Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7766368

DISEASES CLASSIFICATION OF WBC USING DEEP LEARNING

D.Divya[a],*, Logavarshini R[b], Balasubramaniya[c] and Anisha Singh J[d]
[a]Associate Professor, Jerusalem College of Engineering, Chennai, India,
divya21cs@gmail.com
[b]Student, Jerusalem College of Engineering, Chennai, India,
logavarshinir@jerusalemengg.ac.in
[c]Student Professor, Jerusalem College of Engineering, Chennai, India,
balasubramaniyans@jerusalemengg.ac.in
[d]Student Professor, Jerusalem College of Engineering, Chennai, India,
anishasinghj@jerusalemengg.ac.in

ABSTRACT

Acute lymphoblastic leukemia (ALL) is characterized by an abundance of lymphoid blasts in the blood and typically affects adolescents. Due to their ability to undergo rapid differentiation, these cells pose a diagnostic risk. A misdiagnosis can cause additional health problems. Only careful microscopic inspection of these cells will yield an accurate diagnosis. Image analysis for quantitative evaluation of stained blood microscopic images for leukemia detection is an efficient and low-cost method that saves time and allows treatment to begin as soon as the diagnosis is made. Using fuzzy based two stage color segmentation, we can separate leukocytes from other blood components. Support Vector Machine (SVM) classification was used to derive features from a total of 108 images. The current techniques for detecting Leukemia take too long, cost too much money, and rely too heavily on medical professionals. We suggest an automated algorithm for Leukemia detection and classification using the computational framework MATLAB to address these limitations. The microscopic pictures are used as inputs, and a variety of image processing methods including image enhancement, segmentation, feature extraction, and classification are applied to them. Diseases can be detected and recognized at an early stage by analyzing microscopic images of blood cells. Image processing techniques are increasingly being used by hematologists for analysis, detection, and identification of leukemia kinds in patients. Since there is no need for expensive laboratory tools, image-based detection is a quick and low-cost option. Microscopic pictures undergo image processing procedures like enhancement, segmentation, and feature extraction.

KEYWORDS: Image Classification, Blood cells, White blood cells Detection, Deep learning

1.1 INTRODUCTION

Leukemia is a group of hematological neoplasia that typically effects the blood, bone marrow, and lymph nodes. Leukemia is defined as an uncontrollable increase in white blood cells in the bone marrow that does not react to cell growth inhibitors. Anemia, thrombocytopenia, and neutropenia are the end outcomes of hematopoiesis suppression caused by leukemia. Accumulation of white blood cells (WBC) can occur in a number of organs and tissues when the WBC count rises abnormally, including the meninges, gonads, thymus, liver, spleen, and lymph nodes. The increase in lymphatic blasts also causes their release into the bloodstream. [1]. The symptoms of leukemia indicate excess amount of lymphoid blast cells. For

classification and identification of blast cells the hematologists examine the blood smear under the microscope. Acute and chronic can be classified by pathologically the leukemia.

In this article, we focus exclusively on acute lymphoblastic leukemia (ALL), and we try to determine whether a given lymphocyte is a typical or abnormal lymphoblast. Examination of blood cells under a microscope is the gold standard for diagnosing leukemia, independent of the sophistication of the testing being used. The examination of leukemia is done under the microscope the results are standard when the operator is experienced and tiredness which results in inconsistent and subjective reports [2] [3]. The most economical way of leukemia diagnosis is examining under the microscope. So, we need an effective cost technique and accurate method to detect the leukemia which gives the accurate result without involvement of the operator fatigue.

Over the years, numerous methods for automatically segmenting data and detecting leukemia have been suggested. Most of the methods relied on data collected locally. Two-stage segmentation based on the HSV color space is employed. Literature contains studies with comparable aims concerning segmentation and detection. The ability to automatically identify and segment leukemia is crucial. In this paper, we suggest an automated method for analyzing blood smears as an aid to the doctor's diagnosis and treatment [4]. In this paper, we suggest a method to isolate lymphocytes from the rest of the blood components. Fractal characteristics, shape features, and other texture features are extracted from the analyzed lymphocytes. Two novel features were suggested for measuring the roughness of the boundaries of cell nuclei in order to detect leukemia [5]. Support vector machine uses extracted features to classify pictures as healthy or leukemia.

Early detection and diagnosis of the illness are the most important factors in determining the disease's cure rate and overall prognosis. Therefore, the detection of leukemia through the use of image processing is an improved and less difficult method because images are inexpensive, they do not require expensive testing and laboratory equipment, and they give quicker output, all of which surmount the disadvantages of manual testing [6] [7]. Cancer of the early blood-forming cells is known as leukemia. Leukemia most commonly affects the white blood cells, but it can also originate in other blood cell types. The progression of leukemia can be classified as either being rapid (acute) or gradual. Different kinds of leukemia have varying prognoses and can be managed in a variety of different ways.

1.2 RELATED WORKS

LeukoX is an ensemble learning method for classifying leukocytes. In this article [8], the authors suggest using ensemble learning applied to multiple deep neural networks to identify and classify across four types of white blood cells based on morphological features. This approach uses a novel least entropy combiner (LEC) network to successfully combine the individual classifier decisions to minimize the cross-entropy cost, as opposed to the DNN-based WBC classification schemes. In this study, we use a custom-designed convolutional neural network in conjunction with two pretrained DNNs, DenseNet121 and ResNet50, to generate class confidence scores that are then fed into the proposed LEC network for ensemble learning. The proposed network is trained and evaluated using a dataset of approximately 12, 500 images of the four distinct WBC types (eosinophil, lymphocyte, monocyte, and

neutrophil). The suggested model outperforms the component networks. Our findings show that the suggested model has a total test accuracy of 96.67 percent; we've dubbed this comprehensive model LeukoX. (X: Extended).

