

Computer Aided Technique to Separate the Red Blood Cells, Categorize them and Diagnose Sickle Cell Anemia

Vasundhara Acharya¹ and Krishna Prakasha^{2,*}

¹Department of CSE, Manipal Institute of Technology, MAHE, Manipal-576104

²Department of I&CT, Manipal Institute of Technology, MAHE, Manipal-576104.

Received 20 June 2018; Accepted 2 December 2018

Abstract

Anemia is a disease caused due to the lack of healthy Red blood cells. Red blood cells are mainly responsible for carrying oxygen to different parts of the body. It releases generated red blood cells into the bloodstream after the proper maturation. Anemia is primarily classified as chronic or acute. Chronic anemia continues to exist for a long time. Acute anemia occurs rapidly. It causes sickle cell anemia due to red blood cells that are in shape of a sickle or a crescent moon. The irregular shape makes it get stuck in the bloodstream which leads to a reduced flow of oxygen. The primary aim of this work is to overcome the problems involved in the manual process of detection of anemia. Manual processes require proper pipetting skills and hence is imprecise. In the proposed work, an efficient technique using a decision system is proposed that helps the pathologists to identify the disease accurately. The decision system is built by deriving the rules from the decision tree. The granulometric analysis is carried out to separate the Red blood cells from the other components of the blood. Modified Watershed transform algorithm is used to isolate cluttered cells. The work defines a systematic approach to categorize the red blood cells with central pallor and without central pallor. It puts forward the comparison of the work with state-of-the-art methods. It shows the role of the accurate segregation of Red blood cells in getting the blood cell count. The precise separation of the cluttered cells improved the accuracy of the system. Results show that the work achieved an overall accuracy of 98.5% in the identification of Sickle cell anemia and an efficiency of 97.6% in the separation of RBCs.

Keywords: Anemia; Granulometric Analysis; Modified Watershed transform; Red blood cells; Sickle cell Anemia

1. Introduction

The blood molecule consists of several types of cells. Red blood cell forms an essential part of it. Anemia is derived from a Greek term “anaimia” which means the state of lack of blood. It is a sign of disease. It is mainly classified as chronic or acute. In the case of chronic anemia, the symptoms grow gradually. In severe anemia, there will be an abrupt rise in the symptoms. Several reasons such as loss of blood during menstruation, inadequate intake of iron content, reduction in generation of erythropoietin, gain in water weight during pregnancy and production of abnormal-shaped Red blood cells (RBCs) lead to Anemia. Sickle cell anemia also termed as drepanocytosis is a disorder caused due to the presence of sickle-shaped RBCs. The mutation in the hemoglobin gene results in the generation of these type of cells. It is an inheritable disease caused when the person inherits two abnormal genes from each of the parents. Hemoglobin S is mainly responsible for the change in the shape of the RBCs. Due to its irregular shape, it leads to reduced delivery of oxygen and also it gets stuck in the blood vessels. They also tend to break into pieces.

According to data and statistics collected, people from South America, India, Turkey, Greece, and Italy suffer from Sickle cell anemia [30]. According to the estimation nearly 100,00 Americans are affected by it. It also led to a drastic reduction in mortality rates among children who have Sickle cell disease (SCD).

In the market, various techniques exist to count the RBC and diagnose the SCD. The manual method of analyzing the blood smear is highly dependent upon the skills of the pathologist. It is exhaustive and a tedious task. Hemocytometer is used commercially to obtain the blood cell count. The main disadvantage of this device is its cost which makes it infeasible to be used by most of the hospitals. The proposed method tries to overcome these drawbacks by accurately diagnosing the disease with no cost involved. The high variability of the size, shape, location, and edge makes the manual diagnosis complicated. A computer-aided system can help the physicians and the pathologists in identifying the exact type of anemia. In the proposed work, RBCs are accurately isolated from the other components of the blood using the granulometric analysis. A decision system is built by extracting the rules generated by the decision tree algorithm. It identifies the presence of different types of Red blood cells (RBCs). The percentage count of sickle cells computed helps in identifying the frequency of occurrence of the Sickle cell anemia. A

*E-mail address: kkp.prakash@manipal.edu

ISSN: 1791-2377 © 2019 Eastern Macedonia and Thrace Institute of Technology. All rights reserved.
doi:10.25103/jestr.122.10

comparison is made between the segmented image and ground truth determined by the domain expert. Fig. 1 depicts the blood smear extracted from a patient suffering from Sickle cell anemia. Fig. 2 depicts the blood smear extracted from a normal case.

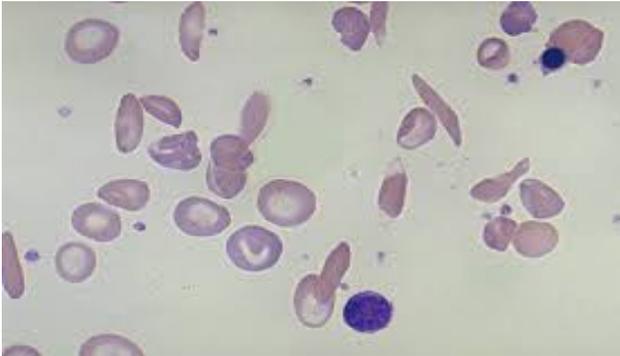


Fig. 1. A case of Sickle cell anemia

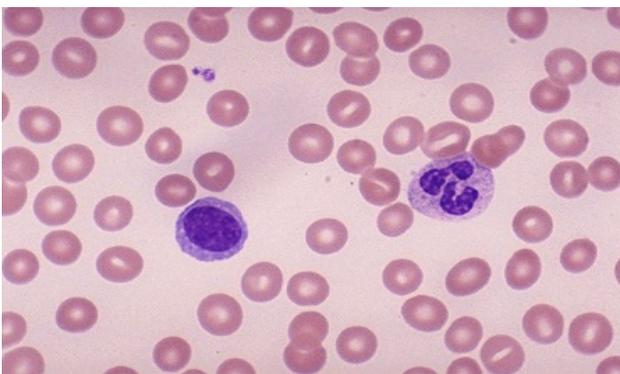


Fig. 2. A normal peripheral blood smear

2. State of the Art

In this section, a review of the research work done to segment the Red blood cells from blood smear is presented. The techniques used by various authors to identify and detect different types of anemia is analyzed.

Patil, Sable, & Anandgaonkar [13] proposed a methodology to identify and count the RBC. Global Otsu's threshold was used to convert the image to binary. Labeling algorithm performed the cell count. A threshold value of 140 was used to convert the RGB image to binary. Circular Hough Transform (CHT) with multiple iterations was used by Alomari et al. [1]. Bhagvathi and Thomas [3] extracted the red component from RGB image and used a FIS editor with rules to detect the circles. The number of circles gave the RBC count. The RBCs were separated from WBCs using the Hough transform. A counter was set up to count the cells by Maitra & Gupta [9]. Thejashwini & Padma [21] extracted only the green component from the RGB image to highlight the RBCs. Circular Hough Transform accomplished the cell count. The input image was cropped to obtain sub-images to compute the minimum and maximum radius by Mazalan, Mahmood, & Razak [10]. The computed radius value was given as input to Circular Hough transform algorithm to count the RBCs(CHT). The second component of YCbCr format was extracted to highlight RBCs present in the blood smear and some morphological operations were applied by Tulsani, Yadav, & Saxena [22]. The watershed transform was used to separate the overlapping cells, and the counter was set up to count the RBCs. Sreekumar & Bhattacharya [20], presented an approach for the counting of the cells using Circular Hough

