DETECTION OF ALL AND CLL BLOOD CELLS IN WBC'S USING BLOOD SMEAR IMAGES

A PROJECT REPORT

Submitted by

P. POORNIMA	**(721719106042)**
M. SHREETHA	**(721719106059)**
SHWATHI R M	**(721719106060)**
D. SOUNDARYA	**(721719106062)**

in partial fulfillment for the award of the degree

of

BACHELOR OF ENGINEERING

in

ELECTRONICS AND COMMUNICATION ENGINEERING

P. A. COLLEGE OF ENGINEERING AND TECHNOLOGY

(Autonomous)

Pollachi, Coimbatore Dt - 642 002

May 2023

P.A. COLLEGE OF ENGINEERING AND TECHNOLOGY

BONAFIDE CERTIFICATE

Certified that this project report **"DETECTION OF ALL AND CLL BLOOD CELLS IN WBC'S USING BLOOD SMEAR IMAGES"** is the bonafide work of **"P.POORNIMA (721719106042), M.SHREETHA (721719106059), SHWATHI.R.M (721719106060), D.SOUNDARYA (721719106062)"** who carried out the project work under my supervision.

Signature of the HoD

Dr. L. Murali, M.E., Ph.D.,
Professor and Head,
Department of Electronics and
Communication Engineering.

Signature of the Supervisor

R. Vishnu Vardhan, M.E.,
Assistant Professor,
Department of Electronics and
Communication Engineering.

Submitted for the Viva-voce examination held on ………………..

…………………….

Internal Examiner

…………….……..

External Examiner

ACKNOWLEDGEMENT

ABSTRACT

WBCs are generated in the bone marrow and account for about 1% of all blood cells. These WBC's proliferate unregulated, which leads to the development of blood cancer. The proposed work proposesa reliable method for classifying Acute Lymphoblastic Leukemia (ALL) and Chronic Lymphoblastic Leukemia using the ALL-IDB and Kaggle datasets (CLL). Acute lymphoblastic leukemia is a cancer that develops when the bone marrow produces an excessive number of lymphocytes (ALL). Contrarily, Chronic Lymphocytic Leukemia (CLL) attacks the spongy tissue inside the bones where blood cells are made. Traditionally, the procedure was completed manually over a long period of time by a qualified expert. By utilizing deep learning techniques, specifically convolutional neural networks,the suggested approach completely eliminates the possibility of errors in the human process. The model first pre-processes the photos and extracts the best features after being trained on images of cells. The model was able to preciselycollect the samples 90 times out of 100 and reproduce all measurements. After classifying the normal and pathological cells using CNN, this task may be implemented utilizing the Python programming environment.

TABLE OF CONTENTS

LIST OF TABLES

TABLE NO.	TITLE	PAGE NO.
4.1	Comparison for detection using machine learning and CNN	34

LIST OF FIGURES

LIST OF ABBREVIATIONS

ALL	Acute Lymphoblastic Leukemia
AI	Artificial Intelligence
ANN	Artificial Neural Network
CAP	Credit Assignment Path
CCA	Controller of Certifying Authorities
CCN	Content - Centric Networking
CLAHE	Contrast Limited Adaptive Histogram Equalization
CLL	Chronic Lymphocytic Leukemia
DBNs	Deep Belief Networks
DCNN	Deep Convolutional Neural Network
DNN	Deep Neural Network
GPU	Graphics Processing Unit
HIS	Hue Saturation Intensity
JPEG	Joint Photographic Experts Group
KNN	K- Nearest Neighbors
NHCs	Neural History Compressors
PIL	Python Imaging Library
RBC	Red Blood Cells
RGB	Red Green Blue
RUS Boost	Random Under Sampling Boost
SVM	Support Vector Machine
WBC	White Blood Cells

CHAPTER 1

INTRODUCTION

1.1 DIGITAL IMAGE PROCESSING

Human can easily detect and identify objects present in an image. The human visual system is fast and accurate and can perform complex tasks like identifying multiple objects and detect obstacles with little conscious thought. With the availability of large amount of data, faster GPUs, better algorithms, we can now easily train computers to detect multiple objects within an image with high accuracy. Image processing is a technique to enhance raw images received from cameras or sensors. Based on satellites, space probes and aircraft pictures taken in normal day-to-day life for various application. Various techniques have been developed in Image processing during the last four to fivedecades. Most of the techniques are developed for enhancing images obtained from unmanned space crafts, space probes and military reconnaissance flights.Image processing systems are becoming popular due to easy availability of powerful personnel computers, large size memory devices, graphics software etc., Image processing is used in various applications such as Remote Sensing, destructive Evaluation, Forensic studies, Textiles, Material science, Military, Film industry, Document processing, Graphic arts, Printing Industry. The common steps in image processing are image scanning, storing, enhancing andinterpretation.

In digital image processing, digital computers are used to process the image. The image will be converted to digital from using a scanner digitizer and then process it. It is defined as the subjecting numerical representations ofobjects to a serious of operations in order to obtain a desired result. It starts

with one image and produces a modified version of the same. It is therefore a process that takes an image into another.

The term digital image processing generally refers to processing of a two-dimensional picture by a digital computer. In a broader context, it implies digital processing of any two-dimensional data. A digital image is an array of real numbers represented by a finite number of bits. The principle advantage ofDigital Image Processing methods is its versatility, repeatability and the preservation of original data precision.

1.2 TECHNIQUES OF DIGITAL IMAGE PROCESSING

The various image processing techniques are:

- Image pre-processing
- Image segmentation
- Feature extraction
- Image classification

1.2.1 Image Pre-processing

In image preprocessing, image data recorded by sensors on a satellite restrain errors related to geometry and brightness values of the pixels. These errors arecorrected using appropriate mathematical models which are either definite or statistical models. Image enhancement is the modification of image by changing the pixel brightness values to improve its visual impact. Image enhancement involves a collection of techniques that are used to improve the visual appearance of an image, or to convert the image to a form which is better suited for human or machine interpretation. Sometimes images obtained from satellites and conventional and digital cameras lack in contrast and brightness because of the limitations of imaging sub systems and illumination

conditions while capturing image. Images may have different types of noise. In image enhancement, the goal is to accentuate certain image features for subsequent analysis or for image display. Examples include contrast and edge enhancement, pseudo-coloring, noise filtering, sharpening, and magnifying. Image enhancement is useful in feature extraction, image analysis and an imagedisplay. The enhancement process itself does not increase the inherent information content in the data. It simply emphasizes certain specified image characteristics.

1.2.2. Image Segmentation

Segmentation is one of the key problems in image processing. Image segmentation is the process that subdivides an image into its constituent parts or objects. The level to which this subdivision is carried out depends on the problem being solved, i.e., the segmentation should stop when the objects of interest in an application have been isolated e.g., in autonomous air-to-ground target acquisition, suppose our interest lies in identifying vehicles on a road, thefirst step is to segment the road from the image and then to segment the contentsof the road down to potential vehicles. Image thresholding techniques are usedfor image segmentation. After thresholding a binary image is formed where allobject pixels have one gray level and all background pixels have another generally the object pixels are 'black' and the background is 'white'. The best threshold is the one that selects all the object pixels and maps them to 'black'. Various approaches for the automatic selection of the threshold have been proposed. Segmentation of images involves sometimes not only the discrimination between objects and the background, but also separation between different regions.

1.2.3 Feature Extraction

The feature extraction techniques are developed to extract features in

syntheticaperture radar images. This technique extracts high-level features needed in order to perform classification of targets. Features are those items which uniquely describe a target, such as size, shape, composition, location etc. Segmentation techniques are used to isolate the desired object from the scene so that measurements can be made on it subsequently. Quantitative measurements of object features allow classification and description of the image, When the pre-processing and the desired level of segmentation has beenachieved, some feature extraction technique is applied to the segments to obtain features, which is followed by application of classification and post processingtechniques. It is essential to focus on the feature extraction phase as it has an observable impact on the efficiency of the recognition system. Feature selection of a feature extraction method is the single most important factor in achieving high recognition performance. Feature extraction has been given as "extractingfrom the raw data information that is most suitable for classification purposes,while minimizing the within class pattern variability and enhancing the between class pattern variability". Thus, selection of a suitable feature extraction technique according to the input to be applied needs to be done with utmost care. Taking into consideration all these factors, it becomes essential to look atthe various available techniques for feature extraction in a given domain, covering vast possibilities of cases.

