

# EFFICIENT WATERSHED BASED RED BLOOD CELL SEGMENTATION FROM DIGITAL IMAGES IN SICKLE CELL DISEASE

T.Revathi<sup>1</sup>, S.Jeevitha<sup>2</sup>

<sup>1</sup>Associate professor, Dept of CS, PSG College of Arts & Science, Coimbatore, Tamilnadu, India

<sup>2</sup>Research scholar, Dept of CS, PSG College of Arts & Science, Coimbatore, Tamilnadu, India

## Abstract

Generally Sickle cell disease is a set of inherited red blood cell disorder i.e. the sickle cell disease patients will have an abnormal haemoglobin on their red blood cells. Haemoglobin is used to pass oxygen all over the body. Normal haemoglobin will be in the shape of disc in which the cells will be flexible so it is easy to deliver oxygen via blood vessels. The abnormal haemoglobin will be in crescent shape or sickle shape. The abnormal haemoglobin creates stiff rods inside the red blood cell so that it stops or slows the blood flow via blood vessels because it will bond to the blood vessels. Sickle cell disease will cause the patient's spleen, brain, eyes, lungs, liver, heart, kidneys, joints, bones or skin. This research paper discusses about the watershed transformation which is used to identify the boundary of blood samples. The main advantage of using watershed transformation algorithm is finding the regional minima on real images and it is based on region processing. Watershed transformation algorithm will provide the result as global segmentation and border closure with high accuracy.

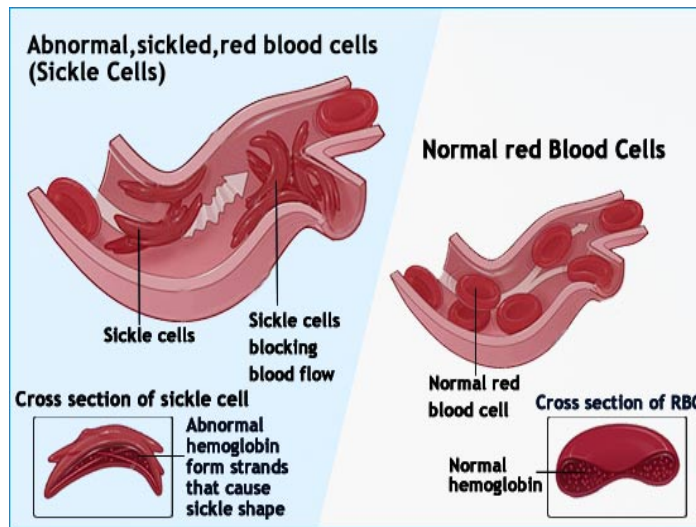
**Keywords:** sickle cell disease, medical image mining, elliptical adjustment algorithm, watershed transformation, comparison of algorithms.

## I INTRODUCTION

### 1.1. Introduction to Sickle cell disease

Sickle cell disease is also known as Sickle cell anaemia which explains a group of red blood cell disorder i.e., the person who are affected by sickle cell disease will have abnormal haemoglobin. Generally normal haemoglobin is a protein which passes oxygen throughout the body and it will have disc shape. And so red blood cells are flexible. It is easy to deliver oxygen

via blood vessels. But the abnormal haemoglobin will have crescent or sickle shape. It forms stiff rods on blood vessels and it reduces the blood flow. Figure 1 represents the normal cells and sickle cells.



*Figure 1: Normal cells and Sickle cells*

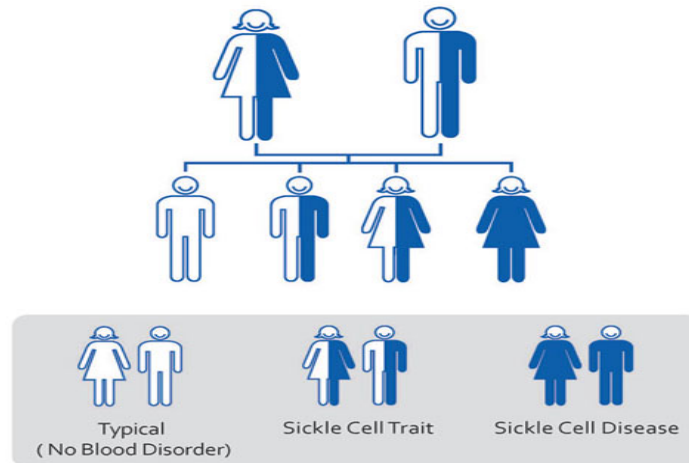
### 1.2. Causes of Sickle cell disease

The abnormal haemoglobin is also known as Haemoglobin S i.e., Sickle haemoglobin. It causes:

- It modifies the shape of red blood cell. So it becomes crescent or sickle shaped cells.
- It passes less oxygen to the body.
- It interrupts healthy blood flow or it may stop the blood flow.

