep learning model and direct feature extraction. Our model used transfer learning Convolutional Neural Network (AlexNet) to be pretrained. The architecture of CNN consists of three types of layers: convolutional, pooling and fully connected layer.

Step 3: Classification with machine learning models to classify the features, we chose different machine learning models

Use case diagram is a simple representation of a user's interaction with the system that shows the relationship between the user and the different use cases in which the user is involved.

Image pre-processing, feature extraction and classification plays major role in detection of white blood cell cancers. The removal of noise is another main factor that yields accuracy to final result. The following flowchart depicts the basic steps involved in automated detection of white blood cancer cells using machine learning and image processing techniques.

The data set used in this study were taken from the Kaggle Dataset. It consists of 195 images (96 normal and 99 abnormal) and divided into 80% training data and 20% test data

1. **Input blood image for detection** The main aim of the proposed system is to Detection of Leukemia cell present in blood image.

2. **De-Noising** By using noise removal techniques or removing any other disturbances present in the data, it is preprocessed to get the fine-tuned image.

3. **Segmentation** is a commonly used technique in digital image processing and analysis to partition an image into multiple parts or regions, often based on the characteristics of the pixels in the image.

4. **Feature Extraction** These features are deals with the extraction of a new low-level feature that combines, in one histogram, color and texture information. This feature help for appropriate retrieving images.

5. **Classification of image** Once the features extracted and normalized, we train a Support Vector Machine model to recognize them using labeled example photos.

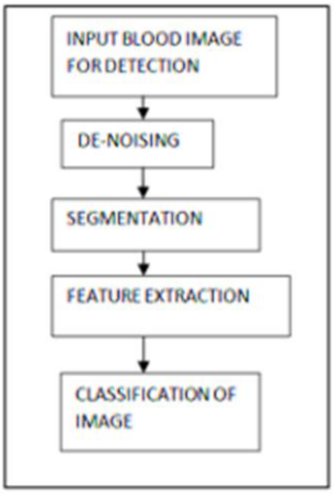

Figure 2: Use Case Diagram

Application areas :

- Medical Industry
- Research based
- Training purpose

Software, hardware & Test Data requirements :

- Software Specifications:
  - Python 3.6+
  - CNN Algorithm
  - WINDOWS OS
  - 2GB Ram
- Hardware Specifications:
  - Processor-i3
  - Hard disk-5GB
  - Memory-2GB RAM

V. REFERENCES

[1] Hend Mohamed, Rowan Omar, Nermeen Saeed, Ali Essam, Nada Ayman, Taraggy Mohiy and Ashraf AbdelRaouf "Automated Detection of White Blood Cells Cancer Diseases", in IEEE 2018 First International Workshop on Deep and Representation Learning (IEDRL).

[2] J. B. Henry, J. P. AuBuchon, Clinical diagnosis and management by laboratory methods, Archives of Pathology and Laboratory Medicine 121 (9) (1997) 1016.

[3] Nikhil R. Pal, "A fuzzy rule based approach to identify biomarkers for diagnostic classification of cancers",2007

[4] Yuchun Tang, Yan-Qing Zhang, Zhen Huang, Xiaohua Hu, Yichuan Zha,“ Recursive fuzzy granulation for gene subsets extraction and Cancer Classification",2008

[5] Muxuan Liang, Zhizhong Li, Ting Chen and Jianyang Zeng,"IntegrativeDataAnalysis ofMulti-platformCancerDatawith a Multimodal Deep Learning Approach",2014

[6] Rasool Fakoor ,Faisal Ladhak ,Azade Nazi, Manfred Huber,"Using deep learning to enhance cancer diagnosis and classication",2014.

[7] Vishwanath S. Mahalle, Ms. Vidhi L. Chawda,"Learning to recommend descriptive tags for Health Seekers using

Deep Learning",2017

[8] Naresh Gowda M, Abdul Imran Rasheed," Hardware Implementation of Hybrid Classifier to Detect Cancer Cells",2017

[9] David AH," A Deep learning approach for cancer detection and relevant gene identification ", 2017

[10] Seema Singh, Sunita Saini , Mandeep Singh , "Cancer Detection Using adaptive Neural Network",2017

[11] SeungJin Lim," FT-IR spectra analysis towards cancer detection", 2014

[12] Li Shen and Eng Chong Tan,"DimensionReduction-Based Penalized LogisticRegressionforCancerClassificationUsing Microarray Data",2005