1.2.4 Image Classification

Image classification is the labeling of a pixel or a group of pixels based on its grey value. Classification is one of the most often used methods of informationextraction. In Classification, usually multiple features are used for a set of pixels i.e., many images of a particular object are needed. In Remote Sensing area, this procedure assumes that the imagery of a specific geographic area is collected in multiple regions of the electromagnetic spectrum and is in good registration. Most of the information extraction

techniques rely on analysis of the spectral reflectance properties of such imagery and employ special algorithms designed to perform various types of spectral analysis.

1.3 OVERVIEW OF LEUKEMIA

Leukemia is a cancer of the blood, characterized by the rapid growth of abnormal blood cells. This uncontrolled growth takes place in your bone marrow, where most of your body's blood is made. Leukemia cells are usually immature white blood cells. There are many types of leukemia. Some are more common in children, while others are more common in adults. Researchers and oncologists have struggled for years to make an early diagnosis of leukemia due to the disease's initial non-obvious symptoms, such as fever, fatigue, and bruising. Moreover, leukemia doesn't form tumors like other cancers do, which makes it very challenging to detect. Manual methods for evaluating blood smears are time-consuming, costly, and prone to error since they rely on costly resources such as the knowledge of qualified medical experts who have specialized in picture interpretation. Patients with leukemia undergo a plethora of blood tests for blood cell counts and analysis after being diagnosed.

White blood cell cancer known as acute lymphocytic leukemia typically affects the bone marrow (ALL). The term "acute" describes the disease's speedy development; if it is not treated when it is first diagnosed, it could swiftly prove fatal. The blood count in this case is normal, unlike leukemia, yet the infected person is found. As acute leukemia will spread more quickly, treatment should begin as soon as the disease is identified.

Chronic lymphoblastic leukemia (CLL) can cause both an excess of and a deficiency of cells. Some of the blood cells can temporarily operate normally, and it involves more completely grown blood cells that mature more gradually. Cancer is caused by a cell growth group that is out of control. The

most common type of cancer, blood cancer, will impede the bone marrow's ability to produce healthy blood cells and will stunt their growth. The majority of illnesses' high death rates could be decreased with early identification.

1.4 DETECTION OF LEUKEMIA

Leukemia means blood cancer which is featured by the uncontrolled and abnormal production of white blood cells (leukocytes) by the bone marrow in the blood and it is the most critical blood disease, common in children and adults. A majority cancer cell begins in body parts but leukemia is the type of cancer which begins and grows in blood cells. Blood is crucial content without which metabolic functions of body severely affects. Human system is like, cellgrows and multiply into new cells. Old cells are destroyed and so that new cellscan take their place. In cancer, an old cell does not die and remains in the bloodso that new cells which are produced cannot get enough space to live. In this way, functioning of blood disturbs and white blood cells production is abnormal and uncontrolled. In the past, a skilled professional would manually execute the process over a long period of time. The suggested approach totally eliminates the chance of human errors by leveraging deep learning techniques, notably convolutional neural networks. After being trained on cell images, the model pre-processes the pictures and extracts the best qualities. Using the model, it ispossible to precisely determine the type of cancer present in the bone marrow.This task can be carried out using the Python programming environment once CNN has classified the normal and diseased cells. Python has a wide range of applications in almost every field. Python is one of the widely used programming languages for this purpose.

1.4.1 Deep Learning

In 1986, Rina Dechter introduced the term "Deep Learning" to the machine learning community, and in 2000, Igor Aizenberg and colleagues did the same for artificial neural networks. In 1967, Alexey Ivakhnenko and Lapa published the first generic, functional learning algorithm for supervised, deep, feed- forward multilayer perceptron's. A deep network with eight layers that had been trained using the group way of data handling algorithm was already reported in a 1971 publication. According to Yann Le Cun, the impact of deep learning on industry started in the early 2000s, when CNNs were processing 10% to 20% of all checks written in the US. Industry uses of deep learning to large-scale speech recognition started around 2010.

An artificial neural network (ANN) with more than two layers between the input and output layers is called a deep neural network (DNN). The DNN finds the correct mathematical manipulation to turn the input into the output, whether it be a linear relationship or a non-linear relationship. DNNs can modelcomplex non-linear relationships. The compositional models produced by DNN architectures express the object as a layered composition of primitives. The extra layers enable composition of features from lower layers, potentially modelling complex data with fewer units than a similarly performing shallow network. Deep architectures include many permutations of a few basic methods. DNNs are typically feed-forward networks in which information moves straight from the input layer to the output layer. At first, the DNN creates a map of virtual neurons and assigns random numerical values, or "weights", to connections between them. An algorithm would change the weights if the network failed to recognize a pattern accurately. That way the algorithm can make certain parameters more influential, until it determines the correct mathematical manipulation to fully process the data.

1.4.2 Deep Learning Architecture

The term "deep learning" refers to how many levels of transformation the data goes through. Although they can also include propositional formulas or latent variables organized layer-wise in deep generative models like the nodes in deep belief networks and deep Boltzmann machines, the majority of contemporary deep learning models are based on artificial neural networks, specifically Convolutional Neural Networks (CNNs). A class of machine learning algorithms known as "deep learning" uses multiple layers to gradually extract higher level features from the input's raw data.

Each level of deep learning learns to turn the data it receives into a representation that is a little bit more abstract and composite. A matrix of pixels may be the initial input for an image recognition application; the first representational layer may abstract the pixels and encode edges; the second layer may compose and encode configurations of edges. A deep learning algorithm can figure out on its own which features to best place in which level.

Deep learning systems have a substantial Credit Assignment Path (CAP)depth. The CAP is the chain of transformations from input to output. CAPS describe potentially causal connections between input and output. For a feed forward neural network, the depth of the CAPS is that of the network and is thenumber of hidden layers plus one. For recurrent neural networks, in which a signal may propagate through a layer more than once, CAP depth is potentiallyunlimited. No universally agreed upon the threshold of depth divides shallow learning from deep learning, but most researchers agree that deep learning involves CAP depth higher than CAP of depth 2 has been shown to be a universal approximator in the sense that it can emulate any function. Beyond that, more layers do not add to the function approximator ability of the network. Deep models (CAP > 2) are able to extract better

features than shallow models and hence, extra layers help in learning the features effectively.

Deep learning architectures can be constructed with a greedy layer-by-layer method. Performance is enhanced by deep learning's ability to separate these abstractions and identify features. Deep learning techniques avoid featureengineering for supervised learning tasks by converting the input into compactintermediate representations to major components and generating layered structures that eliminate representational redundancy. Unsupervised learning tasks can be handled by deep learning algorithms. Since that unlabeled data aremore prevalent than labelled data, this is a significant.

The remaining chapters are organized as follows Chapter 2 - provides an introduction by detailing the survey of literature in existing method. Chapter 3 - gives the information about the detection of leukemia using machine learning and its persistence in the following method. Chapter 4 - describes the proposed work considering the convolutional neural network in deep learning. Chapter 5 - describes the python as the software tool used in the proposed method. Chapter 6 - deals with the experimental results and discussion of the proposed work. Chapter 7 - provides the conclusion and the future scope.

CHAPTER 2

LITERATURE SURVEY

2.1 INTRODUCTION

Many techniques have been proposed by many researchers regarding image processing. Every technique has their own merits and demerits while processing the image from a dataset. Among all other techniques CNN provides efficient output.

2.2 SURVEY ON LITERATURE

Christo Ananth et al. (2022) explains the detection of blood cancer with microscopic images using machine learning. This proposed model used MATLAB to find leukemia cells in healthy blood cells, and it requires no medical equipment or expert and heavily relies on automation. The proposed method correctly identifies WBCs and leukoblasts in images and refines the identification, thresholding, and segmentation phases. To get the most accurateresults, to use the most discriminatory features. This technology can detect anemia, malaria, vitamin B12 deficiency, and brain tumors.

Raheel Baig et al. (2022) works to detect malignant leukemia cells using microscopic images. The approach is practical and does not necessitate image segmentation. The proposed model trained two CCN models in parallel to extract features. The CCA Fused approach is used to concatenate these derived features. The classifier receives fused vectors (SVM, Bagging ensemble, total boost, RUS Boost, Fine KNN, etc.). Using the Bagging ensemble design, we achieved a 97.04 percent accuracy. It will not be used in every platform. As

a result, pathologists may find that this procedure aids in effective diagnosis.

Maryam Bukhari et al. (2022) proposed a detection of leukemia cancer framework using microscopic images. The proposed machine learning architecture emphasizes the channel associations on all levels of feature representation by incorporating the squeeze and excitation learning that recursively performs recalibration on channel-wise feature outputs by modeling channel interdependencies explicitly. The suggested machine learning model exhibits good results and can be utilized to make a reliable computer-aided diagnosis for leukemia cancer.