### 1.3. Sickle cell trait

Sickle cell disease can be inherited from both father and mother. Inheritance of one sickle gene is known as sickle cell trait. But it will not create sickle cell disease. Sickle cells reduces the life expectancy of bloodstream into 10 to 20 days that makes anaemia. But the normal cells will live for 4 months. Figure 2 shows the sickle cell inheritance chart.



**Figure 2: Sickle cell inheritance chart**

The above figure shows that if both parents have normal cells then the child will not have any blood disorders. If anyone of the parent has abnormal cell then the child may have sickle cell trait. If both parents have abnormal cells then the child will be affected by sickle cell disease.

#### 1.4. Symptoms of Sickle cell disease

Some of the symptoms of sickle cell disease are:

- Paleness
- Rapid heart rate
- Shortness of breath
- Yellowing of the eyes and skin (jaundice)
- Dizziness
- Headaches
- Coldness in hands and feet

#### 1.5. Complication of Sickle cell disease

- Stroke
- Gallstones and cholecystitis

- Hyposplenism
- Bacterial bone infection
- Leg ulcers
- Chronic pain
- Acute chest syndrome
- Acute and painful joint crisis

### **1.6. Sickle cell diagnosis**

Sickle cell disease can be diagnosed by simple blood test. But early diagnosis is necessary to check whether the person is having sickle cell disease or not. Frequent blood test will provide better treatment to the patients. Sickle cells can be identified by microscope. So that the simple blood test is not enough to diagnose the disease. Thus this research work provides a tool which is created by using image processing technique.

### **1.7. Introduction to image mining**

Mining knowledge easily from an image is called as Image mining. Image segmentation is main process of image mining. Data mining which is used in image processing field is known as image mining. It grips with the hidden knowledge extraction, image information association with additional patterns that are not obviously gathered in images. Image mining integrates some of the techniques like image processing, data mining, computer vision, machine learning and artificial intelligence. One of the most important processes of mining is to produce all major patterns without using the prior knowledge of the patterns. Mining will be processed according to the collection of images and its relevant data. Most of the researches have been done in image mining. In this research work we examined about medical image mining.

#### **1.7.1. Medical Image Mining**

Partitioning the digital images into several segments is called as medical image mining. The main aim of this segmentation is to transform the image data into some other data that is more meaningful data which is easy to analyze. Image segmentation is used to place objects and its boundaries in images like lines, curves, and etc. We can also say the image segmentation is

the process of setting a label to each and every pixel in an image where the pixels that are having same label will share some visual characteristics.

Image segmentation results will be in set of segments together conceals the whole image or group of contours mined from the image. Each pixel in a region is related to some characteristics like color, intensity or texture of an image. When this technique is used with stack of images, the resulting contours can be used to make 3D reconstructions.

Generally, the image represents thousands of words but the pixel doesn't have that capability. Picture's color and its brightness don't provide in-depth information. These are some helpful information that is processed with a single pixel. But most of algorithms that have been used in image segmentation have a trouble while identifying the helpful region definition, because the image properties may differ from one image to another image.

Segmentation is the process of partitioning the image into meaningful, non-overlapping regions. It is a major task in complex image because it is based on the users' objective of what they need to mine from an image. Nowadays, the focus has been changed from binary labeling problem into multi-label problem. In this proposed work, we also discussed about the experimental comparison of relevant algorithms.

## **II LITERATURE SURVEY**

X.Bai and G.Sapiro has created a framework for soft segmentation that depends on the optimal, linear time and weighted geodesic distances to user-afforded scribbles from which the data is separated. The weights are logically added to the framework. To improve the results, an automatic localized refinement step is processed. One main process of this work is depends on energy formulations that are minimized by discrete optimization techniques. Graph cuts techniques delivers the foreground and background segmentation in images by using max-flow or min-cut energy minimization. The user-afforded scribble gathers the statistical information on pixels with hard constraints. The Grabcut algorithm is used to simplify the user interaction. Scribbles will be added to enhance the initial segmentation. Full color statistics will help and also it will harm the process by creating segmentation errors. So many methods have been developed to extend this framework and it aims at extending it to more dimensions.