transform. They assumed that WBCs were circular than RBCs. Radius range for the Circular Hough transform was obtained using a distance tool. Chintawar et al. [5] computed the mean radius value to detect the sickle cells. 12 different categories of anemia were diagnosed using a system, and a counter was set up to count the proportion of each type of cell. The form factor was computed to detect the presence of different kinds of RBCs such as Elliptocytes, Dacrocytes by Patil and Khot [12]. A specific range of Form factor values was set to classify the RBCs into its different types. Sahu, Biswas, and Uma [19] used fractal dimension approach to identify the sickle cells. The highest, lowest and mean radius were computed to analyze the shape of the Sickle cell. Standard cell size was used as a measure for the comparison. The ratio between the major axis length and minor axis length identified the sickle cells. Component labeling algorithm was used to count the number of cells by Rexcy, Akshaya and Swetha [18]. Rakshita and Bhowmik [15] presented a technique using morphological operations and edge detection to identify the Sickle cells. Sobel operator performed the edge detection. The form factor was computed. Normal cells had a form factor in the range of [0.5-1] and below were considered abnormal. Revathi & Jeevitha [17], proposed a technique to diagnose the Sickle cells. Regional minima in real images were obtained using the gradient watershed transform. The resultant filtered gradient map was subjected to nuclei detection. Valid nuclei were retained by using the shape descriptors. The features such as Root Mean Square Error, Peak Signal to Noise Ratio and Mean square error were computed. The proposed method attained the best ellipse adjustment performance using natural images. S. Bala and A.Doegar [25] introduced a technique to detect the sickle cells using the watershed segmentation. The form factor was computed to identify the abnormal cells. The sickle cells and target cells were highlighted using the contour plot. M.G. Hidalgo et al. [26] proposed a methodology to detect the sickle cells using ellipse adjustment. The concave and convex points present in the contour were identified using a new algorithm. A set of constraints replaced the steps involved in the preprocessing of the image. Parvathy et al. [27] proposed a technique to detect the sickle cells using Otsu's segmentation and watershed segmentation. Form factor and compactness were computed to differentiate a normal cell from an abnormal one. M. Veluchamy et al. [28] proposed a technique to identify normal and abnormal cells using artificial neural networks. The image was segmented using some morphological operations and thresholding technique. The first, second order gray level statistics and algebraic moment invariants were then used to extract the features. The features were then analyzed to differentiate between the normal cell and abnormal cell. J. Poomcokrak and C. Neatpisarnvani [29] proposed a technique to identify the Sickle cells using a neural network. Edge detection algorithm was employed to get the individual RBCs. The neural network was then used to analyze the individual cells and classify them. Fadhel et al. [32] proposed a method using CHT and watershed transform to detect the sickle cells. The normal RBCs were identified using CHT as they were circular. During the watershed segmentation, the effect factor was computed to differentiate between the normal RBCs and the sickle cells. A counter was set up to achieve the count. Sharma et al. [33] proposed a technique to diagnose sickle cell anemia and thalassemia. The touching erythrocytes were separated using the marker-controlled watershed transform. K Nearest Neighbor was used to

classify the extracted features. Rodrigues et al. [34] proposed a technique to differentiate between erythrocytes. The global Otsu's threshold was employed to convert the image to binary. After the application of morphological operations, feature extraction was performed. The features were then classified using some classifiers and the best one was chosen. Chen et al. [35] proposed a method to diagnose hemolytic anemia. The overlapping cells were isolated using the concavity information. The different types of hemolytic anemia were classified using different classifiers.

2.3 Problem statement

Based on the literature, diagnostic framework should incorporate the key necessities mentioned below.

1. Effective algorithm to separate the other components of blood from the RBCs.
2. An Automated and effective algorithm to count the RBCs.
3. An accurate system to diagnose the Sickle cell anemia.

3. Methodology

The proposed framework is described in this section. The images are collected from the internet. The experiment is performed on 15 images. Fig.3 depicts an overview of the proposed method to get the WBC nucleus. The initial process includes converting the image to LAB color space. It achieves color based image segmentation using the K-medoids algorithm. This is followed by identifying the cluster with WBC nucleus. Fig. 4. Depicts the proposed method to separate the Red blood cells from other cells. It then computes Euler number to identify RBCs with central pallor and without central pallor. It describes the proposed algorithm to isolate the RBCs and the Decision system to identify in the following sections. The results got are compared with the ground truth given by the domain experts.

3.1 Identification of WBC nucleus and Nuclear lobes computation

A novel K-medoids algorithm based WBC nucleus detection algorithm with modified watershed transform is proposed. The motivation behind developing such an algorithm is (i) K-medoids algorithm is robust to external noise. ii) Existing watershed transform algorithm has the problem of over-segmentation. The proposed methodology consists of mainly three stages as specified below:

• Image Pre-processing

The image is converted from RGB to LAB color space. It is a color-opponent space which consists of a luminosity layer 'L*', chromaticity-layer 'a*' indicating where the color falls along the red-green axis, and chromaticity-layer 'b*' indicating where the color falls along the blue-yellow axis. It is derived from CIE XYZ tristimulus values. A mixture of the three CIE primaries, X, Y, Z. That mixture may be specified by three numbers X, Y, Z called tristimulus values frames any color on CIE chromaticity diagram. The color is represented by the Eq.1.

$$c = xx + yy + zz \quad (2)$$

The tristimulus values are computed using the CIE primaries [6] and associated color matching functions.

Fig.5 depicts the blood smear image of Sickle cell anemia converted to LAB color space.

• Image Segmentation

Color image segmentation is performed using K-medoids algorithm. Initially, a comparison has been made between K-Means and K-medoids algorithm. The K-Means algorithm takes less execution time compared to K-medoids algorithm but is less robust. Since in the medical field accuracy is of utmost importance we choose K-medoids algorithm to segment the image. The image is reshaped into a vector and is fed to K-medoids algorithm. K-medoids clustering takes two parameters as input. A distance metric to quantify how close two objects are to each other and the number of clusters to be partitioned. Since the color information exists in the a^*b^* space, pixels with a^* and b^* values are chosen. K-medoids is applied to the image with 4 clusters to obtain nucleus, cytoplasm, red blood cells and background separately in each cluster. Euclidean distance is the measure used here. The pseudo code for K-medoids algorithm is given in algorithm 1. The similarity in image regions is considered here. The total cost of configuration denotes the cost involved in swapping a pair of non-selected object 'o,' and selected object 'm.' The cost function is determined by the difference in absolute error-value when a current representative object is replaced by a non-representative object. Clustering is done on color values of the pixels using the distance formula. K-medoids returns an index corresponding to a cluster for every object present in the input. Every pixel in the image is labeled with an index corresponding to a cluster. The objects in the image are separated using the pixel labels. The mean value of the segmented images is obtained along the red channel. Separation of the nucleus from the cytoplasm takes place at this stage. After repeated iterations and analysis it was found that the nuclei in the image will have the minimum intensity along the Red channel. Fig 6. Illustrates the images present in each cluster after the color based image segmentation using K-medoids algorithm. Fig 7. Time consumed by both the algorithms concerning the number of white blood cells is depicted.

Algorithm 1: Pseudo code of K-medoids Algorithm

Input: S (image as vector), K (number of cluster), distance metric.

Output: Clusters

1. Choose the Medoids by selecting k of the n data points
2. Compute the closest Medoid using Euclidean distance and associate each data point
3. for each Medoid p do
4. for each non-Medoid data point a: do
5. Swap p and a and calculate the total cost of the configuration
6. End for
7. End for
8. Find the configuration with the lowest cost and choose it.
9. Repeat steps 2 to 7 until there is no change in the Medoid

• Image Post-Processing

The segmented image having the nucleus is converted to grayscale. The image is converted to binary by using a level value of 0.1. The level value of 0.1 is chosen after analysis of different images. Area open function with a value of 200 is used to remove small particles. Area open function takes

two parameters. The first one is the binary image and the second one is the number of pixels 'P'. It intends to remove all connected components that have fewer than P pixels from the binary image BW, producing another output binary image. The connectivity value for two dimensions is 8 and 26 for three dimensions [2]. Pixels at the border which do not contribute much to the computation are eliminated using the border clear operation. The primary purpose of the clear border operation is to suppress the structures that are lighter than their surroundings and the ones that are connected to the image border [4]. The presence of multiple nuclear lobes is obtained using modified Watershed transform [24]. A mask having small spots placed at the center of each cell to be segmented (to generate local minima only at the desired locations) is used here. The raw watershed transform tends to perform the over-segmentation. It is because even the tiny local minima become the catchment basin. The watershed transform is modified so that no minima occur at the filtered out location. The nuclear lobes are obtained as a result of above operation. The lobes are then labeled using the "bwlabeln" function [8]. It returns a label matrix with labels for connected components in the binary image. The mask values, level value, and area open function value is chosen after generalizing over the images in the dataset. Fig. 8 depicts the two nuclear lobes extracted using the proposed algorithm. Presence of more than four nuclear lobes helps the pathologist to diagnose the presence of the megaloblastic anemia.

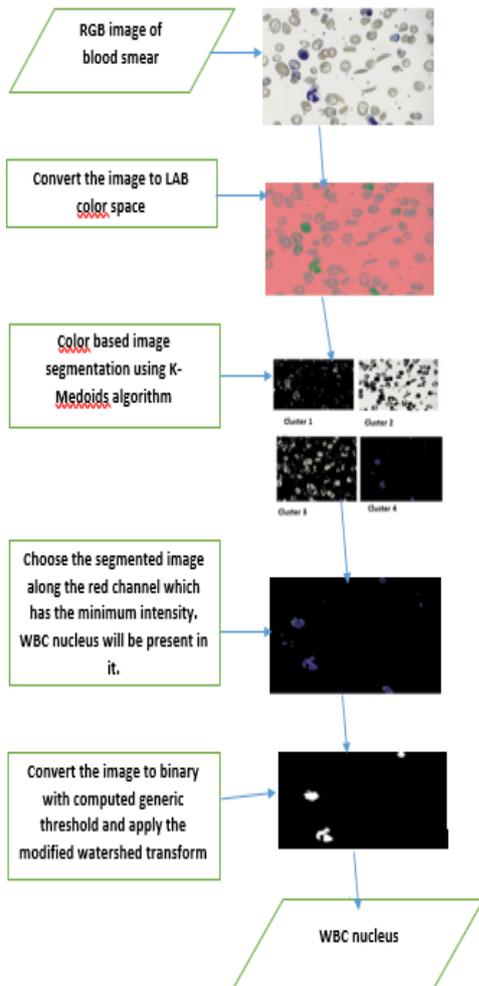


Fig. 3. Proposed methodology to obtain the WBC nucleus.