Mustafa Ghader Zadeh et al. (2022) presented an automated detection model in classification of B-lymphoblast cells from normal B-lymphoid precursors in blood smear microscopic images. In classifying cancerous blood cells from normal cells, the proposed method can achieve high accuracy without the operator's intervention in cell feature determination. It can thus be recommended as an extraordinary tool for the analysis of blood samples in digital laboratory equipment to assist laboratory specialists.

Mohammed Junaid Ahmed et al (2021), this objective of this review is to evaluate a survey on leukemia detection using Image processing technique. This survey paper presents a review on the distinctive conventional strategies towards that have been utilized in leukemia illness diagnosis dependent on platelets images and to analyze between the two methodologies in nature of appraisal, exactness, cost and speed. It gets delegated details; it needed anothermethod called dimension reduction method. At last, in view of the led study, itvery well may be reasoned that the proposed framework CNN was accomplishing immense triumphs in this field.

Nithyaa. A. N et al. (2021), Detection and classification for potent

algorithm to clinically probable WBC cancer. The aim of this approach is to discover the WBC cancer cells in an earlier stage and to reduce the discrepancies in diagnosis, by improving the system learning methodology. This paper presents the potent algorithm, which will eliminate the dubiety, in diagnosing the cancers with similar symptoms. But training theclassifier is the notable difficulty in this research work. It took huge time for anefficient classification. This cancer is categorized by segmentation and feature extraction, which will be further, classified using Random Forest classification(RFC).

Mustafa Ghaderzadeh et al. (2021) presented an image processing technique for the detection and classification of leukemia in smear blood images using machine learning. The proposed methodology was carried out toreview the works intending to identify and classify leukemia by utilizing machine learning. The precision of B cell lymphoma and leukemia classification was 74%. Using this method to process leukemia smear images can improve accuracy, reduce diagnosis time and provide faster, cheaper, and safer diagnostic services.

Alexandra Bodzas et al. (2020), presented an Automated detection of acute lymphoblastic leukemia from microscopic images. This paper proposes anovel approach based on conventional digital image processing techniques andmachine learning algorithms to automatically identify acute lymphoblastic leukemia from peripheral blood smear images. It has extensive false negative classifications.

Deepika Kumar et al. (2020), proposed the patterns of detection of white blood cancer from bone marrow. The proposed model eradicates the probability of errors in the manual process by employing deep learning

techniques, namely convolutional neural networks. The model was able to reproduce all the measurements correctly while it recollected the samples exactly 94 times out of 100. But it can't do work properly on the smooth transition image. Thus, the model can be used effectively as a tool for determining the type of cancer in the bone marrow.

Akshay Belhekar et al. (2019) explains the Leukaemia cancer detection using image analytics. This paper proposes an automated system which uses image analytics. Based on image analytics and classification algorithms performed oncell image samples of patients, the proposed system will give correct output. Inthis comparative study, it was found that K-means clustering performs well forsegmentation phase and also Neural Networks gives better results for classification phase. K-means clustering works well only when dataset size limited.

CHAPTER 3

DETECTION OF LEUKEMIA USING MACHINE LEARNING

3.1 METHODOLOGY

The original SVM algorithm was invented by Vladimir N Vapink and Alexey Ya. Chervonenkis. In 1992, Bernhard E. Boser, Isabelle M. Guyon Vladimir N.Vapnik suggested a way to create non-linear classifiers by applying the kerneltrick to maximum margin hyper planes.

In machine learning, support vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis. In addition to performing linear classification, SVMs can efficiently perform a non-linear classification using kernel trick, implicitly mapping their inputs into high dimensional feature spaces. When data are unlabeled, supervised learning is not possible, and an unsupervised learning approach is required, which attempts to find natural clustering of the data to groups, and then map new data to these formed groups.

The Support vector clustering algorithm, created by Hava Siegelamann and Vladimir Vapnik, applies the statistics of support vectors, developed in the support vector machines algorithm, to categorize unlabeled data. and is one ofthe most widely used clustering algorithms in industrial applications. SVM constructs a hyper plane or set of hyper planes in a high or infinite-dimensional space, which can be used for classification. regression or other task like outliers' detection SVM classifier treated as one of the dominant classificationalgorithms. SVM is a supervised learning technique. When a dataset with features and class labels are both together then SVM can be used. But if the dataset does not have class labels or outputs of the feature set

then it is considered as an unsupervised learning algorithm. In that case, support vectorclustering (SVC) can be used.

```
┌─────────────────────────────┐
│   Input Blood Cell Image    │
│         (Dataset)           │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│      Image Acquisition      │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│    Contrast Enhancement     │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│       Pre-processing        │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│        Segmentation         │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│      Feature Extraction     │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│     SVM Classifier (K-      │
│       means method)         │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│       WBC Detection         │
└─────────────────────────────┘
              │
              ▼
          ◇ Cancer      No
          Detection ─────────┐ (loop back to Image Acquisition)
              │ Yes
              ▼
┌─────────────────────────────┐
│      Leukemia Detected      │
└─────────────────────────────┘
```
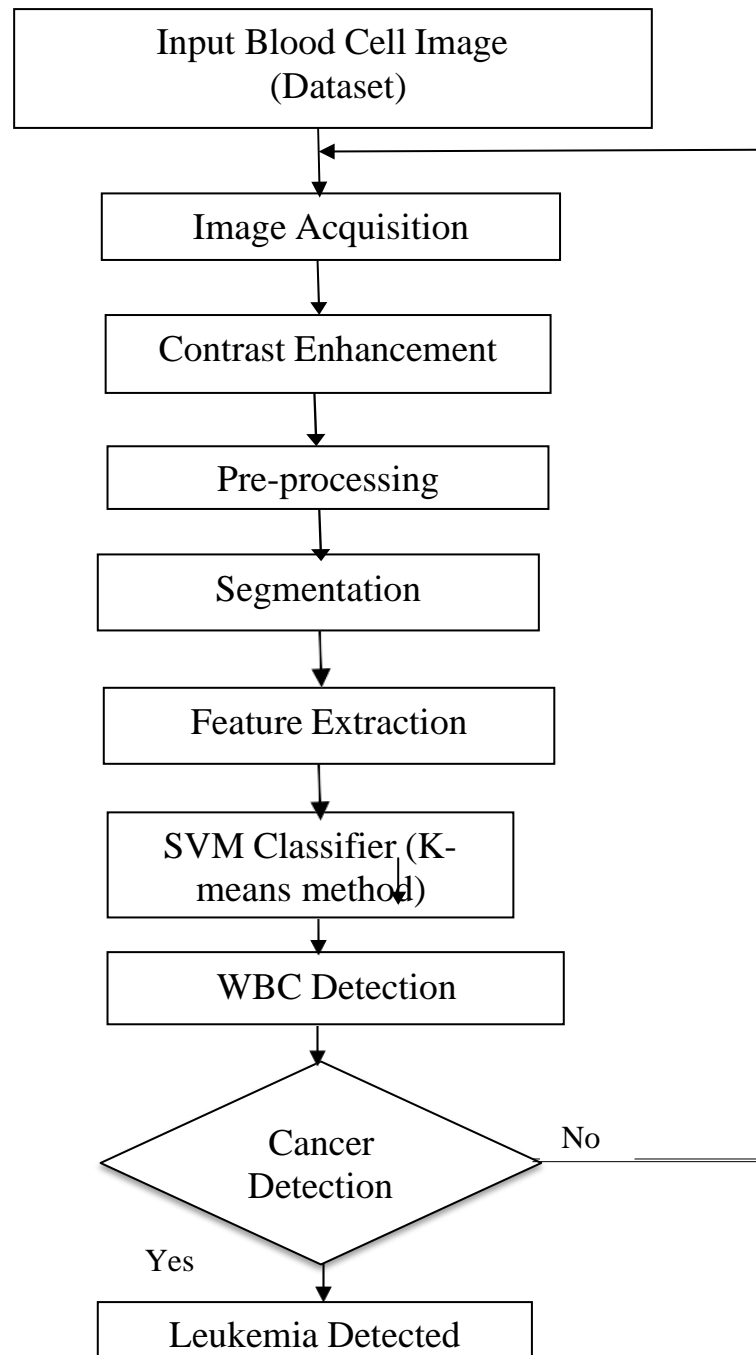
Figure 3.1 Flow diagram for detection of leukemia

3.1.1 Image Acquisition

In image processing, it is defined as the action of retrieving an image from some source, usually a hardware-based source for processing. It is the first step in thework flow because, without an image, no processing is possible.

