Breast Cancer Detection With Optimized Machine Learning Techniques

Apurva Anil Patil[1], Prerana Sanjay Pol[2], Aarti Arjun Kamble[3], Sneha Devanand Dale[4] , Prof. Umesh Anandrao Patil[5]

[1] *Student, Computer Science Engineering, D. Y. Patil technical campus talsande, Maharashtra, India*
[2] *Student, Computer Science Engineering, D. Y. Patil technical campus talsande, Maharashtra, India*
[3] *Student, Computer Science Engineering, D. Y. Patil technical campus talsande, Maharashtra, India*
[4] *Student, Computer Science Engineering, D. Y. Patil technical campus talsande, Maharashtra, India*
[5]*Assistant Professor, Computer Science Engineering, D. Y. Patil technical campus talsande,Maharashtra, India*

ABSTRACT

Cancer is a deadly disease. Initial detection of cancer is the best way to cure the disease. Medical Image Processing plays an essential part in the detection of disease. Leukemia is a kind of blood cancer that happens due to irregular or immature White Blood Cell (WBC) s. In general, WBC is the fighter to fight against infectious cells in the human body. Abnormal growth of WBC from bone marrow will destroy the other cells and affect bone marrow and lymphatic tissues. These cells do not function properly and leads to leukemia. In olden days, identification of disease and cell counting in the blood were very complicated. In the medical sector, they use a device called Haemocytometer which counts the amount of cells in the blood manually. But it takes more time for counting and gives inaccurate results. To overcome these issues, a software based solution is given with the help of microscopic images. With this image processing technique, the number of RBCs, WBCs and platelets are calculated and also whether a person can be affected by leukemia or not is identified.

Keywords— *Blood cancer Detection, CNN, Image processing.*

1. INTRODUCTION

The use of Mobile technology is being used more and more for health purposes. Due to their accessibility, vast audience, and widespread acceptability, apps have the potential to share health care data or monitor patients in real-time. In the Android Play Store, one of the primary download platforms, at the beginning of 2021, there were more than 53,000 medical apps. Numerous medical problems have been the focus of medical applications, including COVID-19, diabetes, pain, rheumatoid arthritis, mental illnesses, and cancer. Although there is little information available about them, the major download platforms also have apps for people with hematological diseases. Hematological conditions include a broad range of illnesses that can be divided into benign (hematological cancers like Hodgkin and non-Hodgkin lymphoma, leukemia, or multiple myeloma, among others) and malignant (anemia, hemorrhagic, or thrombotic disorders and conditions affecting blood-forming organs) categories. With substantial morbidities impacting patients all over the world, these illnesses satisfy the requirements for being classified as highly important public health problems [14–y] of these illnesses, like hemophilia or anemia, are quite common and develop into chronic disorders. Medical apps are becoming an increasingly appealing choice for this use since these patients may benefit from technologies that enhance treatment adherence or self-management instructions. Various hematological disorders include A thorough evaluation is required in light of the abundance of health applications for hematological disease patients already on the market and the growing need for tools that promote patient self-care. Regarding the best way to judge the caliber of health apps, there isn't a clear consensus, but. The most popular

measure for assessing the value and content of health applications is the Mobile App Rating Scale (MARS). By linking to an app's user engagement, functionality, aesthetics, and information quality, it is now possible to evaluate and compare it. Furthermore; it offers a quantifiable and verified approach that enables consumers and medical experts to steer clear of incorrect information.

2. NEED OF PRESENT WORK

Giving the patient the best care possible based on an analysis of their medical history, lifestyle decisions, and any biological characteristic variability is one of the most crucial requirements of a healthcare system. To solve these issues, a number of intelligent solutions built on machine learning and data-driven techniques have been created. In this study, statistical and machine learning techniques are examined. Additionally, we provide a secure Android-based architecture and a reliable paradigm for cloud-based data storage to collect patient data. The article describes Leu-Life, an android m-health app that employs machine learning techniques to identify leukemia cancer in addition to offering a set of features that aid in managing and improving the lives of leukemia cancer patients. The talk will end with a CNN algorithm.

