

A Hybrid Methodology for Automation the Diagnosis of Leukemia Based on Quantitative and Morphological Feature Analysis

Hussam N. Fakhouri¹ & Saleh H. Al-Sharaeh¹

¹ King Abdullah II School of Information Technology, The University of Jordan, Amman, Jordan

Correspondence: Hussam N. Fakhouri, King Abdullah II School of Information Technology, The University of Jordan, Amman, Jordan. E-mail: h.fakhouri@ju.edu.jo

Received: January 27, 2017

Accepted: January 8, 2018

Online Published: February 27, 2018

doi:10.5539/mas.v12n3p56

URL: <https://doi.org/10.5539/mas.v12n3p56>

Abstract

Recent year's witnessed a huge revolution for developing an automated diagnosis for different disease such as cancer using medical image processing. Many researches have been dedicated to achieve this goal. Analyzing medical microscopic histology images provide us with large information about the status of patient and the progress of diseases, help to determine if the tissue have any pathological changes. Automation of the diagnosis of these images will lead to better, faster and enhanced diagnosis for different hematological and histological tissue images such as cancer. This paper propose an automated methodology for analyzing cancer histology and hematology microscopic images to detect leukemia using image processing by combining two diagnosis procedures initial and advance; the initial diagnosis depend on the percentage of the white blood cells in microscopic images affected by leukemia as indicator for the existence of leukemia in the blood smear sample. Whereas, the advance diagnosis classifying the leukemia according into different types using feature bag classifier. The experimental results showed that the proposed methodology initial diagnosis is able to detect leukemia images and differentiate it from samples that do not have leukemia. While, advance diagnosis it is able to detect and classify most leukemia types and differentiate between acute and chronic, but in some cases in the chronic leukemia where the percent of blast cells and shape are similar; it gave a diagnosis of the type of leukemia to the most similar type.

Keywords: leukemia, digital image processing, acute leukemia diagnosis, chronic leukemia diagnosis

1. Introduction

Many novel researches and efforts have been devoted for developing automated systems for detecting and analyzing of microscopic histology images. However, diagnosis traditionally depends on the qualified eye of a pathologist to make judgment from a qualitative perspective. computer automation and diagnosis is now possible with digital image processing (Gonzalez and Woods, 2002). Digital image processing means the process of images by digital computer. This includes detection, sensing, analysis of digital images (Jensen, 1996) (Alhadidi et al., 2006). Which contains limited number of elements named as pixels; these elements have values that represent image (Gonzalez and Woods, 2002). It begins with image acquisition and image enhancement, for the reason that irrelevant details shall be shown and potential highlight parts of details or interest features must be displayed. A lot of digital image steps then applied to digital image such as object classification, segmentation, morphological processing ...etc. (Jensen, 1996) (Ablameyko and Nedzved, 2005) (Alhadidi et al., 2007).

Developing an efficient and reliable algorithmic and automated methods, will give a powerful tool which aids in the collection of data, assists researchers in further studies and researches (Hudaib et al., 2017). It ultimately helps with the diagnosis of abnormal tissue changes such as leukemia visual examinations of blood samples are often slow and are also limited by subjective interpretations and less accurate diagnosis. (Brothwell et al., 2003) (Long, et al., 2010) (Adwan et al., 2013) (Alhadidi et al., 2008).

Leukemia is the general term for some different types of blood cancer. The term Leukemia comes from the Greek - leukos which means "white" and aima which means "blood". It refers to the cancer of blood or bone marrow (a place where blood cells are produced). Blood comprises of many components such as red blood cells, white blood cells and platelets. White blood cells are produced in human body to provide immunity. In the case of leukemia, the white blood cells produced in the bone marrow are immature. In other words they are incapable of providing immunity to the body. These immature cells are termed as 'blasts'.

The DNA of immature cells becomes damaged leading to uncontrollable proliferation. Cells produced in the bone marrow are regularly replaced by new cells. In the case of leukemia, the lymphoblasts formed do not die and end up accumulating. Due to this, there is not enough room for the normal healthy cells to occupy. There are four main types of leukemia called: Acute lymphoblastic (lymphocytic) leukemia (ALL), Acute myeloid (myelogenous) leukemia (AML), Chronic lymphocytic leukemia (CLL), Chronic myeloid (myelogenous), leukemia (CML).

Leukemia is further sub-divided based on the marrow that is affected. each type of leukemia:

Acute: Acute leukemia usually develops quickly. The number of leukemia cells increases rapidly, and these abnormal cells don't do the work of normal white blood cells. A bone marrow test may show a high level of leukemia cells and low levels of normal blood cells. People with acute leukemia may feel very tired, bruise easily, and get infections often.

Chronic: Chronic leukemia usually develops slowly. The leukemia cells work almost as well as normal white blood cells. People may not feel sick at first, and the first sign of illness may be abnormal results on a routine blood test. For example, a blood test may show a high level of leukemia cells. If not treated, the leukemia cells may later crowd out normal blood cells.

Automation the analysis of histopathology images have been a very important research subject with the revolution of computer and image processing development. The new tools in image processing have allowed the investigators and scientist to develop techniques that support pathologists in disease diagnosis and classification (Ilyich et al, 2002) (Leong et al., 2003). An overview for the automation of histological image processing is shown in Figure 1.

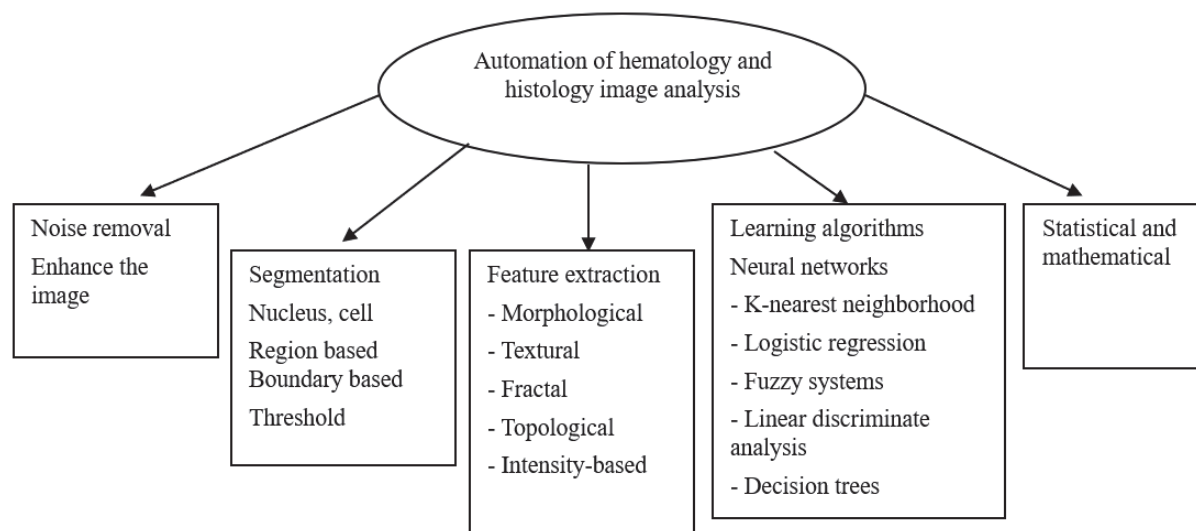


Figure 1. Overview of automation of hematology and histology image analysis

The objective of this paper is to develop automated methodology for the diagnosis of different leukemia types using image processing. It introduces an automated method for diagnosis of blood smear microscopic images for different types of leukemia and this method is tested on samples consisted of 100 microscopic images, the proposed method's starts with the laboratory preparation and ends with the diagnosis of blood smear image whether it contains leukemia or not, after the laboratory preparation is done, the image is preprocessed to enhance them and remove the noise then the color based segmentation applied, after that a set of mathematical calculations are applied to the image to calculate the percent of leukemia cells in the image and finally the diagnosis of the image is done based on the previous phases. It also introduces a new mathematical calculation method for leukemia microscopic images diagnosis that depends on comparing the percentage of leukemia cells.