Classification of WBCs in Blood Smear Images: A Comparison of Conventional and Deep Learning Models. [9] [10] Traditional machine learning and deep learning techniques in computer vision have greatly aided the progress of medical image analysis by improving prediction accuracy, which in turn has led to more precise planning and diagnosis. These techniques can aid the hematologist and doctors by giving a second opinion and have significantly improved the automatic diagnosis of brain tumors and leukemia/blood cancers. This article presents a comprehensive study of current TML and DL methods for MIA. This review paper extensively explores the cutting-edge DL techniques, especially the developing convolutional neural networks-based models in the MIA domain, with a specific emphasis on leukocytes classification in blood smear images and other medical imaging domains. According to the research in this area, standard TML techniques are widely used for white blood cell analysis in microscopic blood smear images.

Leukemia diagnosis support via white blood cell segmentation using multiscale information fusion. Numerous individuals every year fall victim to leukemia, one of the deadliest forms of blood cancer. The presence of abnormalities in white blood cells is strongly linked to the detection of leukemia. The correct segmentation of WBCs allows for the detection of morphology and WBC count, both of which aid in the diagnosis and prediction of leukemia. Manual ways of determining WBC count are time consuming, subjective, and less precise. Our study [11] proposes a multi-scale information fusion network for WBC segmentation as a solution to these issues. The cytoplasm in WBC images has poor contrast in comparison to the background, and the shape and boundary of the nuclei can be difficult to make out in some instances. Our network's ability to fuse data from multiple scales aids in maintaining border information and boosts segmentation accuracy. We tested our network on four open-source datasets, and it outperformed all other methods in terms of segmentation accuracy. The suggested architecture also uses only 2.67 million trainable parameters, demonstrating excellent computational efficiency.

Microscopical Image Analysis of Acute Myeloid Leukemia Blood Smears for Feature Extraction of White Blood Cells Using CMYK-Moment Localization and Deep Learning. Aggregate classification accuracies of 97.57% and 96.41% were obtained using the main and secondary datasets, respectively [12], demonstrating that the proposed method obtained outstanding results in accuracy, generalizations, and stability using all the classifiers. This technique provides a new option for enhancing WBC detection, which may ultimately contribute to more accurate AML diagnosis.

White blood cell subtype detection and categorization was the topic of this presentation's [13] presentation. The suggested model has achieved 98.23% accuracy during training and 84.64% accuracy during testing on a dataset of 12,515 subtyped blood microscopic images. Utilizing localized convolutional neural networks, we can identify and categorize leukocytes. In this research, we suggest a computer-assisted, automated system for locating and identifying WBC types in blood images. As a technique, we've been employing R-CNNs, or regional convolutional neural networks. This detector has allowed for the simultaneous

classification of multiple cell types within the same picture. [14] AlexNet, VGG16, GoogLeNet, and ResNet50 designs have all been tried out with full learning and transfer learning while training CNN, the foundation of R - CNN architecture. By the conclusion of the experiment, the system had a perfect record of identifying WBCs. One of the CNN architectures, ResNet50, has demonstrated the highest transfer learning ability. The accuracy rates for distinguishing lymphocyte, monocyte, basophil, eosinophil, and neutrophil ranged from 99.52% to 98.48%, 96.16% to 95.04%, respectively.

The diagnosis of cancer in white blood cells through the use of an image processing. Within the scope of this research article, we have suggested a novel method in which the input image is comprised of microscopic images of blood. [15] A dataset consisting of 100 images, of which 62 are used for training and 38 are used for assessment. Following that, we proceeded to convert the picture into the appropriate format for segmentation, which was YCbCr. We used the combination of Gaussian Distribution and Otsu Adaptive Thresholding for segmenting, and we used the K-Means technique for clustering the data. The features are extracted with the help of the Gray Level Co-occurrence Matrix (GLCM), and a Convolutional Neural Network was utilized for the categorization process. After processing, we were able to acquire an accuracy rating of 97.3% for the system as a whole.

The deep learning analysis of white blood cells, red blood cells, and platelets are discussed. The objective of the newly suggested method [16] is to achieve a higher level of accuracy with YOLOv5 in comparison to earlier iterations of the YOLO model, which is based on the automatic detection, segmentation, and counting of individual blood cells extracted from blood smear images. Additionally, real-time implementation is possible, and the findings can be transmitted for additional patient diagnosis as soon as they are obtained. The primary objective of this study is to identify three primary groups of blood cells, and enhanced detection and segmentation of blood cells is demonstrated to be possible as a result of these efforts. The findings of the experiment conducted on YOLO v5s led the researchers to the conclusion that the greatest mAP was seen for 8 batches, 75 epochs, and had a map value of 93%.

TABLE 1 Method, Features and Challenges of Existing System

REFERENCE	METHODOLOGY	FEATURES	CHALLENGES
Sourodip Ghosh et al	DNN, CNN	1. A technique for bringing back blurred edges and details. 2. To hasten medical professionals' ability to diagnose	Out of the five trained CNNs and achieved an average recognition rate of 96.67%.
SIRAJ KHAN et al	TML	1. Methods are widely used for analysing WBCs in blood smear depicts under the microscope. 2. Assist in the diagnosis of hematological conditions like AIDS and blood malignancy.	This study's comprehensive data set will help researchers get off to a running start on developing better TML and DL frameworks for MIA.
Nadeem Akram et al	MIF-Net	The proposed technique has the potential to reliably aid medical professionals and lessen the strain on the diagnostics industry.	Manual WBC testing is time-consuming, unreliable, and can lead to false results.