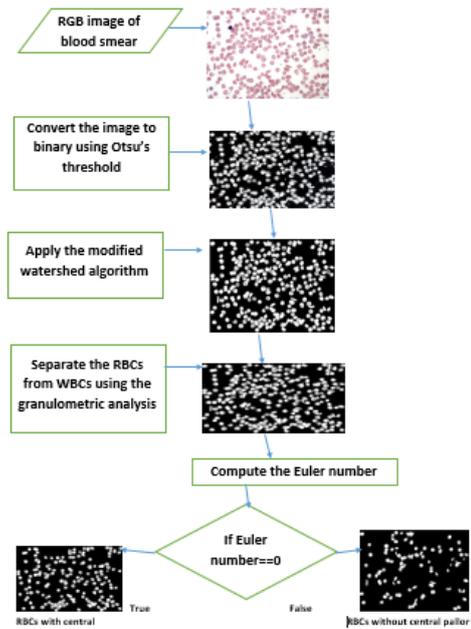


Fig. 4. Proposed system to separate the RBCs from WBCs and Identify RBCs with central pallor and without central pallor.

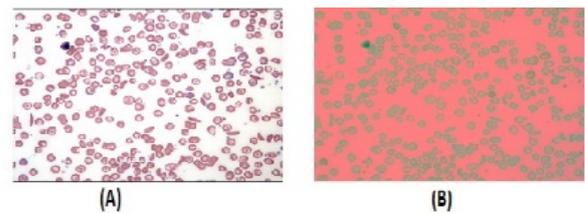


Fig. 5. (A) Original Image. (B). Image converted to LAB color space

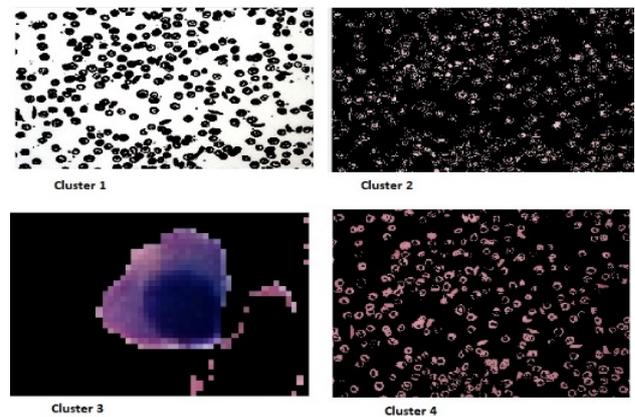


Fig. 6. Four clusters obtained as a result of color based segmentation using K-medoids algorithm.

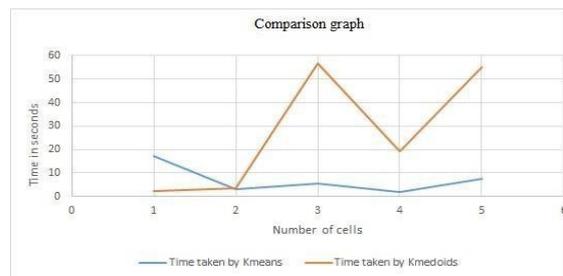


Fig.7. Comparison between K-medoids and K-means algorithm with respect to the time taken for execution.

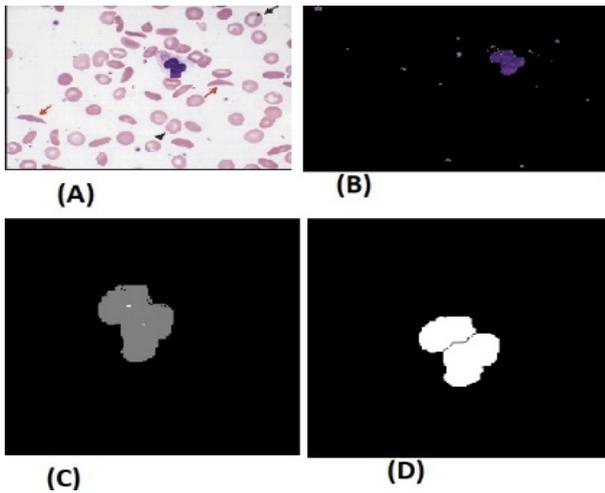


Fig 8. (A) Original Image (B). WBC nucleus. (C). Spots generated using extended minima function (D). WBC nucleus with two lobes after the application of modified watershed transform.

3.2 Separation of RBCs from other components in the blood smear

The steps followed in the separation of RBCs from other components are given below:

•Image Pre-Processing

The input to the system is the blood smear images stained with Wrights stain. The images are collected from normal patients and patients with blood disorders. The image is converted to grayscale by eliminating the Hue and Saturation element and luminance is retained.

• Image Segmentation

Global Otsu's threshold is used to convert the image to binary as it minimizes the intra-class variance. The level is calculated using this algorithm. The resulting image is converted to binary using the global Otsu's threshold and is complemented. The global Otsu's threshold uses a threshold that helps to reduce the intraclass variance of the black and white pixels. It is the weighted sum of variances of the two classes. It is given in Eq. 2.

$$\sigma_w^2(t) = w_0(t)\sigma_0^2(t) + w_1(t)\sigma_1^2(t) \quad (2)$$

The probabilities of the two classes is given by the weights w_0 and w_1 . The threshold is represented by t . Variances of the two classes is given by σ_0^2 and σ_1^2 . The class probability is computed from the L bins of the histogram. It is computed using the Eq. 3 and Eq.4.

$$w_0(t) = \sum_{i=0}^{t-1} p(i) \quad (3)$$

$$w_1(t) = \sum_{i=t}^{L-1} p(i) \quad (4)$$

The intra class variance is minimized as given in Eq.5.

$$\sigma_b^2(t) = \sigma_2 - \sigma_w^2(t) = w(\mu_0 - \mu_T^2)_+ w_1(\mu_1 - \mu_T^2) \quad (5)$$

Where w is class probabilities and μ is the class mean. The class mean is computed using Eq. 6, Eq.7 and Eq. 8.

$$\mu_0(t) = \sum_{i=0}^{t-1} i.p(i)/w_0 \quad (6)$$

$$\mu_1(t) = \sum_{i=t}^{L-1} i.p(i)/w_1 \quad (7)$$

$$\mu_T = \sum_{i=0}^{L-1} i.p(i)/w_1 \quad (8)$$

Platelets are eliminated using an erosion operation with a disk shaped structuring element of size 1. Pixels at the border which do not contribute much to the computation are eliminated using the border clear operation.

•Image Post-Processing

The blood smear image will have cells overlapping each other and will be tightly attached to each other. Touching cells if not separated generate inaccurate results. Modified watershed transform is applied to it as explained in section 2.1. The noise is filtered using a Median filter with a 2*2 neighborhood. The procedure employed to separate the RBCs from other components is explained in algorithm 2. Fig. 9 depicts the parameters passed to XOR operation.

Algorithm 2: Pseudo code for the separation of RBCs from other components in the blood

Input: Binary image having all the cells obtained after image post-processing.

Output: Image with only RBCs

1. Acquire the number of WBC nucleus present in the input image.
2. Using the "regionprops" function, compute the area of the cells in the image
3. Use an array to store the area values in descending order.
4. for 1 to Number_of_WBC_nucleus do
5. Apply XOR operation on the binary image with lower bound value as an area open function with two parameters. One is the binary image, and other is value one. Upper bound value as another area open function with two parameters. One is the binary image, and the other is the area of the cells that is computed.