In order to achieve the objectives of this research, the following steps and procedures were performed; First step was to investigate and study previous methods and researches done in the field of analyzing microscopic histology images and to determine the field that needs more researches and investigations. Second step was data collection from previous patient record of 100 patient, a sample a 100 digital images for both non leukemia and

leukemia cases have been studied. Thirdly A new algorithm was developed to process and automatically diagnose microscopic image of different leukemia types. Fourthly An implementation of the algorithm was done using matlab image processing toolbox. Finally Testing the implemented system with all samples that was selected in step one, storing the results and analyzing them.

The rest of this paper is organized as follows In Section2 we present related work. Section 3 describe the proposed methodology. Section 4 provides some experimental results obtained by the implementing the proposed methodology Finally, Section 5 contain the conclusion and future work

2. Related work

S. Jagadeesh et al. (2013) proposed an image processing based approach to cancer cell prediction in blood samples there proposed solution include the segmentation of the bone marrow aspirate by applying the watershed transformation, selection of individual cells, and feature generation on the basis of texture, statistical and geometrical analysis of the cells. H. B. Kekre et al, 2013 proposed a vector quantization technique for segmentation of blast in acute leukemia images. This method is applied on 115 microscopic images and succeeds with specificity of 90% and sensitivity of 60% to detect abnormal white blood cells. Salim Arslan, 2014, proposed a model color and shape characteristics of white blood cells by defining two transformations and introduce an efficient use of these transformations in a marker-controlled watershed algorithm. Particularly, these domain specific characteristics are used to identify markers and define the marking function of the watershed algorithm as well as to eliminate false white blood cells in a post processing step Subrajeet Mohapatra et, al. 2013 proposing a quantitative microscopic approach toward the discrimination of lymphoblasts (malignant) from lymphocytes (normal) in stained blood smear and bone marrow samples and to assist in the development of a computer-aided screening of ALL. Automated recognition of lymphoblasts is accomplished using image segmentation, feature extraction, and classification over light microscopic images of stained blood films. Accurate and authentic diagnosis of ALL is obtained with the use of improved segmentation methodology, prominent features, and an ensemble classifier, facilitating rapid screening of patients. Experimental results are obtained and compared over the available image data set.

Den et al, (1999) proposed a method to localize WBC by using a simple thresholding approach. Canny edge detector was used followed by a gradient vector flow (GVF) active contour to detect the nucleus and then Zak threshold was used to define the cytoplasm component Liao.

Foran et al. (2013) have reported a method to discriminate among lymphoma and leukemia with a classification accuracy of around 83%. they have developed a distributed, clinical decision support prototype for distinguishing among hematologic malignancies. The system consists of two major components, a distributed telemicroscopy system and an intelligent image repository. The hybrid system enables individuals located at disparate clinical and research sites to engage in interactive consultation and to obtain computer-assisted decision support. The method is reported to have successfully worked on 19 lymphoproliferative cases, which is a very small data set to evaluate the performance of the system. Further, the presented method is yet to be validated on ALL cases. Markiewicz et al. (2005) presented a system for automatic recognition of the leukemia blast cells on the basis of the image of the bone marrow aspirate. The recognizing system uses support vector machine (SVM) as the classifier and exploits the features of the image of the blood cells related to the texture, geometry and histograms. Belsare et, al 2012 reviews computer assisted histopathology image analysis for cancer detection and classification. reviews and summarize the applications of digital image processing techniques for histology image analysis mainly to cover segmentation and disease classification methods. He studied different steps to automatically analyze histopathological images for objective diagnosis which assists pathologist in diagnosis and lessen their time for reviewing large number of tissue slide per day. He developed algorithms for automated analysis and evaluation of histology images assists the pathologists in disease diagnosis and also reduces human error.

Mohapatra, et al (2014) improved the all diagnostic accuracy by analyzing morphological and textural features from the blood image using image processing. by proposing a quantitative microscopic approach toward the discrimination of lymphoblasts (malignant) from lymphocytes (normal) in stained blood smear and bone marrow samples and to assist in the development of a computer-aided screening. Madhloom et. al (2011) presented a new method that integrates color features with the morphological reconstruction to localize and isolate lymphoblast cells from a microscope image that contains many cells. Described a method for lymphoblast cells localization and segmentation. Presented algorithm can also be used to detect normal WBC like lymphocytes and monocytes, so it can be used for differential blood count systems. From an end-user point of view, this work can facilitate the laboratory work by reducing the time and cost. Huang et al. 2014 focuses on investigating the potential

correlations between ALIP and AML relapse for early prediction. proposed an ALIP detection method using biopsy image processing in order to investigate the relevance with AML relapse. Thirty-seven patients with AML are examined. The results shows ALIP can be efficiently detected by our proposed method. This research reveals the strong correlations of AML relapse with ALIP. Sadeghian et al. 2009 segment the WBC to its two dominant elements: nucleus and cytoplasm. The segmentation is conducted using a proposed segmentation framework that consists of an integration of several digital image processing algorithms. Twenty microscopic blood images were tested, and the proposed framework managed to obtain 92% accuracy for nucleus segmentation and 78% for cytoplasm segmentation. The results indicate that the proposed framework is able to extract the nucleus and cytoplasm region in a WBC image sample. Demonstrated a proposed framework for segmenting white blood cells using integration of concepts in digital image processing.

Joshi et al (2013) proposed automatic Otsu's threshold blood cell segmentation method along with image enhancement and arithmetic for WBC segmentation. kNN classifier has been utilized to classify blast cells from normal lymphocyte cells. The system is applied for 108 images available in public image dataset for the study of leukemia. This method gives 93% accuracy. Putzu et.al 2013 presents a complete and fully automatic method for WBCs identification and classification from microscopic images. The proposed method firstly individuates WBCs from which, subsequently, are extracted morphological features necessary for the final stage of classification.

Vaghela et, al., (2015) discusses about methods for detection of leukemia. Various image processing techniques are used for identification of red blood cell and immature white cells.

proposed method: shape based features finding is more accurate than other methods for counting leukemic cells and it also gives highest accuracy 97.8 %. To detect different types of geometrical shape of cells like basophils, eosinophil, lymphocytes, monocytes etc. shape based features are used and according to count of immature cells, disease can be diagnosed. Further Optimization can be to enhance image processing as will as to apply processing in the cloud (Al-Sayyed et al,2017)

3. Proposed Methodology

In this paper, an automated methodology for the diagnosis of leukemia is presented. The proposed methodology is divided into four main phases the preprocessing for the image to enhance it and the second is the processing of the image with different image processing technique in order to segment, detect and diagnose different leukemia types in the image. and the second phase is the processing of the microscopic image, third phase is the initial diagnosis which is based on leukemia cells percentage presented in the image and provide initial classification and the last phase is the advance diagnosis which provide a classification of the type of the leukemia depending on the cell features described in Figure 2.

The proposed methodology is discussed by studying different cases of leukemia images and normal tissue images; a sample of one hundred images has been experimented. 50 sample where taken for normal cases and the other 50 images for cases which was previously diagnosed that have different kind of leukemia.

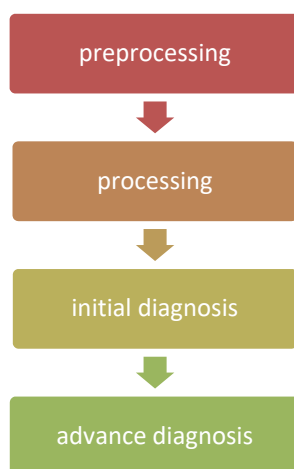


Figure 2. Proposed methodology main phases

3.1 Preprocessing

The preprocessing is an essential phase in microscopic histology image analysis because of the nature and circumstances of the slides preparation, staining and image shooting, which affect the quality of the image as seen in Figure 3 and Figure 4 which shows poor quality images for both normal and leukemia images respectively. The algorithm flow chart that explains preprocess phases is shown in Figure 5. The first step is noise removal by applying fuzzy filter for the colored image to remove noise. The second step is image sharpening and the last step is enhancing the contrast.