TUSNEEM AHMED M. ELHASSAN et al	CNN, ROI	The main dataset achieved an overall classification accuracy of 97.57%, while the secondary dataset achieved 96.41%.	Good when multiple illnesses can be diagnosed and managed.
Hüseyin Kutlu et al	CNN, R-CNN	WBC cell determination by the method has been 100% accurate.	One of the CNN designs, ResNet50, has performed particularly well.

1.3 METHODOLOGY

Microscopical images must undergo preprocessing, segmentation, feature extraction, and categorization before leukocytes can be identified. There are various types of blood cells visible in the micrograph, including RBCs, WBCs, and Platelets. Our goal is to isolate the cytoplasm and nucleus of white blood cells using a color picture segmentation-based technique. Due to the limited amount of cytoplasm in acute leukemia cells, we must rely on the nucleus to extract the essential features of this disease.

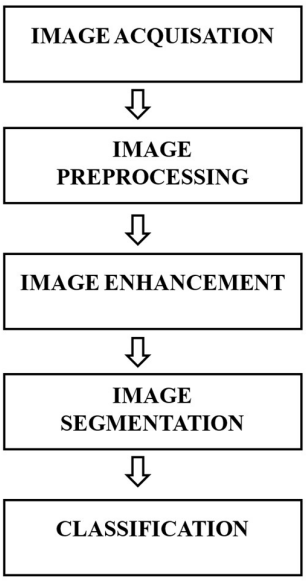

Fig. 1.1 Proposed Framework

This examination is carried out in the hope of discovering abnormal white blood cells, which would be an indication of the presence of cancer cells. Throughout the years, experienced operators have been responsible for conducting the examination. These operators will typically perform two analyses, including the classification and counting of cells (which are now carried out by cytometer). Due to the fact that the morphological analysis can be performed using just a single picture, it is not typically necessary to provide a blood sample. As a result, this analysis does not require any matching to be done because the expense of doing so is extremely high, the accuracy is typically the same across a variety of images, and the screening process is done remotely. Image acquisition, pre-processing, image segmentation, feature extraction, and detection of leukemia cells are the five phases that make up the process of detecting leukemia using microscopic image processing in the computer vision technique. The final stage is the detection of leukemia cells.

1.3.1 IMAGE ACQUISTION

Most video microscopy uses used analog video cameras, often just closed-circuit TV cameras, for image acquisition until the early 1990s. Video cameras gave images at full video frame rate (25-30 frames per second), enabling live video recording and processing [17], though this required the use of a frame grabber to digitize the images. While there were benefits to the development of solid-state detectors, the real-time video camera was far better. The action of retrieving a picture from some source, typically a hardware-based source, in order to prepare it for further processing can be referred to as "image acquisition." This is one of the more general definitions of what "image acquisition" means in the field of image processing. The process of acquiring an image in preparation for image processing is always the first stage in the workflow sequence. This is due to the fact that processing cannot take place in the absence of an image.

In some areas, having a consistent starting point from which to work is crucial, and the acquired image is entirely unprocessed and the result of whatever hardware was used to generate it. One of the ultimate aims of this process is to have an input source that works within such controlled and measured parameters that the same picture can, if necessary, be nearly perfectly reproduced under the same conditions so that anomalous factors are easier to locate and eliminate.

In this configuration, the digital microscope is connected to a computer, and the images acquired by the microscope are saved in digital format. The suggested system was trained and validated on a local dataset that was supplied by the Kaggle website. The dimensions of each image are 256 pixels by 256 pixels.

1.3.2 PRE-PROCESSING

The procedures performed on images at the most fundamental level of abstraction are collectively referred to as image pre-processing. If information is measured in terms of entropy, then these procedures do not make the image more informative; rather, they make it less informative. [18] The purpose of pre-processing is to improve the image data by eliminating unwanted distortions or enhancing certain image characteristics that are important for the subsequent processing and analysis tasks. This is the goal of the process. The preprocessing of pictures makes use of the redundancy that exists in the images. The brightness levels of neighboring pixels that each represent the same physical object are the same or very comparable. If a distorted pixel can be isolated from the picture, its original value can be recreated by taking the average of the values of the pixels that surround it. The size of the pixel neighborhood that is used for the calculation of a new pixel's brightness can be used to categorize the various image pre-processing techniques into a number of different categories.
The steps to be taken in preprocessing are:

- Read Image
- Resize Image
- Remove Noise

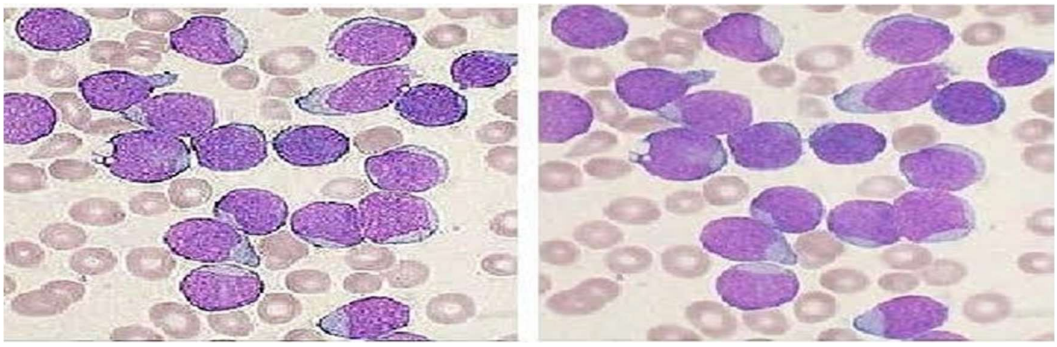

Fig 1.1 Noisy image and Enhanced image

1.3.3 IMAGE ENHANCEMENT

Imagery can be understood and interpreted visually more easily thanks to the enhancements that have been implemented. The ability to modify the values of the digital pixels that make up a picture is one of the benefits that come with using digital imagery. Before the data are given to the user, radiometric corrections may be made to account for factors such as illumination, atmospheric influences, and sensor characteristics; however, the picture may not be optimized for visual interpretation in spite of these corrections. [19] Devices for remote sensing, particularly those that are controlled from satellite platforms, need to be designed so that they can withstand levels of target/background energy that are typical of all conditions that are likely to be encountered during routine use. No generic radiometric correction could adequately account for and display the optimal brightness range and contrast for all targets due to the large variations in spectral response from a diverse range of targets (such as forests, deserts, snowfields, water, etc.). This is because there are so many different types of targets. Therefore, it is typically essential to make a bespoke adjustment of the range and distribution of brightness values for each application and image individually.