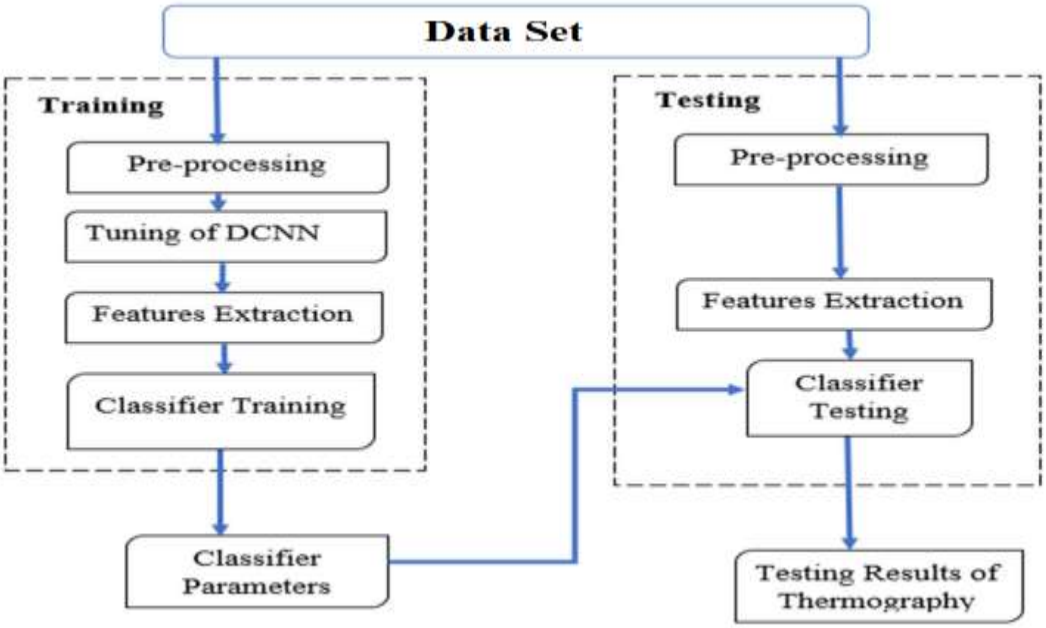
3. OBJECTIVES

In future research we are planning to implement the CNN model for classification and detection of the primary goal of cancer treatment is to cure the illness, followed by palliation (life extension and pain alleviation) in cases when cure is impossible owing to advanced disease. 30% of malignancies are now regularly cured nowadays. When feasible, treatment should result in a cure while maintaining an adequate standard of living. While symptom alleviation may occur after curative therapy, it becomes crucial to do so quickly when a cure is not attainable.

4. SYSTEM ARCHITECTURE

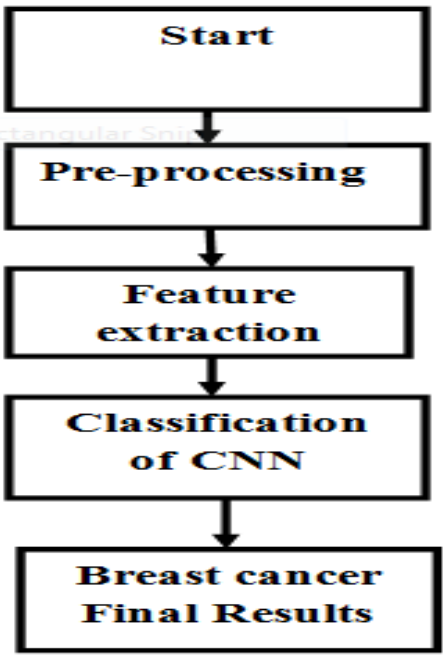
Data pre-processing is the process of transforming raw data into an understandable format. It is also an important step in data mining as we cannot work with raw data. The quality of the data should be checked before applying machine learning or data mining algorithms. Examples of data pre-processing include cleaning, instance selection, normalization, one hot encoding, transformation, feature extraction and selection, etc. The product of data pre-processing is the final training set.

Data pre-processing is the process of transforming raw data into an understandable format. It is also an important step in data mining as we cannot work with raw data. The quality of the data should be checked before applying machine learning or data mining algorithms. Examples of data pre-processing include cleaning, instance selection, normalization, one hot encoding, transformation, feature extraction and selection, etc. The product of data pre-processing is the final training set.



Proposed System Architecture Diagram

4.1 Flowchart



4.2 Algorithm - CNN

CNN algorithm used for the detection of breast cancer to gives the input as a image of breast cancer and detect the output.