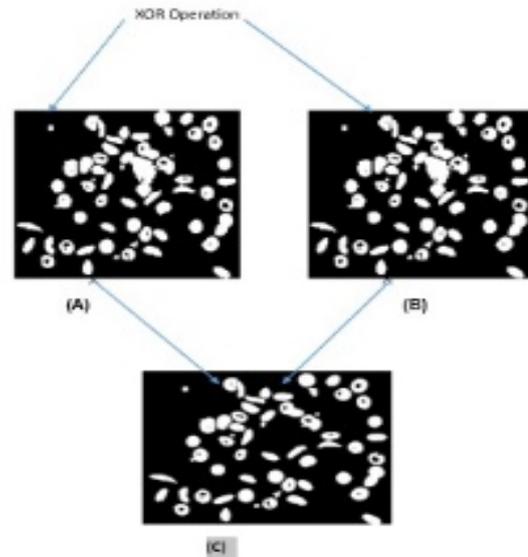


Fig 9. (A) Binary Image with RBC given as first parameter to area open function (B). Binary Image with RBC given as second parameter to area open function along with maximum area (C). Binary image with only RBC

3.3 Categorization of RBCs in to with central pallor and without central pallor

The proposed work intends to categorize the RBCs in to with central pallor and without central pallor. Normal RBCs will

Attribute	Id
Area	1
Perimeter	2
Shape geometric feature	4
Area proportion	5
Deviation	6
Form Factor	7
Erythrocyte color value	8
Eccentricity	9

have a central pallor equal to one-third of its diameter. Hypochromic RBCs will have a central pallor greater than the normal RBCs. Hence, to differentiate between them the cells with central pallor are first separated from the cells without central pallor. Features are computed for each case to check if they are normal or hypochromic [16]. Euler number is first computed to separate them [7]. It is the difference between the total number of objects present in the image and the total number of holes in that object. The computation is shown in Eq. 9 and Eq.10.

$$E = 1 - H \tag{9}$$

Where H is the holes present in the binary image
E is the Euler number

The number of holes can be computed using Eq.10.

$$H = NCC - 1 \tag{10}$$

Where NCC represents the four or eight connected objects in the binary image.

Hence, the RBCs with central pallor would have a Euler number equal to zero. Fig. 10 shows RBCs with central pallor and RBCs without central pallor. Fig. 11 depicts categorization of RBCs after the computation of the Euler number.



Fig 10. (A) RBC with central pallor. Euler number computed is zero (B). RBC without central pallor. Euler number computed is one

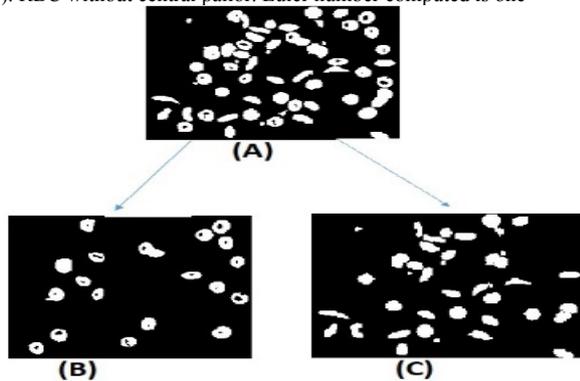


Fig 11. (A) Image with RBCs. (B). RBCs with central pallor. (C). RBCs without central pallor.

3.4 Features Computation

The proposed work intends to extract features that helps in categorizing the RBCs and also helps in diagnosing Sickle cell Anemia. The features are extracted using the “regionprops” function [11]. Tab.1 gives the list of features extracted. The description of each feature is given below.

Table 1. List of Features

Area: It represents the actual number of pixels in the region.

1. Perimeter: Distance around the boundary of the region. The area and perimeter are required for the calculation of the form factor.
2. Diameter: It is the ratio between the Area and perimeter.
3. Shape Geometric Feature: The Red blood cells are going to be sometimes oval in form. Normal cells, Hypochromic red blood cell are going to be circular in form that may be a special oval with equal diameter. The procedure for the computation of diameter values is given in algorithm 3.

Algorithm 3: Pseudo Code for the Calculation of Diameter Values

Input: Binary image having the RBCs

Output: Maximum and Minimum diameter value of cells.

1. Use an area open function with value 200 to eliminate the unwanted particles. Label the objects and identify their boundary.
2. For each object measure the centroid value and compute the number of Blobs.
3. for each BLOB do
4. Obtain the coordinate values of the centroid.
5. Measure the distances of boundary pixels from the centroid.
6. Identify the pixel that is farthest and closest from the centroid by scanning the boundary.
7. Compute the maximum and minimum radius values by calculating the maximum and minimum value of the distance.
8. Use an array to record the above values.
9. End for
10. for each entry in the array do
11. Multiply the maximum distance with 2 as the diameter is twice the value of radius to obtain the large diameter. Similarly, multiply the minimum distance with 2 to calculate the small diameter value.
12. End for

4. Area Proportion: It is the ratio between the central pallor area to the cell area. It is used to differentiate a Hypochromic cell from the other cell as the Hypochromic cell will have larger central pallor.

The procedure to compute the central pallor area and cell area is given in algorithm 4.

Algorithm 4: Pseudo code for the Calculation of ratio between central pallor area and the cell area

Input: Binary image having the RBCs

Output: Central pallor area and cell space.

1. Extract every cell one by one from the image utilizing the bounding box technique.
2. Use ”bweuler” function to calculate the Euler number of every cell object.
- 3 if Euler number is equal to zero then
4. Use the” imfill” function to fill the opening within the object and then compute the outer blob cell area.
5. Label the blob within the image using the labeling algorithm and then compute its area.

6. Get the label of the biggest blob and compare the label with others.
7. Get the logical mask of the biggest blob.
8. Use the "regionprops" function to calculate its area.
9. Deduct the original image from the outer blob to get the central pallor region and repeat steps from 5-7.
10. Calculate its area using the "regionprops" function.
11. Calculate the ratio between areas of pallor and cell.
- 12 End if

- 5.Deviation: It is the ratio between the shape geometric factor and area of the object.
- 6.Form Factor: It is a measure of irregularity. It is given as the ratio between Area and perimeter.
- 7.Erythrocyte color value: The green and red scale value of center point of the cell gives details about the variations in the color values. It Is measured using the "impixel" [14] function.
- 8.Eccentricity: The sickle cells are identified with this feature. The circular RBCs will have an eccentricity value of 0, and non-circular RBCs will have an eccentricity value nearing to one. The average of eccentricity values of all RBCs is computed. The "mean" function is used to obtain the average value.

3.5 Decision System

The proposed work intends to detect the presence of Sickle cell Anemia. The extracted features are fed into the decision tree classifier [36]. The tree classifies the RBCs into different types. The count of cells that fall under each category of RBC is computed. The percentage of each type of RBC is then calculated. To achieve accurate results, the different classes of RBCs and their count is taken into consideration. The different categories of RBCs are given in Tab 2.

Table 2. Types of RBCs

RBC name	Class_Id
Normocyte	1
Normocyte without central pallor	2
Macrocyte without central pallor	3
Spherocyte	4
Normochromic macrocyte	5
Hypochromic microcyte	6
Hypochromic macrocyte	7
Hypochromic normocyte	8
Normochromic microcyte	9
Elongated cell	10
Sickle cell	11

The description of each type of RBC is given below:

- 1.Normocyte: A RBC with normal size and rich hemoglobin content.
- 2.Normocyte without central pallor: RBC with normal size with no central area of pallor.
- 3.Macrocyte without central pallor: RBCs larger than normal RBCs in terms of size and without central area of pallor.
- 4.Spherocyte: RBCs lose their biconcave shape and become spherical. They lack central area of pallor.

5. Normochromic macrocyte : RBC with normal hemoglobin content but size is larger than normal RBC.
- 6.Hypochromic microcyte: Small RBC with reduced hemoglobin content.
- 7.Hypochromic macrocyte: Large RBC with reduced hemoglobin content.
- 8.Hypochromic normocyte: RBC with normal size and reduced hemoglobin content.
- 9.Normochromic microcyte: RBC smaller than normal RBC and with normal hemoglobin content.
- 10.Elongated cell: RBCs with normal central pallor and having hemoglobin content concentrated near the ends.
- 11.Sickle cell: Crescent shaped RBCs with reduced ability to carry oxygen.

Algorithm 4 shows the pseudocode with set of rules to categorize the RBCs. The features extracted is fed in to decision tree algorithm. The rules are generated by the decision tree classifier [37]. The counter is maintained to keep track of RBC count under each category. It helps in calculating the percentage in later stage.

Algorithm 5: Pseudo Code with rules to categorize RBCs

Input: Binary image having the RBCs

Output: Cells classified in to its categories.

- 1.Compute the Euler number of each object present in the image.
- 2.if Euler number==0 then
- 3.for Each object do
- 4.if f1 < 1.2 && 6.5 <f2<8.5 && f3 >0.2 then
- 5.Classify as class 9
- 6.if 6.5<f2<8.5 && f3<0.2 && f1 <1.2 then
- 7.Classify as class 1
- 8.if f2>8.5 && f3>0.2 && gvalue>140 then
- 9.Classify as class 8
- 10.if f1 >1.2 && f4<0.2 && 3<f2<8.5 && gvalue>140&& f3>0.2 then
- 11.Classify as class 7
12. if f1>1.2 && 3<f2<8.5 && f3>0.2 then
- 13.Classify as class 10
- 14.if f1>1.2&& 6.5<f2<8.5 && f3<=0.2 && f4<0.2 && gvalue>140 then
- 15.Classify as class 6
- 16.if f1>1.2 && f4<0.2 then
- 17.Classify as class 5
- 18.End for
19. if Euler number is not equal to 0 then
20. for Each object do
21. if f1>1.2 && f2>8.5 && 100<gvalue<200 then
22. Classify as class 3
- 23.if f1>1.2 && 3<f2<7.5 && gvalue<140 && rvalue>140 then
24. Classify as class 4
- 25.if f1<1.2 && 7.5<f2<12 then
26. Classify as class 2
27. if f5<0.6 && f6>= threshold then
28. Classify as class 11
- 29.End for

The term f1 represents the shape geometric feature.