1.3.3.1 Median Filtering

The median filter is a method of non-linear digital filtering that is frequently utilized in the process of removing noise from either a picture or a signal. This kind of noise reduction is a typical stage in the preprocessing that is done in order to improve the results of the processing that comes later. Because, under certain circumstances, median filtering can maintain edges while simultaneously removing noise and having applications in signal processing, it is used extensively in digital image processing. Median filtering also has applications in signal processing.

The fundamental premise behind the median filter is that it should be applied to the signal in such a way that it is traversed entry by entry, with each entry being substituted with the median of the entries in the surrounding neighborhood. Both median filtering and linear Gaussian filtering are examples of techniques that can be used to normalize data. The edges of a signal are negatively impacted by all smoothing techniques, despite the fact that these techniques are all successful at removing noise from smooth patches or smooth regions of a signal. However, in many cases, it is essential to keep the boundaries while simultaneously cutting down on the noise that is present in a signal [20]. Because of this, median filtering is utilized quite frequently in the process of digital picture creation.

1.3.4 IMAGE SEGMENTATION

In digital image processing and analysis, image segmentation is a technique that is frequently used to partition an image into multiple sections or regions. This partitioning is typically done on the basis of the characteristics of the pixels contained within the image. In image segmentation, the foreground and background may be separated from one another, or regions of pixels may be grouped together based on similarities in color or structure [25], [26]. Over the course of time, numerous algorithms and methods for image segmentation have been developed using domain-specific knowledge to successfully solve segmentation problems in that particular application area. These were developed in order to segment images in a more precise manner are illustrated in Fig. 1.2. Imaging for medicinal purposes, automated driving, video surveillance, and machine vision are some examples of these applications. When conducting a medical diagnosis for cancer, pathologists use hematoxylin and eosin staining of body tissue to differentiate between the various kinds of tissue. After that, they employ a method of image segmentation known as clustering in order to recognize the various kinds of tissue present in their images.

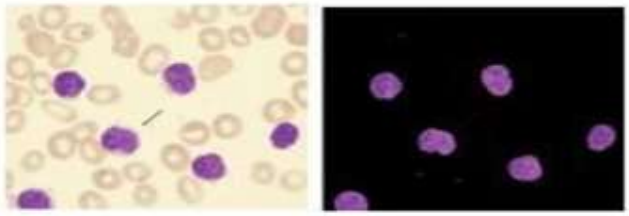

Fig. 1.2 Segmented Image

1.3.4.1 FUZZY C-MEANS CLUSTERING:

Fuzzy C-Means (FCM) Clustering is the most common approach for segmenting images due to its high accuracy and efficiency in classifying data. Soft segmentation techniques like feature correspondence mapping (FCM) can be used on medical pictures. The effectiveness of this approach to finding the best answer is conditional on a variety of factors, such as the locations of the cluster centers to begin with, the degree of participation measured for each data point, etc. In the conventional FCM, the centers are seeded at random, and the gray characteristic is the only one used in the membership measure [21, 22]. As a result, it takes a long time and is easily distracted by background sounds. Many researchers have spent a lot of time and money developing efficient methods for segmenting FCM images. Some methods for quickening the segmentation process pay special attention to how to initiate the centers of necessary groups. Since noise is inevitable in medical images, we also show some FCM segmentation methods that are less susceptible to it. Some researchers use a hybrid of the FCM and other mathematical methods to avoid the trap of a locally optimal answer [23].

Algorithm steps for fuzzy c-means clustering

The algorithm of fuzzy c-means image segmentation are given as follows.

1. Initialize the random cluster centers vec_j for the variable $t = 0$.
2. Initialize fuzzy membership functions l_{ij}.

3. Compute new clusters vec_jusing the centroid equation by incrementing the variable $t = t + 1$.

4. Repeat step2 and step 3, until converges.

Therefore, the Fuzzy c mean approach converges to a local minimum. This is employed as a local search to resolve the problem and also modeled to differentiate the actual positions of the cluster centers.

1.3.5 CLASSIFICATION

The process of assigning a label from one of the known classes to an unknown test vector is what we mean when we talk about classification. Support vector machines, also known as SVMs, are utilized for categorization purposes because the patterns are relatively similar in the feature space. The Support Vector Machine (SVM) is a potent tool for data classification that is built on the hyper plane classifier. A separating surface, either linear or nonlinear, is utilized in the input space of the data collection in order to accomplish this classification [24]. They are, in essence, two-class classifiers that work to maximize the difference in threshold between the classes. Finding the support vectors is the task that the classification training algorithm is responsible for. The significance value model (SVM) receives its input from relevant extracted features.

1.3.5.1 DEEP LEARNING TECHNIQUES

Deep learning is an AI technique that attempts to mimic the way that humans acquire new information. Learning by example is second nature to humans, and deep learning is a machine learning method that trains computers to do the same. Driverless cars rely heavily on deep learning technology, which allows them to do things like identify stop signs and tell people from lampposts. It's essential for voice-activated functions on smartphones, tablets, TVs, and wireless speakers. These days, deep learning is all the rage, and for good cause. It means accomplishing goals that were previously out of reach.

