Keywords: database and finally classified as AML and ALL which includes contrast enhancement and image segmentation techniques. Finally the extracted features are compared with those stored in the subtree and terminal set for inputting the leaf node instructor flexible neuron instructor for joining the non-leaf node’s Cancer. Two types of sets are used in the model namely hybrid Approach have been used for classifying the Leukemia hematological analyzer. Hybrid Flexible Neural Tree red blood cells count. The results were benchmarked using a projection of Artificial Neural Network’s done. Parameters utilized are number of hemoglobin, packed cell volume and red blood cells count. The results were benchmarked using a hematological analyzer. Hybrid Flexible Neural Tree Approach have been used for classifying the Leukemia Cancer. Two types of sets are used in the model namely hybrid flexible neuron instructor for joining the non-leaf node’s subtree and terminal set for inputting the leaf node instructor [3]. For proper identification of diseases it is very important to identify which information is relevant and which is irrelevant. ‘Ask expert’ approach is used in [4]. Proposed approach cannot deal with the unknown diseases and the development cost is also high. In [5] an open source based system is developed for the retrieval, analysis and interpretation of Leukemia images. Quad Tree was applied to represent successive subdivisions of the image due to its capability of handling the disjoint structures. Open source software ‘Imgsseek’ was used to evaluate the performance of retrieving images. The proposed system clearly marks a distinction between AML and CML images whereas ‘Imgsseek’ cannot. Vacuoles are used to identify a specific type of AML but it was not considered in the proposed system in order to reduce the complexity (only cytoplasm, nucleus and nucleolus were considered). In [6] another approach for classification of haematologic diseases viz malaria and thalassemia is proposed using artificial intelligence. Before preparing dataset for training and testing various image processing tasks were performed including size normalization, color normalization, specifying ROI, extracting ROI, edge erosion etc. Three layer and four layer neural networks are employed for classification and the three layer structure has the better performance. [7] presents literature based on granulometry. It also considers agglutination and edge fracture as the side effects of preprocessing. DCC Technique along with IFS is used to compare set of images for object recognition. Performance of the proposed technique is compared with the existing methods for the three parameters: space, time and execution efficiency. Performance for space and execution efficiency is higher for the proposed method but it takes longer time to execute [8]. In [9] cancer and blood disorders are predicted using MLP with BP learning algorithm. [10] discuses automatic diagnosis of leukemia using a Hybrid Fuzzy SVM Classifier and concludes that filtered gene classifier more accurately classifies as compared to the unfiltered classifier since redundant information is excluded. [11] uses a simple method for estimation of hemoglobin in blood by first taking the photograph of the thumb then a rubber band is used to increase the flow of blood in the thumb and again it is photographed and both images are

INTRODUCTION

Leukemia is the cancer of blood. It can be classified as acute or chronic. Acute leukemia is further classified into acute lymphocytic leukemia (ALL) and acute myelogenous leukemia (AML). Abnormal white blood cell known as blast consists of cytoplasm and nucleus. Both need to be separated properly for the proper diagnosis of leukemia. Identification of lymphoid or myeloid series can be done by the shape of the nucleus. Myeloid has a regular shape that is it has circular nucleus whereas lymphoid doesn’t have a regular nucleus. Nucleus of lymphoid is sparse whereas that of myeloid is dense. Image segmentation is done to separate WBC from the image. Contrast enhancement techniques ease the classification of leukemia into ALL and AML. Bright stretching, dark stretching and partial contrast are the techniques to enhance the contrast of the image. Partial contrast can be combined with bright or dark stretching to improve the results. Before any diagnosis can be done the counting area has to be identified. Counting area refers to the area where cells do not interfere with each other and are properly separated. Once the area is identified then extraction of the area of interest can be performed.