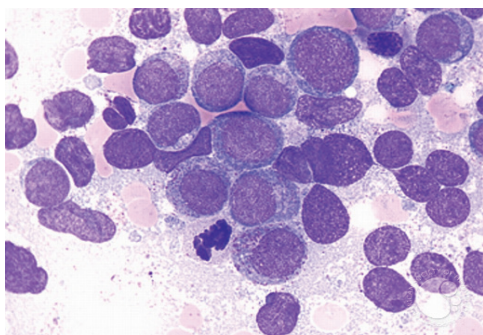


Figure 3. Poor quality microscopic image for acute myeloid leukemia without Maturation

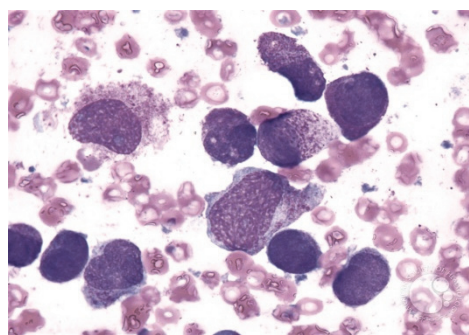


Figure 4. Poor quality microscopic image for hypergranulated acute promyelocytic leukemia

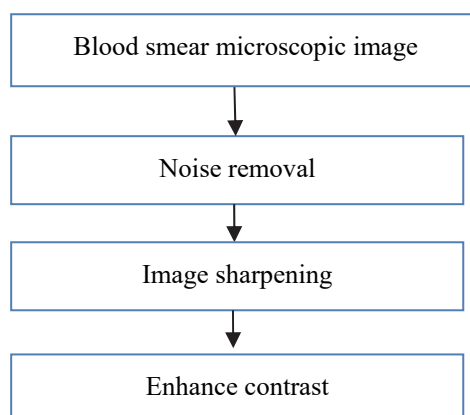


Figure 5. Preprocess of the image from preparation phase

3.1.1 Noise Removal

Noise in microscopic images can occur during image transmission, capture or preparation of the slide images. That's why we need to remove the noise and enhance the image, we have applied fuzzy filter to remove the noise, applying fuzzy filters for noise removal gives great results in many digital image analyses such as microscopic image which overcome some malfunctions of classical filters. In the situation of the microscopic histology images fuzzy filter is very helpful for removing noise as shown in Figure 6. Fuzzy filter that takes the nearest data to remove the noise, it also performs edge preservation.

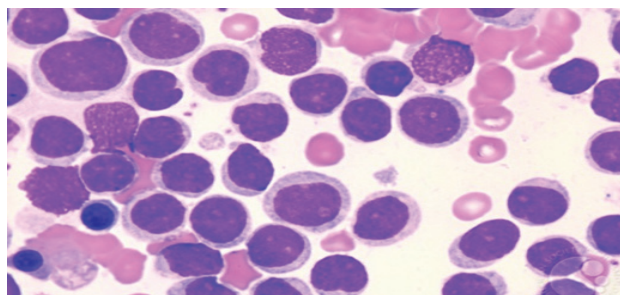


Figure 6. Applying fuzzy filter

3.1.2 Image Sharpening

Because of the fact that all camera's produce soft images as though digital image for the microscopic tissue images that are soft, so there is a need for sharpening these image and to do this we have applied the shock filter as shown in Figure 7, we have chosen shock filter because it is based on the idea to apply it locally either dilation or erosion process, depending on whether the pixel belongs to the influence zone of a maximum or a minimum and this is the case in the leukemia images where there are variation in the image structure. Applying this filter have enhanced the image and produced a sharp discontinuity called shock at the borderline between the objects and the background (Gonzalez and Woods, 2002).

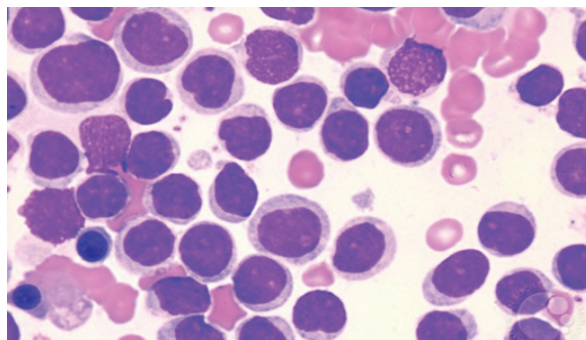


Figure 7. Image after applying shock filter

3.1.3 Enhance Contrast

In light microscopy, the specimen quality does not always lend itself to easy observation and image recording that's why we need to enhance the contrast of the image as the studied images are taken by light microscope. Poor light absorption by the specimen results in extremely small variations in the intensity distribution difference between the specimen and the background. So we need to adjust image intensities to enhance contrast of the images and this is done by applying histogram equalization to the image which computes the probabilities per pixel and further spreads them over the whole band of [0,255] to obtain enhanced contrast, sample images that shows the results after enhancing the contrast are shown in Figure 8.

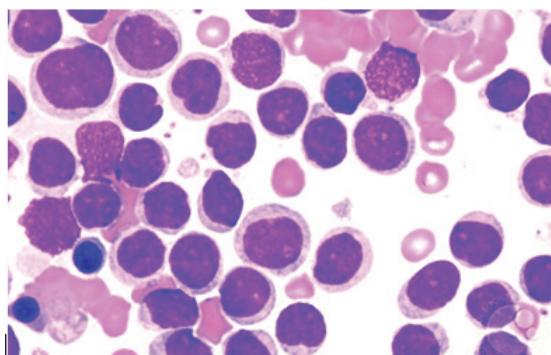


Figure 8. Image with enhanced contrast

3.2 Processing Phase

This phase consist of several steps as shown in Figure 9. the main process in this phase relay on color based segmentation for the purple color which is the stain color for the blast cells that we concern about in leukemia diagnosis

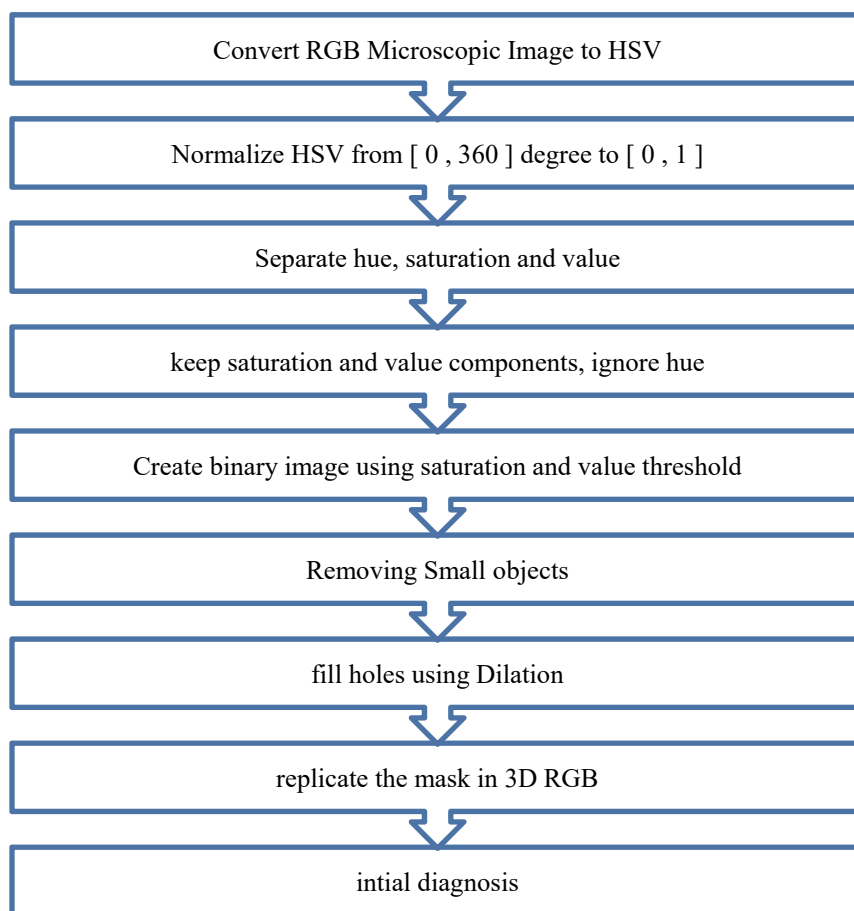


Figure 9. Processing phase main steps

3.2.1 Convert RGB Microscopic Image to HSV

This will return a 3D matrix that has the hue, saturation and value as 2D slices in a 3D matrix. We aim to described colors by their dominant color, followed by attributes such as how washed out or how pure the color is, and how bright or dark the color is. The dominant color is represented by the Hue, the appearance of how washed out or how pure the color is is represented by the Saturation and the intensity of the color is represented by the Value

3.2.2 Normalize HSV from [0 , 360] degree to [0 , 1]

After converting the image into the HSV colour space. It will be converted to double precision and then normalize each component to [0,1]

3.2.3 Separate Hue, Saturation and Value

Each of the hue, saturation and value are separated, Figure 10. Illustrate this step where each component of the image are shown in a single columns. The first image represents the hue, second image the saturation and finally the last image being the value.