A computer model can learn to execute classification tasks directly from images, text, or sound using a technique known as "deep learning." Models trained with deep learning have the potential to achieve a precision that is comparable to or even superior to that of humans. When training a model, a sizable amount of data that has been labeled is utilized, as well as neural network architectures that have many levels.

The three primary kinds of cells that are found in blood are known as red blood cells, platelets, and white blood cells. Red blood cells are essential for the movement of oxygen from the heart to the rest of the body's tissues, as well as the removal of carbon dioxide from those tissues. They account for up to half of the total amount of blood in the body. White Blood Cells, also known as WBCs, are the primary component of the immune system that the body uses to fight off infections and diseases. As such, they play an essential role in the immune system. The accurate categorization of WBCs is therefore necessary and is becoming an increasing area of focus and interest. The composition of the cytoplasm can be used to differentiate between two distinct kinds of white blood cells (WBCs). Granulocytes are the first form of white blood cell and are comprised of three subtypes: basophils, eosinophils, and neutrophils. The lymphocytes and monocytes that make up the second group are referred to collectively as agranulocytes.

A subclass of blood cells that are produced in the bone marrow and can be discovered in the blood as well as the lymph tissue. White blood cells are a component of the defense system that the body possesses. They assist the immune system in warding off infections and other illnesses. It is possible to use it to screen for conditions such as infections, inflammations, allergic reactions, and even malignancy, also known as leukocytes and white blood cells.

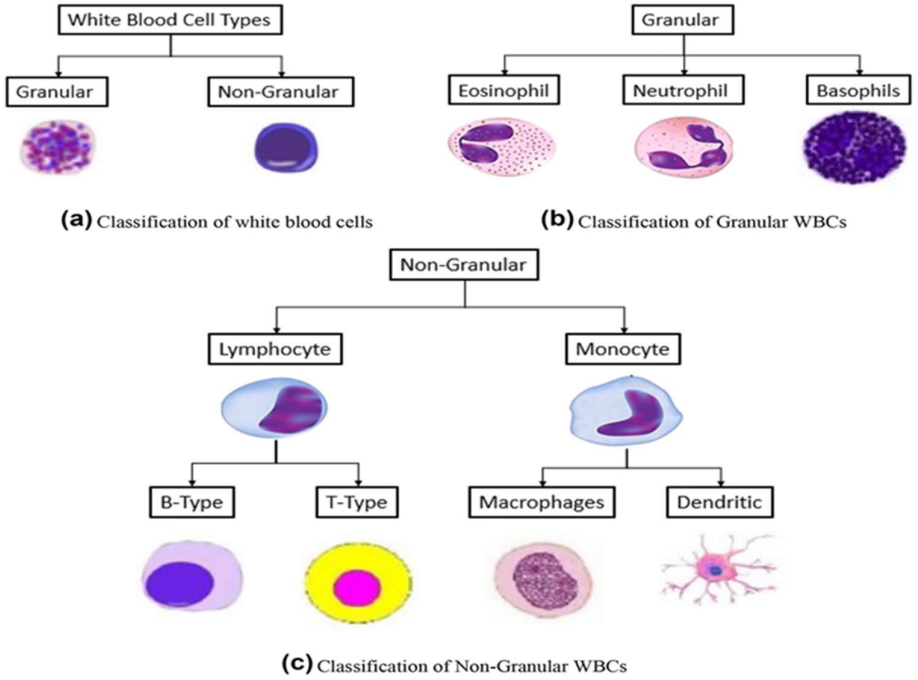

Fig 1.3 WBC Classification

The phase of classification is one of the most essential stages in image processing and techniques involving machine learning. Additionally, this is a field that is in high demand in this sector of the industry. A collection of previously unclassified data can be assigned and classified through the use of classification. Both supervised and unstructured classifications are examples of different kinds of classifiers. Supervised classifications are the more common of the two. The collection of potential results or classes is already known in advance when performing supervised classification. During the unsupervised classification process, the set of classes that will be used is not known in preparation. Classification of data can be accomplished through a variety of approaches, which are collectively referred to as the classifications. These classifiers, such as support vector machines (SVM), artificial neural networks (ANN), random forests (RF), k-nearest neighbor (KNN), naive bayes (NB), multilayer perceptrons (MLP), hybrids, and others, can be used to identify the types of objects. A summary of these classifiers is provided below. Following the selection and extraction of features from the segmented picture, the stage that follows involves the process of recognizing and determining the type of the object. Image classification is a technique of image processing that differentiates between various categories of objectives based on the various characteristics of images. Pattern identification and computer vision are two fields that make extensive use of it. The Support Vector Machine, or SVM, is a new technique for machine learning that is based on statistical learning theory. It has a rigorous mathematical foundation and is constructed on the structural risk minimization criterion. In this paper, we develop an image classification algorithm based

on SVM. To extract the image feature, we use the Gabor wavelet transformation. Principal Component Analysis (PCA) is used to reduce the dimension of the feature matrix.

1.4 RESULT AND DISCUSSION

1.4.1 Experimental Setup

During the course of this project, we will collect a fresh collection of characteristics in order to improve our classification. We have made an effort to determine such characteristics, the majority of which are followed by hematologists. An authority, examines the findings obtained in terms of the features to ensure their accuracy. For training we considered out of 108 blood cell images, we have extracted 45 different types of leukemia cells and non leukemia cells. 90 images are used for feature extractions. For a better classification between leukemia and non leukemia we have extracted 9 different features from this image. The characteristics extracted from 25 cells with leukemia and 25 cells without leukemia are used for training, and the remainder cells are used for testing. The result of the leukemia data was considered to be 1, while the output of the other data was considered to be 0. The suggested plan has an advantage over other plans already in place because it takes into consideration smear images that contain a large number of lymphocytes. Most of the time, the existing schemes take into account only the images that have a single lymphocyte visible within the field of vision.