3.1.2 Pre-processing

The image pre-processing is done on gathered image for improving the image quality. It removes the background noise as well as to suppress the undesired distortion. In this, image is resized to size 300*300, Gaussian filtering is carried out to remove noise in the image. Pre-processing is a technique used to analyze real time problems in images. Attack of micro-organism's damages fruits and vegetables. Fruit inspection is only way to investigate disease present within fruit. In order to accomplish this image of fruit can be captured and then analyzed using pre-processing technique. It is utilized for this purpose if RGB to different color space conversion. This section removes noise, smoothen the image and also perform resizing of images. RGB images are converted to grayimages also contrast of image is increased at certain level.

To remove noise in image or other object removal, different pre- processing technique is considered. Image smoothening is done using smoothing filter. Image enhancement is carried out for increasing the contrast.

RGB image into gray images using color conversion equation (1),

$$F(x)=0.2989*R+0.5870*G +0.114*B \qquad ………\ (3.1)$$

Then histogram equalization which distributes the intensities of image is applied on image to enhance fruit disease images. The cumulative distribution function used to distribute intensity values. Pre-processing is prior step which is considered to be one of important steps in world of image processing. It

mainly deals with various picture enhancement like adjusting brightness of pixels, applying rotation, scaling another transformation.

3.1.3 Contrast Enhancement

Image enhancement is usually considered to be an interesting field of image processing. Here, a comparative analysis has been which involves CLAHE (Contrast Limited Adaptive Histogram Equalization), simple histogram and morphological equalization.

3.1.4 HIS Image Conversion

Image quality can be increased or improved by adjusting its pixel value that bychanging color. So, a real time implementation of it has been proposed. A conversion strategy has been done by implementing color conversion algorithm and also suggested with help of XSG.

3.2 SEGMENTATION

In image segmentation the separation of given image into different regions with respect to some features. Clustering is a method by which large set of data are grouped into clusters of smaller sets or segments of similar data. In this work, k-means clustering is used for segmenting an image. The RGB image is converted into HSI model for segmenting. For boundary detection 8 connectivity of pixel is considered and boundary detection algorithm is applied.

3.3 K-MEANS CLUSTERING

It is used for classification of object based on set of features into 'k' number ofclasses. The classification of object is done by minimizing sum of square of distance between object and corresponding cluster.

3.3.1 Properties of K-mean Algorithm

1. There is k number of clusters always.

2. There is minimum one item in each of given cluster.

3. The cluster will never overlap with each other.

4. Each member of single cluster is nearer to its cluster than any other cluster.

3.3.2 Process of K-means Algorithm

1. First divide the data set into k number of clusters and assign the data pointsrandomly to the clusters.

2. Then for each data point, calculate the Euclidean distance, from datapoint to every cluster.

3. If data point is closest to its own cluster than leave it where it is.

4. Shift it into the nearby cluster, if data point is not closest to its own cluster.

5. Repeat all steps until an entire pass through all the data points. 6. Now clusters become stable and process of clustering will stop.

3.4 FEATURE EXTRACTION

It is used to extract the information that can be used to find out the significance of given sample. The main types of features are shape, color and texture which are mostly used in image processing technique.

3.4.1 Collection of Database Image

A database is an organized collection of data, generally stored and accessed electronically from a computer system. Databases are more complex they are often developed using formal design and modelling techniques. Images are required for training of the model. For this model, the dataset is taken from

theleukemia classification dataset from Kaggle. This dataset contains the following:

- The training set of 15,135 images
- Testing set of 2586 images
- The total number of patients considered are 118 with two labelled classes:

  i.   Normal Cell
  ii.   Leukemia Blast

3.4.2 Classification

Two image databases square measure needed for detection of fruit sickness, one for coaching purpose and alternative for testing. Within the coaching section, first of all input image is non heritable then image pre- processing is finished for resizing the photographs. Then feature extraction is allotted. Next, clump is performed by applying K-means algorithmic rule and eventually classification is performed victimization SVM classifier. Within the testing section input image are going to be non-heritable from user, then pre- processing, feature extraction is allotted and eventually the image is going to be classified as normal or abnormal blood cells. Leukemia is a heterogeneous group of hematologic malignancies that arise from the dysfunctional proliferation of developing leukocytes. It is classified as either acute or chronicbased on the rapidity of proliferation and as myelocytic or lymphocytic based on the cell of origin. Treatment depends on the type of leukemia but generally involves chemotherapy. Multiple genetic and environmental risk factors are identified in the development of leukemia.

3.5 SVM CLASSIFIER

Support Vector Machine (SVM) is a supervised machine learning algorithm that

can be used for both classification or regression challenges. However, it is mostly used in classification problems. In this SVM algorithm, we plot each data item as a point in n-dimensional space (where n is the number of features you have) with the value of each feature being the value of a particular coordinate. The problem to be solved in this project is one of supervised binary classification. That is, to categorize new unseen objects into two separate groups based on their properties and a set of known examples, which are already categorized. A support vector models the situation by creating a featurespace, which is a finite-dimensional vector space, each dimension of which represents a "feature" of a particular object.

The goal of the SVM is to train a model that assigns new unseen objects into a particular category. It achieves this by creating a linear partition of the feature space into two categories. Based on the features in the new unseen objects it places an object "above" or "below" the separation plane, leading to a categorization. This makes it an example of a non-probabilistic linear classifier. It is non-probabilistic, because the new projects fully determine its location in feature space. However, much of the benefit of SVM's comes from the fact that they are not restricted to being linear classifiers. Utilizing a technique known as the kernel trick. they can become much more flexible by introducing various type of non- linear decision boundaries.

Formally, in mathematical language, SVMs construct linear separating hyper planes in high-dimensional vector spaces. Data points are viewed as (->x, y) tuples ->x = (x1 to xp) where the xj are the feature and y is classification(usually given as +1 or -1). Optimal classification occurs such hyper planes provide maximal distances to the nearest training data points. Intuitively, this makes sense, as if the points are well separated, the classification between twogroups is much clearer.

3.6 PROPERTIES

- Flexibility in choosing a similarity function
- Sparseness of solution when dealing with large datasets only supportvectors are used to specify the separating hyperplane
- Ability to handle large feature spaces
- Over fitting can be controlled by soft margin approach
- Complexity does not depend on the dimensionality of the featurespace
- Feature selection

3.7 LIMITATIONS

- Difficult to predict k-values
- It is used only for classification of fruits
- Different initial partitions can result in different final clusters

3.8 RESULTS

With the help of this strong training set and potent algorithm, we canextend this work to classify all other blood cancers by adding the features. In future work, the classification could be extended in the detection of all other types of blood cancer irrespective of WBC

CHAPTER 4

DETECTION OF ALL AND CLL CELLS USING CNN

4.1 PROPOSED SYSTEM

The name "convolutional neural network" indicates that the network employs a mathematical operation called convolution. Convolution is a specialized kindof linear operation. In deep learning, a CNN is a class of deep neural networks,most commonly applied to analyzing visual imagery. They are also known as shift invariant or space invariant artificial neural networks, based on the shared- weights architecture and translation invariance characteristics. Convolutional networks are simply neural networks that use convolution in place of general matrix multiplication in at least one of their layers. A convolutional neural network consists of an input and an output layer, as well as multiple hidden layers. Convolutional layers convolve the input and pass its result to the next layer. Each convolutional neuron processes data only for its receptive field. CNNs are regularized versions of multilayer perceptron's. Multilayerperceptron's usually mean fully connected networks, that is, each neuron is onelayer is connected to all neurons in the next layer. CNNs use relatively little pre-processing compared to other image classification algorithms.

