Sudan University of Sciences & Technology

College of Graduate Studies

Detection of Sickle Cell Disease Based on an Improved Watershed Segmentation

A Dissertation submitted for Partial Fulfillment of Requirement for Master Degree in Biomedical Engineering

Submitted by:

Hala Algailani Mutwali Osman

Supervisor:

Dr.Musab Elkhair

July 2018

الآيـــــــــــــــة

قال تعالى:

(قالو سبحانك لا علم لنا الا ما علمتنا انك انت العليم الحكيم)

صدق الله العظيم

سورةالبقرة

الآية (32)

Acknowledgement

Firstly I would like to thank my dearest person in my life who supported me and he was being beside me step by step my father, and I would also thank a great women for her motivation and encouragement my dear mother.

I would like to express my thanks to my supervisor Dr.MusubAlkhair for his knowledge, guidance, patience and support, who helped me get the expected results through continuous assessment of my work.

Special thanks for Sawsan Ali my close friend and advisor, I would like to thank her for everything she was done her help, knowledge, suggestions, time and efforts.

Finally all thanks for everyone who helped and supported me during my master studies.

Abstract

Sickle cell disease (SCD) is the most popular inherited blood disease, that red blood cells change its shape form circular shape to sickle shape and loses its main job which carries oxygen throughout the body. The watershed segmentation method has become highly developed for automated analysis of overlapping red blood cell microscopic images. The aim of this work is to suppress over segmentation problem which is a major drawback of the watershed algorithm. The experimental results showed that, watershed is most effective when done on filtered image using non local means de-noising method. The effectiveness of the proposed method is validated by analyzing the image segmentation quality measures. The proposed method provides higher performance in term of accuracy, sensitivity and specificity factors.

المستخلص

الانيميا المنجلية من اكثر امراض الدمشيوعا وهو مرض وراثي وسمي المرض بهذا الاسم لان خلايا الدم الحمراء يتغير شكلها من الشكل الكروي الي الشكل المنجلي(المنجل اداة من ادوات الزراعة). خلايا الدم الحمراء الطبيعية تحتوي علي جزئ يعرف بالهيموقلوبين والذي بدوره يقوم بنقل الاكسجين الي كل اجزاء الجسم ،اما عندما يكون الشخص مصاب بالانيميا المنجلية يتغير شكل خلايا الدم الحمراء الي شكلي منجلي مما يفقد الهيموقلوبين وظيفته الاساسية. الكشف عن هذا المرض بالطريقة التقليدية عملية مرهقة وتستهلك الكثير من الوقت بالاضافة الي انها غير دقيقة وقد تسبب خطأ في التشخيص. في هذا البحث تم استخدام تقنيات معالجة الصور لايجاد طريقة آلية للكشف عن هذا المرض وذلك للحصول علي نتائج دقيقة بطريقة فعالة وبالتالي تحسين جودة التشخيص. لازالة التشويش عن الصورة استخدم مرشح الوسائل غير المحلية و لتقسيم الصورة تم استخدام خوارزمية مستجمعات المياه المحسنة ايضا تم حساب عامل الشكل الذي عن طريقه يتم التعرف علي شكل الخلية المنجلية. هذه الطريقة سريعة دقيقة فعالة و اقل تكلفة.

Table of Content:

List of Figures

List of Tables

Chapter One

Introduction

1.1. Sickle Cell Disease

Sickle cell disease (SCD) is the most common inherited blood disorder. That means it’s passed down through families. The disease called sickle cell anemia because red blood cells change its shape from circular shape to sickle shape .Red blood cells contain a molecule called hemoglobin, which carries oxygen throughout the body. In a healthy person, hemoglobin is smooth, round, and flexible. That allows red blood cells to glide easily through bloodstream but SCD change the hemoglobin’s shape. It forms rods that clump together and block blood stream. About 100,000 people in the United States have sickle cell disease. Most of them are African-Americans. [2]

The figures below show normal red blood cells and sickle cells:

