Morphological Features Analysis for Erythrocyte Classification in IDA and Thalassemia

Izyani Ahmad[1], Siti Norul Huda Sheikh Abdullah[2]

Centre for Artificial Intelligence Technology
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia Bangi, Malaysia

Raja Zahratul Azma Raja Sabudin[3]

Pathology Department
Hospital Canselori Tuanku
Mukhriz Cheras, Malaysia

Abstract—**Iron Deficiency Anemia (IDA) and Thalassemia is a common disease in the world population. In hospital routine, those diseases are being recognized based on level of hemoglobin in Complete Blood Count (CBC) result. Then, visual experts will conduct examination under the light microscope which is subjected to human error. In this research, we suggested a methodology via machine learning to classify and characterize erythrocyte related with IDA and Thalassemia. We employ some image pre-processing techniques on the blood smear images to enhance edges and reduce image noise such as gamma correction and morphological processing. Then, every single erythrocyte image will segment the background and foreground by using Otsu's threshold method. Here, we have considered nine types of erythrocyte such as teardrop, echinocyte, elliptocyte, microcytic, hypochromic, target cell, acanthocyte, sickle cell and normal cell to be classified and portray based on their morphological features. Later, these 24 and 31 features from Hue's moment, Zernike moment, Fourier descriptor and geometrical features are confirmed as potential features for each condition by calculating one-way ANOVA. Next, the rank of subset features is done based on their information gain value from maximum to minimum. Each of subset is separated by incremental of five features. Here, we compare the performance for each subset with five selected classifiers namely logistic regression, radial basis function network, multilayer perceptron, Naïve Bayes Classifier and Classification and Regression Tree. The best subsets from 31 features provide the highest result of classification with 83.5% accuracy, 83.5% sensitivity and 83.3% positive predictive value respectively via logistic regression compared to other classifiers. This study could be extended by using image dataset from other blood based disease for future work.**

Keywords—(Iron Deficiency Anemia) IDA; Thalassemia; erythrocyte; morphological features; classifier; information gain; logistic regression

I. INTRODUCTION

Red blood cells (RBC) also termed as erythrocyte, forms the main part of human body system. They deliver oxygen to the different body tissues via blood flow through the circulatory system. It also contains haemoglobin which provide red colour and protein lipids to maintain their stability and deformability.

In hospital routine, pathologist detects erythrocyte abnormality in blood smear images under the light microscope. This process is very subjective evaluation that lead to tedious, time consuming and error prone work [1], [2] The pathologist accomplishes judgement based on his clinicopathological understanding in disease diagnostics. Thus, expose it to a higher variability of intra-observer due to small disparity in morphological feature.

Amongst various microscopic of blood smear test, most common disease detection including anaemia, thalassemia, leukaemia, red cell haemolysis, etc. In specific, anaemia and thalassemia frequently happens in our population. According to [3], approximately 1 in 10 Malaysian is a carrier for Thalassemia. Currently, there are 6753 Thalassemia patients that require medical attention.

The detection of certain type of anaemia and thalassemia can be based on microscope evaluation of erythrocyte in terms of their shape and size. So, the classification of abnormal and normal erythrocytes from microscopic images of blood smears has a large influence in developing a rapid and efficient tool. Hence, feature descriptors play an important role in classification. In this study, we are focusing on nine types of erythrocytes (teardrop, target, elliptocyte, hypochromic, microcytic, normocytic, burr and keratocyte) shapes that are commonly seen in iron deficiency anaemia (IDA) and thalassemia patient (Fig 1).
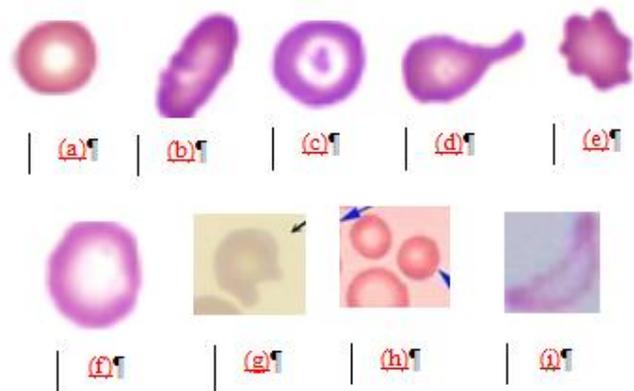
Fig. 1. Erythrocyte Cell Seen in IDA and Thalassemia Patient (a) Normal (b) Elliptocyte (c) Target (d) Teardrop (e) Burr (f) Hypocromic (g) Keratocyte (h) Microcyte and (i) Sickle.