LITERATURE SURVEY

Identification of counting area in hematology is very important and slide maker devices are very expensive to be afforded by medium sized labs. To facilitate the identification of counting area an approach is proposed in [1]. Many papers can be found in literature for counting red blood cells. In [2] comparison of two algorithms i.e. ‘Connected component labeling’ and ‘Back projection of Artificial Neural Network’ is done. Parameters utilized are number of hemoglobin, packed cell volume and red blood cells count. The results were benchmarked using a hematological analyzer. Hybrid Flexible Neural Tree Approach have been used for classifying the Leukemia Cancer. Two types of sets are used in the model namely hybrid flexible neuron instructor for joining the non-leaf node’s subtree and terminal set for inputting the leaf node instructor [3]. For proper identification of diseases it is very important to identify which information is relevant and which is irrelevant. ‘Ask expert’ approach is used in [4]. Proposed approach cannot deal with the unknown diseases and the development cost is also high. In [5] an open source based system is developed for the retrieval, analysis and interpretation of leukemia images. Quad Tree was applied to represent successive subdivisions of the image due to its capability of handling the disjoint structures. Open source software ‘Imgsseek’ was used to evaluate the performance of retrieving images. The proposed system clearly marks a distinction between AML and CML images whereas ‘Imgsseek’ cannot. Vacuoles are used to identify a specific type of AML but it was not considered in the proposed system in order to reduce the complexity (only cytoplasm, nucleus and nucleolus were considered). In [6] another approach for classification of haematologic diseases viz malaria and thalassemia is proposed using artificial intelligence. Before preparing dataset for training and testing various image processing tasks were performed including size normalization, color normalization, specifying ROI, extracting ROI, edge erosion etc. Three layer and four layer neural networks are employed for classification and the three layer structure has the better performance. [7] presents literature based on granulometry. It also considers agglutination and edge fracture as the side effects of preprocessing. DCC Technique along with IFS is used to compare set of images for object recognition. Performance of the proposed technique is compared with the existing methods for the three parameters: space, time and execution efficiency. Performance for space and execution efficiency is higher for the proposed method but it takes longer time to execute [8]. In [9] cancer and blood disorders are predicted using MLP with BP learning algorithm. [10] discuses automatic diagnosis of leukemia using a Hybrid Fuzzy SVM Classifier and concludes that filtered gene classifier more accurately classifies as compared to the unfiltered classifier since redundant information is excluded. [11] uses a simple method for estimation of hemoglobin in blood by first taking the photograph of the thumb then a rubber band is used to increase the flow of blood in the thumb and again it is photographed and both images are

Abstract: Blood is a very important component of our body. The main components in the blood are RBCs (Red Blood Cells), WBCs (White Blood Cells), plasma and platelets. Blood serves as an indicator for identifying many diseases namely anemia, malaria, thalassemia, Leukemia etc. Abnormality in the shape and count of white blood cells causes different diseases. Leukemia is one of them. This paper deals with the identification of Leukemia. Leukemia is the cancer of blood. Many methods have been proposed for the same in the literature. Complete review on the papers on Leukemia detection is done in the literature survey. For proper diagnosis of Leukemia image preprocessing has to be done which includes contrast enhancement and image segmentation techniques. Finally the extracted features are compared with those stored in the database and finally classified as AML and ALL