A.Blake, C.Rother, M.Brown, P.Perez and P.Torr has described the interactive foreground and background segmentation problems created while image editing. Boykov and jolly's graph cut algorithm uses both color and contrast information with prior knowledge for region definition. At first, a general, probabilistic formulation of this model is established in the form of Gaussian mixture markov random field (GMMRF). After that a pseudo like algorithm is developed and also it studies the foreground and background's color mixture and coherence parameters. GMMRF segmentation error rates are computed by using a image database which is provided by the human segmenter. This model provides better object segmentations with little segmentation. But the pseudo limits the complexity of working models.

Y.Boykov and M.Jolly developed a new method for segmenting N-dimensional images. Here, the user indicates particular pixels as objects to produce hard constraints for segmentation. In addition to hard constraints, soft constraints like boundary and region information are produced. Graph cut algorithms are helps to identify the globally optimal segmentation of N-dimensional image. The result provides the best balance of boundary and region properties which satisfies the constraints. Interactive segmentation divides an image into two segments: objects and background. Here the user selects some hard constraints by pointing certain pixels have to be the part of an object and certain pixels have to be the part of background. Generally the hard constraints produces some hint on what the users wants to segment. Rest of the image will be segmented automatically by calculating a global optimum and it satisfies the hard constraints. Boundary and region properties are soft constraints for segmentation. This segmentation is very efficient because here the user can easily add or remove their hard constraints. It will help to get desired segmentation results quickly.

Y.Boykov, O.Veksler and R.Zabih has described the problem of minimizing a huge class of energy functions. The major limitation of energy function is that the energy function's smoothness that should involve only in pair of pixels. There are two algorithms are used to calculate a local minimum even while very large moves are allowed. First move is swap. It exchanges the labels between an arbitrary group of pixels and another arbitrary group labelled. The first algorithm creates labelling where there is no swap which decreases the energy. And second move is an expansion for a label. It allocates an arbitrary group of pixels for the label. Here the second algorithm needs the smoothness to be metric that creates a labelling where there

is no expansion which decreases the energy. One of the problems of energy minimization is enormous of computational costs.

T.Brox and D.Cremers has developed Mumford-Shah function for image segmentation. It produces the possibility to denote the regions by its smooth approximations. It computes a statistical interpretation of Mumford-shah function by associating it to recent works on local region statistics. It has more implications. At first, one can extract extended versions of Mumford-shah functional with more general distribution models. After that it directs to faster implementations. At last, the analytical expression of the smooth approximation by Gaussian convolution, the coordinate descent will be replaced by a true gradient descent.

### **III PROBLEM SPECIFICATION**

The existing research work uses an elliptical matching to identify the circular and elliptical objects in cellular cluster that has normal and elongated erythrocytes. To detect normal and elongated erythrocytes efficiently, the constraints are proposed. The main aim of existing research work is to identify the shape of erythrocytes in images of sickle cell disease peripheral blood smear samples. Here the first step is to identify the cells and cluster of overlapping cells in blood smear samples, where the segmentation is used. After segmentation the borders of analyzed objects are attained by using an automatic initialization technique depends on the gray scale intensities. Segmentation process handles the noise in the images and efficiently identifies the contours without having prior knowledge. It also removes the internal holes that have been found because of the composition of the erythrocytes. If the experiment uses overlapping circular objects then it uses circumference adjustment algorithm. If it uses overlapping elliptical objects then it uses an ellipse adjustment algorithm.

#### **Circumference adjustment**

Circumference adjustment algorithm is used to analyze the circular objects. After identifying the concave points the arches are obtained between the pair of concave points and the circular objects are detected with the best fit to the arch. Thus, all possible combinations are checked to find out the valid existing objects.