Diameter is represented using f2. Area Proportion is represented using f3. The deviation is shown using f4. Form factor is represented using f5. The eccentricity is represented using f6.

4. Results

The experiment is performed on a dataset with 15 images. The smear was colored with different stains. The ground truth of the images is determined by the domain expert. The ground truth refers to the diagnosis done by the domain expert which helps in measuring the efficiency of the proposed system. Out of the 10 images, 8 images were labeled as Sickle cell anemia and one was normal and the other was megaloblastic anemia. The accuracy of the system is determined by computing the precision and the recall. Fig 13-19 shows the segmentation on a blood smear with Sickle cell anemia. Fig. 20 shows the segmentation on a blood smear of a normal case. Fig. 21 shows the segmentation on a blood smear of a megaloblastic anemia case.

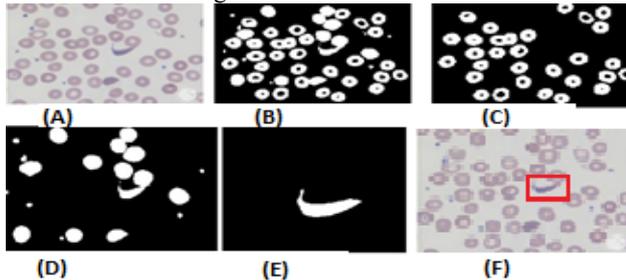


Fig.12. (A) Blood smear image of patient with Sickle cell anemia. (B) Binary image having all the cells. (C) RBCs with central pallor. (D) RBCs without central pallor. (E). Sickle cell diagnosed by the system. (F). Sickle cell labeled by the expert.

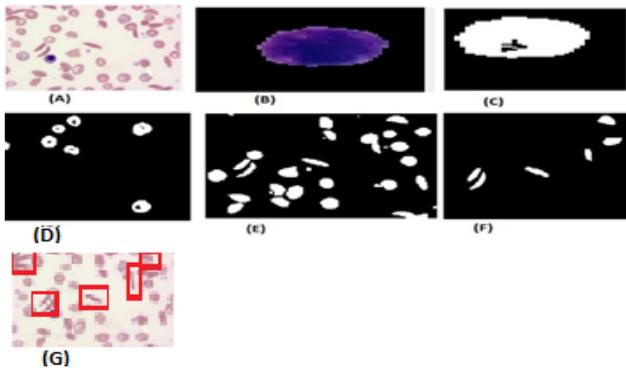


Fig.13. (A) Blood smear image of patient with Sickle cell anemia. (B) WBC nucleus. (C) Lobes of nucleus. (D) RBCs with central pallor. (E) RBCs without central pallor. (F). Sickle cell. (G). Cells labeled by the domain expert.

Fig. 14 and Fig.15 shows an instance where the blood smear image does not have a WBC. The system can accurately identify the sickle cells, but some sickle cells are lost which are located near the border.

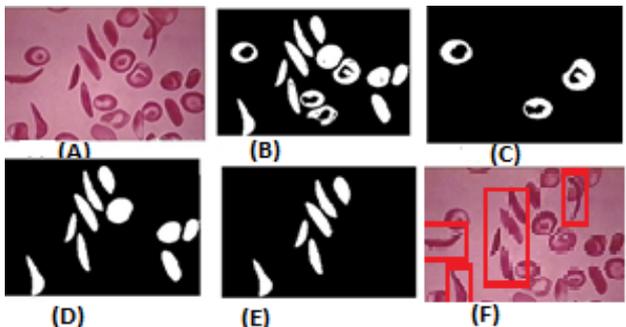


Fig.14. (A) Blood smear image of patient with Sickle cell anemia. (B) Binary image having all the cells. (C) RBCs with central pallor. (D) RBCs without central pallor. (E). Sickle cells diagnosed by the system. (F) Sickle cells labeled by the expert.



Fig.15. (A) Blood smear image of patient with Sickle cell anemia. (B) Binary image having all the cells. (C) RBCs with central pallor. (D) RBCs without central pallor. (E). Sickle cells diagnosed by the system. (F) Sickle cells labeled by the expert.

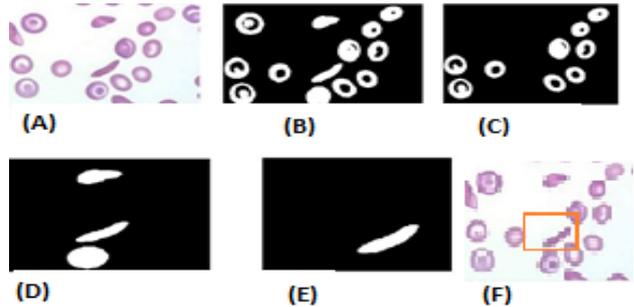


Fig.16. (A) Blood smear image of patient with Sickle cell anemia. (B) Binary image having all the cells. (C) RBCs with central pallor. (D) RBCs without central pallor. (E). Sickle cell diagnosed by the system. (F) Sickle cell labeled by the expert.

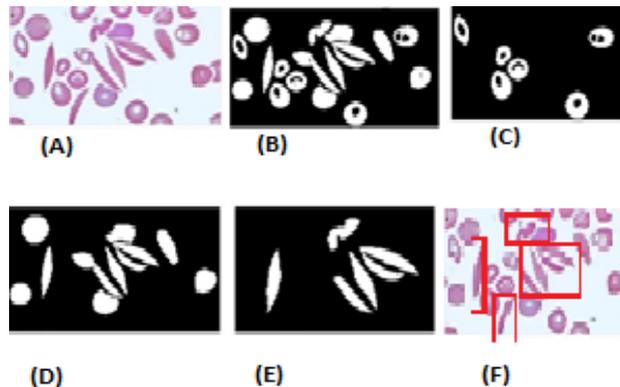


Fig.17. (A) Blood smear image of patient with Sickle cell anemia. (B) Binary image having all the cells. (C) RBCs with central pallor. (D) RBCs without central pallor. (E). Sickle cells diagnosed by the system. (F) Sickle cells labeled by the expert.

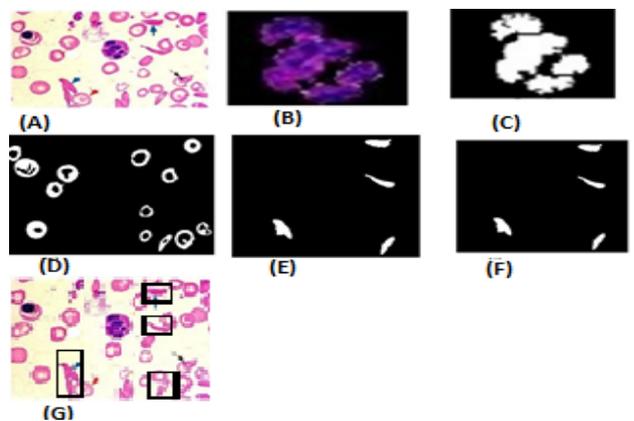


Fig.18. (A) Blood smear image of patient with Sickle cell anemia. (B) WBC nucleus. (C) Lobes of nucleus. (D) RBCs with central pallor. (E) RBCs without central pallor. (F). Sickle cells diagnosed by the system. (G) Sickle cells labeled by the expert.

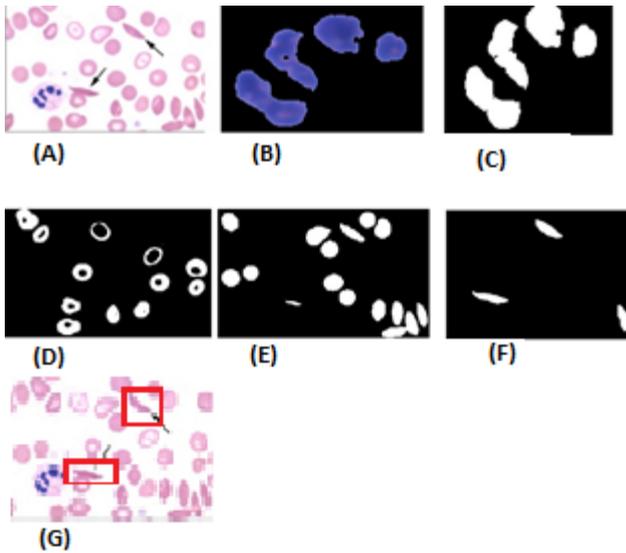


Fig.19. (A) Blood smear image of patient with Sickle cell anemia. (B) WBC nucleus. (C) Lobes of nucleus. (D) RBCs with central pallor. (E). RBCs without central pallor. (F). Sickle cell. (G).