The proposed method is made to digitally identify ALL and CLL cells inleukemia. Although AI has been increasingly used to medicine in recent years, using machine learning techniques can encourage more effective, accurate andcost-efficient diagnostic and blood tests. This system is made up of many building components, including image capture, image preprocessing, image gathering for data sets, and deep neural network classification.
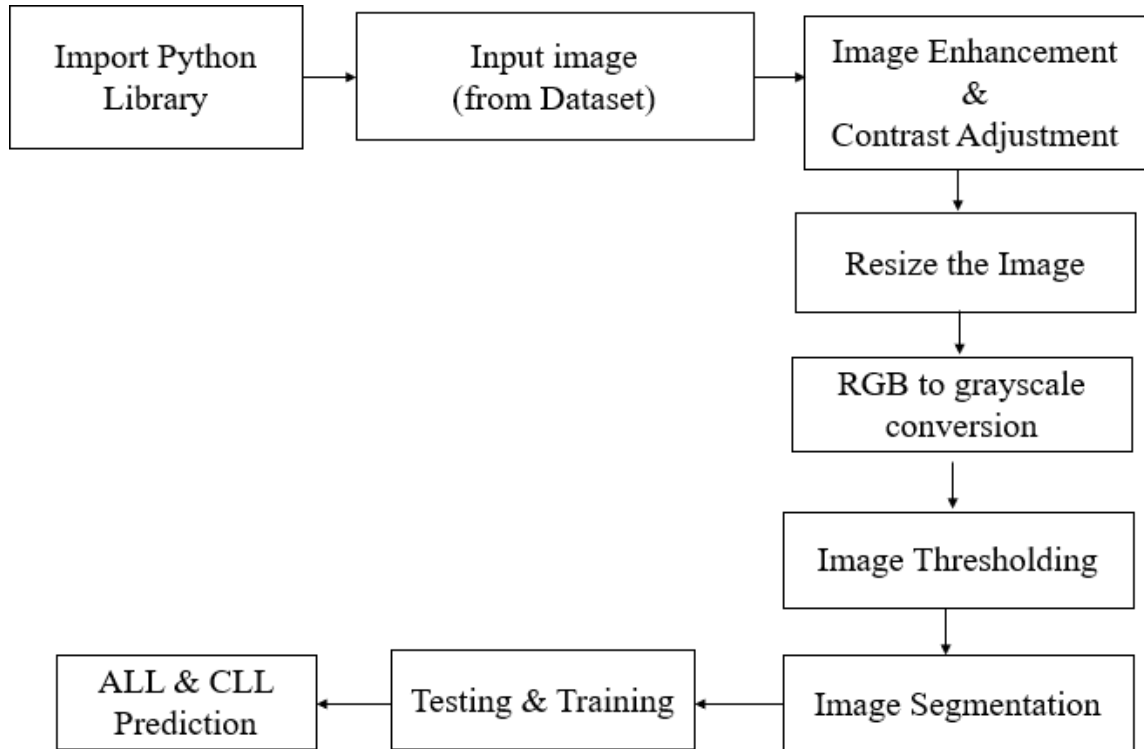
Figure 4.1 Block diagram for detection of ALL and CLL using CNN

4.2 Collection of Database

The ALL-IDB and Kaggle-Malignant Blood Cells public datasets, which were created expressly for testing segmentation and image classification algorithms,provided the data for this work. An optical laboratory microscope and a digitalcamera were used to take the microscopic pictures of the blood samples in the dataset. The dataset is divided into two parts: the ALL-IDB1 dataset is used totrain classification systems. The dataset also includes, if applicable, the blast cell centroid coordinates for each image, indicating the location of ALLlymphoblasts. The ALL-IDB2 dataset, which was created to evaluate the performance of the classification algorithms, is the second component. It consists of cells from the ALL-IDB1 dataset that have had their area of interest cells cropped. We did not use the ALL-IDB2 dataset because our study aims to improve doctors' productivity by classifying complete photos and finding individual ALL cells straight from an actual lab image.

Although Kaggle supports a number of different dataset publication formats, we firmly advise authors to release their data in an open, non-proprietary format whenever possible. Open, accessible data formats are not only more widely accepted on the platform, but they are also simpler for more users to work with regardless of the tools they use. Kaggle Datasets lets you to privately or publicly post andexchange datasets.

4.2.1 Classification

Image classification is the process of taking an input (like image of blood cell)and outputting a class (acute or chronic leukaemia or normal blood cell) or a probability that the input is a particular class. A CNN Convolves learned features with input data and uses 2D convolutional layers. This means that thistype of network is idle for processing 2D images. Compared to other image classification algorithms, CNNs actually use very little pre-processing. This section analyses numerical property of image features and organize its data intocategorizes. It uses neural network which performs training and classification of leukaemia.

4.3 CONVOLUTIONAL NEURAL NETWORK

The Image segmentation is the process of splitting a picture into numerous layers that are each represented by a clever, pixel-by-pixel mask. An image's integration level determines how it is mixed, blocked, and separated.

The first stage in image processing is to separate a photo into a collection of Image Objects with related features. Image segmentation, which divides an image into several portions or areas sometimes depending on the characteristics of the image's pixels, is a commonly used technique in digital image processingand analysis.

Our goal was to reduce runtime without considerably sacrificing

accuracy on high-resolution, therefore we constructed a unique CNN rather than utilising a pre-trained model. To compensate for the short dataset, weimproved the data by randomly flipping, rotating, and zooming in the images.After initial image resizing and rotation, the bulk of cancer cells were kept inthe image. In order to assess whether a big collection of cells contained anyALL and CLL cells, we avoided over-zooming the data. All of the photos thatwere labelled as malignant retained cancer cells and could still be categorized as positive because the images were thoroughly classified, therefore accuracyvalues did not significantly change if a few cells were removed from the image.Our CNN model consisted of three convolution-pooling layers, one dropout layer, a flatten layer, three dense layers, and one dropout layer inbetween. We stopped adding layers after the validation and training losses stabilized simultaneously.

Figure 4.2 illustrates a general CNN model that has four parts: (a) convolution layer, (b) pooling layer, (c) activation function, and (d) fully connected layer. The functionality of each component is shown in the following diagram.
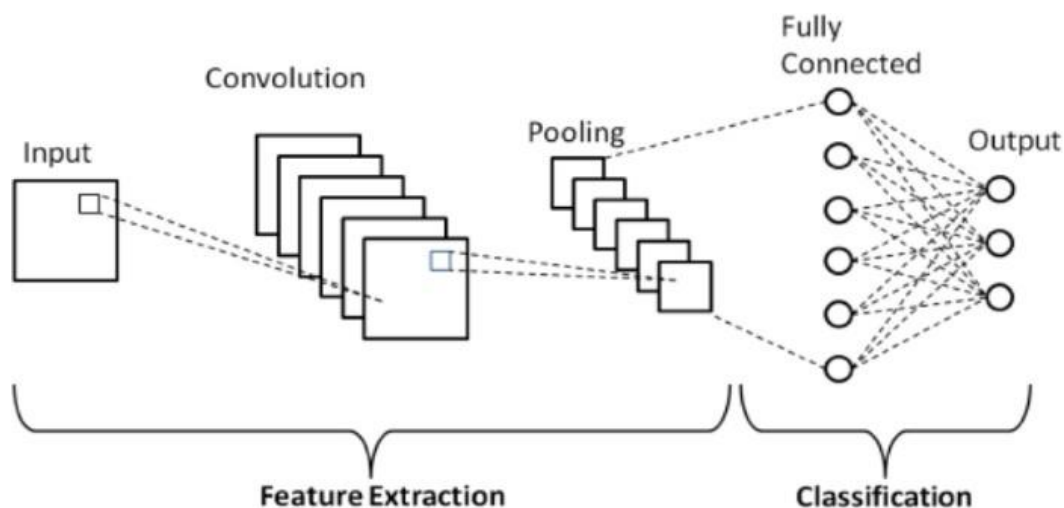

Figure 4.2 CNN Architecture

4.4 TYPES OF CNN LAYERS

4.4.1 Input Layers

In this layer of a CNN model, images are used as inputs. The size of the inputs is also handed then. The input layer of a neural network is composed of artificial input neurons and brings the initial data into the system for further processing by subsequent layers of artificial neurons. The input layer is the very beginning of the workflow for the artificial neural network. Input layer being constituted of "passive" neurons that do not take in information from previous layers because they are the very first layer of the network.

The input layer in deep learning is the first layer of a neural network, where the data is inputted into the network for processing. The input layer is responsible for accepting the raw input data, which could be in the form of images, audio, text, or any other type of data. The input data is then transformed and processed by the network's hidden layers, which perform feature extractionand abstraction.

The input layer is usually a rectangular or square grid of neurons, where each neuron corresponds to a pixel or a feature of the input data. The number of neurons in the input layer is determined by the size of the input data. For example, if the input data is a 32x32 image, then the input layer will have 1,024 neurons (32x32=1,024).

The input layer is typically followed by one or more hidden layers, which perform the bulk of the processing in a deep neural network. The output layer of the network generates the final output, which could be a classification label, a regression value, or some other type of output.

In addition to accepting raw input data, the input layer can also perform some preprocessing or normalization of the data before it is inputted into the network. For example, in image processing tasks, the input layer may perform

operations such as normalization, centering, or scaling of the image data to improve the performance of the network.

In summary, the input layer in deep learning is the first layer of a neural network that accepts raw input data and performs some initial preprocessing before the data is passed through the hidden layers for feature extraction and abstraction.

4.4.2 Convolutional Layers

Convolutional layer is the first layer to extract the features from an input image. Convolutional preserves the relationship between pixels by using small squaresof input data. It is a mathematical operation that takes two inputs such as imagemonitor and a filter. Convolution of an image with different filters can performoperations such as edge detection, blur and sharpen by applying filters.