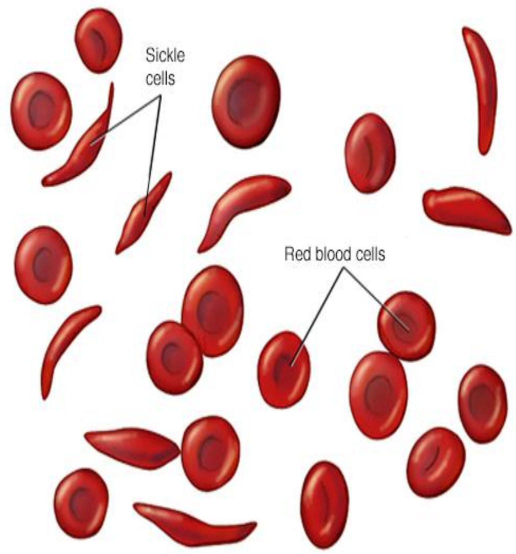

Fig1.1. Normal and sickle red blood cells

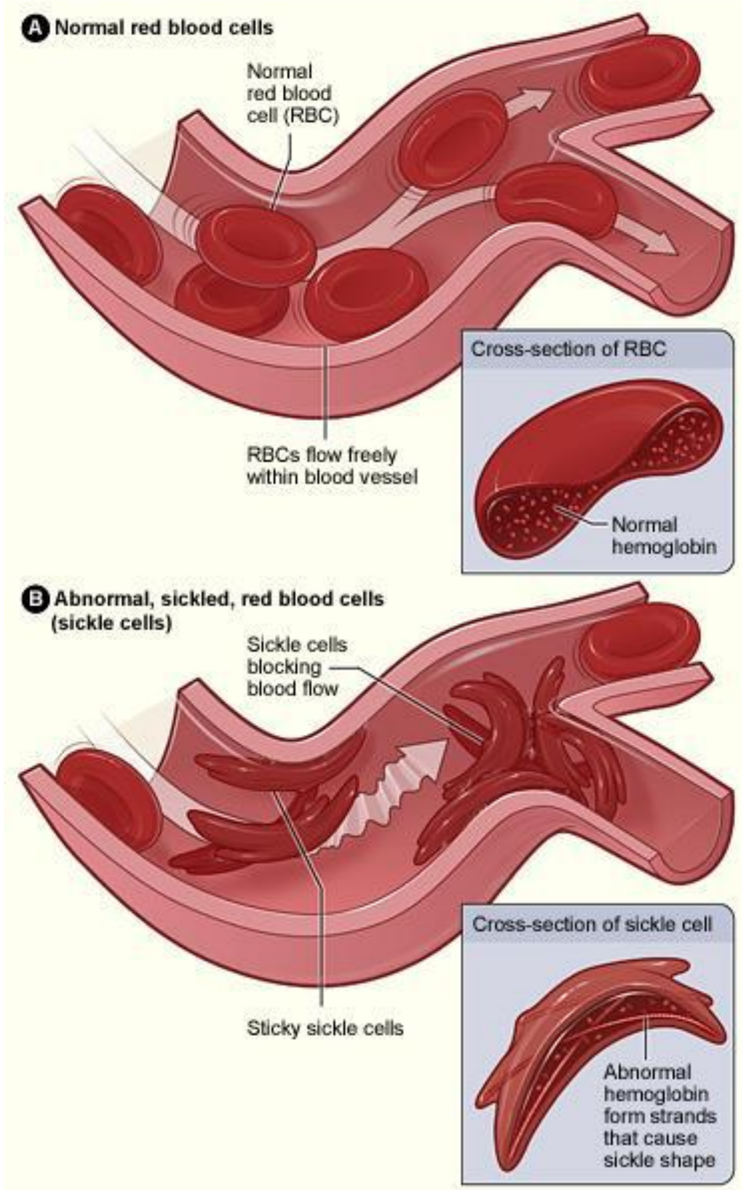

Fig1.2. Normal red blood cells and sickle cells in blood stream

1.1.1. Causes of Sickle-Cell Disease

Sickle cell Anemia is caused by an abnormal type of hemoglobin called Hemoglobin S. Hemoglobin S changes the shape of red blood cells. The red blood cells become shaped like crescents or sickles, sickle-shaped cells deliver less oxygen to the body's tissues. They can also get stuck more easily in small blood vessels, as well as break into pieces that can interrupt healthy blood flow. These problems decrease the amount of oxygen flowing to body tissues even more.

1.1.2. Symptoms of sickle cell

Patients with sickle cell disease may develop severe pain in the chest, back, arms, legs, and abdomen. Pain can occur anywhere in the body. Sickle red blood cells in the lungs can cause severe illness with chest pain, fever, and difficulty breathing. Sickle cell disease can also cause permanent damage to the brain, heart, kidneys, liver, spleen, and bones. The severity and symptoms vary greatly from person to person, even within the same family.

Symptoms of sickle cell crisis include:

- Severe pain
- Anemia
- Chest pain and difficulty breathing
- Strokes
- Joint pain and arthritis and bone infractions
- Blockage of blood flow in the spleen or liver
- Severe infection [2]

1.1.3. Blood Test Diagnosis:

Sickle cell disease can be diagnosed with a blood test, which is analyzed with hemoglobin electrophoresis and other techniques to determine if the defective type of hemoglobin is present in the blood. A trace amount of defective hemoglobin in the blood sample suggests that the individual carries the genetic trait, whereas a high level represents the presence sickle cell disease [4]. Detection of sickle cell anemia is visual assessment of color and shape of red blood cells to recognize normal red blood cells and sickle cells. The detection by this method is not accurate enough, caused misdiagnosis, consumed time, exhaustive process, when detection and counting process stopped it needs to start from beginning, and there is no quick reference for analyzed image. In addition to previous problems, there are some cases that normal blood cells are connected together or with sickle cells which makes the visual detection so difficult because the shape is changed. Recently, image processing techniques play a significant role in blood diseases diagnosis [5].Segmentation one of the image processing techniques that subdivides an image into its constituent regions the level to which the subdivision is carried depends on the problem being solved. The segmentation stopped when the objects of interest have been isolated. In segmentation process a good segmentation algorithms have been chose, then measuring their performance, and understanding their impact on the scene analysis system. Using of image segmentation in blood diseases detection improve the quality of diagnosis, cost effective, reduce time consuming for this process, and give accurate results.

1.1.4. Problem Statement

This study was proposed to reduce over-segmentation which is a main problem in watershed segmentation. Watershed is a good choice for detection of overlapping cells, it depends on a good estimation of image gradients but, background noise tends to produce spurious gradients, causing over-segmentation, also low-contrast edges generate small magnitude gradients, causing distinct regions to be incorrectly merged.

1.1.5. Objectives

- General objectives:

1) To propose effective algorithm by using image processing techniques to convert visual detection to automatic detection.
2) To improve the quality of diagnosis.

- Specific objectives:
  1) To detect overlapping cells using an improved watershed segmentation by applying effective de-noising method.
  2) To increase the accuracy of diagnosis.

1.1.6. Thesis Layout

The Thesis consists of five chapters, chapter one includes introduction, problem statement, objectives, and thesis layout. Chapter two contains theoretical background and literature review. Chapter three includes research methodology (materials and methods). Chapter four consists of results and discussion. Chapter five contains conclusion and recommendations.

Chapter Two

Theoretical Background & Literature Review

2.1. Theoretical Background

Due to the development of Computer Science, computer vision, digital image processing and pattern recognition several automated methods have been developed by engineers and scientists which makes the analysis process accurate, faster, continuous and interactive within a very small interval of time. Computer based methods are now increasingly used to improve the quality of medical services. Artificial intelligence is the area of computer science focusing on developing machines that can engage on behaviors that humans consider intelligent. In microscopic image analysis we observe the image of blood cells through a microscope attached with a camera interfaced with a computer. Using these arrangements we can acquire the image in the computer. Using suitable algorithms, necessary software can be developed for the detection and diagnostic process, without human errors and relatively faster. Identification of disease is a critical step for curing the disease. Microscopic analysis of peripheral blood smears by medical professionals is an important test in the procedures for the diagnosis of any blood related disease. Accurate diagnosis of disease is critical for curing and controlling the disease. In olden days the process was only carried out by expert medical professionals alone. Now a day's these works are carried out by automated systems based on digital image processing algorithms and computer vision methods. Development of information technology affects in image processing technology which is becoming an essential and effective tool in scientific research. It is widely used and effective in the field of biomedical engineering. Digital image processing techniques are important in the analysis of medical images. It includes image enhancement, image filtering, segmentation, image masking, edge detection etc. Image enhancement consists of contrast stretching, un-sharp masking, edge detection etc. Image filtering is mainly used for removing noise from an image. There are several filtering methods which include, mean filter, median filtering, Gaussian filter, Laplacian filter etc. Segmentation subdivides an image into its constituent parts it includes region growing, splitting and merging and thresholding etc. Medical fields like bioinformatics and biomedical imaging are using several machine vision techniques require image processing components with sufficient accuracy. This is true in biomedical image processing which has experienced vigorous growth. Digital image processing techniques are used today in a wide range of applications; share a common need for methods capable of enhancing pictorial information for human interpolation and analysis. In the image processing after acquiring a digital image, the main tasks are enhancement, segmentation, measurement and data analysis. Image enhancement methods are often used to emphasize certain features and to remove artifacts respectively. Two types of measurements are made; feature measurements are defined by a segmentation process and field measurements are obtained globally from complete images, these features and field measurements must be analyzed.

2.1.1. Image filtering

Image filtering is used mainly to improve the quality of input image that corrupted with noise in image acquisition, variation in intensities, or poor contrast; to be easily process in later stages.

- Non-local means filter

The Non local Means has been proposed by Buade s*et al*. [16] it is mainly used for the suppression the white Gaussian noise. Non Local Means filter estimates noise-free pixel intensity as a weighted average of all pixel intensities in the image and the weights are proportional to the similarity between the local neighborhood of the pixel being processed and local neighborhoods of surrounding pixels [17].