### **Ellipse adjustment**

Circumference adjustment algorithm is not fit for clusters which has elliptical shaped cells. So the system uses an elliptical adjustment algorithm to identify the circular or elliptical objects. The ellipse that relates to the adjacent concave points has been analyzed and it has been attained by processing a least square adjustment of the arch created by the  $n$  points of the contour between the two concave points. The ellipse adjustment algorithm is given as follows:

#### **Algorithm:**

- Step 1:** At first, we need to generate accumulator space that has a cell for each pixel; initially each of these pixels is set to 0.
- Step 2:** For each and every edge point of an image  $(i, j)$ : Increment all cells based on the equation of a circle  $((i-a)^2 + (j-b)^2 = r^2)$  that can be the centre of a circle, these objects are represented by 'a' in the equation.
- Step 3:** For all possible value of 'a' obtained in the previous step, discover all possible values of 'b' which satisfies the equation.
- Step 4:** Look for the local maxima cells; these can be any cells whose values are greater than every other cell in its neighbourhood. These cells can be the one with the higher probability of being the location of the circle in which we need locate.

### **3.1. Proposed System**



In our proposed work, we are using gradient watershed transform depends on the contextual information of pixels in an image. The main aim of watershed transformation is to detect regional minima in real images that are generally the regions of nuclei to be segmented. At first, we have to compute the gradient map to show the nuclei boundary. The homogeneous regions in the gradient map related to the regional minima in the given image. We need to refrain from processing nuclei detection straightforwardly on the gradient map because its noise may create over-segmentation. In nuclei initialization, important nuclei and its related cell boundaries can be attained and it is useful to design a filter that eliminates the noise from the gradient map. The resultant filtered gradient map will be subjected to nuclei detection using watershed transformation. Nuclei validation by using a shape descriptor to eliminate the outliers and it preserves the valid nuclei. If we have user interaction in analysis process, the results will be good for medical diagnosis.

Watershed transformation algorithm is depends on region processing. Result of this algorithm is global segmentation with border closure and high accuracy. It can attain one-pixel wide, closed, connected and exact place of outline. Basic concept of this algorithm is depends on visualizing a gray scale images into topographical representation which has minima, catchment basins and watershed lines. Watershed transformation algorithms have been used mostly in image processing. Some of the advantages of watershed transformation is given as follows:

- Simple
- Fast
- It can be parallelized
- It provides complete division of an image.

### **Watershed Implementation Methods**

Watershed transformation has three techniques that are:

- Distance Transform Approach
- Gradient Approach
- Marker Controlled Approach

**Distance Transform Approach:** a technique that is used normally in combination with the watershed transformation is distance transform approach. It is the distance between the pixels to the nearest non-zero valued pixel.

**Gradient Method:** this technique is helpful to preprocess a gray scale image prior to using watershed transformation. It has greater pixel values beside the object edges and it has low pixel values in rest of the image. Watershed transformation algorithm will result in watershed ridge lines beside an object edges. The topological gradient produces a global analysis of an image and it reduces the unwanted contours. In this method, we need to identify the main edges of the image processed. And we need to calculate the watershed of the gradient identified.

**Marker Controlled Methods:** watershed transform application on gradient image will result in over-segmentation because of noise. We can use markers concept to control over segmentation. Generally, a marker is a connected component fit into an image. It is used to change the gradient image. There are two types of markers that are: internal and external. Internal marker is used for an object and external marker is used for boundary. This method is robust and flexible for segmentation of objects with closed contours in which the boundaries are expressed as ridges. Normally, the markers are located inside an object of interest; internal markers related with objects of interest, external markers related with the background. After processing the segmentation, the boundaries of watershed regions are organized on the desired ridges, which divide each and every object from its neighbors. Watershed transformation algorithm is given as follows:

**Algorithm:**

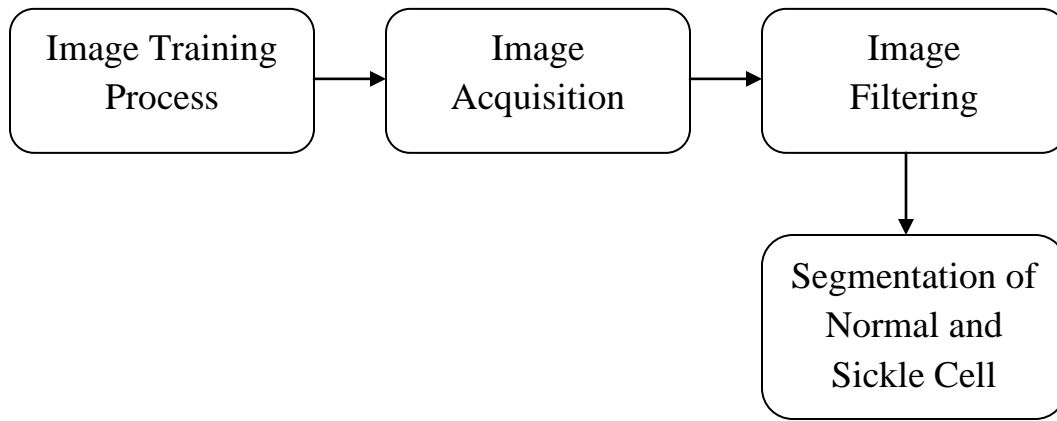
- Step 1:** At first, set of markers and pixels where the flooding will start that are chosen. Each will be given as a different label.
- Step 2:** Neighbouring pixels of each and every marked are added into a priority queue with a priority level related to the gray level of the pixel.
- Step 3:** The pixel that has lowest priority level is discovered from the priority queue. If the neighbours of discovered pixel that have been labelled as same, then the