Convolutional layers are a fundamental component of deep learning architectures, particularly in computer vision tasks. A convolutional layer is designed to detect and learn spatial features from the input data, typically images or other forms of multidimensional arrays.

During training, the convolutional layer applies a set of filters (also known as kernels) to the input data, producing a set of output feature maps. Each filter is a small matrix that slides across the input data, computing the dotproduct between its weights and a small patch of input data at each position. The resulting output feature maps capture different aspects of the input data, such as edges, textures, and other spatial patterns. Convolutional layers are typically followed by non-linear activation functions, such as ReLU, to introduce non-linearity into the network and allow it to learn complex representations of the input data. The output of the convolutional

layer is then fed to additional layers in the network, such as pooling layers, fully connected layers, or more convolutional layers, to perform increasingly complex computations and make predictions based on the input data.

Convolutional layers are highly effective in computer vision tasks because they are able to capture and learn spatial features in a translation-invariant way. This means that the network can recognize the same pattern in different locations of the input data, making it more robust to variations in the input. Convolutional layers have been used with great success in a variety of applications, such as image classification, object detection, and image segmentation. Convolutional neural networks (CNNs) have been used in the detection and classification of different types of leukemia, including acute andchronic lymphocytic leukemia (ALL and CLL).In medical imaging, such as blood smear images, CNNs can learn to detect and classify the specific featuresand patterns of leukemia cells. The input data to the CNN would be the digitalimages of the blood smears, and the output would be a classification of whetherthe sample contains leukemia cells or not, and if so, whether they are of the acute or chronic lymphocytic type. The CNN architecture typically consists ofseveral convolutional layers, followed by pooling layers and fully connected layers. During training, the weights of the filters in the convolutional layers areoptimized to learn the important features and patterns in the input images that are relevant to detecting and classifying leukemia.

Several studies have demonstrated the effectiveness of CNNs for detecting and classifying leukemia. For example, a study published in the Journal of Digital Imaging in 2020 used a CNN to classify blood smear images into three categories: ALL, CLL, and normal. The study achieved an accuracy of 96.3%, sensitivity of 93.3%, and specificity of 99.2%. Overall, CNNs have shown promising results for the detection and classification of acute and chronic lymphocytic leukemia in blood smear images. However, further

research is needed to validate these findings and evaluate the generalizability of CNNs for detecting leukemia in different populations and settings.

4.4.3 Pooling Layers

Pooling layers are commonly used in convolutional neural networks (CNNs) for image processing tasks, such as detecting Acute and Chronic lymphocytic leukemia from microscopic images. The main purpose of pooling layers is to reduce the spatial dimensions of the input feature maps while retaining the most important features. In the context of detecting Acute and Chronic lymphocyticleukemia, pooling layers can be used to reduce the size of the feature maps obtained from the convolutional layers, while retaining the most salient information. This can help in reducing the number of parameters in the network and prevent overfitting. Pooling layer section would reduce the number of parameters when the images are too large. Taking the largest element could also take average pooling. Pooling layers are a type of layer commonly used in convolutional neural networks (CNNs) for image processing tasks. They are typically inserted after convolutional layers and before fully connected layers. The main purpose of pooling layers is to reduce the spatial dimensions of the input feature maps, while also retaining the most important features. Spatial pooling also called sub-sampling or down sampling which reduces the dimensionality of each map but retains important information. Spatial pooling can be of different types namely Max pooling, Average pooling, Sum pooling. Max pooling takes the largest element from the rectified feature map.

- **Max Pooling:** The value from the image region to which the pooling kernel is applied is handed at the loftiest position by this option. Its job is to reject

any noisy activations in order to lower the dimensionality while contemporaneously dampening the noise. This is the most common type of pooling layer, where the maximum value within a certain neighborhood of the input feature map is selected and passed on to the next layer. The size of the neighborhood and the stride (the amount of shift between each neighborhood) are usually specified by the user. Max pooling is the most commonly used pooling operation, where the maximum value within a local neighborhood of the input feature map isselected and passed on to the next layer. The size of the neighborhood and the stride (the amount of shift between each neighborhood) are usually specified by the user. Average pooling can also be used, where the average value within the neighborhood is selected instead of the maximum value.

- **Average Pooling:** In this type of pooling layer, the average value withina certain neighborhood of the input feature map is selected and passed on to the next layer. This can be useful for reducing the impact of outliers. Additionally, average pooling can be used, which involves pooling overthe entire feature map, resulting in a single output value for each featuremap. Average pooling has been found to be particularly effective in some deep learning architectures, such as the Global Average Pooling (GAP) layer, which has been shown to improve the accuracy and efficiency of CNNs. Overall, pooling layers can play a crucial role in reducing the spatial dimensions of feature maps while retaining the most important information, helping CNNs to detect Acute and Chronic lymphocytic leukemia with higher accuracy and efficiency. This system presents the normal of the data from the area of the image where the pooling kernel was applied. Its job is to reject any noisy activations in order to lower the dimensionality while contemporaneously dampening the noise.

4.4.4 Fully Connected Layers

Fully connected layers in a CNN are not to be confused with fully connected neural networks – the classic neural network architecture, in which all neurons connect to all neurons in the next layer. Convolutional neural networks enable deep learning for computer vision. Neural networks are a set of dependent non-linear functions. Each individual function consists of a neuron (or a perceptron). In fully connected layers, the neuron applies a linear transformation to the inputvector through a weight's matrix. A non-linear transformation is then applied to the product through a non-linear activation function. The classic neural network architecture was found to be inefficient for computer vision tasks. Images represent a large input for a neural network (they can have hundreds orthousands of pixels and up to 3 colour channels). In a classic fully connected network, this requires a huge number of connections and network parameters. It is simple feed forward neural network. It is one of the last few layers in CNN.

The input to these layers is output from final pooling or conventional layer which is flattened and fed into fully connected layer. The objective of thefully connected layer is to take the results of the convolutional or pooling process and use them to classify the image into a label. The output of the convolutional or pooling is flattened into a single vector of values, each representing a probability that a certain feature belongs to a label.

A convolutional neural network leverages the fact that an image is composed of smaller details, or features, and creates a mechanism for analyzing each feature in isolation, which informs a decision about the image as a whole. As part of the convolutional network, there is also a fully connected layer that takes the end result of the convolution/pooling process and reaches a

classification decision. Convolutional Neural Network or CNN is a type of artificial neural network, which is widely used for image/object recognition and classification.

4.4.5 CNN Algorithm

CNN is a well- liked and effective pattern discovery and image processing approach. It has a number of benefits, including as rigidity, a simple structure, and reduced training conditions. CNN is a powerful algorithm for image processing. These algorithms are currently the best algorithms we have for theautomated processing of images. Many companies use these algorithms to do things like identifying the objects in an image.

4.5 COMPARISON

The below table 4.1 represents the comparison between the existing method and proposed method.

Table 4.1 Comparison of machine learning and CNN

S.NO	DETECTION USING MACHINE LEARNING	DETECTION USING CNN
1	The existing method is based on Potent Algorithm.	The proposed method is implemented using deep learning.
2	The software language used here is MATLAB.	The software language used here is PYTHON.
3	The SVM classifier is used here.	CNN Layers are used here for better accuracy and time redundancy.
4	The image will not be resized here.	The image will be resized after the enhancement process.

5	This method took huge time for better accuracy such as 25 minutes.	This method took less time for better accuracy when compared to existing method.
6	It only detects whether the leukemia is existed or not.	It detects both the Acute Lymphoblastic Leukemia and Chronic Lymphoblastic Leukemia.
7	Training the images took more time.	Training the images took less time when compared to existing method.
8	MATLAB Software should be installed for this method.	Google Collab website is used for the compilation.
9	Software installation took more time.	No need of software installation.
10	Accuracy is 95%	Accuracy is 99.7%
11	Implementation process is difficult.	Implementation process is simple.

CHAPTER 5

THE PYTHON

5.1 INTRODUCTION

Python is an interpreted, high-level, general-purpose programming language. Created by Guido van Rossum in 1991, Python's design philosophy emphasizes code readability with its notable use of significant white space. Its language constructs and object-oriented approach aims to help programmers write clear,logical code for small and large-scale projects. Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, includingprocedural, object-oriented, and functional programming. Python is often described as a "batteries included" language due to its comprehensive standardlibrary. Python was conceived in the late 1980s as a successor to the ABC language. Python 2.0, released 2000, introduced features like list comprehensions and a garbage collection system capable of collecting reference cycles. Python 3.0, released 2008, was a major revision of the language that is not completely backward-compatible, and much Python 2 codedoes not run unmodified on Python 3. Due to concern about the amount of codewritten for Python 2, support for Python 2.7 (the last release in the 2.x series) was extended to 2020. Language developer Guido van Rossum shouldered sole responsibility for the project until July 2018 but now shares his leadership as amember of a keyperson steering council. Python interpreters are available for many operating systems. A global community of programmers develops and maintains Python, an open-source reference implementation. A non-prot organization, the Python Software Foundation, manages and directs resources for Python and Python development.