2.1.2. Image segmentation

Segmentation subdivides an image into its constituent regions or objects. The level to which the subdivision is carried depends on the problem being solved. That is, segmentation should stop when the objects of interest have been isolated.
Segmentation algorithms for monochrome images generally are based on one of two basic properties of image intensity values: discontinuity and similarity.
In the first category, the approach is to partition an image based on abrupt changes in intensity, such as edges. The principal approaches in the second category are based on partitioning an image into regions that are similar according to a set of predefined criteria such as region growing, region splitting and merging. [14]

The segmentation process can be based on finding the maximum homogeneity in gray levels within the regions identified. One of the common problems encountered in image segmentation is choosing a suitable method for separating different objects from the background. The segmentation doesn't perform well when the gray levels of different objects are similar. Image enhancement techniques seek to improve the visual appearance of an image. They emphasize the salient features of the original image and simplify the task of image segmentation. The type of operator chosen has a direct impact on the quality of the resultant image. It is expected that an ideal operator will enhance the boundary differences between the objects and their background making the image segmentation task easier. Issues related to segmentation involve choosing good segmentation algorithms, measuring their performance, and understanding their impact on the scene analysis system. [6] .There is a various types of segmentation techniques:

A. Segmentation Based on Edge Detection:

This method attempts to resolve image segmentation by detecting the edges or pixels between different regions that have rapid transition in intensity are extracted and linked to form closed object boundaries result is a binary image. Based on theory there are two main edge based segmentation methods gray histogram and gradient based method.

B. Thresholding Method

Thresholding operation converts a multilevel image into a binary image it chooses a proper threshold value, to divide image pixels into several regions and separate foreground from background.

C. Region Based Segmentation Methods

Segmentation algorithms based on region are simple and more immune to noise. Edge based methods partition an image based on rapid changes in intensity near edges whereas region based methods, partition an image into regions that are similar according to a set of predefined criteria.

- Region Growing:

  Region Growing is a procedure that groups pixels or sub regions into larger regions based on predefined criteria for growth. The basic approach is to start with a set of "seed" points and from these grow regions by appending to each seed those neighboring pixels that have predefined properties similar to the seed.

- Region Splitting and Merging:

  Rather than choosing seed points, user can divide an image into a set of arbitrary unconnected regions and then merge the regions in an attempt to satisfy the conditions of reasonable image segmentation. Region splitting and merging is usually implemented with theory based on quad tree data.

D. Clustering Methods:

Clustering based methods have been widely used in segmentation of gray level images. Since the clustering methods are either directly applicable or easily extendable to higher dimensional data, their application in segmentation of color and multispectral images is a natural choice. Some of the clustering techniques rely on knowing the number of clusters prior. In that case the algorithm tries to partition the data into the given number of clusters K-means and Fuzzy C-means clustering.

- K-means clustering

The K-means clustering, or Hard C-means clustering, is an algorithm based on finding data clusters in a data set such that a cost function of dissimilarity measure is minimized. K-means clustering is a partitioning method. The function k-means partitions data into k mutually exclusive clusters, and returns the index of the cluster to which it has assigned each observation. Also, k-means clustering operates on actual observations and creates a single level of clusters.

- Fuzzy C-means clustering

Fuzzy C-means (FCM) is one of the commonly used methods for image segmentation and its success is mainly due to the introduction of fuzziness for the belongingness of each image pixels. Compared with crisp or hard segmentation methods, FCM is able to retain more information from the original image. However, one disadvantage of FCM is its sensitivity to noise and other imaging artifacts. The reason may perhaps stems from the non-unimodal property of its membership functions and the use of the squared Euclidean distance.

Fuzzy c-means (FCM) is a data clustering technique in which a dataset is grouped into n clusters with every data point in the dataset belonging to every cluster to a certain degree.

2.1.3. Feature Extraction

Feature extraction is a reduction of data when is too large to be processed so the input data will be transformed into a reduced set of features then the input data will be transformed into a reduced representation set of features. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy. [6]

There are two types of feature extraction geometrical feature extraction and textural feature extraction .The textural feature extraction include either Herralicks features such as energy, entropy, angular second moment, inverse difference moment, etc. [13]

2.2. Literature Review

There are many studies presented in sickle cell anemia detection; [6] Proposed method using image processing techniques to improve detection of sickle cell anemia used microscopic image captured by camera connected to light microscope. The image was processed and then used clustering based segmentation (k mean clustering and fuzzy C mean clustering) to differentiate between normal red blood cells and infected cells. Finally feature extraction was computed (geometrical and textural features) to identify the types of sickle cells. [7] Presented correct identification of aberration in normal parameters of RBCs in an anemic blood sample has been presented using different image processing tools and techniques. Preprocessing was done using Weiner filter and Sobel Edge detection method was used to find the boundary of the corpuscles. Then region properties were used to calculate a metric value that determines abnormal shape of the corpuscles to diagnose the disease. [8] Developed an automated image processing system to identify the shape of infected blood cells. To help people in the rural areas where the expert in microscopic analysis may not be available, the microscopic image acquired from live video or loaded from memory, used edge detection segmentation with different types of operators (Canny, Sobel, Roberts, Prewitt, and The Laplacian of Gaussian) and compared between them. From all operators it found that canny operator was more proper for detection because it gave more details of original image, Roberts operator was unable to identify sickle cell properly. [9] Proposed method used fractal dimension approach to recognize the sickle shaped red blood cells present in the blood smear. K means clustering segmentation is used to identify red blood cell and sickle cell, this method is rapid and cost effective than earlier methods. [10] Proposed to detect sickle cell anemia and plasmodium parasites, the segmentation based on simple idea work directly on gray level histogram. This system is 87% accurate and reduces time consuming. [11] Presented method to detect sickle blood cells in blood samples using image segmentation (Otsu's segmentation) and shape detection, method was based on calculating the max axis and min axis of the cell. The form factor was computed using these properties to determine whether the cell is sickle or not. This method is 90 percent more accurate than the existing method.

All the previous methods were good for detection of separated cells, but for images contained overlapping cells, methods failed to detect them. [12] Proposed techniques used in counting of infected and normal erythrocytes for purpose of estimating parasitemia, thresholding segmentation is used to segment foreground from background and watershed transform is applied to separate overlapping cells, in addition to SVM classifier which used for the recognition and classification of cells. This algorithm implemented using MATLAB. [13] Proposed real time system for sickle cell analysis by using image processing techniques, they used Otsu segmentation and watershed transform. Otsu segmentation is simple for implementation, watershed transform is a fast, robust and widely used in image processing and analysis, but it suffer from over-segmentation. From comparison between two techniques found that watershed segmentation gave better result than Otsu segmentation. [14] Proposed an algorithm to detect the sickle cells, tear drop cells and ovalocytes using marker controlled watershed segmentation, feature extraction (Metric value, aspect ratio, radial signature) and KNN classifier. The segmentation of different objects in the image helps in improving the efficiency of the algorithm as features are extracted for only the selected cells rather than the entire group of erythrocytes present. The method helped in detection and classification of sickle cells, dracocytes and elliptocytes with an accuracy of 80.6% and sensitivity of 87.2%.

These methods improved the quality of detection by using watershed segmentation for separating overlapping cells; watershed is more effective for this work.