pixel is labelled with their label. All non-marked neighbours that are not so far in priority queue those are inserted into the priority queue.

**Step 4:** Redo step 3 until the queue is empty.

Therefore, the non-labelled pixels are the watershed lines.

### Process Flow Diagram of Proposed System



*Figure 3: Classification of Normal and Sickle Cell*

## IV EXPERIMENTAL RESULTS

A segmentation result gives improved performance by comparing the existing methods. The proposed technique applies the following measures:

1. Mean Square Error (MSE): it will be calculated by averaging the squared intensity of the input image and the output images.
2. Peak Signal to Noise Ratio (PSNR): it is the ratio among the maximum possible power of a signal and the power of corrupting noise which concerns the fidelity of its representation. It can be stated by the logarithmic decibel scale. It is estimation to human perception of reconstruction quality. A high value of PSNR represents the reconstruction is of good quality.
3. Root Mean Square Error (RMSE): it concerns about the quality of the removed noise. It can be defined as  $RMSE = \sqrt{MSE}$ . RMSE worth must be low for an efficient filter.

PSNR is most easily described by the mean squared error (*MSE*). For example, a noise-free  $m \times n$  monochrome image  $I$  and its noisy approximation  $K$ , *MSE* can be defined as:

$$MSE = \frac{1}{m \cdot n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2$$

Here,  $m$ : numbers of rows of original image

$n$ : number of column of original image.

The PSNR (in dB) is termed as:

$$\begin{aligned} PSNR &= 10 \cdot \log_{10} \left( \frac{MAX_I^2}{MSE} \right) \\ &= 20 \cdot \log_{10} \left( \frac{MAX_I}{\sqrt{MSE}} \right) \\ &= 20 \cdot \log_{10} (MAX_I) - 10 \cdot \log_{10} (MSE) \end{aligned}$$

Here,  $MAX_I$  is the maximum potential pixel value of the given image. When the pixels are characterized by using 8 bits per sample, this is 255. For colour images with three RGB values per pixel, PSNR can be defined as the same except the *MSE*. *MSE* will be defined as the sum over all squared value differences divided by the image size and by three.

PSNR values in lossy image and video compression are between 30 and 50 dB, produced the bit depth is 8 bits, where higher is good. 16-bit data values for PSNR are between 60 and 80 dB. Suitable values for wireless transmission quality loss are judged to be about 20 dB to 25 dB.