The count of different categories of RBCs obtained for each image is tabulated in table 3. “a” denotes the count of Normocyte. Normocyte without central pallor count is depicted by “b.” Macrocyte without central pallor count is shown by “c.” Spherocyte count is depicted using “d.” “e” denotes the count of Normochromic macrocyte. Hypochromic microcyte count is denoted by “f.” “g” denotes the count of Hypochromic macrocyte. Hypochromic normocyte count is shown using “h.” Count of Normochromic microcyte is shown using “I.” Elongated cell count is depicted using “j.” “k” denotes the Sickle cell count. In order to verify the prediction capacity of the proposed system, one image of a normal case and another image with megaloblastic anemia is fed in to the system. The results are shown in Fig.20 and Fig.21.

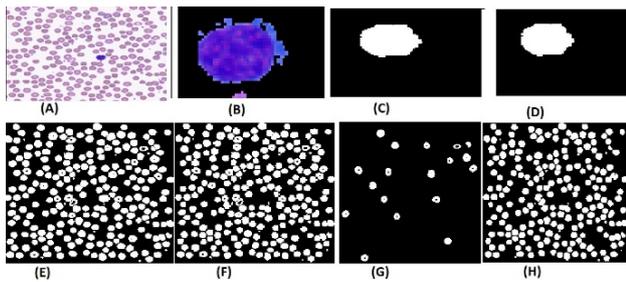


Fig. 20. (A). Blood smear image of a normal case. (B). WBC nucleus. (C). Nuclear lobes obtained using the proposed method. (D). WBC (E). RBCs with WBC after separating touching objects. (F). RBCs. (G). RBCs with central pallor. (H). RBCs without central pallor.

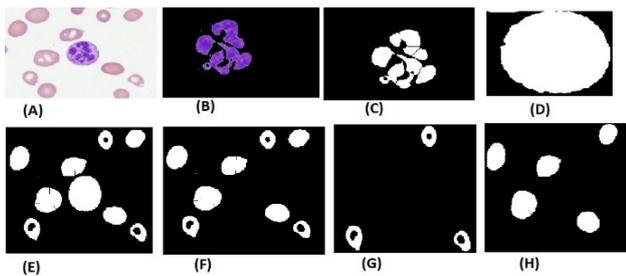


Fig 21. (A). Original Image. (B). WBC nucleus. (C). Nuclear lobes obtained using the proposed method. (D). WBC (E). RBCs with WBC after separating touching objects. (F). RBCs. (G). RBCs with central pallor. (H). RBCs without central pallor.

Table 3. Count of RBCs

Id	a	B	c	d	e	f	g	h	i	j	k
1	5	11	01	0	22	0	0	0	0	0	01
2	2	6	12	1	8	0	0	0	0	0	06
3	0	0	4	0	3	0	0	0	0	0	06
4	0	0	4	1	4	0	0	0	0	0	04
5	0	0	0	4	9	0	0	0	0	0	01
6	0	0	6	0	06	0	0	0	0	0	06
7	0	7	0	0	6	0	0	0	0	0	04
8	0	2	13	0	10	0	0	0	0	0	03
9	18	212	0	5	0	0	0	0	0	0	0
10	3	0	5	0	0	0	0	0	0	0	0

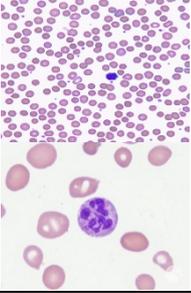
The results diagnosed by the proposed system is mapped against the ground truth in Tab.4. The percentage of Sickle cells present in the image is also shown. It is computed using the Eq. 19.

$$\% = a/b \tag{19}$$

a denotes the count of Sickle cells
b denotes the count of all the types of RBCs

Table 4. Prediction and Diagnosis

A	B	C	D	E	F
	0	0	2%	S	S
	01	01	11%	S	S
	0	0	46.15%	S	S
	0	0	33.33%	S	S
	0	0	7.14%	S	S
	0	0	35.29%	S	S
	01	04	23.5%	S	S
	01	04	11.11%	S	S

	01	01	0%	N	N
	01	10	62.5%	M	M

“A” indicates the blood smear image. “B” shows the WBC count. “C” represents the number of lobes in WBC nucleus. “D” indicates the percentage of sickle cells. “E” shows the diagnosis made by the proposed system.” F” represents the prediction made by the domain expert. The sickle cell anemia is represented using “S”. The label “M” denotes Megaloblastic anemia and the label “N” denotes Normal case.

5. Discussion

In the proposed work, an automatic algorithm to separate and categorize the RBCs is proposed. The discussion section is divided into three parts; the first part indicates the effect of Modified Watershed Transform on segmentation accuracy, the second part deals with the impact of granulometric analysis over the counter value and the third part provides a comparative study between the existing methods and the proposed method.

4.1 Effect of Modified Watershed Transform on the image

In the proposed work, a modified watershed transform algorithm is illustrated. The original watershed transform chooses a regional minima point and starts growing from there. Each of this creates a single region. Both discontinuity and similarity features are taken into consideration successfully. It readily distinguishes the foreground and the background object. But, it suffers from the problem of over-segmentation. The modified watershed transform eliminates this problem by filtering out the tiny local minima using the “imextendedmin” function. It generates small spots in the middle of every cell. The distance transform is then modified to have minima at those desired locations. Watershed is then applied to the resulting image to obtain accurate separation. Fig. 22 depicts the nuclear lobes extracted using the existing watershed transform method and the proposed method. The number of lobes obtained using current watershed transform was very high due to over-segmentation. The results obtained using proposed method were shown to the domain expert. It was successful in extracting the correct number of lobes. Fig. 23 shows another such instance.

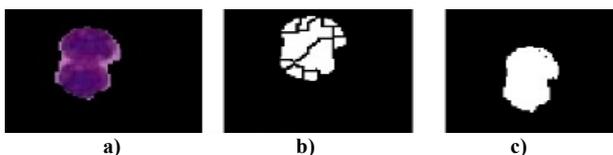


Fig. 22. (A) WBC nucleus. (B). Incorrect nuclear lobes obtained due to over segmentation. (C). Nuclear lobes obtained using the proposed method.

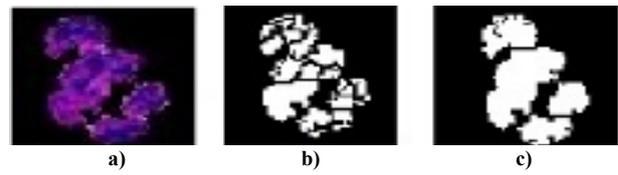


Fig 23. (A) WBC nucleus. (B). Incorrect nuclear lobes obtained due to over segmentation. (C). Nuclear lobes obtained using the proposed method.

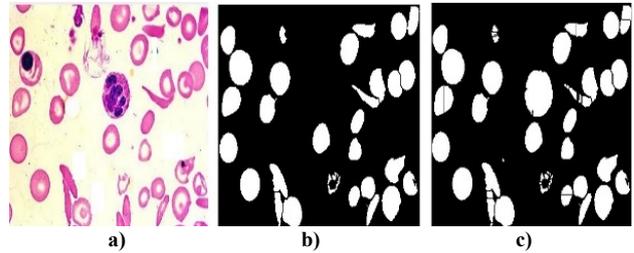


Fig 24. (A) Blood smear image of a patient suffering from Sickle cell anemia. (B). Overlapping RBC separated using the proposed method. (C). Overlapping RBC separated using the existing watershed transform.

Fig 24 illustrates the red blood cells separation achieved using proposed method and existing watershed transform. Fig 25 illustrates another such instance.

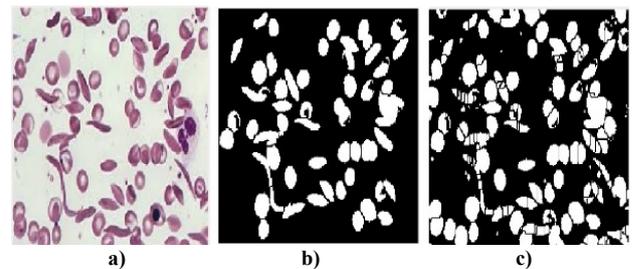


Fig 25. (A) Blood smear image of a patient suffering from Sickle cell anemia. (B). Overlapping RBC separated using the proposed method. (C). Overlapping RBC separated using the existing watershed transform.