5.2 LIBRARIES USED

5.2.1 Open CV

OpenCV is an open-source software library for computer vision and machine learning. The OpenCV full form is Open-Source Computer Vision Library. It was created to provide a shared infrastructure for applications for computer vision and to speed up the use of machine perception in consumer products. OpenCV, as a BSD-licensed software, makes it simple for companies to use and change the code. There are some predefined packages and libraries that make our life simple and OpenCV is one of them.

5.2.2 NumPy

NumPy also called Numerical Python is an amazing library open-source Python library for data manipulation and scientific computing. It is used in the domain of linear algebra, Fourier transforms, matrices, and the data science field. which is used. NumPy arrays are way faster than Python Lists.

5.2.3 Skimage

Scikit-image or skimage, is an image processing Python package that works with NumPy arrays which is a collection algorithm for image processing. It is a simple and efficient tool for image processing and computer vision techniques. It is built on the top of NumPy, SciPy and matplotlib.

5.2.4 Keras Preprocessing

Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Keras contains numerous implementations of commonly used neural-network building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to makeworking with

image and text data easier to simplify the coding necessary for writing deep neural network code.

5.2.5 Python Imaging Library (PIL)

PIL stands for Python Imaging Library, and it's the original library that enabled Python to deal with images. Python Imaging Library is a free and open-source additional library for the Python programming language that adds support for opening, manipulating, and saving many different image file formats.

5.2.6 Image Data Generator

The Image Data Generator class allows your model to receive new variations of the images at each epoch. It is used to take the inputs of the original data and then transform it on a random basis, returning the output resultant containing solely the newly changed data.

CHAPTER 6

RESULTS AND DISCUSSION

6.1 INTRODUCTION

The efficiency and the accuracy of the tumors are tested with the image of acute lymphocytic leukemia and chronic lymphocytic leukemia. The input images is taken from the digital microscope which is initialized in the algorithm. After initializing the JPEG format image, the complexity of the image is reduced andthe pixels are evenly distributed to enhance the image quality. The similar information in the processed image is, further, clustered for effective segmentation. Then the morphological operation is performed to get the binaryimage. With the help of binary and complement image, we can clearly figure out acute lymphocytic leukemia and chronic lymphocytic leukemia in the image. The boundaries and the unwanted objects around the WBC cells were fixed, and the numbers of WBC cells are calculated using the bounding box. Then, geometric, texture and color features are extracted in the matrix format for the classification process.

6.2 STAGES OF DETECTION

Firstly, the input image is captured from the camera and loaded into the PYTHON software or in any digital media as shown in Figure 5.1.
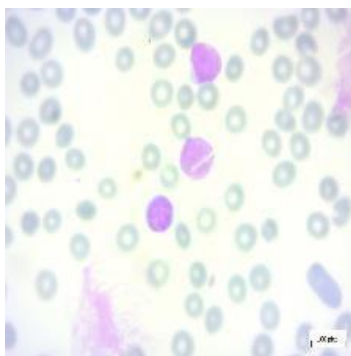

Figure 5.1 Input Image

The Figure 5.2 represents the enhancement and resizing of original image to get the clarity of an image. Image enhancement is the process of improving thequality and appearance of an image.
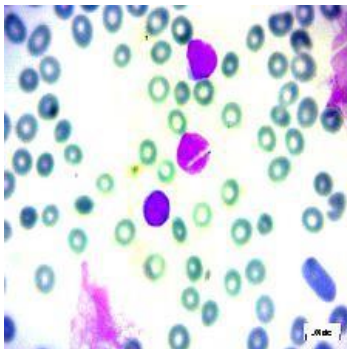

Figure 5.2 Enhancement and Resized Image

After the enhancement and resizing process, the image is converted into grayscale image using grayscale conversion. Then, the image will be processedin segmentation method to get the deep view of an image as shown in Figure 5.3.
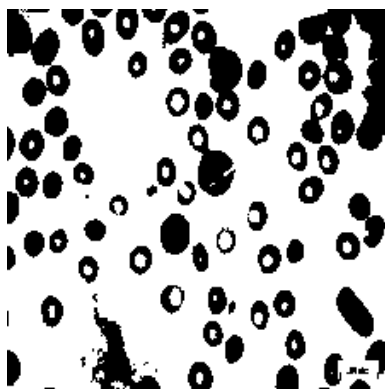

Figure 5.3   Segmented Image

The purpose of training is to modify the network's weights in order to improve the network's ability to correctly identify images based on their attributes. During training, the CNN learns to recognize crucial visual elements including edges, textures, forms, and colors and use those elements to generate predictions as shown in Figure 5.4.

During testing, the CNN makes predictionsfor each image in the testing set, and the accuracy of those predictions is testedby comparing them to the true labels of the images.
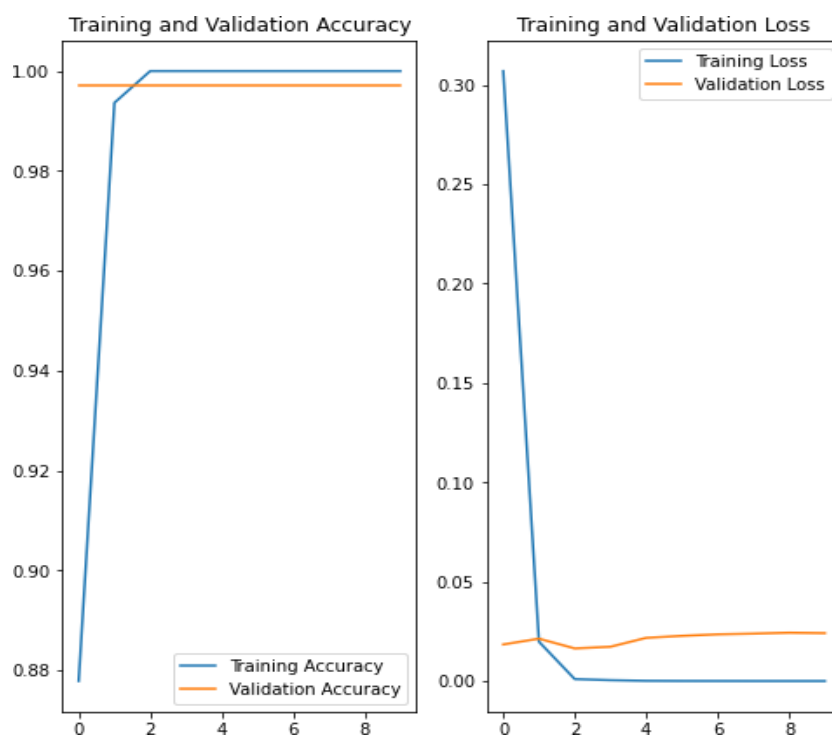

Figure 5.4 Training and Testing of an image

6.3 OUTPUT

The categorization model was developed using TensorFlow, an open-source framework that runs from top to bottom. An algorithm for binary classification was created using at least 1000 images. The accuracy and effectiveness of diagnosis can be greatly enhanced by the employment of deep learning techniques in the identification of acute and chronic leukemia. These methods are very helpful when analyzing medical images like bone marrow biopsies and blood smears. It is feasible to develop strong image processing pipelines for the identification of acute and chronic leukemia by combining the functions in various ways. To accurately identify acute and chronic leukemia, the performance of the most popular convolutional and