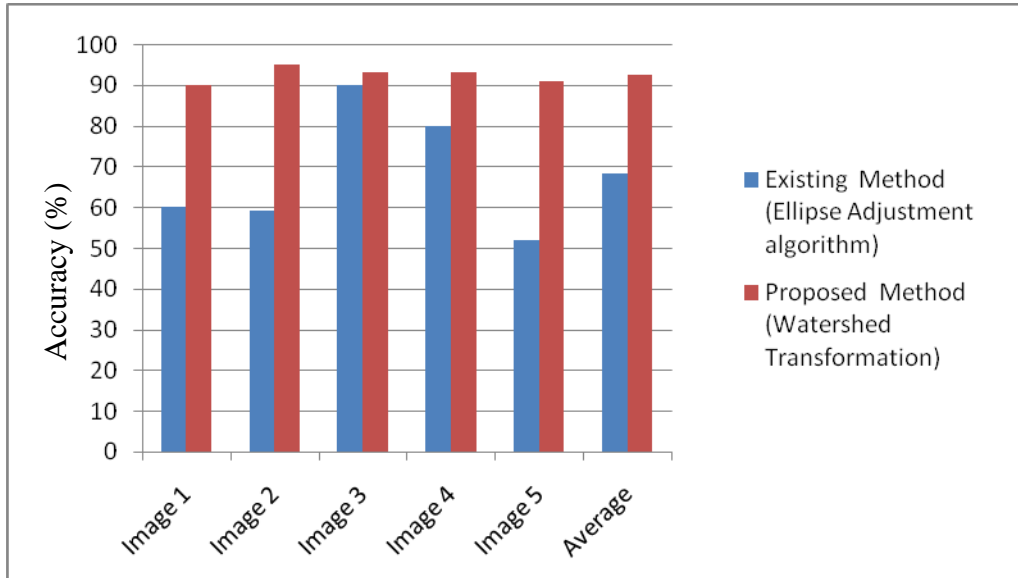


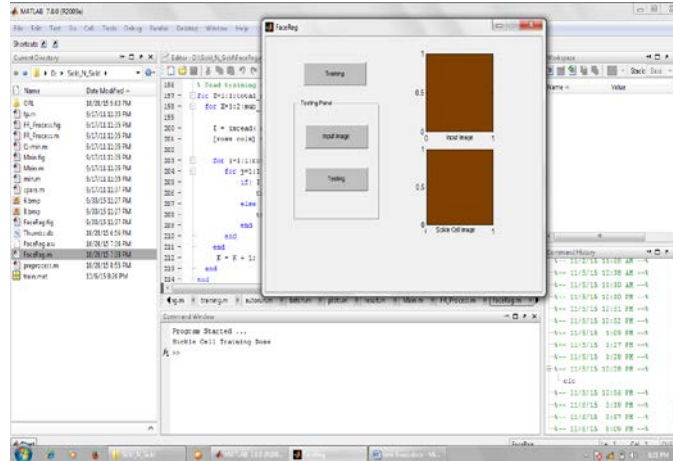
Figure 4: Accuracy

Table 1: Comparison of Precision values

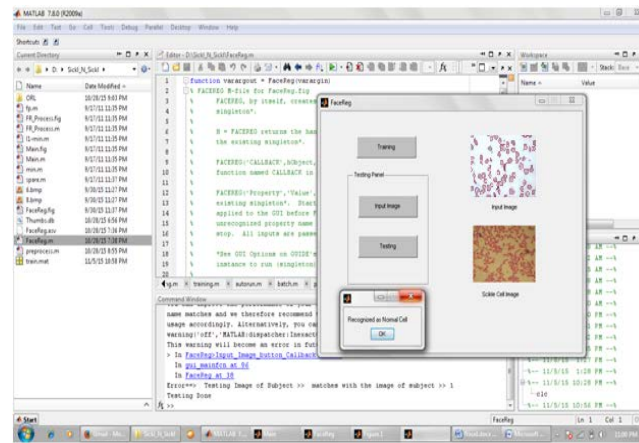
Category	Existing Method (Ellipse Adjustment algorithm) in %	Proposed Method (Watershed Transformation) in %
Image 1	60	90
Image 2	59	95
Image 3	90	93
Image 4	80	93
Image 5	52	91
Average	68.2	92.4

The comparison results of the table and chart represents the proposed method does better than the existing and it has higher accuracy. Our proposed method attained the best ellipse adjustment performance using natural images. The efficiency of proposed method is 100% while comparing with the existing methods.