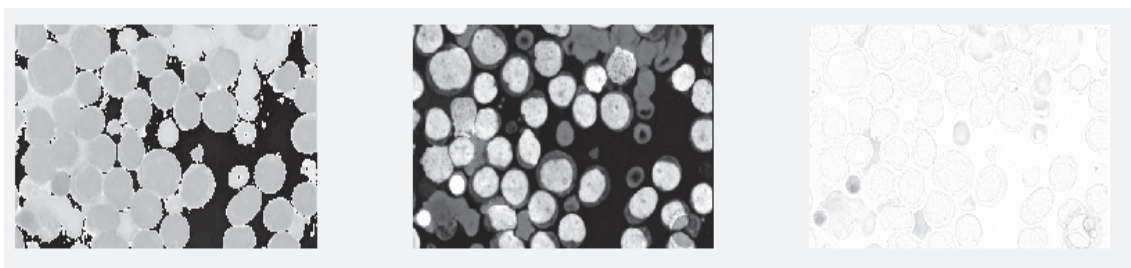


Figure 10. HSV Separate representation for microscopic leukemia image

Since the dominant color in leukemia staining is purple, whether it's a light shade or a dark shade of the colour, so the hue won't help us here. If you look at a HSV color wheel Figure 11.

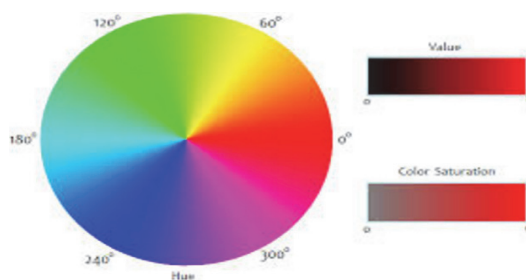


Figure 11. HSV Color wheel

We have Normalize the wheel so that it falls between $[0,1]$ instead of 0 to 360 degrees. The hue is actually represented as degrees due to the nature of the color space, but MATLAB normalizes this to $[0,1]$. The color purple falls within a hue of $[0.6, 0.8]$, which corresponds image leukemia image shown in previous step the pixels of leukemia cells in the image, they fluctuate between this range. Therefore, we ignore the hue.

What will certainly care about when analyzing leukemia images is the saturation and value components. The purple pixels have a higher saturation than the rest of the background, because the deep purple has a much more pure version of purple than the rest of the background. For moreover the brightness of the dark purple is darker than the background. these two points are used as an exploit to segment out the purple cells of leukemia in the image. The next step then will be to threshold using the saturation and value planes so that any values that are within a certain range are kept while those that are outside are ignored.

3.2.4 Create Binary Image Using Saturation and Value Threshold

As indicating by Ander Biguri 2015 the purple regions have a saturation value between 0.6 and 0.9, while the value component has values between 0.4 and 0.65. Two binary threshold masks are created if the pixel located in the color range. Logical OR are then used to masks them together. The resulted image have two main parts one is above and the other is below the threshold value, these two values will define the foreground and background of the image and each will represent a studied object according to what the image represents.

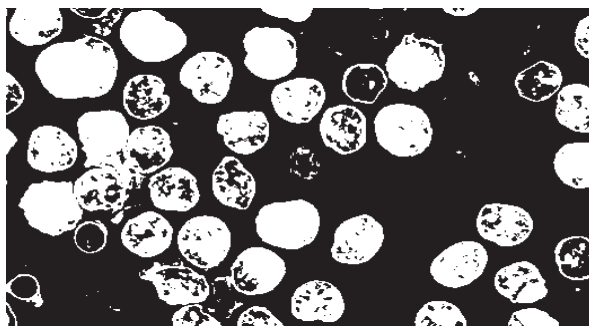


Figure 12. Binary image after applying threshold for microscopic leukemia image

3.2.5 Removing Small Objects

This step aim's to remove the small objects To do this we used a an opening filter of a small window so that we don't affect the pixels that we want as much. A morphological opening removes isolated pixels that appear in the image, then using structuring element any pixel regions that are as small as the shape that is contained within the structuring element get removed. Because we want to preserve the shape of the leukemia cells and remove any other shape, we used a 3×3 disk structuring element to clean these pixels up without affecting the cells of leukemia. See Figure 13.

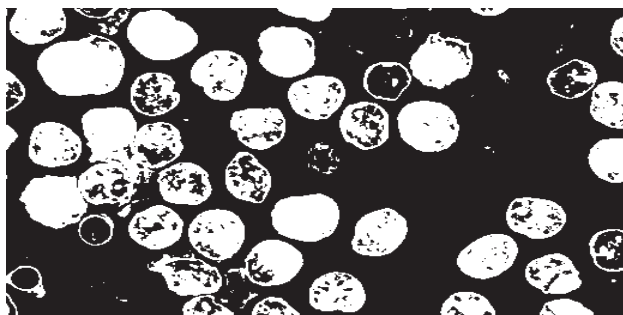


Figure 13. Binary image after removing small objects for microscopic leukemia image

3.2.6 Filling Holes Using Dilation

The Dilation is used to fill the holes because dilation operation uses a structuring element for probing and expanding the shapes contained in the input image. It aims to gradually enlarge the boundaries of regions of foreground pixels to fill holes in leukemia cells. Thus areas of foreground pixels grow in size while holes within those regions become smaller as shown in Figure 14.

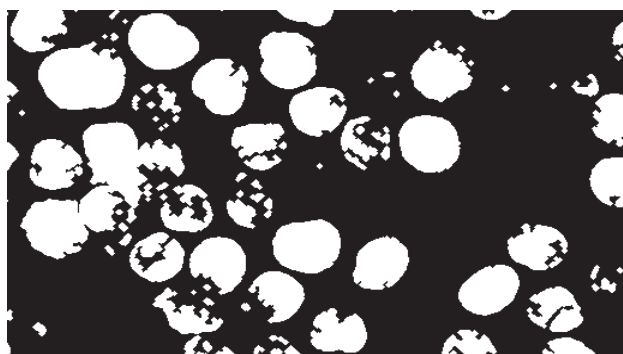


Figure 14. Binary image after filling holes using dilation for microscopic leukemia image

3.2.7 Replicate the Mask in 3D

This step is done so that to mask out the unwanted RGB pixels and only keep the ones we are looking for which is the leukemia cells as shown in Figure 15

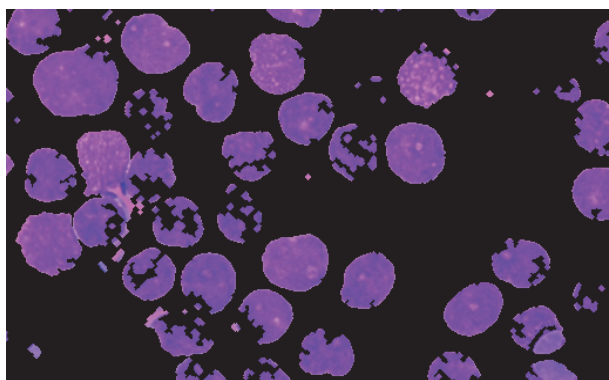


Figure 15. Binary image replicate the mask in 3D for microscopic leukemia image

3.3 Initial Leukemia Diagnosis

The initial leukemia diagnosis relay on the percentage of white blasts percent in the image , To calculate the percentage of leukemia cells we use the binary image because the microscopic image structure that contains many values and this is related to the stain procedure and its values differ from one location to another. Binary image overcomes the limitation because it has only two color values 0 for black and 1 for white. Figure16

The idea of the calculation is based on using the structure of the whole microscopic image and the percentage of the black pixels to the white pixels as an indicator for the presence of leukemia cells. The blood that have leukemia will have a changes in the number of white blood cells as seen in the microscopic images, and it can be noticed that the number stained cells in purple colors in the cases of leukemia is vary according to the type of leukemia, but the major fixed point that any increase in the number of cells indicating the presence of leukemia. thus the white area is larger in case of leukemia than the normal blood cells and as a result the number of pixels that contain the value of one will be larger in leukemia images and smaller in the normal blood smear images,

This step mainly consist of counting the white pixels to be used later in the next step for tissue classification as shown in Figure 16.