A research reported [4] an approach to classify between malaria, thalassemia and normal patient using backpropagation neural network, while [5] proposed 271 features of color, texture and geometrical with PCA adopted. Author in [6] and [7] presented a rule-based technique for four classes of erythrocyte such as elliptocyte, macrocyte,

microcytic, sphrerocyte and thalassemia beta major. Other author proposed cancer cell classification using geometric mean transform and dissimilarity metrics [8]. While some others, proposed shape geometric features into 8 features [9], 4 features [6] and 6 features [7] including of area, perimeter, diameter, shape geometric feature, area proportion, deviation, central pallor, form factor, shape area factor and diameter area factor. Meanwhile, other author proposed moment invariant [10] and geometric features in their classification [11], [12]. Recently, [13] proposed a new shape descriptors using generalized support functions to describe convex figures. In conclusion, recent cell detection and recognition research are still focusing on morphological features to boost up the performance of overall digital pathology image analysis.

Our objective is to classify nine different categories of abnormal and normal erythrocyte in IDA and Thalassemia using a set of 24 and 31 features (Hue's moment, Zernike moment, Fourier descriptor and geometrical features). The gold standard analysis of the images is performed by two experts. Then, evaluation and statistical analysis are conducted for both set of features and compared. The rest of this paper is arranged as follows: the proposed morphological feature analysis in IDA and Thalassemia blood smear images in Section II, results and discussion in Section III, and the conclusion of this work in Section IV.

II. MATERIAL AND METHODS

A. Data Collection and Image Acquisitions

In this study, we have considered about seven healthy, nine Iron Deficiency Anaemia and seven Thalassemia blood smear images. These patients' peripheral blood smear were captured under light microscope from Pathology Department, Hospital Canselori Tuanku Muhriz (HCTM), Cheras, Malaysia. The process for displaying RBC image involved digitization of image from optical image with 40 times (40X) objective which equals to approximately 400 magnifications. Image acquisition has been guided by Hematologist (Fig. 2) in HCTM. From this study, the effective resolution and pixel size were 150dpi and 1920 X 1440 respectively. These images were captured using Olympus Digital Camera model DP22.
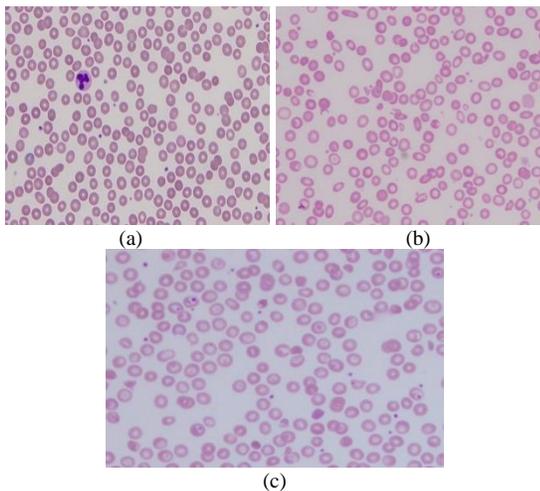