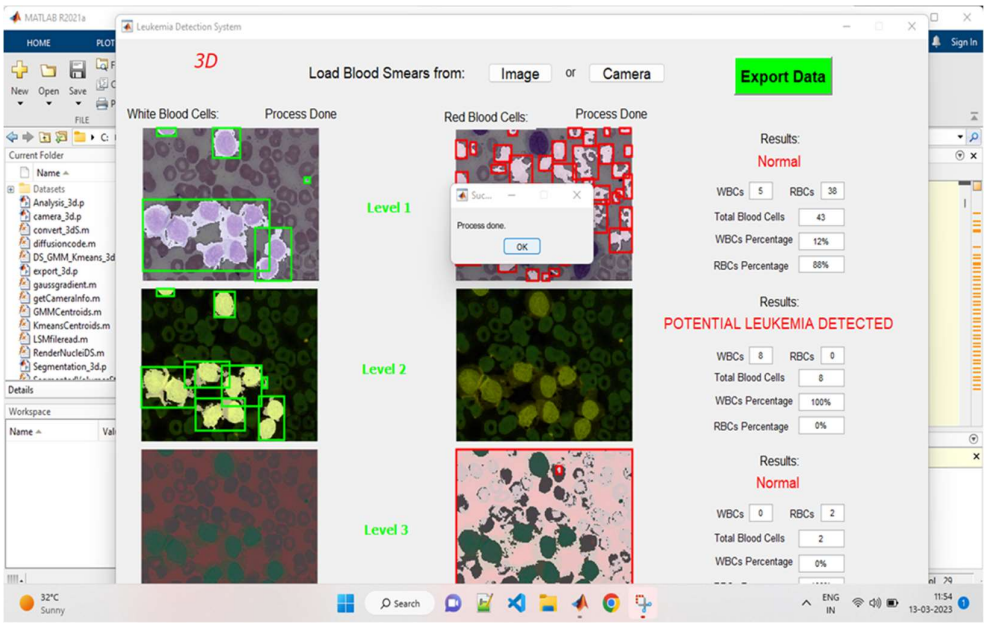

Fig 1.4 Output

1.4.2 Performance Metrics

In this context, performance measurements are taken into consideration for the extraction of shape features, color features, and textural features.

(a) Area:

Counting the overall number of pixels within the image region that are not zero allowed us to calculate the area.

(b) Perimeter:

The calculation of the distance between each succeeding boundary pixel was how it was calculated.

(c) Compactness:

The degree to which a nucleus is compact or spherical can be thought of as its measure. The degree to which the object is compressed into its space. The structure of the nucleus can look very different depending on both the age of the leucocytes and the sort of leucocytes they are. In mature leucocytes, the nuclei typically have more than two lobes and are connected by thin threads between the lobes. The outlines of the nucleus can sometimes take the form of a kidney bean in very specific circumstances. On the other hand, the nuclei of leukemic cells are round and have a greater overall compactness than the nuclei of mature cells. The following equation represents the compression measure we are interested in.

Compactness = Perimeter Sq \ Area

(d) Solidity:

Solidity is defined as the ratio of an object's actual area to the area of its convex hull, and it is an important characteristic for the classification of blast cells. The standard for this measurement is.

$$Solidity = \frac{Area}{Convex\ Area}$$

(e) Eccentricity:

With the help of this parameter, one can determine the degree to which the structure of a nucleus departs from being circular. Due to the fact that lymphocytes are more circular than blast cells, this is an essential characteristic. In order to quantify this, a relation has been specified in.

$$Ecc = \frac{\sqrt{a^2 - b^2}}{a}$$

where ‘a’ is the major axis and ‘b’ is the minor axis of the equivalent ellipse representing the nucleus region.

(f) Form factor:

This is a dimensionless parameter that responds differently depending on the surface irregularities, and its definition is as.

$$Form\ factor = \frac{4 \times \pi \times area}{Permimeter^2}$$

(g) Elongation:

Leukemia can also be identified by the abnormal bulging of the nucleus, which is a characteristic of the disease. As a result, the degree to which the nucleus bulges is evaluated in terms of a relation known as elongation.

$$Elongation = \frac{R_{max}}{R_{min}}$$

(h) Nuclear- cytoplasmic ratio

It is measured as the proportion of the total area of the cell nucleus to the total area of the cytoplasm. This metric plays a very significant role in determining whether or not the cell has reached its mature state.

$$nuclear\ cytplasmic\ ration = \frac{area\ of\ the\ nucleus}{cytoplasm\ 9\ ratio}$$

(i)Homogeneity:

It is a measurement of the amount of difference between values. The homogeneity of image textures is reflected by the homogeneity of homogeneity, and the local variations in image texture are scaled. The absence of intra-regional variation is indicated by high levels of homogeneity.

(j) Energy:

It is utilized in the process of measuring consistency. Small energy profiles were produced when all of the co-occurrence matrices had the same values; on the other hand, high energy profiles could be anticipated when the values were not identical.

(k) Correlation:

This is a representation of the correlation between the pixel numbers and its surrounding area. The value of correlation is a reflection of the uniformity of the image texture.

(l) Entropy:

Typically applied in order to evaluate the degree of unpredictability. The non-uniformity and complexity of an image's structure are reflected in its entropy value.