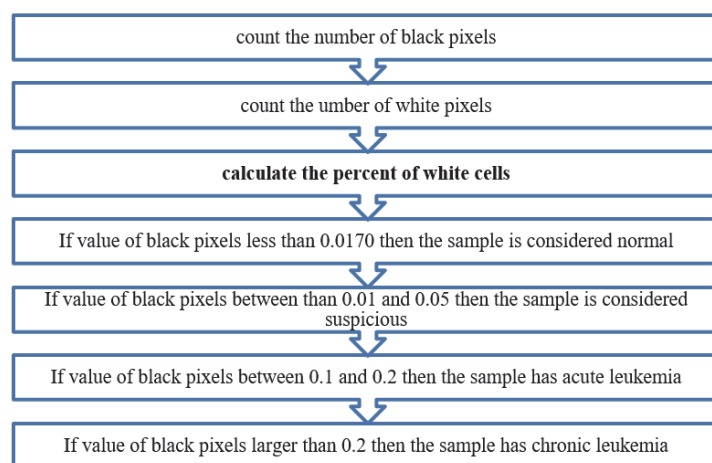


Figure 16. Initial leukemia diagnosis main steps

3.3.1 Count Number of White Background Pixels

The binary image gave us clear values of white pixels that have one value and this makes it easier to count these pixels, the number of white pixels represents the background in image.

3.3.2 Count Number of Black Pixels

The binary image gave us clear values of dark pixels labeled with zero values, and this makes it easier to count these pixels, the number of dark pixels represents the objects in the image.

3.3.3 Initial Diagnosis

The classification is based on the percent of leukemia cells in the microscopic image. The percent is calculated by calculating the percentage of white pixels compared to the percentage of black pixels as shown in equation 1 described below:

$$\text{Percentage of black pixels} = \frac{\text{Black Pixels}}{\text{Total number of pixels}} * 100\% \quad \text{Equation 1}$$

According to the experimental results of normal blood cells analysis and leukemia cells analysis the following categories has been stated for initial diagnosis

If value of black pixels less than 0.0170 then the sample is considered normal

If value of black pixels between than 0.01 and 0.05 then the sample is considered suspicious

If value of black pixels between 0.1 and 0.2 then the sample has acute leukemia

If value of black pixels larger than 0.2 then the sample has chronic leukemia

After the initial diagnosis and in order to determine the specific type of leukemia further analysis is done in the next step

3.4 Advance Leukemia Diagnosis

After determining that the blood smear image has a leukemia and according to the previous steps and after the initial diagnosis performed using the cell percentage in the image the next step is to classify the type of leukemia

according to the features extracted.

Since the leukemia cells differ in shape from one type to another and relay of the leukemia cell shape then, The method that we have used for leukemia classification is Image Category Classification Using Bag of Features; This technique is also often referred to as bag of words. Visual image categorization is a process of assigning a category label to an image under test. The Categories in our experiments contained images representing different types of leukemia.

The first step in the classification phase is to Load Image Sets , the second step is to Construct an array of image sets based on the following leukemia types, to do this we have Used image Set class to help to manage the image data. Since image Set operates on image file locations, and therefore does not load all the images into memory, it is safe to use on large image collections

Each element of the image Sets variable now contains images associated with the particular leukemia type category. E.g. the number of images for leukemia types as well as labels shown in Figure 17, the labels was derived from directory names used to construct the image sets

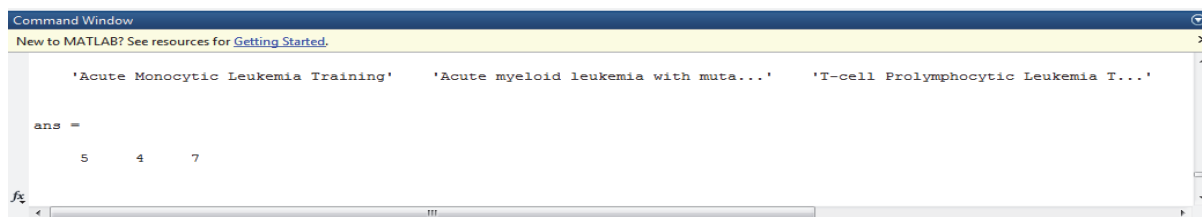


Figure 17. Load leukemia types

The second step is Training and Validation of the Image Sets, Since image Sets above contains an unequal number of images per category, then the first thing here to do was to adjust it, so that the number of images in the training set is balanced. the next step is to Separate the sets into training and validation data. 30% of images from each set for the training data and the remainder, 70%, for the validation data. The split is done Randomly to avoid biasing the results. The result of this step is two arrays of image Set objects ready for training and validation.

The third step is to Create a Visual Vocabulary and Train an Image Category Classifier, since Bag of words is a technique adapted to computer vision from the world of natural language processing. And images do not actually contain discrete words, we first construct a "vocabulary" of SURF features representative of each image category. This step was done using to bag Of Features function, which: extracts SURF features from all images in all leukemia image categories, then constructs the visual vocabulary by reducing the number of features through quantization of feature space using K-means clustering, Figure 18.

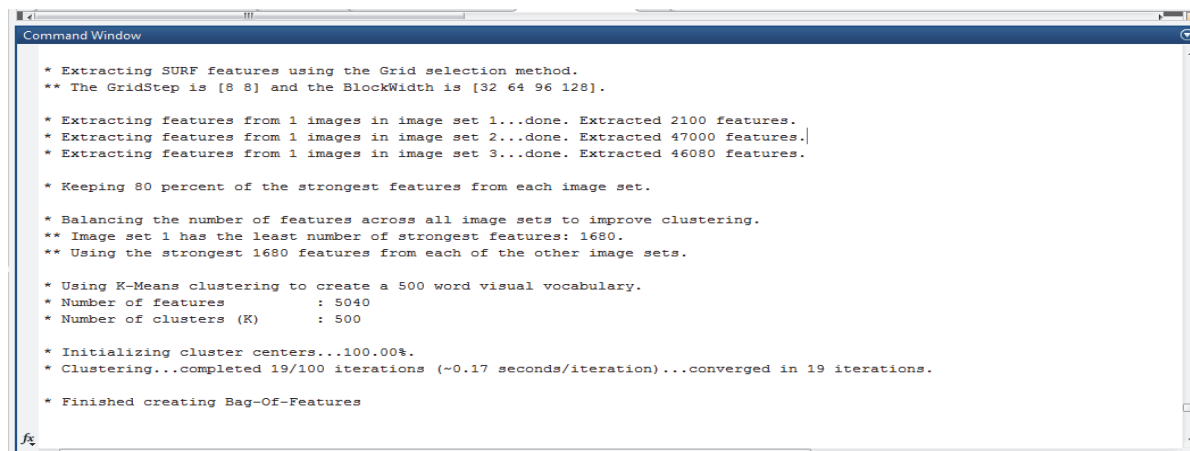


Figure 18. Training and validation of the image sets using bag of features

In the next step a histogram is produced using bag Of Features object that becomes a new and reduced representation of an image. The histogram forms the basis for training a classifier and for the actual image

classification. In essence, it encodes an image into a feature vector. As shown on Figure 19.

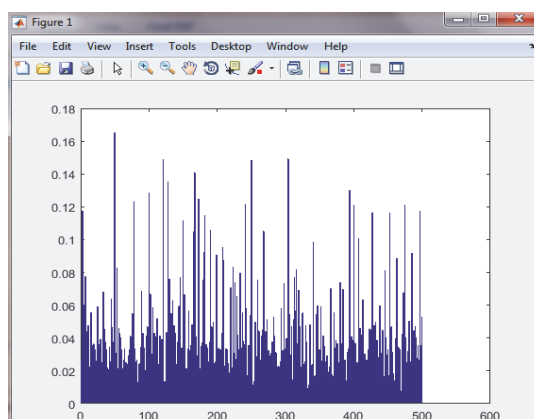


Figure 19. Histogram is produced using bag of Features for microscopic extracted feature

Encoded training images from each category are fed into a classifier training process invoked by the train Image Category Classifier function. this function relies on the multiclass linear SVM classifier from the Statistics and Machine Learning Toolbox™ in MATLAB . The next step is to see the evaluation of the classification according the 30% value of divided leukemia images described above see Figure 20.