Chapter Three

Research Methodology

3.1. Proposed method

The proposed algorithm was developed to detect blood smear image contains sickle cell anemia and overlapping cells by using nonlocal means filter for noise removal, watershed segmentation to separate overlapping cells, in addition to calculation of form factor and aspect ratio to identify the sickle cell shape. Figure (3.1) below shows the steps of proposed method:

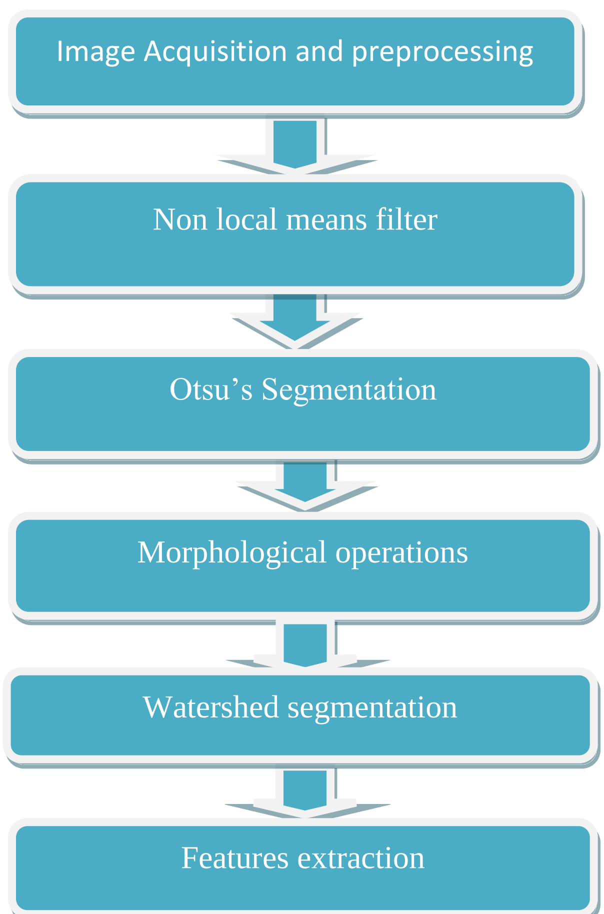

Fig3.1. Block diagram of proposed algorithm

3.1.1. Image acquisition and preprocessing

The image of blood sample contains sickle cell anemia [2], 279×414×3 RGB image. The main purposes of preprocessing are color transformation, contrast adjusting and noise removal. First step RGB image converted to gray scale image, then un-sharp masking applied to improve the quality of input image, for image filtering non local means filter was applied. Test images which used corrupted with Gaussian noise, fig 3.3.below illustrates the noise type:

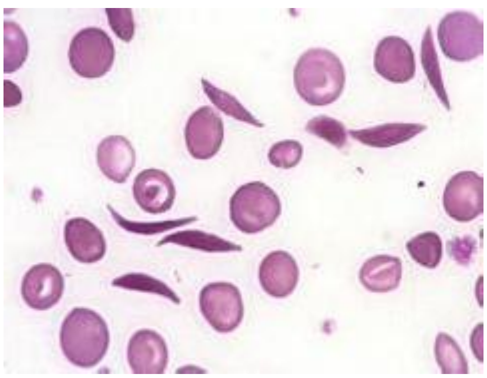

Fig.3.2. input image

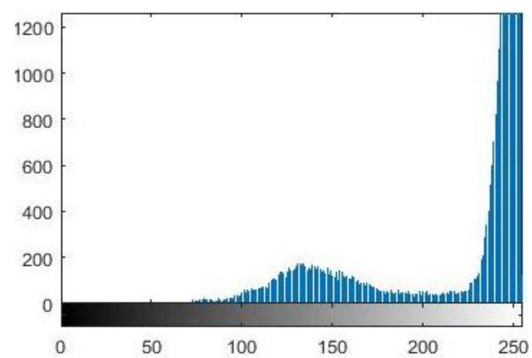

Fig 3.3. Histogram of input image

- Non local means filter

  The idea of the de-noising methods replacing the color of a pixel with an average of the colors of nearby pixels. The non-local Means proposed by Buades is used mainly for white Gaussian noise removal the principle of work for this filter computes the mean of all pixels in the image and then weighted by how similar they are to the pixel being processed.

3.1.2. Image segmentation

- Otsu's segmentation

Otsu's segmentation used to automatically perform clustering-based image thresholding, or the reduction of a gray level image to a binary image. The segmentation process assumes that the image contains two classes of pixels following bi-modal histogram (foreground pixels and background pixels), it then calculated the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal, or equivalently (because the sum of pair wise squared distances is constant), therefore their inter-class variance is maximal. [18]

- Watershed segmentation

Watershed is a transformation defined on a gray scale image. The idea of watershed can be view as a landscape immersed in a lake; catchment basins will be filled up with water starting at each local minimum. Dams must be built where the water coming from different catchment basins may be meeting in order to avoid the merging of catchment basins. The water shed lines are defined by the catchment. Basins divided by the dam at the highest level where the water can reach in the landscape. As a result, watershed lines can separate individual catchment basins in the landscape [19]. The tool that used with watershed is distance transform. Distance transform needed to preprocess the image to make it suitable for watershed segmentation it's a distance between every pixel and the nearest non-zero valued pixel [15]. Watersheds may also be defined in the continuous domain. There are also many different algorithms to compute watersheds. Watershed algorithm is used in image processing primarily for segmentation purposes. The watershed transform can be classified as a region-based segmentation approach. When simulating this process for image segmentation, two approaches may be used: either one first finds basins, then watersheds by taking a set complement; or one computes a complete partition of the image into basins, and subsequently finds the watersheds by boundary detection.

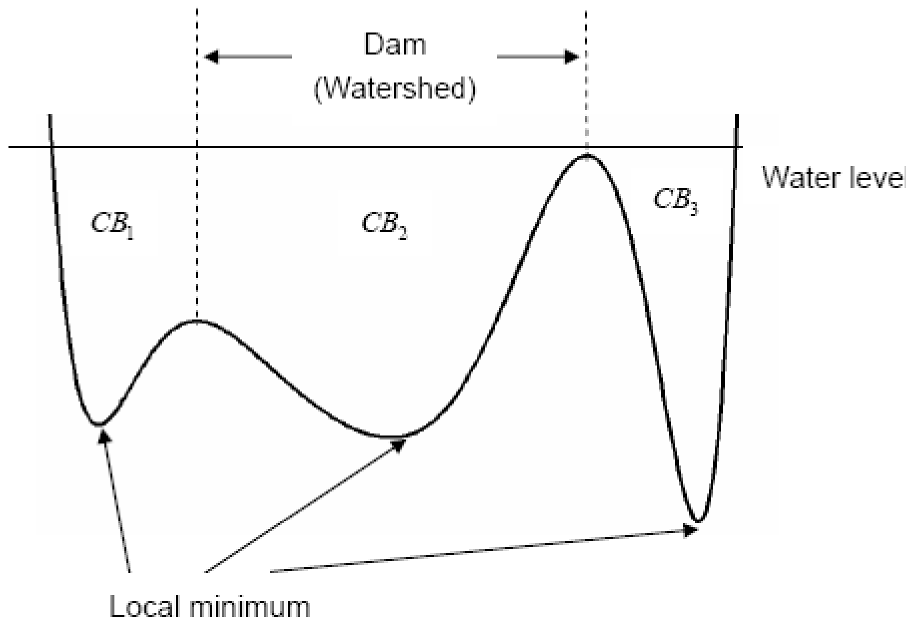

Fig3.4. Immersion process of Watershed Transform

3.13. Image segmentation evaluation

A. Segmentation Time

The time taken in complete segmentation includes the time taken in preprocessing step, time taken for the application of the watershed and time taken by all the processes.