Fig. 2. Image Acquisition of (a) Normal (b) IDA (c) Thalassemia Source Real Blood Smear Dataset from Hospital Canselori Tuanku Muhriz (HCTM).
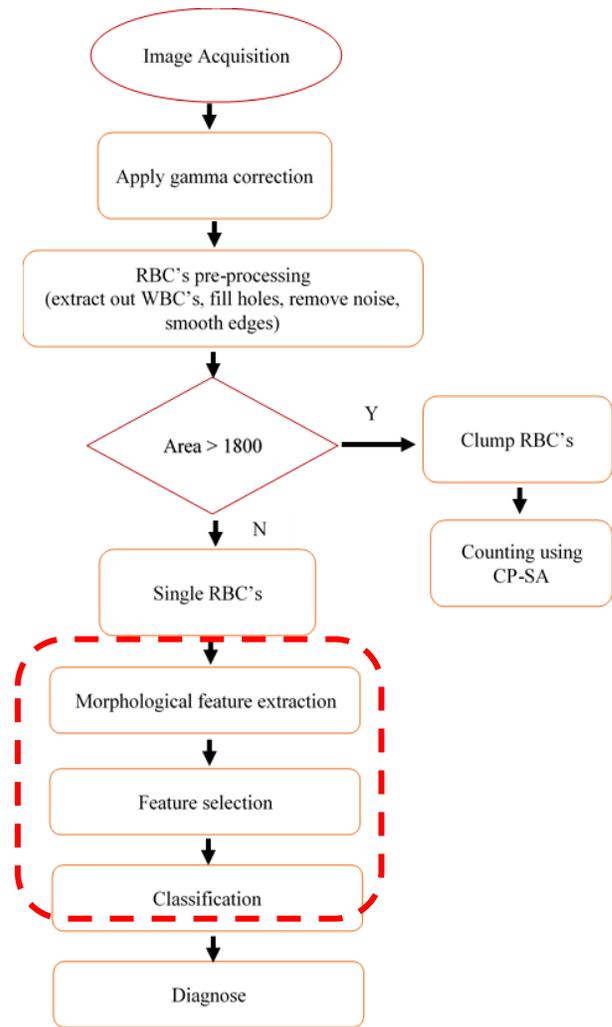

Fig. 3. General Methodology for the Proposed Method.

B. Erythrocyte Segmentation

Based on Fig. 3, we apply gamma correction to get sharpen edges. Followed by a few pre-processing step for erythrocyte segmentation. Hence, erythrocyte segmentation is vital for distinguishing abnormality present in IDA and Thalassemia disease. We have used *connected component labelling* method to segment erythrocyte from peripheral blood smear images. Later, each segmented erythrocyte was separated to clump and single cell based on their area. Here, clump cells will be counted by CP-SA method. CP-SA used skeleton algorithm in getting the backbone of each clump RBC, and analysed it with the concavity point pixel to detect and count the RBC. While single cell will be input to next level.

C. Morphological Feature Extraction

In IDA and Thalassemia condition, erythrocyte morphology varies from the normal cell. Thus, mathematical model was chosen to characterizing abnormal RBC's. Here, two separated experiment was conducted. First experiment consists of 24 morphological features in Table I, while another experiment includes 31 morphological features as shown in Table II.

TABLE I.   24 EXTRACTED FEATURES FOR IDA AND THALASSEMIA CATEGORIZATION

NO. OF FEATURES	FEATURES
7	HUE'S MOMENT
1	ZERNIKE MOMENT
4	FOURIER DESCRIPTOR
12	PERIMETER, AREA EQUIVALENT DIAMETER(AED), ECCENTRICITY, COMPACTNESS, MAJOR AXIS, MINOR AXIS, SOLIDITY, PERIMETER EQUIVALENT DIAMETER (PED), AREA, ROUNDNESS, CONVEX AREA, CONCAVITY