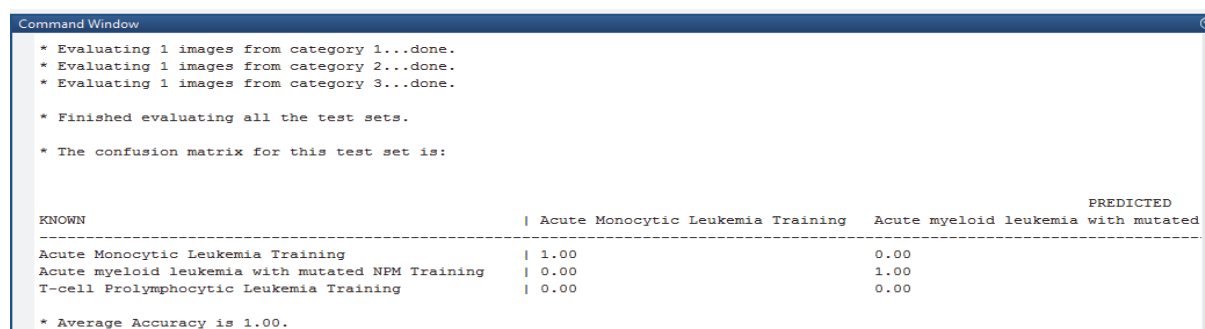


Figure 20. Classification according the 30% value of divided leukemia images

And finally to diagnose the tested image to see the diagnosis results and this will give us the type of leukemia by comparing the features of the image with the trained one, as shown in Figure 21

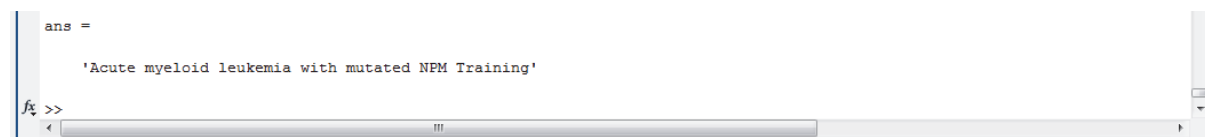


Figure 21. Final classification of leukemia types by comparing the features of the images

4. Experimental Results and Analysis

4.1 Data Description

In order to evaluate the proposed method A dataset of one hundred images for both normal tissue and abnormal tissue have been taken from a well known and specialized pathology and histology university libraries and websites for pathology images e.i. library.med.utah.edu, imagebank.hematology.org, cord.edu, pathpedia, atlases.muni.cz, pathologyoutlines, pathologystudent, eclinpath have been tested by our proposed method. All studied images as documented in the references that were taken from have been clinically tested and examined using blood and bone marrow examination. And correctly diagnosed and classified. The dataset that were chosen to cover cases of tissues and blood film for both male and female and for different age periods.

Fifty images of them were taken for normal cases and the other fifty images contain different leukemia types including the following types: Flower leukemia cells, Chronic Lymphocytic Leukemia, Acute Monocytic

Leukemia, T-prolymphocytic leukemia, Acute myeloid leukemia with mutated NPM1, T-cell prolymphocytic leukemia type 4, T-cell prolymphocytic leukemia – type 3, T-cell Prolymphocytic Leukemia type 2, T-cell Prolymphocytic Leukemia type 1, Hairy cell leukemia, Chronic Myelomonocytic Leukemia, Acute erythroid leukemia, Aggressive NK Cell Leukemia, Acute undifferentiated leukemia-Peripheral Blood, sample image are for Erythrocyte hemophagocytosis Leukemia microscopic image are shown in Figure 22.

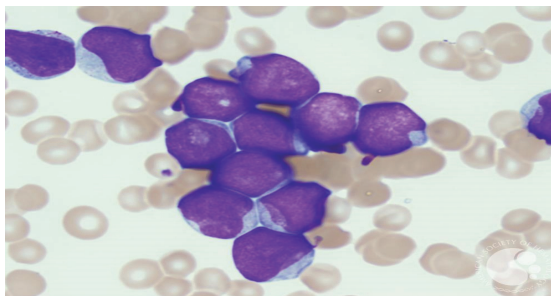


Figure 22. Acute myeloid leukemia with mutated NPM

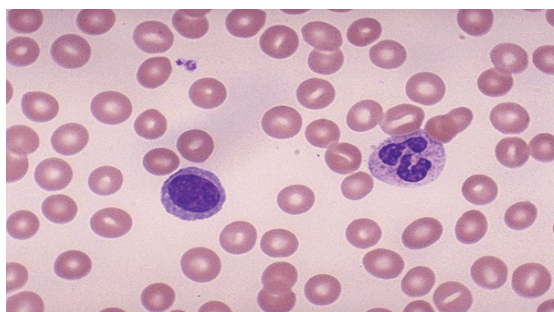


Figure 23. Normal blood sample

4.2 Initial Diagnosis Experiment

The initial diagnosis experiment concern about testing the methodology to be able to distinguish between images that has leukemia and those that don't have leukemia and the ability and accuracy of the system to give the initial diagnosis. The experiment have been performed with different samples and plotting the experimental results for the percentage of leukemia cells in comparing with samples that don't have leukemia have shown that there is a gap between the percentage of white pixels taken from samples that has leukemia and those that don't have leukemia Figure 24 shows the plotted results of 20 sample of both cases.

The system was able to correctly diagnose all tested cases that were correctly stained in the laboratory and taken by digital microscope with both magnification 40x and 100x. after the initial classification and diagnosis the tested images have been tested for the advance diagnosis proposed by this methodology and according to the feature bag.

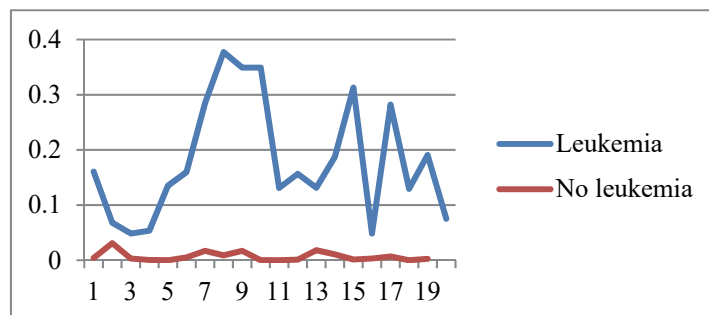


Figure 24. Percentage of black color in both leukemia and non leukemia for the tested 20 images

According to the experimental results of normal blood cells analysis and leukemia cells analysis the following

categories has been stated for initial diagnosis

If value of black pixels less than 0.03 then the sample is considered normal

If value of black pixels between than 0.03 and 0.05 then the sample is considered suspicious

If value of black pixels between 0.1 and 0.2 then the sample has acute leukemia

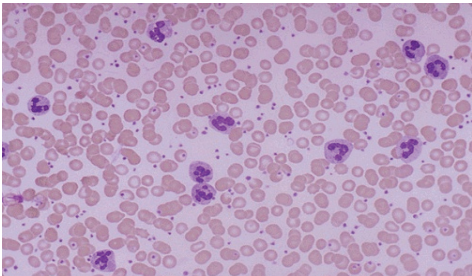

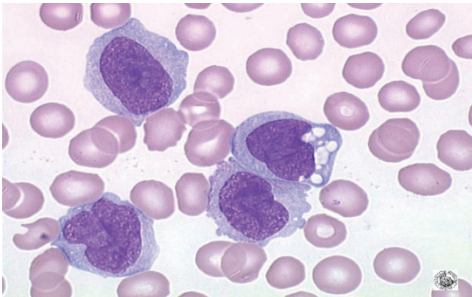
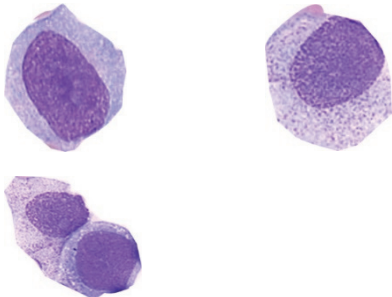
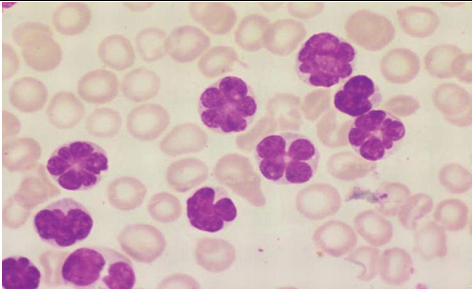