The count of RBCs is tabulated in table 3. The count obtained were verified with domain expert. The precision was then computed using the Eq. 20. True positives (TP) indicate the number of RBCs that is diagnosed as a RBC by proposed system and the domain expert also diagnoses it as a RBC. False positives(FP) indicate the number of RBC diagnosed by the proposed method but not by the domain expert. The recall value is computed using Eq. 21. False negative (FN) indicate the number of RBC detected by the domain expert but not by the proposed method. The accuracy is computed using the Eq. 22. The precision and recall values are tabulated in table 5.

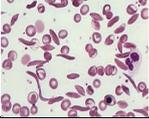
$$Precision = \frac{TP}{TP+FP} \tag{20}$$

$$Recall = \frac{TP}{TP+FN} \tag{21}$$

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{22}$$

Table 5. Precision and Recall

Image	F1	F2	F3	F4	F5	F6	F7
	31	59	30	01	0	96.77%	100%

	48	193	45	2	1	95.74%	97.82%
---	----	-----	----	---	---	--------	--------

F1 is the RBC count obtained using the proposed method. F2 is the RBC count obtained using existing watershed transform. F3 indicates the true positive. F4 indicates the False positive. F5 indicates the false negative. F6 indicates the precision value. Recall value is shown by F7.

5.2 Effect of Granulometric analysis on the RBC count

In the proposed study, the RBCs are separated from the surrounding cells in the blood smear using the granulometric analysis. The procedure is discussed in section 3.2. This technique helps in achieving the accurate count of RBCs. The accuracy of the proposed method was compared with state of the art methods. Fig. 26 and Fig. 27 illustrates the RBC obtained using the state of the art methods and the proposed method. The global threshold was used to convert the image to binary. The RBCs were then counted using Labeling algorithm by Patil, Sable, & Anandgaonkar [13]. The image with label “B” shows the result. The WBC is not separated; hence it leads to erroneous count. The WBC is depicted with red color. The image with label “C” shows the RBC obtained after applying a threshold of 140. The cells were counted using the circular Hough transform by Alomari et al. [1]. But, the threshold value did not satisfy all the images. The image with label “D” shows the green component extracted from the RGB plane. Thejashwini & Padma [21] converted the image in green plane to binary and counted the cells using CHT. The image with label “E” shows the red component extracted from the RGB plane. The image was later converted to binary and RBCs were counted by Bhagvathi and Thomas [3].

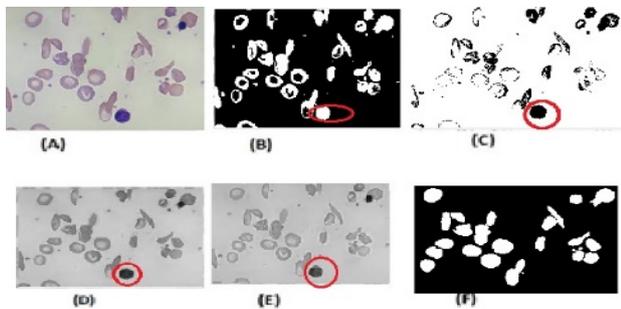


Fig. 26. Comparative analysis of proposed method with state of art methods. (A). Blood smear of Sickle cell anemia. (B). RBC obtained using Global Otsu’s threshold. (C) RBC obtained using a threshold value of 140. (D). RBC obtained by extracting the green plane of image. (E) RBC obtained by extracting the red plane of image. (F) RBC obtained using the proposed method.

Table 7. Comparative Analysis

Sl. No	Method	Accuracy	Advantages	Disadvantages	Reference
1	Segmentation using Threshold Technique	95.3%	The method was successful in determining irregular cells and detection of overlapping cells.	Threshold values used failed to satisfy the images with different stains	[1]
2.	Segmentation using the Fuzzy Logic	-	The method was successful in developing a system to count the RBCs.	WBC masks are generated accurately in some cases that resulted in imprecise results.	[3]
3.	Segmentation using threshold and identification using edge detection	-	The method was able to diagnose the presence of Sickle cells.	The method did not consider the feature extraction stage which would help the pathologists in	[5]

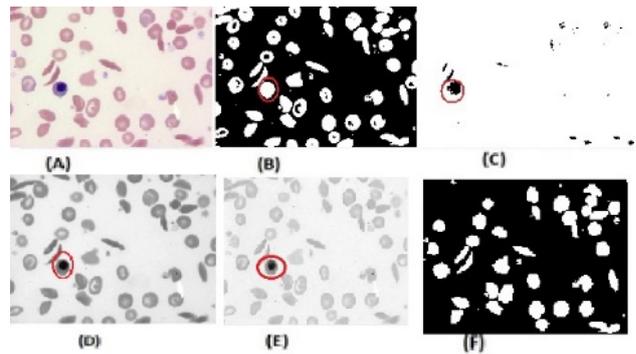
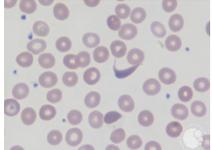
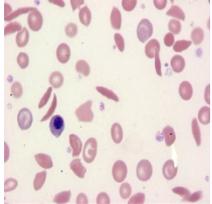


Fig.27. Comparative analysis of proposed method with state of art methods. (A). Blood smear of Sickle cell anemia. (B). RBC obtained using Global Otsu’s threshold. (C) RBC obtained using a threshold value of 140. (D). RBC obtained by extracting the green plane of image. (E) RBC obtained by extracting the red plane of image. (F) RBC obtained using the proposed method.

Table 6. RBC Count

Image	A1	A2	A3	A4	A5	A6
	44	32	65	70	50	52
	101	64	200	210	53	50

A1 is the RBC count obtained from image labeled as “B”. A2 is the RBC count obtained from image labeled as “C”. A3 is the RBC count obtained from image labeled as “D”. A4 is the RBC count obtained from image labeled as “E”. A5 is the RBC count obtained from image labeled as “F”. A6 is the RBC count given by the domain expert.

5.3 Comparative Analysis

In this section, the accuracy of the proposed method is compared with the existing techniques to segment the RBCs and detect the Anemia. The advantages and drawbacks of current technologies are discussed in this section. The different segmentation approaches involved in identifying and counting RBC along with identification of the disease and their accuracy is tabulated in Tab 7.

4.	Segmentation using CHT	-	The method was able to count the RBCs with different sizes.	the diagnosis The technique involved human intervention in identifying the radius range values.	[9]
5.	Segmentation and Detection of Red blood cells using Hough transform method	91.87%	The method provided a cost-effective way to analyze and count the cells.	The technique involved the cropping of the individual cells in obtaining the minimum and maximum radius.	[10]
6.	Segmentation and Identification of Red blood cells using Sobel edge detection technique	83%	The method can differentiate between the different abnormal cells.	The technique did not separate the WBCs from RBCs which led to incorrect count.	[12]
7.	Segmentation using Gray thresholding	94.58%	The method was successful in determining the RBC and WBC count.	The technique did not take into consideration the separation of overlapping cells.	[13]
8.	Segmentation using thresholding and detection using Sobel edge operator	95.8%	The method was able to detect the RBCs with different shapes.	The process did not specify any step to separate the other blood components from RBCs.	[15]
9.	Segmentation using gradient watershed transform	92.4%	The method was able to diagnose the disease with decent accuracy.	The technique suffers from the problem of over-segmentation.	[17]
10.	Segmentation using Otsu's threshold and detection using feature extraction	87%	The method was able to diagnose anemia and malarial fever.	The step to separate the touching objects was not taken into consideration.	[18]
				The method considered single cells for computation.	
		94.6%			[34]
11.	Segmentation and Detection of Red blood cells using Circular Hough transform method	93.1%	The method was able to diagnose the overlapping cells and gave a separate count for each.	The technique misinterpreted the WBCs as RBCs.	[21]
12.	Segmentation using morphological watershed transform	-	The method was able to segment the blood smear to its different components.	The method did not take into consideration the separation of touching objects.	[22]
13.	Segmentation using the Diffused expectation maximization algorithm	-	The method was successful in identifying the sickle cells.	The method did not take into consideration the separation of touching objects.	[25]
14.	Segmentation using the ellipse adjustment	99.35% for circular objects and 94.53% for elongated objects	The method was successful in differentiating an elongated cell from a circular one.	The images considered for the experimentation had only RBCs, and no other component was present.	[26]
15.	Segmentation using watershed algorithm and thresholding algorithm	-	The recorded feature values of cells helped in the diagnosis of the disease	The segmentation results generated by Otsu's method was not accurate.	[27]
16.	Segmentation using artificial neural network and feature extraction using algebraic moments	66.6%	Identification of significant features helped in differentiating a normal cell from an abnormal one.	The step to separate WBC from RBC was not undertaken. The touching cells were also not	[28]