Various image processing tools like MATLAB provide options, commands or other means to execute and obtain a detailed description of time consumption of the major commands, their processes and the overall time taken by an algorithm to complete the processing and generate output. [20]

B. Result Accuracy

The accuracy of a segmentation algorithm is generally evaluated by comparison with a manual segmentation, as there is no gold standard. Ideally, all methods should be evaluated for performance on data from a patient, but this is not practical. So the final result of the algorithm is generated as an overlay of the segmentation result and the original image which is considered as the ground truth image for the purpose of accuracy scrutiny. The following terms are evaluation metrics used for verifying quality of a segmented image:

True positive (TP): pixels correctly segmented as foreground.

False positive (FP): pixels falsely segmented as foreground.

True negative (TN): pixels correctly detected as background.

False negative (FN): pixels falsely detected as background.

Then the metrics used to calculate sensitivity, specificity and accuracy as:

Accuracy: (TP+TN)/ (TP+FP+TN+FN).

Sensitivity: TP/ (TP+FN).

Specificity: TN/ (TN+FP). [20]

3.1.4. Feature extraction

Feature extraction is the final step in sickle cell detection, form factor and aspect ratio calculated to use in recognition of the sickle cell shape. Features extracted to calculate these metrics (form factor and aspect ratio) are:

1. Area: The actual number of pixels in the region.

2. Major Axis Length: The length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region.
3. Minor Axis Length: The length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the region.
4. Perimeter: The distance around the boundary of the region. Region props compute the perimeter by calculating the distance between each adjoining pair of pixels around the border of the region.

3.1.4.1. Form Factor:

The form factor is a metric used to measure a roundness of an object, the form factor for a perfect circle is 1, reducing in value with decrease in the roundness of an object. As healthy RBCs are circular in shape, the form factor for a healthy RBC is equivalent to 1. But for a sickle cell the value decreases to 0.7 to 0.

$$Form\ factor\ = 4 *\ pi *\ area\ /(perimeter)2 \quad (1)$$

3.1.4.2. Aspect Ratio:

The ratio of the major axis (d1) to minor axis (d2) of the cell, the aspect ratio of a circle is 1 and as the circularity of the shape decreases the aspect ratio increases.

$$Aspect\ ratio\ =\ major\ axis\ (d1)\ /\ minor\ axis\ (d2) \quad (2)$$

Chapter Four

Results and Discussion

4.1. Results

In study [14] the researchers proposed an algorithm for sickle cells, tear drop cells and ovalocytes detection using marker controller watershed segmentation, feature extraction and KNN classifier. The steps followed by this algorithm are: Image acquisition, preprocessing (median filter was applied to remove noise), marker controller watershed segmentation used for overlapping cells, morphological operations applied to enhance the image, extracted features for shape recognition, and finally training the K-nearest neighbor classifier(KNN) to test the images.

The method above is simple, fast and detects sickle cells, dracocytes and elliptocytes with an accuracy of 80.6% and sensitivity of 87.2%. In fig.4 (d) one of sickle cells was not detected and the accuracy of detection needed to increase to improve the quality of diagnosis.

The new algorithm proposed to increase the accuracy of detection, so instead of using median filter nonlocal means filter is used, besides applying Otsu's segmentation before watershed segmentation to automatically perform clustering-based image thresholding and watershed segmentation to separate overlapping cells. The figures below show the results of the study [14]:

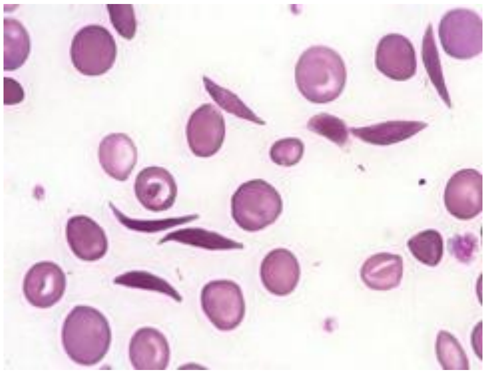

Fig.4 (a) Original Image of existing algorithm

Fig.4 (b) Image obtained applying watershed

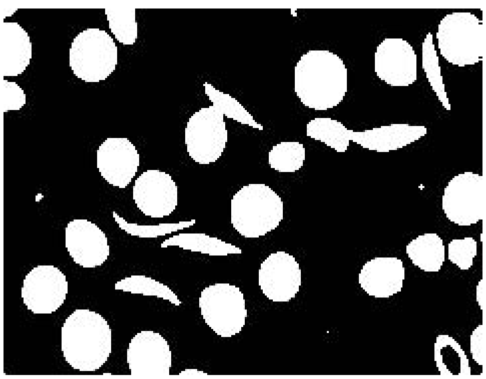

Fig.4 (c) Image obtained on applying morphological operations

Fig.4 (d) Separated sickle cell

- In figures below the new algorithm applied to two samples of blood smear images contain sickle cells anemia and detected the shape of cell by using metric value (form factor). Detection steps of sickle cells in blood smear image of first sample:

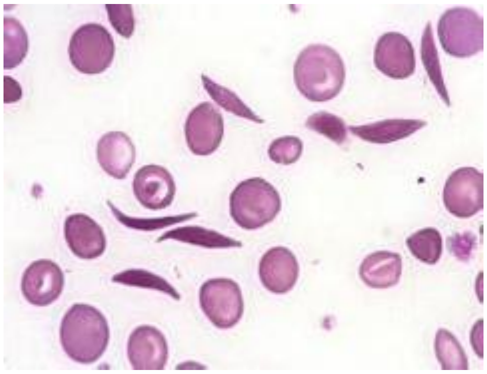

Fig.5 (a) Input image

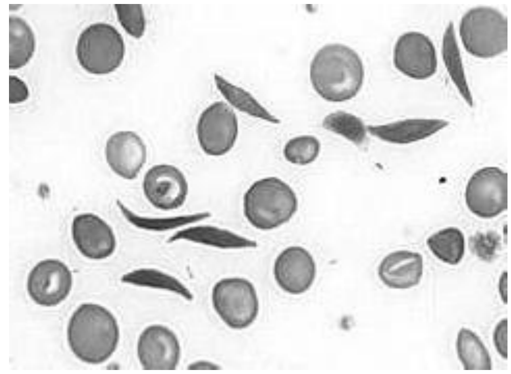

Fig.5 (b) Un-sharp masking

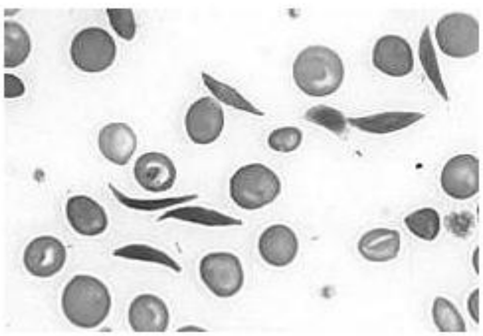

Fig.5 (c) Filtered imge

Fig.5 (d) Otsu's segmentation image

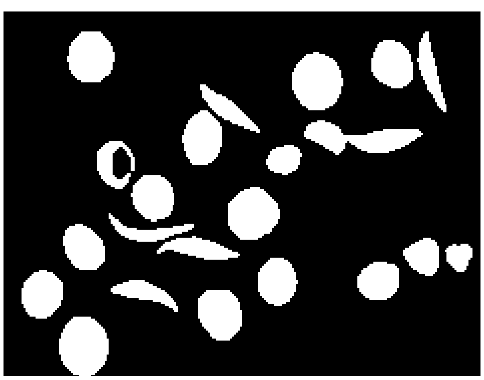

Fig.5 (e) Morphological operations image

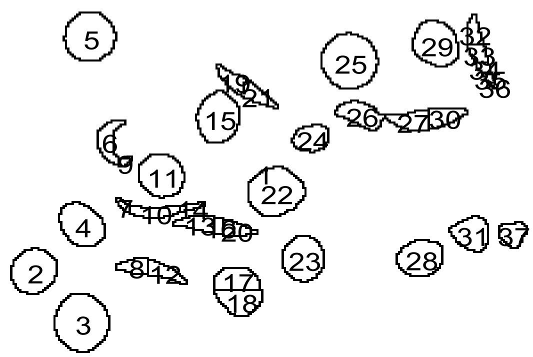

Fig.5 (f) Watershed segmentation image using median filter

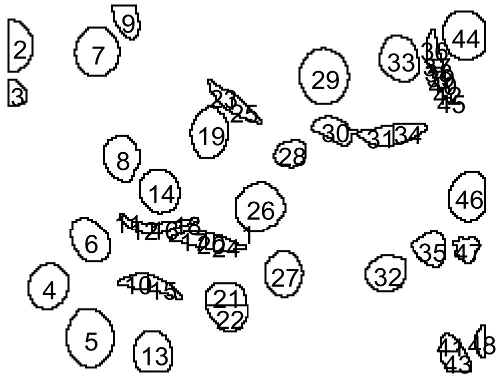

Fig.5 (g) Watershed segmentation image using median filter

Fig.5 (a) shows the input image, Fig.5 (b) shows preprocessed image after applying un-sharp masking which used to improve the quality of input image, Fig.5 (c) shows filtered image after applying nonlocal means filter for noise removal and image smoothing, Fig.5 (d) shows image after Otsu's segmentation used to convert a gray level image to a binary image, Fig.5 (e) shows morphological operations image, Fig.5 (f) shows watershed segmentation image after using median filter, Fig.5 (g) shows watershed segmentation image after using nonlocal means filter, in Fig.5 (f) some of objects in input image lost in output image, in Fig.5 (g) all of the objects in the input image preserved in the output image, that means nonlocal means filter kept the detail in the input image as its after image smoothing and noise removal. when compared nonlocal means filter compared to other image de-noising filters such as Weiner filter, Gaussian filter, Median filter, the nonlocal means filter gave the better results.

Cell No.	**Area**	**Perimeter**	**Minor axis**	**Major axis**	**Form factor**	**Aspect ratio**
1	217	57.982	11.8228	24.8967	0.8111	2.1058
2	76	30.674	8.4405	12.3778	1.0150	1.4665
3	314	60.687	18.5667	21.6266	1.0714	1.1648
4	489	76.508	22.5279	27.7367	1.0498	1.2312
5	291	59.471	16.1771	23.0464	1.0339	1.4246
6	385	67.328	21.0677	23.3550	1.0673	1.1086
7	275	58.156	16.5128	21.4655	1.0218	1.2999
8	141	44.08	10.8014	17.3567	0.9119	1.6069
9	66	32.642	6.8780	13.9741	0.7784	2.0317
10	11	13.307	2.1522	8.0915	0.7806	3.7597
11	33	20.948	4.2930	10.5344	0.9450	2.4539
12	291	58.561	18.5168	20.2689	1.0663	1.0946
13	304	59.835	18.4498	21.0562	1.0670	1.1413
14	69	39.424	5.7460	18.3889	0.5579	3.2003
15	39	24.223	4.2590	12.4442	0.8353	2.9219
16	79	38.522	6.7305	17.1384	0.6690	2.5464

Table.4.1. Geometrical features of the cells in blood smear image for first sample

Table (4.1) showed the properties extracted from test image (area, perimeter, minor axis, and major axis) besides form factor and aspect ratio that calculated to recognize sickle cell shape, In image labeling the image background labeled with number (1) and the cells labeled from number (2) to the last one, in figure Fig.5 (g) cell (2) had a form factor (0.8111) this cell was not considered as a sickle cell,

cell(4), cell (5), Cell (6) and cell (7) had a form factor (1.0714, 1.0498, 1.0339, and 1.0673) respectively these cells considered as normal cells, cell (10) had form factor (0.7784) so the cell considered as a sickle cell and so on.

Watershed transform produced closed and adjacent contours included all image boundaries (edges) which arranged on desired ridges and separated each object from its neighbor.

Fig.6 (a) and Fig.6 (b) showed the watershed lines after using median filter and nonlocal means filter.

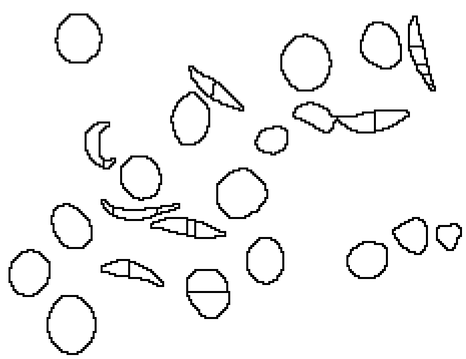

Fig.6 (a) Watershed lines using median filter

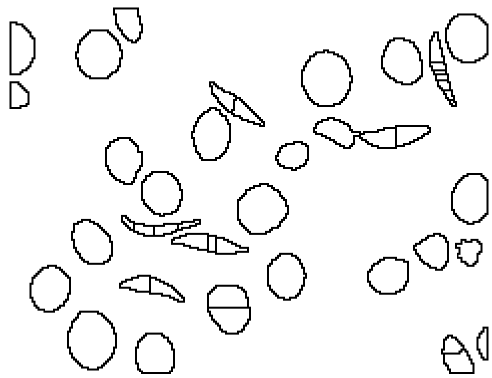

Fig.6 (b) watershed lines using nonlocal means filter

Table 4.3: below shows the filter type and its accuracy:

Type of filter	Accuracy
Non local means filter	0.9321
Median filter	0.8141

Table.4.2. Filter's accuracy for first sample

Table (4.2) showed the accuracy of the proposed algorithm when median filter and nonlocal means filter were used, the main difference between nonlocal means filter and median filter that nonlocal means filter takes the mean of all pixels in the image weighted by how similar they are to the target pixel it didn’t depend on spatial proximity to current pixel it based on intensity similarity. The median filter replaces the value of a pixel by the median of the gray levels in the neighborhood of that pixel depend on spatial proximity. The principle of work made nonlocal means filter is more effective in smoothing and noise removal than median filter because it took all pixels in the image in consider beside weighted average that gave more importance to pixel.