TABLE II.   31 EXTRACTED FEATURES FOR IDA AND THALASSEMIA CATEGORIZATION

NO. OF FEATURES	FEATURES
7	HUE'S MOMENT
1	ZERNIKE MOMENT
4	FOURIER DESCRIPTOR
19	PERIMETER, AREA EQUIVALENT DIAMETER(AED), ECCENTRICITY, COMPACTNESS, MAJOR AXIS, MINOR AXIS, SOLIDITY, PERIMETER EQUIVALENT DIAMETER (PED), AREA, ROUNDNESS, CONVEX AREA, CONCAVITY, FILL AREA, CENTRAL NODE, CENTRAL PALLOR, AREA PROPOTION (AP), AP EQUIVALENT CIRCLE (APEC), AREA MINOR, RADIAN

a) Hue's moments

This moment was introduced by Hu (1962). It manipulates the moment invariant within translation, scaling and rotational. This method calculate central moments, and the next seven invariant moments were created based on it [12].

b) Zernike moments

Zernike moments are based on a unit circle of a complete orthogonal set of complex polynomial. It is better than other methods in terms of information redundancy, reconstruction capability and noise resilience.

In this method, region of interest is plotted using polar coordinate in a unit circle. The origin of circle is also the center of the region of interest. Here, Θ is the angle of polar coordinate and r is the radius of polar coordinate. The plotting to the polar coordinate is:

$$x = r \cos \emptyset \; and \; y = r \sin \emptyset \qquad (1)$$

where

$$r = \sqrt{x^2 + y^2} \; and \; \emptyset = tan^{-1}\left(\frac{y}{x}\right)$$

Next, Zernike moments A(x,y), was calculated by normalized the region of interest using polar coordinates. Equation as following is considered for achieving both properties.

$$A(x,y) = f\left(\frac{x}{a} + \bar{x}, \frac{y}{a} + \bar{y}\right) \qquad (2)$$

Where

$$a = \sqrt{\frac{\beta}{m_{00}}}, \; \bar{x} = \frac{m_{10}}{m_{00}} \; and \; \bar{y} = \frac{m_{01}}{m_{00}}$$

Note that, m_{00}, m_{10}, m_{01} are taken from Hue's moments. In binary image, $m_{00} = \beta$.

c) Fourier descriptor

A region of interest image consists of *M*-point pixel boundary [14]. The boundary points coordinate is $(x_0, y_0), (x_1, y_1), \ldots\ldots. (x_{k-1}, y_{k-1})$. These coordinates can be stated as x(k) = x_k and y(k) = y_k.. It can be defined in complex number as Eq. (3).

$$s(k) = x_k + jy_k \qquad (3)$$

where $k = 0, 1, 2, 3 \ldots M - 1$. In complex number plane, real axis is stated as x-axis and imaginary axis as y-axis

The Fourier descriptor of $s(k)$ is defined in Eq. (4) as below,

$$a(u) = \frac{1}{M}\sum_{k=0}^{M-1} s(k)e^{-j2\pi uk/M} \qquad (4)$$

For $k = 0, 1, 2, 3 \ldots M - 1$.

Fourier descriptor of the boundary is defined as *a(u)*. The four features of fd1, fd2, fp1 and fp2 were calculated based on this descriptor [15].

d) Other morphological features

Other morphological features consist of eccentricity, roundness, area, fill area, area proportion have also been considered in [16], [17]. Part of them are detailed as following

Fill Area: Specifies the number of pixels in the region, with all holes filled in.

Centre Node (CN): Specifies value of center point in binary (1 or 0).

Central Pallor (CP)

$$CP = Fill \; Area - Area \qquad (5)$$

Area Proportion (AP)

$$AP = \frac{CP}{Fill \; Area} \qquad (6)$$

(i) Area Proportion Equivalent Circle (APEC):

$$APEC = \frac{Fill \; Area}{EC} \qquad (7)$$

where, Equivalent Circle, $EC = \pi r^2$

(i) Area Minor (AM)

$$AM = \pi \left(\frac{MinorAxisLength}{2}\right)^2 \qquad (8)$$

where, Minor Axis Length = length of the minor axis of the ellipse.