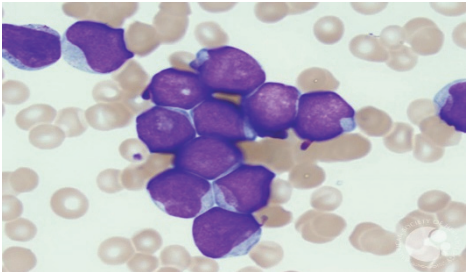
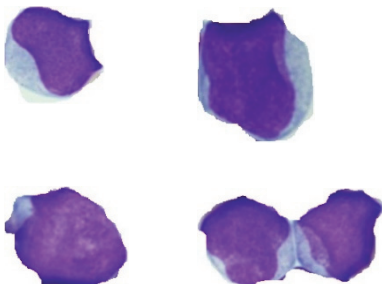
If value of black pixels larger than 0.2 then the sample has chronic leukemia

After the initial diagnosis and in order to determine the specific type of leukemia further analysis is done in the next step.

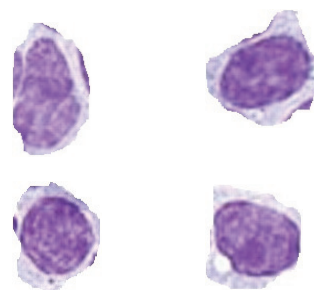
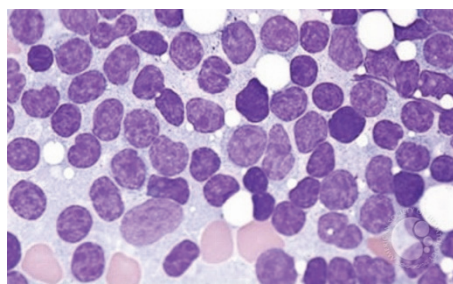
4.3 Advance Diagnosis Experiment

In order to test the ability of the system to classify leukemia a set of trained features has been prepared for the studied leukemia types table 1 illustrate a sample of the trained data and the images take from for the leukemia types.

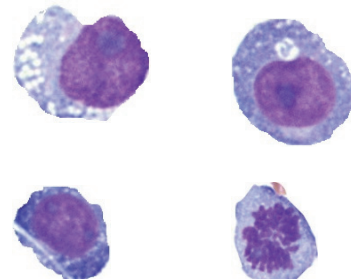
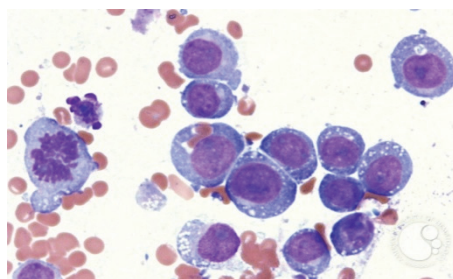
Table 1. types of leukemia and normal studied and the trained features extracted

Type	Original image	Trained feature
Normal blood smear		
Acute Monocytic Leukemia		
Flower cell Leukemia		
Acute myeloid leukemia with mutated NPM		

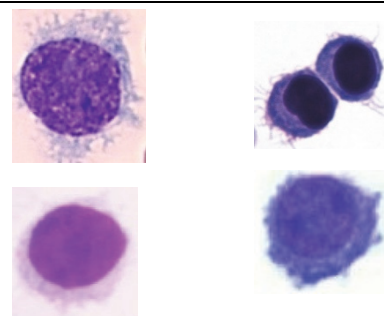
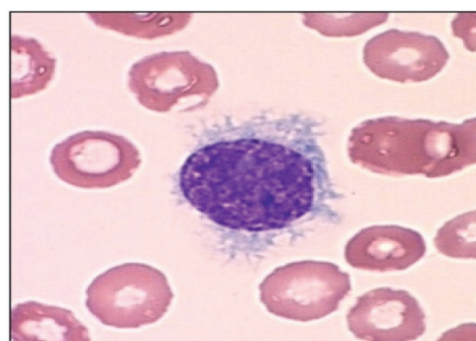
**Chronic
Lymphocytic
Leukemia Training**



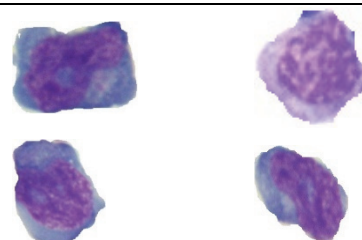
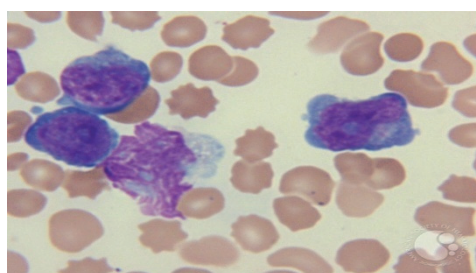
**Erythrocyte
hemophagocytosis
Training**



hairy cell leukemia



**T-cell
Prolymphocytic
Leukemia**



The implemented methodology was able give indication for the diagnosis of leukemia in images. The experimental results showed that the proposed methodology was also able to diagnose microscopic images with low resolution as shown in Figure 25. Leukemia.

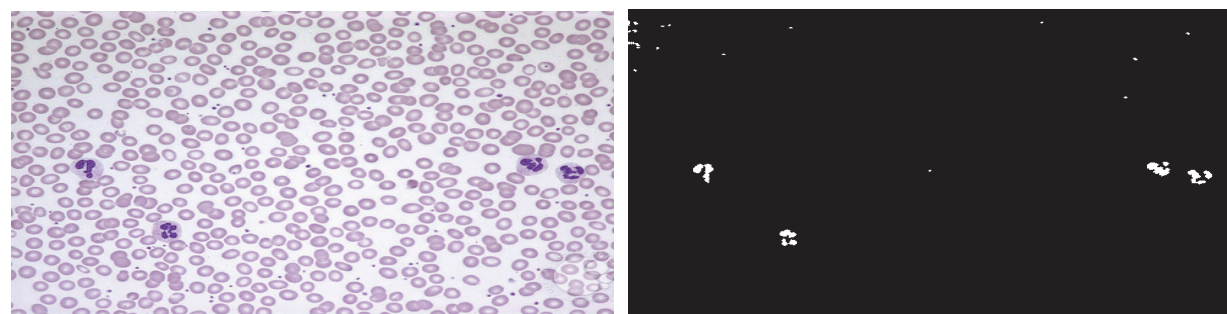


Figure 25. Low resolution microscopic image

The experimental results showed that the proposed methodology was unable to diagnose microscopic images with stained with different color stain's and has connected cells as shown in Figure 26. And gave a value of 0.3677 for white cells percentage

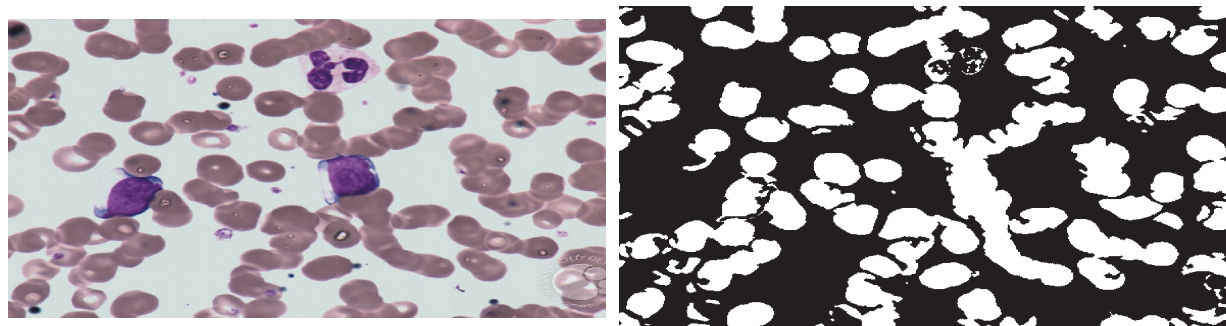


Figure 26. Microscopic image for connected RBC's stained with different stain

5. Conclusion

This paper propose an automated methodology for analyzing cancer histology and hematology microscopic images to detect leukemia using image processing by combining two diagnosis procedures initial and advance; The experimental results showed that the proposed methodology initial diagnosis was able to detect leukemia images and differentiate it from samples that don't have leukemia. in the advance diagnosis it was able to detect and classify most leukemia types and differentiate between acute and chronic but in some cases in the chronic leukemia where the percent of blast cells and shape are similar it gave a diagnosis of the type of leukemia to the most similar type. The initial diagnosis depend on the percentage of the white blood cells in microscopic images affected by leukemia as indicator for the existence of leukemia in the blood smear sample. Whereas the advance diagnosis classifying the leukemia according into different types using feature bag classifier.