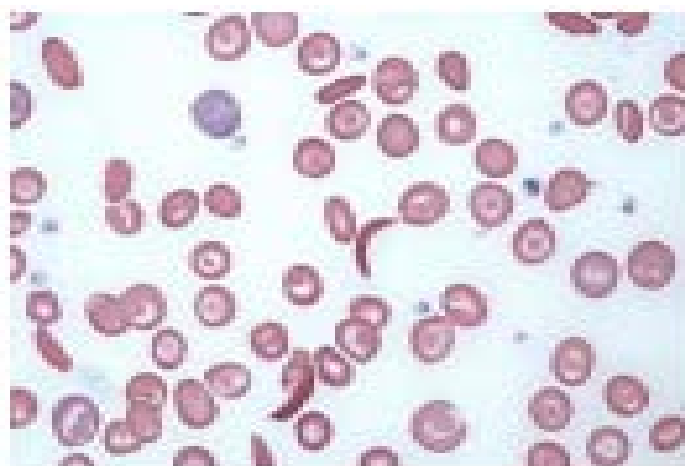
## V SIMULATION RESULTS



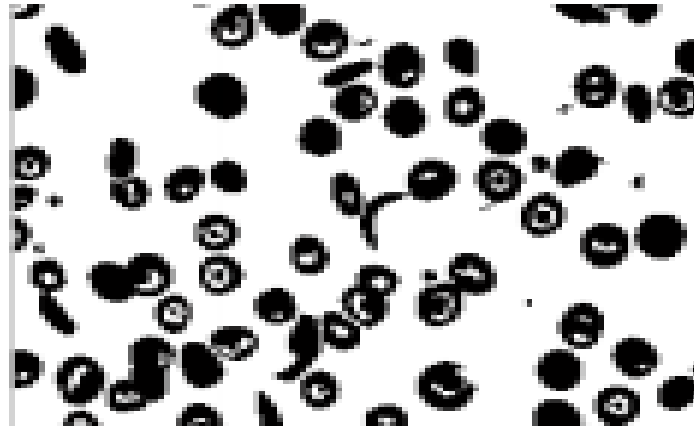
**Training Process of Sickle cells**



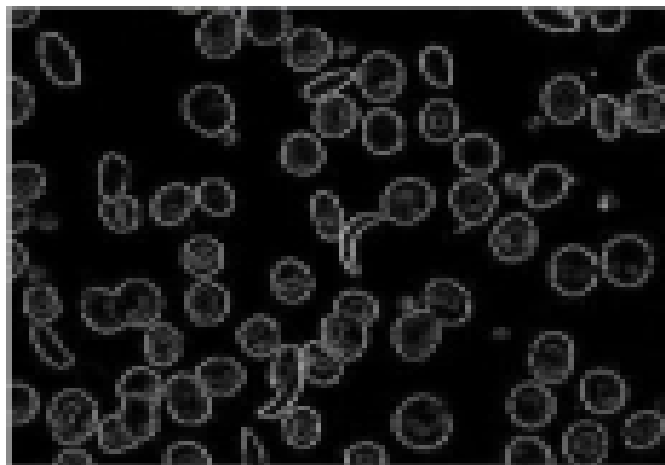
**Checks whether the given image is having Sickle cells or not**