- Detection steps of sickle cells in blood smear image of second sample:

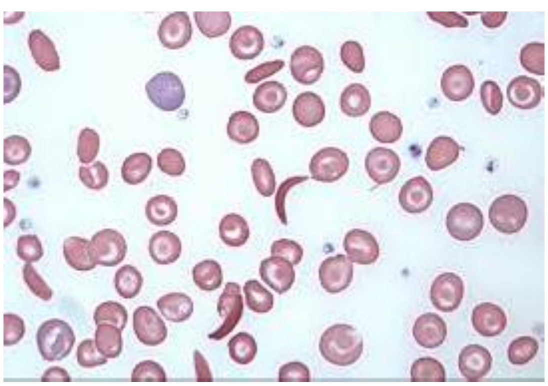

Fig.7 (a) Input image

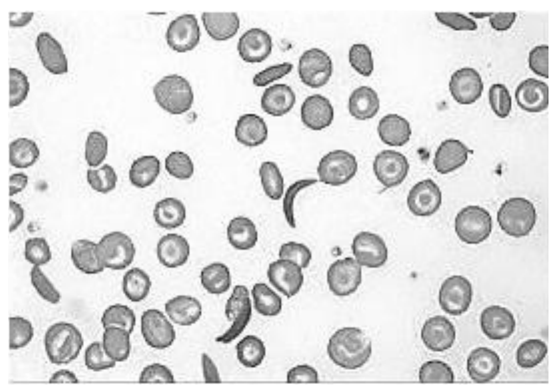

Fig.7 (b) Un-sharp masking

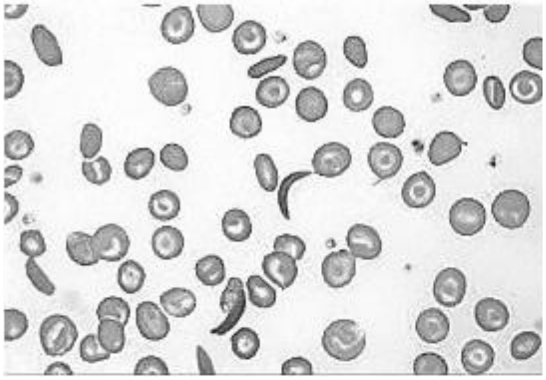

Fig.7 (c) Filtered image

Fig.7 (d) Otsu’s segmentation image

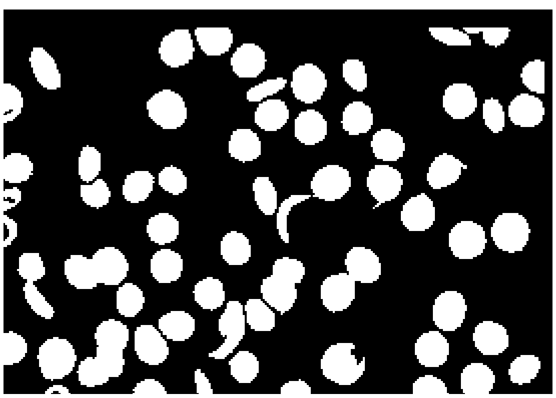

Fig.7 (e) Morphological operations image

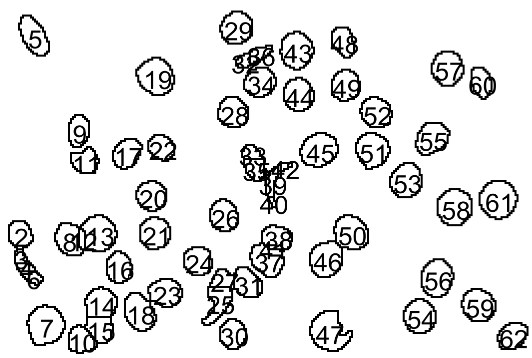

Fig.7 (f) Watershed segmentation image using median filter

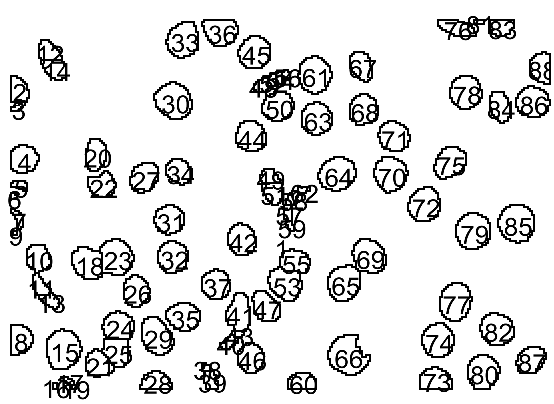

Fig .7 (g) Watershed segmentation image after using non local means filter

Cell NO.	Area	Perimeter	Minor Axis	Major Axis	Form Factor	Aspect Ratio
1	72	31.534	8.0297	12.369	0.6945	1.5943
2	7	11.992	1.7989	7.6293	0.9099	1.5405
3	129	37.836	12.594	13.215	0.6117	4.2410
4	16	14.696	3.2046	7.7675	1.1324	1.0493
5	3	3.093	1.7638	2.582	0.9310	2.4239
6	17	18.914	3.2712	9.7438	3.9407	1.4639
7	110	36.86	10.357	14.316	0.5972	2.9787
8	1	0	1.1547	1.1547	1.0174	1.3822
9	103	34.743	10.52	12.707	Inf	1.0000
10	43	24.91	5.4571	11.315	1.0723	1.2079
11	67	29.475	7.8638	11.887	0.8708	2.0734
12	33	21.172	4.8767	9.4178	0.9691	1.5116
13	77	30.889	8.1495	12.546	0.9251	1.9312
14	241	53.558	16.452	18.797	1.0141	1.5395
15	2	1.96	1.1547	2.3094	1.0558	1.1425
16	7	10.958	1.768	7.4197	6.5423	2.0000

Table.4.3 Geometrical features of the cells in blood smear image for second sample

Type of filter	**Accuracy**
Non local means filter	0.9186
Median filter	0.7818

Table.4.4. Filter's accuracy for second sample

4.2. Discussion:

In the proposed algorithm non local means filter was used instead of using median filter, the principle of work of median filter is consider each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings and then replaces the pixel value with the median of those values. The nonlocal means filter takes the mean of all pixels in the image weighted by how similar they are to the target pixel, the weight gives this pixel more importance in the calculation of the average and makes nonlocal means filter can result in greater post-filtering clarity, minimal loss of detail, preserved features present in the image even those features are small and thin, preserving object's edges in addition to image smoothing.

The new proposed algorithm developed for overlapping cells and sickle cells detection obtained results with high accuracy and consume less time for implementation, which can help physicians to get a good detection by simple and effective method. This algorithm improved the quality of sickle cell detection as compared to the previous methods. Using of nonlocal means filter for noise removal enhanced the quality of the input image besides watershed segmentation helped in overlapping cells detection, so improved the quality of the output image.

Chapter Five

Conclusion & Recommendation

5.1. Conclusion

Sickle cell disease (SCD) is genetic alteration causes the body produces abnormal sickle- or crescent-shaped red blood cells. Which make the blood loses its main job. Diagnosis of sickle cell anemia disease by visual assessment is exhaustive process and obtains inaccurate results. In this research image processing techniques were used to propose automated method for sickle cell detection. Watershed segmentation is a proper choice for this method because it works well with overlapping cells but it suffers from over-segmentation problem which affect in detection quality, to avoid this problem nonlocal means filter was applied to input image. This method is rapid, reliable, and more accurate as compared to previous methods, because nonlocal means filter is an efficient de-noising method preserved all image details after removing noise even they are small. This helped in detection of all sickle cells in test image therefore resulted in accurate detection.