deep learning neural network approaches was tested. The trained model was then used to estimate the type of cancer depicted in the photograph**s.** Comparisons and analyses of the suggested approach with deep learning models are also described to gather the data in Python using tensor flow and Keras. The outputs of the suggested model are first described in the given Figure 5.5.

```
Epoch 1/10
89/89 [==============================] - 63s 697ms/step - loss: 0.5172 - accuracy: 0.7733 - val_loss: 0.0490 - val_accuracy: 0.9943
Epoch 2/10
89/89 [==============================] - 62s 691ms/step - loss: 0.0295 - accuracy: 0.9887 - val_loss: 0.0185 - val_accuracy: 0.9943
Epoch 3/10
89/89 [==============================] - 62s 691ms/step - loss: 0.0544 - accuracy: 0.9816 - val_loss: 0.0207 - val_accuracy: 0.9915
Epoch 4/10
89/89 [==============================] - 60s 674ms/step - loss: 0.0104 - accuracy: 0.9958 - val_loss: 0.0196 - val_accuracy: 0.9972
Epoch 5/10
89/89 [==============================] - 62s 691ms/step - loss: 0.0093 - accuracy: 0.9972 - val_loss: 0.0204 - val_accuracy: 0.9972
Epoch 6/10
89/89 [==============================] - 66s 741ms/step - loss: 2.6212e-04 - accuracy: 1.0000 - val_loss: 0.0236 - val_accuracy: 0.9943
Epoch 7/10
89/89 [==============================] - 62s 700ms/step - loss: 1.3186e-04 - accuracy: 1.0000 - val_loss: 0.0259 - val_accuracy: 0.9943
Epoch 8/10
89/89 [==============================] - 62s 699ms/step - loss: 8.5439e-05 - accuracy: 1.0000 - val_loss: 0.0277 - val_accuracy: 0.9943
Epoch 9/10
89/89 [==============================] - 63s 702ms/step - loss: 6.2733e-05 - accuracy: 1.0000 - val_loss: 0.0291 - val_accuracy: 0.9943
Epoch 10/10
89/89 [==============================] - 63s 701ms/step - loss: 4.8668e-05 - accuracy: 1.0000 - val_loss: 0.0302 - val_accuracy: 0.9943
```

Figure 5.5 Output Image

CHAPTER 7

CONCLUSION AND FUTURE SCOPE

7.1 CONCLUSION

In conclusion, the use of deep learning techniques in the recognition of acute and chronic leukemia can significantly improve the accuracy and efficacy of diagnosis. When examining medical images like bone marrow biopsies and blood smears, these techniques are particularly beneficial. The effectiveness of the most widely used convolutional and deep learning neural network algorithms was evaluated for the distinction of acute and chronic leukemia. CNN produced a detection accuracy rate of 99.72%. Wecan detect leukemia with 99.50% accuracy and classify its subtypes with 96.06% accuracy thanks to data augmentation.

7.2 FUTURE SCOPE

Leukemia can be diagnosed early with the use of this automated diagnosis technology, allowing for effective treatment. In order to make this diagnostic system usable in daily life and aid pathologists and oncologists in making a more accurate diagnosis of leukemia, we can also employ deep learning models to train from scratch with larger image data sets. By specifying the approach's input and output parameters and consolidating itas a component of a sub module, we can further enhance it to a fully automated system. The study shows that the suggested DCNN is more accurate at classifying leukemia and improving low-intensity photos than many other machine learning classifiers utilized in this field of study. The experimental finding shows that the suggested framework outperforms competing networks with the highest accuracy.

REFERENCES

1. Avci. K and Ozyurt. F (2020) 'White blood cells discovery and bracket grounded on indigenous convolutional neural networks', Med. suppositions,vol. 135.

2. Banik. P.P, Saha. R and Kim. K.D (2020) 'An automatic nexus segmentation and CNN model- grounded bracket system of whiteblood cell', Expert Syst.Appl.,vol. 149.

3. Dhal. K.G, Gálvez. J, Ray. S, Das. A (2020) 'Acute lymphoblastic leukemia image segmentation driven by stochastic fractal hunt', Multimedia Tools Appl.,vol. 2020,pp. 1 – 29.

4. Fan. H, Zhang. F, Xi. L, Li. Z, Liu. G and Xu. Y (2019) 'Leukocyte Mask An automated localization and segmentation system for leukocytein blood smear images using deep neural networks',J. Bio photonics,vol. 12,no. 7.

5. Hegde. R.B, Prasad. K, Hebbar. H, and Singh.B.M.K (2019) 'Point birth using traditional image processing and convolutional neural network styles to classify white blood cells a Computational and Mathematical styles in Medicine 7 study', Australasian Physical & Engineering lores in Medicine,vol. 42.

6. Hebber. H, Hegde. B (2019) 'Development of a robus t algorithm for discovery of capitals of white blood cells in supplemental blood smear images', Multimedia ToolsAppl., vol.

7. Jha. K.K and Dutta. H.S(2019) 'Collective information- grounded mongrel model and deep literacy for acute lymphocytic leukemia discovery in single cell blood smear images', cipher. styles Programs Biomed., vol.

8. Liu. Y and Long. F (2019) 'Acute lymphoblastic leukemia cells image analysis with deep bagging ensemble literacy', in CNMC Challenge Bracket in Cancer Cell Imaging, Singapore Springer.

9. Madhumalini. M, Meera Devi. T (2022) 'Discovery of Glaucoma from Fundus Images Using Novel Evolutionary - Grounded Deep Neural Network', Journal of Digital Imaging Volume 35.

10. Mustafa Ghaderzadeh, Farkhondeh Asadi and Arash Roshanpour (2021) 'Machine Learning in Discovery and Bracket of Leukemia Using Smear Blood Images a Methodical Review', Vol. 2021.

11. Ozyurt (2019) 'A fused CNN model for WBC discovery with MRMR point selection and extreme literacy machine', Soft Compute., vol.24, pp. 8163 – 8172.

12. Ritter. N and Cooper. J (2019), 'Segmentation and Border Identification of Cells in Images of Peripheral Blood Smear Slides', 30th Australasian Computer Science Conference, Conference in Research and Practice in Information Technology, Vol. 62, Pp. 161– 169.

13. Sharif. M, Amin. J, Siddiqa. A (2020) 'Recognition of different types of leukocytes using YOLOv2 and optimized Bag- of- Features', IEEE Access, vol. 8, pp. 167448 – 167459.

14. Togaçar. M, Ergen. B and Comert. Z (2020) 'Bracket of white blood cells using deep features attained from convolutional neural network models grounded on the combination of point selection styles', Applied Soft Computing, vol. 97.

15. Xu Gongwen, Zhang Zhijun, Yuan Weihua and XuLi Na (2014), 'On Medical Image Segmentation based on Wavelet Transform' IEEE.

DETECTION OF ALL AND CLL BLOOD CELLS IN WBC'S USING BLOOD SMEAR IMAGES

R. Vishnu Vardhan[1], P. Poornima[2], M. Shreetha[3], Shwathi R.M[4], D. Soundarya[5]

[1]Assistant professor, [2,3,4,5]UG Scholar

Department of Electronics and Communication Engineering

P.A. College of Engineering and Technology, Pollachi, TamilNadu, India

Abstract

White blood cells (WBC) are generated in the bone marrow and account for about 1% of all blood cells. These white blood cells proliferate unregulated, which leads to the development of blood cancer. The proposed work proposes a reliable method for classifying Acute Lymphoblastic Leukemia (ALL) and Chronic Lymphoblastic Leukemia using the ALL-IDB and Kaggle datasets (CLL). Acute lymphoblastic leukemia is a cancer that develops when the bone marrow produces an excessive number of lymphocytes (ALL). ALL is one of the most common types of cancer in kids. Contrarily, Chronic Lymphocytic Leukemia (CLL) attacks the spongy tissue inside the bones where blood cells are made. CLL is a type of blood and bone marrow malignancy. Older people are more frequently affected by CLL.Traditionally, the procedure was completed manually over a long period of time by a qualified expert. By utilizing deep learning techniques, specifically convolutional neural networks, the suggested approach completely eliminates the possibility of errors in the human process. The model first pre-processes the photos and extracts the best features after being trained on images of cells. The model was able to precisely collect the samples 90 times out of 100 and reproduce all measurements. In order to accurately identify the type of cancer in the bone marrow, the model can be employed. After classifying the normal and pathological cells using CNN, this task may be implemented utilizing the Python programming environment.

Keywords:Thresholding methods, image processing, image segmentation, image enhancement, and Convolution Neural Network (CNN).

I. Introduction

A blood cancer called leukemia is caused by the rapid proliferation of aberrant white platelet cells in the bone marrow. The uncontrolled growth of white platelets damages the body's immune system. The immune system provides sickness resistance. There are more white blood cells because of the malignancy leukemia's irregular cell structure. Researchers and oncologists have struggled for years to make an early diagnosis of leukemia due to the disease's initial non-obvious symptoms, such as fever, fatigue, and bruising. Moreover, leukemia doesn't form tumors like other cancers do, which makes it very challenging to detect.Manual methods for evaluating blood smears are time-consuming, costly, and prone to error since they rely on costly resources such as the knowledge of qualified medical experts who have specialized in picture interpretation. Patients with leukemia undergo a plethora of blood tests for blood cell counts and analysis after being diagnosed.