Neural networks are a set of algorithms. They interpret sensory data through a kind of machine perception, labeling or clustering raw input. The patterns they recognize are numerical, contained in vectors, into which all real-world data, be it images, sound, text or time series, must be translated. Neural networks help us cluster and classify. They help to group unlabeled data

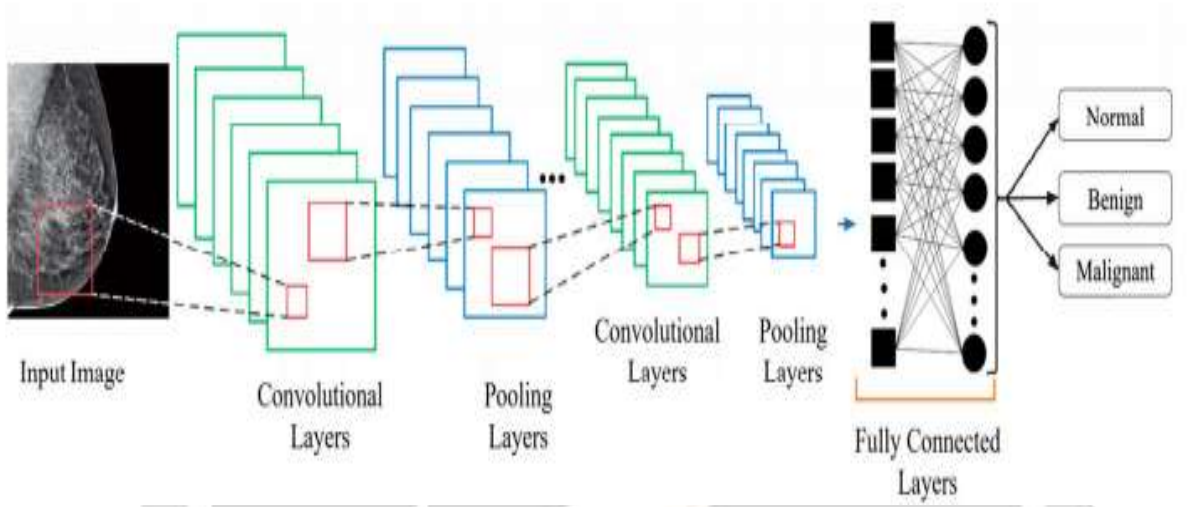


Fig 4.1.2: CNN Algorithm Steps

5. **CONCLUSION**

The dataset collected from the Kaggle public dataset is made available to researchers for experimental purposes in the suggested approach of diagnosing blood cancer. One of the simple ways to extract the function from the image is through image processing. MATLAB 2018a is used to implement the suggested system. The suggested approach identifies blood cancer with stages and delivers 90% accuracy. The suggested strategy, which relies on grouping k-means, has the following drawbacks. 1) The number of groups, or k value that can form empty groups for greater values. 2) It shrinks significantly over time and is slow.3) Different beginning divisions can lead to different end groupings. 4) There are local optimums that are worse. Other algorithms can be used to enhance the findings. Accuracy may be increased by doing more research with more data. K- Means had the best accuracy compared to other algorithms, as seen in the comparison between various algorithms by several researches in the chart below.

6. **REFERENCES**

[1] R. Adollah, M.Y. Mashor, N.F.M. Nasir, H. Rosline, H. Mahsin & H. Adilah (2018), " Blood Cell Image Segmentation: A Review", Proceedings of Biomed, Pp. 141 144.

[2] A. Belhekar, K. Gagare, R. Bedse, Y. Bhelkar, K. Rajeswari, M. Karthikeyan, Proceeding - IEEE Int. Conf. Comput. Commun. Autom. ICCCA, (2019).

[3] R. G. Bagasjvara, I. Candradewi, S. Hartati, A. Harjoko, Proc. - 2nd Int. Conf. Sci. Technol.- Comput. ICST, (2017).

[4] M. B. N. Shaikh, S. Deshpande, IEEE Int. Conf. Recent Trends Electron. Inf. Commun. Technol. RTEICT 2017 - Proc., (2017)

[5] S. Mandal, V. Daivajna, R. V, INDICON 2019 - 16th IEEE India Counc. Int. Conf., 1-4, (2019)

[6] R. G. Bagasjvara, I. Candradewi, S. Hartati, A. Harjoko, Proc. - 2nd Int. Conf. Sci. Technol.- Comput. ICST, (2018).

[7] R. D.Labati, V. Piuri, F. Scotti, " All-Idb: the acute lymphoblastic leukemia image database for image processing",2017.