5.2. Recommendation

This research showed significant improvement in sickle cell detection and accuracy of diagnosis, but the accuracy needed to increase to improve quality of diagnosis, suggestion for future work:

- Using more improved de-noising method to enhance input image and avoid over-segmentation problem.

- The future work may include detection of other blood cells diseases beside sickle cells in one algorithm.
- Using proper classifier for cells classification.

References

[1] Patra D. and Mohapatra S., "Automated Cell Nucleus Segmentation and Acute Leukemia Detection in Blood Microscopic Images", NIT Rourkela, 2010.

[2] https://www.webmd.com/a-to-z-guides/understanding-anemia-basics.

[3] http://www.webmd.com/a-to-z-guides/what-is-sickle-cell-disease.

[4]http://www.mayoclinic.org/diseases-conditions/sickle-cell-anemia/symptoms-causes/dxc-20303269.

[5] http://www.news-medical.net/health/Sickle-Cell-Disease-Diagnosis.aspx.

[6] SiddharthSekharBarpanda, "Use of Image Processing Techniques to Automatically Diagnose Sickle-Cell Anemia Present in Red Blood Cells Smear", Department of Electrical Engineering National Institute of Technology Rourkela-769008 (ODISHA) May-2013

[7] PranatiRakshita, KritiBhowmikb, Detection of Abnormal Findings in Human RBC in Diagnosing Sickle Cell Anaemia Using Image Processing, aHOD, Dept. of CSE, JIS College of Engineering , Kalyani, PIN-741235, India bM.Tech Scholar, Dept. of CSE, JIS College of Engineering , Kalyani, PIN-741235, India,2013.

[8] Aruna N.S. , Hariharan S., Edge Detection of Sickle Cells in Red Blood Cells, Aruna N.S et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 4140-4144, 2014.

[9] MenikaSahu, Amit Kumar Biswas, K. Uma, Detection of Sickle Cell Anemia in Red Blood Cell, International Journal of Engineering and Applied Sciences (IJEAS) ISSN: 2394-3661, Volume-2, Issue-3, March 2015

[10] Mrs. S.Mary Asha Rexcy. (M. E)1, V.S.Akshaya2, S.Abinaya3, K.S.Swetha4 ,February-2016, Effective use of image processing techniques for the detection of sickle cell anemia and presence of Plasmodium parasites, IJARIIE-ISSN(O)-2395-4396, Vol-2 Issue-2 2016

[11] KholoudAlotaibi, Sickle Blood Cell Detection Based on ImageSegmentation, South Dakota State University, Open PRAIRIE: Open Public Research Access Institutional, Repository and Information Exchange, 2016

[12] S. S. Savkare, S. P. Narote, Automatic Classification of Normal and Infected Blood Cells for Parasitemia Detection, IJCSNS International Journal of Computer Science and Network Security, VOL.11 No.2, February 2011

[13] Parvathy B.H.1, Hariharan S2, Aruna N.S.3, A Real Time System for the Analysis of Sickle Cell Anemia Blood Smear Images Using Image Processing, International Journal of Innovative Research in Science, Engineering and Technology, Vol. 5, Issue 4, April 2016

[14] Vishwas Sharma, AdhirajRathore, Garima Vyas,Detection of sickle cell anaemia and thalassaemia causing abnormalities in thin smear of human blood sample using image processing, Department of Electronics & Communication Engineering Amity University , U.P, Noida-201303, 2016.

[15] Gonzalez RG, RE Woods and SL Eddins, “Digital Image Processing”, Pearson Education, Inc. NJ. 2007.

[16] Buades, B. Coll& J. M. Morel, “Image Denoising Methods: A New Nonlocal Principle,” M REVIEW 2010 Society for Industrial and Applied Mathematics Vol. 52, No. 1, pp. 1, 13–147,
2010.

[17] B. Goossens, Q. Luong, A. Pizuric& W. Philips, “An improved non-local denoising algorithm,”, In Local and Non-Local Approximation in Image Processing, International Workshop, Proceedings , pp. 143-156, 2008.

[18]https://en.wikipedia.org/wiki/Otsu%27s_method.

[19] Jos B.T.M. Beordink, Arnold Meijster, “The Watershed Transform: Definition, Algorithms and Parallelization Strategies”, FundamentaInformatcae 41187-228, IOS press, 2001.

[20] Vartika Sharma, ANALYSIS OF WATERSHED ALGORITHM FORSEGMENTATION OF MEDICAL IMAGES, International Journal of Advances in Engineering Science and Technology, ISSN: 2319-1112.

Appendix

MATLAB Implementation:

For proposed algorithm implementation MATLAB R2015a was used, because MATLAB platform is easy to use, good for image visualization and integer's computation.

The functions were used to implement this algorithm are showed below:

1. **(imread) for image reading:**

Read image from graphics file as a matrix imread can read many different formats such as JPG, PNG, GIF and TIF.

Syntax:

A=imread('file name');

2. **(imshow) for image display:**

Display image in handle graphics figure.

Syntax:

Imshow(A)

3. **(rgb2gray) for color conversion:**

Converting RGB images to gray scale images.

Syntax:

B=rgb2gray(A)

4. **(unshrap masking) for image sharpening:**

Returns an enhanced version of the grayscale or true color input image.

Syntax:

C =imsharpen (B);

5. **(im2bw) for converting gray scale image to binary image:**

Converts this grayscale image to binary by thresholding. The output binary image BW has values of 1 (white) for all pixels in the input image with luminance greater than thresold and 0 (black) for all other pixels.

Syntax:

level=graythresh(C);

x=im2bw(C,level);

The graythresh function uses Otsu's method, which chooses the threshold to minimize the intra-class variance of the black and white pixels.

6. **(imfill) for filling images holes:**

Fills holes in the input image, a hole is a set of background pixels that cannot be reached by filling in the background from the edge of the image.

Syntax:

filled = imfill (~x, 'holes');

7. **(imclearborder) for borders clearing:**

Suppresses structures that are lighter than their surroundings and that are connected to the image border.

Syntax:

D=imclearborder (filled);

8. **(bwareaopen) for deleting small objects:**

Removes from a binary image all connected components (objects) that have fewer than specific pixel value.

Syntax:

E=bwareaopen (D, 60);

9. (bwlabel) for objects labeling:

Label connected components in 2-D binary image.

Syntax:

[L,num1] = bwlabel(E);

10.(bwdist) distance transform computing:

Computes the Euclidean distance transform of the binary image.

Syntax:

S = bwdist(~E);

11.(watershed) watershed transform computing:

Computes a label matrix identifying the watershed regions of the input matrix

Syntax:

L = watershed(S);

12.(regionprops) for region properties calculation:

Measures a set of properties for each connected component (object) in the binary image.

Syntax:

```
stats=regionprops('table',L,'Centroid','MajorAxisLength','MinorAxisLength','Perimeter', 'Area');
```