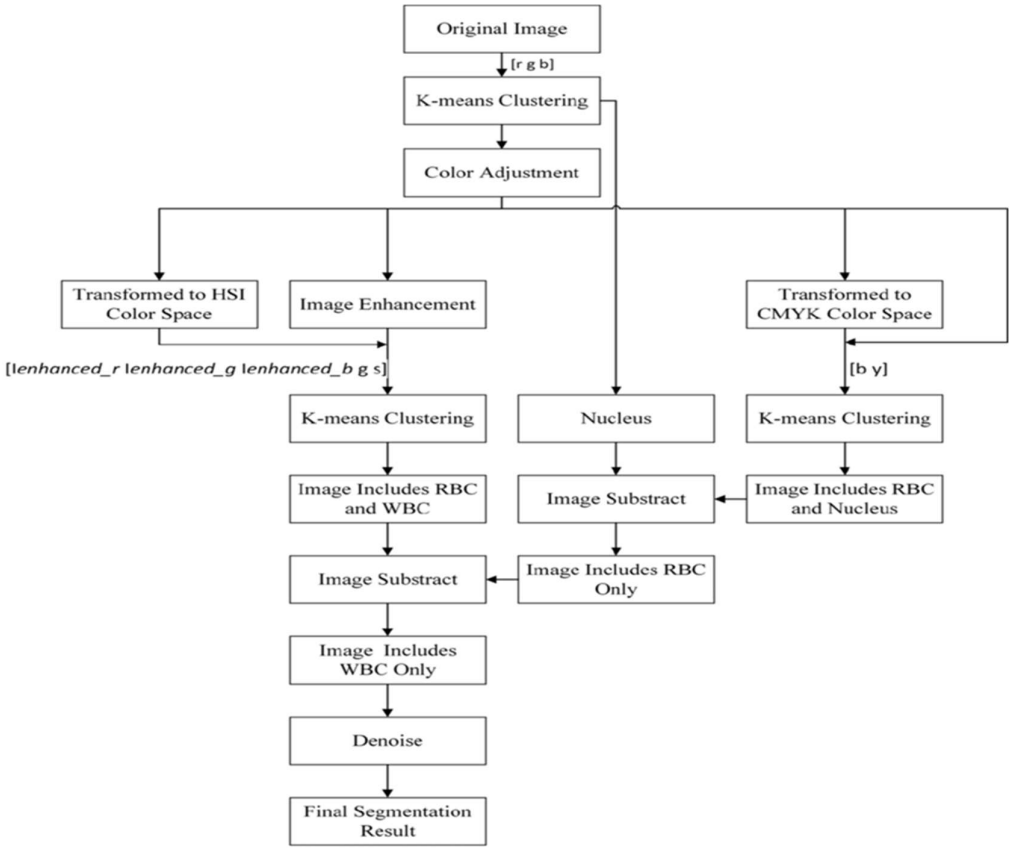

Fig 1.4 Proposed flowchart

Table-2 Results for Features

Measure	Leukemia	Non-Leukemia
Area	481	102
Perimeter	93.0480	36.4220
Eccentricity	0.8579	0.8146
Compactness	17.9999	13.0055
Solidity	0.9007	0.9273
Formfactor	0.6981	0.9662
Homogeneity	0.6822	0.6732
Energy	0.1928	0.2017
Entropy	0.1784	0.1992

1.5 OVERALL PERFORMANCE ANALYSIS

Our objective is to acquire a fresh group of characteristics for the purpose of improved categorization. We have made an effort to determine such characteristics, the majority of which are followed by hematologists. An authority, examines the findings obtained in terms of the features to ensure their accuracy. For training we considered out of 108 blood cell images, we have extracted 45 different types of leukemia cells and non leukemia cells. 90 images are used for feature extractions. For a better classification between leukemia and non leukemia we have extracted 9 different features from this images.

The feature extracted from 25 leukemia and non leukemia cells are used for training and remaining are used for testing. We considered the output of the leukemia data as 1 and non leukemia data as 0. We are taking into consideration smear images with a large number of lymphocytes, which is a significant benefit of the suggested scheme in comparison to the other schemes that are currently in use. The majority of the images in which there is only one lymphocyte visible under the field of vision are the ones that are taken into consideration by the existing schemes.

Obtaining images from a microscope is typically challenging and always requires the participation of a human observer, which is undesirable in an automatic system. The plan that has been suggested, which makes use of the new features, is most definitely moving in the direction of an automated system. The features that were extracted from the images that were available using our suggested method were put through the SVM classifier training process, and the results showed an accuracy of 94.7368%.

1.5 CONCLUSION

The suggested plan has an advantage over other plans already in place because it takes into consideration smear images that contain a large number of lymphocytes. Most of the time, the existing schemes only take into account images that have a single lymphocyte visible within the field of vision. Obtaining images from a microscope is typically challenging and always requires the participation of a human observer, which is undesirable in an automatic system.

The plan that has been suggested, which makes use of the new features, is most definitely moving in the direction of an automated system. An accuracy of 94.7368% was achieved when utilizing the features extracted from the images that were made available according to the plan suggested.

1.6 FUTURE SCOPE

The primary focus of this work is a two-step process for segmenting white blood cell nuclei from stained blood smear pictures, followed by feature extraction useful for identifying leukemia. The main focus of the project is to use the Hausdorff dimension and the contour signature to measure the irregularities of the nucleus's border. Better detection accuracy is achieved by also taking into account color and texture characteristics in addition to shape. SVM classifier was used for leukemia detection using the suggested features.

As a consequence of the findings, it is possible to take into consideration the possibility of conducting additional research in this field, such as the classification of lymphoblasts into a variety of subtypes. Additionally, various techniques can be developed or enhanced for touching cells, classifying leukemia types, and image segmentation without the use of staining.

REFERENCES

Rezatofighi SH, Khaksari K, Soltanian-Zadeh H. Automatic recognition of five types of white blood cells in peripheral blood. Proceedings of the International Conference of Image Analysis and Recognition; 2022; pp. 161–172.