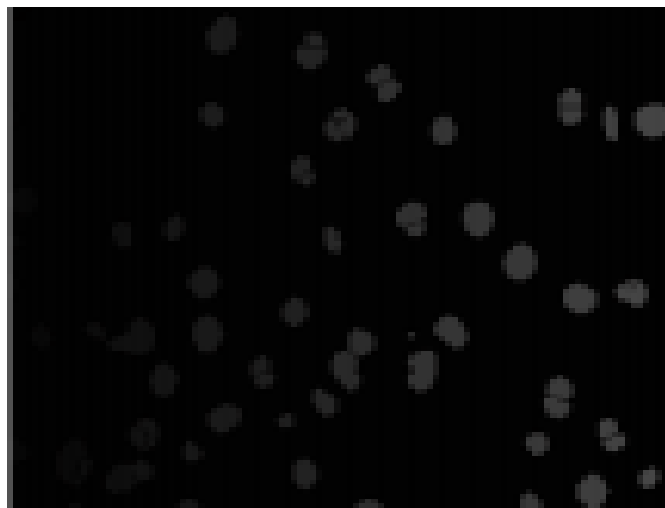
**Input Image**



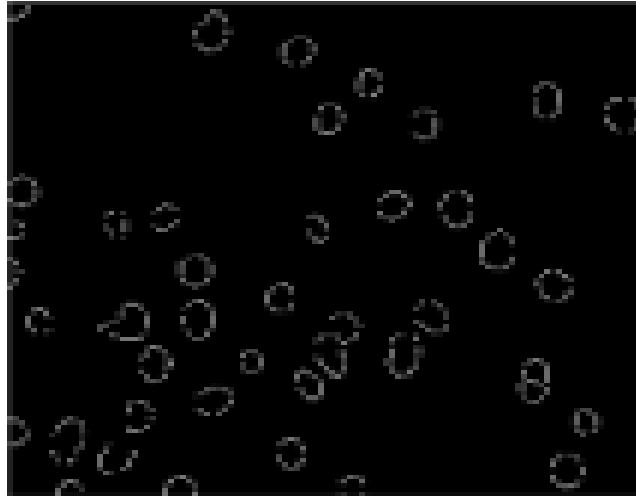
**Filtered Image**



**Gradient Magnitude Image**



**Watershed Segmentation**



**Watershed Ridge Lines**

## **VI CONCLUSION**

Computer-aided design for sickle cell mining was proposed in this work. The experimental results show that the proposed method provides objective mining results with high efficiency and consistent accuracy. In addition, the evaluated ratio values are very close to the results of manual cell selection, indicating that the proposed work considering not only the performance of analysis procedure but also practical criteria as well as clinical requirement has major potential for biomedical imaging analysis and medical values in a variety of applications. For example, it can help medical doctors non-invasively and immediately to identify early symptoms of diseases. The proposed algorithm is more objective and robust than that using manual approach, and hence medical doctors can diagnose potential diseases without the influence of any subjective factor, such as the subjective judgment of analyzer and the fatigue of the medical personnel.

## **FUTURE WORK**

The proposed method indeed saves much time and provides convincing results with interpretation and discussion by medical doctors. Therefore, the method has significant potential for biomedical imaging analysis which can be treated for noises added for future analysis and in future, we can add more images for training process of sickle cells. We can improve the accuracy



of cell segmentation and computational speed of segmentation methods as well as we can reduce the amount of over segmentation.

## **REFERENCES**

- [1] A.Blake, C.Rother, M.Brown, P.Perez and P.Torr, “Interactive image segmentation using an adaptive GMMRF model”, Microsoft Research Cambridge UK, 2004.
- [2] Amandeep Kaur, Aayushi, “Image Segmentation Using Watershed Transform”, International Journal of Soft Computing and Engineering, Vol.4, Issue 1, 2014, pp. 5-8.
- [3] Amruta Pandit, Shrikrishna Kolhar, Pragati Patil, “Survey on Automatic RBC Detection and Counting”, International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol.4, Issue 1, 2015, pp.128-131.
- [4] D.Comaniciu and P.Meer, “Mean shift: a robust approach to feature space analysis”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, pp.603-619.
- [5] Kantilal P. Rane, Gauri D. Zope, Atul Rane, “Blood Cell Segmentation – A Review”, International Journal of Advanced Electronics & Communication Systems, February, 2014.
- [6] M.Mumthaj Begam,R. Geetha and A. Sagaya selvaraj, “Red Blood Cell Identification Using Watershed Technique”, International Journal for Research in Applied Science & Engineering Technology, Volume 3 Issue IV, April 2015, pp.41-51.
- [7] M.R. Bharat Kumar and Edison Prabu, “Skin cancer detection by automatic cell segmentation and nuclear-to-cytoplasmic ratio evaluation”, International Journal of Innovative Research & Studies, Vol.3, Issue 1, 2014, pp. 347-361.
- [8] Niket Amoda, Ramesh K Kulkarni, “Image Segmentation and Detection using Watershed Transform and Region Based Image Retrieval”, International Journal of Emerging Trends & Technology in Computer Science (IJETTCS), Vol.2, Issue 2, 2013, pp. 89-94.
- [9] Subrajeet Mohopatra, Dipti Patra, Sanghmitra Satpathi, “Image analysis of blood microscopic images for acute leukemia detection”, IEEE International Conference on Industrial Electronics Control and Robotics, 2010, pp. 215-219.
- [10] T.Brox and D.Cremers, “On local region models and a statistical interpretation of the piecewise smooth Mumford-Shah functional”, International journal of computer vision, 2009, pp. 184-193.

- [11] T.Chan, S.Esedo\_glu and M.Nikolova, “Algorithms for finding global minimizers of image segmentation and denoising models”, SIAM journal on applied mathematics, 2006, pp. 1632-1648.
- [12] X.Bai and G.Sapiro, “A geodesic framework for fast interactive image and video segmentation and matting”, IEEE International Conference on Computer Vision, 2007.
- [13] Y.Boykov and M.Jolly, “Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images”, IEEE International Conference on Computer Vision. Vol. 1, 2001.
- [14] Y.Boykov, O.Veksler and R.Zabih, “Fast approximate energy minimization via graph cuts”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, Issue 11, 2001, pp. 1222-1239.
- [15] Y.Y.Chuang, B.Curless, D.H.Salesin and R.Szeliski, “A bayesian approach to digital matting”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 2, 2001.