References

- A B Nobuyuki Otsu (1979). A threshold selection method from gray-level histograms. *IEEE Trans. Sys., Man., Cyber*, 9(1), 62-66. <https://doi.org/10.1109/TSMC.1979.4310076>.
- Amjad, A., Hudaib, H. N. F., Fatima, E. A. A., & Sandi, N. F. (2017). A Survey about Self-Healing Systems (Desktop and Web Application. Scientific Research Publishing.
- Basim, A., & Hussam, N. F. (2008). Automation of iron deficiency anemia blue and red cell number calculating by intictinal villi tissue slide images enhancing and processing, *Computer Science and Information Technology*, 2008. ICCSIT'08. International Conference.
- Basim, A., Hussam, N. F., Omar, S. A. (2006). CDNA microarray genome image processing using fixed spot position. *American Journal of Applied Sciences*, 3(2), 1730-1734.
- Basim, A., Mohammad, H. Z., & Hussam, N. S. (2007). Mammogram breast cancer image detection using image processing functions. *Information Technology Journal*, 6(2), 217-221.
- Belsare, A. D., & Mushrif, M. M. (2012). Histopathological Image Analysis Using Image Processing Techniques: An Overview. *Signal & Image Processing: An International Journal (SIPIJ)*, 3(4), August 2012.
- Brothwell, D. J., Lewis, D. W., Bradley, G., Leong, I., Mock, D., & Leake, J. L. (2003). Observer agreement in the grading of oral leukemia, *Community Dent Oral Epidemiol*.
- Camargo, J., Caicedo, C., & Gonzalez, F. (2009). Content-based Histopathology Image Retrieval Using a Kernel-based Semantic Annotation Framework.
- Capdevielle, O., & Dalle, P. (1994). Image Processing Chain Construction by Interactive Goal Spec_iation, *Proceedings of the First IEEE Int. Conf. on Image Processing*, 3.
- Chan, F., Lam, F., & Zhu, H. (1998), Adaptive thresholding by variation method. *IEEE Transactions on Image Processing*.
- David, J. F., Wenjin, C., & Lin, Y. (2013). Automated image interpretation and computer-assisted diagnostics. *Anal Cell Pathol (Amst)*, 34(6), 279-300. <https://doi.org/10.3233/ACP-2011-0046>
- Deng, Y., Kenney, C., Moore, M. S., & Manjunath, B. S. (1999), Peer group filtering and perceptual color image

- quantization, to appear in Proc. of ISCAS.
- Deng, Y., Kenney, C., Moore, M., & Manjunath, B. (1999). Peer group filtering and perceptual color image quantization.
- DeSimone, P. A., Cohen, A. M., & Garofalo, M. C. et al. (2008). Cancer of the rectum. In: Abelloff MD, Armitage JO, Niederhuber JE. Kastan MB, McKenna WG, eds. Abelloff's Clinical Oncology. Pa: Elsevier, pp. 1420–1430.
- Dougherty, E., & Lotufo, R. (2003). Hands-on morphological image processing. *Society of Photo Optical*, 20(2).
- Fatma, M., & Jaibir, S. (2014). Identification and classification of acute leukemia using neural network." In Medical Imaging, m-Health and Emerging Communication Systems (MedCom), 2014 International Conference on, pp. 142-145. IEEE.
- Fernandez, F., & Gonzalez, et al. (2010). Histopathological findings in oral lichen planus and their correlation with the clinical manifestations, *Med Oral Patol Oral Cir Bucal*.
- Gonzalez, R., & Woods, R. (2002). Digital Image Processing, Prentice Hall, Upper Saddle River.
- Gorman L. O. (1994), Binarization and multithresholding of document images using connectivity, *Graphical Model and Image Processing*, 350-365.
- Haralick, Shapiro, L. (1992). Survey image segmentation techniques, computer vision graphics and image processing.
- Himali, P., Vaghela, L. (2011). Detection using Digital Image Processing Techniques. *International Journal of Applied Information Systems (IJ AIS)*. ISSN: 2249-0868
- Huang, C., Veillard, A., Roux, L., Loménie, N., & Racoceanu, D. (2011). Time-efficient sparse analysis of histopathological whole slide images, *Computerized medical imaging and graphics: The official journal of the Computerized Medical Imaging Society*, 20.
- Ilyich, J., Karapetian, G., Nalibotsky, B., Nedzved, A., & Podoprighora, A. (2002). System of image processing and analysis. Bioscan-IW. Informatics for health care.
- Iranli, A., Fatemi, H., & Pedram, M. (2005). Histogram Equalization for Backlight Scaling, University of Southern California Los Angeles.
- Jensen, J. R. (1996) Introduction to Digital Image Processing: A Remote Sensing Perspective. Practice Hall, New Jersey.
- Johnston, D. A., & Ruifrok, A. C. (2001) Quantification of histochemical staining by color deconvolution, *Journal of Urology*, Ruifrok AC, Johnston.
- Kekre, H. B. et al. (June 2013). Segmentation of Blast using Vector Quantization Technique. *International Journal of Computer Applications (0975 – 8887)*, 72(15).
- Leong, F. M., Brady, M., & McGee, J. (2003). Correction of uneven microscopy images, *J Clin Pathol*, 520-530.
- Long, X., Cleavel, L., & Yao, Y. (2010). Multiclass detection of cells in multicontrast composite images. *Computers in Biology and Medicine*, 20(3), 150-165.
- Markiewicz, T., Osowski, S., & Marianska, B. (2005). Automatic recognition of the blood cells of myelogenous leukemia using SVM, *Neural Networks, 2005. IJCNN '05. Proceedings. 2005 IEEE International Joint Conference*. <https://doi.org/10.1109/IJCNN.2005.1556295>
- Omar, A., Amjad, H., Hussam, N., & Fakhouri, L. A. (2013). Microarray Gene Expression Extraction and Archiving Based on Local Spot Locations, *IJCSI Int. J. Comput. Sci*.
- Rizik, M. H., Al-Sayyed, H. N., Fakhouri, A. R., & Colin, P. (2017) Polar Particle Swarm Algorithm for Solving Cloud Data Migration Optimization Problem. *Modern Applied Science*, 11(8), 98.
- Salim, A., Emel, O., & Cigdem, G. D. (2014). A Color and Shape Based Algorithm for Segmentation of White Blood Cells in Peripheral Blood and Bone Marrow Images, *Cytometry Part A* 85A: 480490, 2014.
- Sezgin, M., & Sankur, B. (2004). Survey over image thresholding techniques and quantitative performance evaluation. *Journal of Electronic Imaging*, 13(1), 146-165. <https://doi.org/10.1117/1.1631315>
- Subrajeet, M., Dipti, P., & Sanghamitra, S. (2013) An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images. Springer-Verlag London 2013.

Supardi, N. Z., Mashor, M. Y., Harun, N. H., Bakri, F. A., & Hassan, R. (2012). Classification of blasts in acute leukemia blood samples using k-nearest neighbour. In *Signal Processing and its Applications (CSPA), 2012 IEEE 8th International Colloquium on*, pp. 461-465. IEEE, 2012.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).