Tabrizi PR, Rezatofighi SH, Yazdanpanah MJ. Using PCA and LVQ neural network for automatic recognition of five types of white blood cells. Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10); September 2021; pp. 5593–5596.

Yampri P, Pinnation C, Daochai S, Teartulakarn S. White blood cell classification based on the combination of eigen cell and parametric feature detection. Proceedings of the 1st IEEE Conference on Industrial Electronics and Applications (ICIEA '06); May 2020; pp. 1–4.

Gautam, A., Bhadauria, H. (2014). Classification of white blood cells based on morphological features. In 2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 2363-2368. https://doi.org/10.1109/ICACCI.2014.6968362

Macawile, M.J., Quinones, V.V., Ballado, A., Cruz, J.D., Caya, M.V.: White blood cell classification and counting using convolutional neural network. In: 2018 3rd International Conference on Control and Robotics Engineering (ICCRE) (2018) Author Proof12 I. Singh et al.

Ghosh M, Das D, Mandal S, et al. Statistical pattern analysis of white blood cell nuclei morphometry Proceedings of the IEEE Students' Technology Symposium (TechSym '10); April 2021; pp. 59–66.

K. S. Ravichandran and B. Ananthi. Color Skin Segmentation using K–Means Cluster. International Journal of Computational and Applied Mathematics, 4(2):153 – 157, 2009.

A. P. Pentland. Fractal based description of natural scenes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 6:661 – 674, 1984.

B. B. Mandelbrot. How long is the coast of Britain? Statistical self similarity and fractional dimension. Science, 156:636 – 638, 1967.

Daniel Graves and Witold Pedrycz Fuzzy C-Means, Gustafson Kessel FCM, and Kernel Based FCM: A Comparative Study Analysis and Design of of Intelligent Systems using Soft Computing Techniques, Springer, 2007.

B. T. Milne. Measuring the fractal geometry of landscapes. Applied Mathematics and Computation, 27:67 – 79, 1988.

N. Sinha and A. G. Ramakrishnan. Automation of Differential Blood Count. In Proceedings Conference on Convergent Technologies for AsiaPacific Region, 2:547 – 551, 2003.

C. Haworth, A. Hepplestone, P. Morris Jones, R. Campbell, D. Evans, M. Palmer. Routine Bone Marrow Examination in the Management of Acute Lymphoblastic Leukaemia of childhood. . Journal of Clinical Pathology, 34: 483 – 485, 1981.

G. Ongun, U. Halici, K. Leblebicioglu, V. Atalay, M. Beksac, and S. Beksak, An Automated Differential Blood Count System.. In Int. Conf. of the IEEE Engineering in Medicine and Biology Society , volume 3, pages 2583 - 2586, 2001

Subrajeet Mohapatra Development of Impulse Noise Detection Schemes for Selective Filtering Master Thesis, National Institute of Technology Rourkela, 2001.

K Ananthajothi, K Karthikayani, R Prabha 'Explicit and implicit oriented Aspect-Based Sentiment Analysis with optimal feature selection and deep learning for demonetization in India', Data Knowledge and Engineering, DOI:doi.org/10.1016/j.datak.2022.102092, Vol. 142.

K Ashokkumar, S Parthasarathy, S Nandhini, K Ananthajothi, Prediction of grape leaf through digital image using FRCNN, Measurement: Sensors, Volume 24, 2022, 100447, ISSN 2665-9174, https://doi.org/10.1016/j.measen.2022.100447.

Ananthajothi, K., Rajasekar, P. & Amanullah, M. Enhanced U-Net-based segmentation and heuristically improved deep neural network for pulmonary emphysema diagnosis. Sādhanā 48, 33 (2023). https://doi.org/10.1007/s12046-023-02092-5

Prameeladevi Chillakuru, D. Divya & K. Ananthajothi (2022) Enhanced Segmentation with Optimized Nine-Layered CNN-Based Classification of Leaf Diseases: An Automatic Approach for Plant Disease Diagnosis, Cybernetics and Systems, DOI: 10.1080/01969722.2022.2151173

Sangeetha, S. K. B., and K. Ananthajothi. "Machine Learning Tools for Digital Pathology-The Next Big Wave in MEDICAL SCIENCE." Solid State Technology 63.4 (2020): 3732-3749.

Ananthajothi, K., and M. Subramaniam. "Efficient classification of medical data and disease prediction using multi attribute disease probability measure." Applied Mathematics & Information Sciences 13.5 (2019): 783-789.

Ananthajothi , K and Subramaniam M, 'Efficient Classification of Medical Data Using Class Level Disease Convergence Divergence Measure', International Journal of Innovative Technology amd Exploring Engineering, 8(10) 2019.

Divya D and Ganeshbabu T R 'Fitness Adaptive Deer Hunting-Based Region Growing and Recurrent Neural Network for Melanoma Skin Cancer Detection', in the International Journal of Imaging Systems and Technology, ISSN 0899-9457, E.ISSN 1098-1098, DOI: 10.1002/ima.22414, 2020.

Divya D and Ganeshbabu T R 'An Empirical Study of Melanoma Detection and its Deep Learning Classifiers', in the International Journal of Critical Review, ISSN 2394-5125, Vol. No. 7, Issue No. 15, 2020.

D. Divya 'A Survey on Image Segmentation Techniques' in the Second International Conference on Emerging Current Trends in Computing and Expert Technology (COMET 2019), which is held on March 22 nd and 23 rd in the year 2019, and organized by Panimalar Engineering College, Chennai.

D. Divya 'A Comprehensive Survey on Medical Image Classification Methods' is presented in the International E-Conference on Information Communication and Networking, which is held on July 10 th and 11 th in the year 2020, and organized by Easwari Engineering College, Chennai.