Keywords: Leukemia, AML, ALL, CML, CLL, Lymphoid, Artificial Neural Network
then compared. Both images have to be taken under same standard conditions. In [12] the Hb levels in human blood were estimated using the BPN Architecture consisting of five layers and was compared with the standard cyanmethemoglobin method. It was also verified that the results do not offer more accuracies with more than three layers but it took more time to train the network. [13] discusses Flow Cytometry in the form of a game for the recognition of cancer. Technique determines two diseases closer to the goal then it is the task of the doctor to diagnose the exact disease between the two results. In [14],[15],[16],[17] detection of Acute Lymphocytic Leukemia is done and classification is done using Support Vector Machine. Color image segmentation is done and RGB Color space is converted to L*a*b* before clustering. Two stage segmentation is done first using K means clustering then by nearest neighbor classification. Images with many lymphocytes are considered in the proposed scheme. [18] proposed three contrast enhancement techniques namely bright stretching, dark stretching and partial contrast for the color images (using RGB components) and concludes partial contrast as the best technique. For improving the accuracy of segmentation a combination of image segmentation (HSI based) and contrast enhancement techniques is proposed in [19]. Combination of partial contrast and bright stretching achieves fully segmented blast whereas partial contrast and dark stretching obtains fully segmented nucleus. [20] also does color segmentation of WBC’s to obtain acute leukemia images. It combines color segmentation (HSI based) and linear contrast technique. Segmentation of WBC is done based on H and S components of the image. Finally median filter and region growing are applied on the resultant images thus improving the segmentation accuracy. Comparison of segmentation based on HSI and RGB color space is done in [21] and concluded HSI as a better approach for segmenting color images. Subtyping of leukemia is done in [22] using novel compressive sensing based approach. Leave-one-out is used for cross validation for evaluating the accuracy of detection. 100% accuracy is achieved for ALL group but due to the difference in sample size misclassification occurs in AML group. [23] illustrates that SVM Model gives highest accuracy for the classification of leukemia cancer. Diagnosis of leukemia from the bone marrow blood cell images is done in [24]. Detection rate is below the usable value for the promyelocyte and is identified as the further work to be done.

METHODOLOGY

Algorithm:
1) Identify the counting area (area where cells do not interfere with each other and are properly separated).
2) Perform color normalization on the image.
3) Separate white blood cells from the image by performing image segmentation.
4) Specify the region of interest.
5) Extract the region of interest.
6) Apply edge erosion filter to deal with the fractured edges.
7) Feature vector analysis is performed which includes extraction and selection of features of the cell images including size of cell, circularity of cells, count of cells, area of cytoplasm and nucleus, nucleus/cytoplasm ratio etc.
8) Initial classification into lymphoid or myeloid series is done by the shape of the nucleus.
9) The extracted features are then compared with those stored in the database and classification is done into AML or ALL categories.
10) K nearest neighbor classifier is used for the classification.

RESULTS & DISCUSSION

Table I. Percentage of Cell Segmentation Accuracies for Different Methods

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Technique used</th>
<th>Result (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Watershed Algorithm</td>
<td>99.85%</td>
</tr>
<tr>
<td></td>
<td>Optimal Thresholding</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[Error! Bookmark not defined.]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Scale space filtering</td>
<td>98.9%</td>
</tr>
<tr>
<td></td>
<td>Watershed clustering [25]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Morphological operations</td>
<td>95.09%</td>
</tr>
<tr>
<td></td>
<td>Watershed transform [26]</td>
<td></td>
</tr>
</tbody>
</table>

Table II. Results of segmentation of lymphocytes using different techniques

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Technique used</th>
<th>Result of segmentation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Nucleus (%)</td>
</tr>
<tr>
<td>1</td>
<td>Morphological analysis using snake algorithm</td>
<td>92%</td>
</tr>
<tr>
<td></td>
<td>Pixel intensity thresholding [27]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Watershed algorithm</td>
<td>99.92%</td>
</tr>
<tr>
<td></td>
<td>Optimal Thresholding [28]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Image arithmetic</td>
<td>98.81%</td>
</tr>
<tr>
<td></td>
<td>Automatic thresholding using contrast stretching [29]</td>
<td></td>
</tr>
</tbody>
</table>

CONCLUSION AND FUTURE WORK

Color feature must be added to improve the classification rate of the leukemia images. It is difficult to segment the images in RGB color space. So, it is better to convert the images into L*a*b* or HSI color space. If in addition to cytoplasm, nucleus and nucleolus; vacuoles are also considered then a specific type of AML could be determined though the complexity will increase.

REFERENCES