17.	Segmentation using artificial neural network	74%	The method was successful in identifying the different components of the blood smear along with overlapping cells.	separated. The system must be trained to recognize the collapsed RBCs.	[29]
18.	Segmentation using CHT and watershed transform	-	The method was able to differentiate between normal and abnormal cells in a short period.	The method did not consider the step to separate the WBC from the RBCs	[32]
19.	Segmentation using marker-controlled watershed transform	80.6%	The method was able to achieve a decent accuracy in identification of disease.	The method did not consider the step to separate the WBC from the RBCs	[33]
20.	Segmentation using 8-connection chain codes technique	-	The method was able to achieve accurate separation of RBCs.	The method did not consider the step to separate the WBC from the RBCs	[35]
21.	Proposed Method	97.6%	-	-	-

6. Conclusion

Anemia can be diagnosed by understanding the morphology of RBCs. The inherent issues with RBCs are its visual interpretations are tedious and subjective prone to inter-observer variations. To help the pathologists, a new approach to separate the Red blood cells from other components in the blood smear and a system to capture the change in the morphology is proposed. The proposed system takes into consideration the size value analysis to achieve the separation. The efficiency of the proposed algorithm is shown using the precision and recall values. In a variation to the existing technique, the new approach uses K-medoids algorithm to achieve the color based segmentation. The work identified that the presence of clustered RBCs leads to the

generation of false positive cases. Hence, a modified watershed transform is used to separate them. It proposed a wholly automated decision system with various extracted features to diagnose the Sickle cell anemia. The work achieved an overall accuracy of 98.5 % in identification of Sickle cell anemia and an accuracy of 97.6% in separation of RBCs. The future work involves the process of accurate extraction of RBCs near the border for improvisation of accuracy.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License



References

[1]. Y.M. Alomari, S.N.H.S. Abdullah,, R.Z. Azma, K. Omar, Automatic Detection and Quantification of WBCs and RBCs using iterative structured circle detection algorithm, Computational and Mathematical Methods in Medicine, 2014 (2014).

[2]. Area open function. <https://in.mathworks.com/help/images/ref/bwareaopen.html>.

[3]. S.L. Bhagvathi, and N. S. Thomas, An Automatic System for Detecting and Counting RBC and WBC using Fuzzy Logic, ARPN Journal of Engineering and Applied Sciences, 11(11), pp. 6891-6894 2016.

[4]. Border clear. <https://in.mathworks.com/help/images/ref/imclearborder.html>

[5]. I.A. Chintawar, A . Mishra, P. Narnaware, and, C. Kuhikar, Detection of sickle cells using image processing, International Journal of Science Technology Engineering, 2(9), pp. 335-339 (2016).

[6]. C.I.E primaries. <http://hyperphysics.phyastr.gsu.edu/hbase/vision/cieprim.html>.

[7]. Euler Number. <https://in.mathworks.com/help/images/ref/bweuler.html>.

[8]. Label the components. <https://in.mathworks.com/help/images/ref/bwlabel.html>.

[9]. M. Maitra, R.K. , Gupta, Mukherjee, Detection and Counting of Red Blood Cell in Blood Cell Images using Hough Transform. International Journal of Computer Applications. 53(16), pp. 18-22, (2012).

[10]. S..M. Mazalan , N.H. Mahmood, and M. A. A, Razak. Automated Red Blood Cells Counting in Peripheral Blood Smear Image using Circular Hough Transform. International Conference on Artificial Intelligence, Modelling & Simulation, pp. 285-289, (2013).

[11]. Measurement properties. <https://in.mathworks.com/help/images/ref/regionprops.html>.

[12]. D.N. Patil and U.P. Khot, Image processing based abnormal blood cells detection, International Journal of Technical Research and Applications, 31, pp. 37-43, (2015).

[13]. P. R. Patil, G. Patil and, G. Anandgaonkar, Counting of WBCs and RBCs from Blood Images using Gray Thresholding, International Journal of Research in Engineering and Technology, 3(4), pp. 391-395 (2014).

[14]. Pixel color values . <https://in.mathworks.com/help/images/ref/impixel.html>.

[15]. P. Rakshita, and K. Bhowmik, Detection of abnormal findings in human RBC in diagnosing Sickle cell Anaemia using image processing, In International Conference on Computational Intelligence: Modeling, Techniques and Applications, pp. 28-36, (2013)

[16]. RBCs morphology. <https://laboratoryinfo.com/variations-in-red-blood-cell-morphology/>.

[17]. T. Revathi, and S. Jeevitha, Efficient watershed based red blood cell segmentation from digit al images in sickle cell disease. International Journal of Scientific Engineering and Applied Science, 2(4), pp. 300-317, (2016).

[18]. S.M.A. Rexcy , V.S. Akshaya , and K.S.Swetha, . Effective use of image processing techniques for the detection of sickle cell anemia and presence of plasmodium parasites. International Journal of Advance Research and Innovative Ideas in Education. 2 (2), pp. 701-706, (2016).

[19]. M.K.Sahu ,A.K.Biswas, and Uma. Detection of sickle cell anemia in red blood cell: A review. International Journal of Engineering and Applied Sciences. 2(3), pp.45-48, (2015).

[20]. A.Sreekumar and A. Bhattacharya, Identification of sickle cells from microscopic blood smear image using image processing. International journal of Emerging Trends in Science and Technology, 1(5), pp. 783-787, (2014).

[21]. M, Thejashwini, and M. C. Padma, Counting of RBCS and WBCs using image processing technique. International Journal on Recent and Innovation Trends in Computing and Communication, 3(5), pp. 2948-2953, (2014).

- [22]. H. Tulsani, N. adav, and S. Saxena, Segmentation using morphological watershed transformation for counting blood cells. *International Journal of Computer Applications and Information Technology*, 2(3), pp. 28–36, (2013).
- [23]. A. Vasundhara, K. Preetham, Identification and red blood cell classification using computer aided system to diagnose blood disorders. In *Proc. 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. (2017) <https://doi.org/10.1109%2Ficacci.2017.8126155>.
- [24]. Watershed Transform. <https://blogs.mathworks.com/steve/2013/11/19/watershed-transform-question-from-tech-support/>.
- [25]. S. Bala and A. Doegar, Detection of Sickle Cell Anemia and Thalassemia using Image Processing Techniques, *International Journal of Innovative Science and Research Technology*, 2(12),(2017).
- [26]. M.G. Hidalgo, F.A. Guerrero-Peña, S. Herold-García, A. Jaumei-Capó and P.D. Marrero-Fernández, sharma, *IEEE Journal of Biomedical and Health Informatics*, 19, pp. 1514-1525 (2015)
- [27]. B.H. Parvathy, S. Hariharan and N.S. Aruna, 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*, 5(4),pp.6200-6207 (2016).
- [28]. M. Veluchamy, K.Perumal and T. Ponuchamy, Feature Extraction and Classification of Blood Cells Using Artificial Neural Network, *American Journal of Applied Sciences*, 9 (5),pp. 615-619 (2012).
- [29]. J. Poomcokrak and C. Neatpisarnvani, RED BLOOD CELLS EXTRACTION AND COUNTING, In *proc. The 3rd International Symposium on Biomedical Engineering*, pp.199-203 (2008).
- [30]. Survey Data. <https://www.cdc.gov/ncbddd/sicklecell/data.html>
- [31]. Symptoms of Sickle Cell Anemia. <https://www.chw.org/medical-care/macc-fund-center/conditions/hematology-and-blood-disorders/blood-disorders/sickle-cell-disease>.
- [32]. M.A. Fadhel, A. J. Humaidi and S. R. Oleiwi, "Image processing-based diagnosis of sickle cell anemia in erythrocytes," In *Proc. 2017 Annual Conference on New Trends in Information & Communications Technology Applications (NTICT)*, pp. 203-207 (2017). doi: 10.1109/NTICT.2017.7976124
- [33]. V. Sharma, A. Rathore and G. Vyas, "Detection of sickle cell anaemia and thalassaemia causing abnormalities in thin smear of human blood sample using image processing," 2016 International Conference on Inventive Computation Technologies (ICICT), Coimbatore, pp. 1-5 (2016).
- [34]. L.F. Rodrigues, M.C. Naldi and J. F. Mari, "Morphological analysis and classification of erythrocytes in microscopy images", In *Proc. 2016 Workshop de Visão Computacional* pp. 1-6. (2016).
- [35]. H.M.Chen, Y.T.Tsao, and S.N.Tsai. Automatic image segmentation and classification based on direction texton technique for hemolytic anemia in thin blood smears. *Machine Vision and Applications*, 25(2), pp. 501–510 (2014).
- [36]. Decision Tree classifier. http://mines.humanoriented.com/classes/2010/fall/csci568/portfolio_exports/lguo/decisionTree.html
- [37]. Decision Tree Rules and Pruning. http://www2.cs.uregina.ca/~dbd/cs831/notes/ml/dtrees/4_dtrees3.html.