D. Selection of Significant Feature for IDA and Thalassemia

The feature selection is playing an important role in machine learning process. It has impact to increase the potentiality of each classifier. Some are suitable to differentiate textural and morphological features for describing normal and abnormal erythrocytes. Though, for predicting this situation, a set of features needs to be identified, in which is given the highest prediction accuracy. Plus, reduced the computational complexity and non-selected features. Here, statistical process were employ to the extracted features using one-way ANOVA (analysis of variance) for pinpointing a set of significant features [18]. Next, the information gain value is computed for each significant features using Weka version 3.7.5 and ranked based on it [19]. One way ANOVA calculated and compared the means of two or more instance using F-test statistic [18] as denoted below

$$F = \frac{between\ class\ variance}{within\ class\ variance} > 0 \qquad (9)$$

The lower F value indicates lower discrimination potentiality and bigger F value indicates the better discrimination potentiality. Table III displays the analysis of one-way ANOVA from the extracted morphological erythrocyte images for both 24 and 31 features. Each feature was modeled for nine types of erythrocyte in IDA and Thalassemia representation. From this table, we can observe that all 24 and 31 features are statistically significant.

TABLE III.   ONE WAY ANOVA ANALYSIS OF IDA AND THALASSEMIA FEATURES

FEATURES	P VALUE	FEATURES	P VALUE
ROUNDNESS	< 0.0001	MAJOR	< 0.0001
APEC	< 0.0001	CP	< 0.0001
ECCENTRICITY	< 0.0001	HUE5	< 0.0001
MINOR	< 0.0001	HUE4	< 0.0001
AREA MINOR	< 0.0001	AP	< 0.0001
COMPACTNESS	< 0.0001	AREA	< 0.0001
SOLIDITY	< 0.0001	CONVEXAREA	< 0.0001
PED	< 0.0001	HUE3	< 0.0001
CN	< 0.0001	HUE1	< 0.0001
RADIAN	< 0.0001	ZERNIKE	< 0.0001
AED	< 0.0001	FD1	< 0.0001
HUE2	< 0.0001	FP2	< 0.0001
HUE6	< 0.0001	FP1	< 0.0001
FILL AREA	< 0.0001	FD2	< 0.0001
HUE7	< 0.0001	PERIMETER	< 0.0001
CONCAVITY	< 0.0001		

a) Features ranking

Here, process for ranking the significant features based on their information gain was done. This study compared the ranked features into two conditions: (i) for 24 features (Table IV) and (ii) for 31 features (Table V). These ranked features were allocated into five subsets (24 features) and seven subsets (31 features) based on their ranks.

b) Information gain ranking

Information gain is designed via entropy theory of Shannon's [19]. The top-ranked features have the highest value of information gain. All significant features were set from maximum to minimum value of their information gain.

TABLE IV.   FEATURE RANKING BASED ON INFORMATION GAIN FOR 24 FEATURES

RANK	INFO GAIN VALUE	RANK FEATURES	RANK	INFO GAIN VALUE	RANK FEATURES
1	0.959	COMPACTNESS	13	0.42	HUE7
2	0.959	SOLIDITY	14	0.395	ZERNIKE
3	0.882	ECCENTRICITY	15	0.389	CONVEXAREA
4	0.875	ROUNDNESS	16	0.38	HUE5
5	0.823	CONCAVITY	17	0.35	HUE4
6	0.8	MINOR	18	0.347	HUE3
7	0.55	MAJOR	19	0.303	FD2
8	0.511	HUE2	20	0.268	FP1
9	0.511	PED	21	0.266	FP2
10	0.511	AED	22	0.252	AREA
11	0.479	HUE1	23	0.247	FD1
12	0.426	HUE6	24	0.236	PERIMETER

TABLE V.   FEATURE RANKING BASED ON INFORMATION GAIN FOR 31 FEATURES

RANK	INFO GAIN VALUE	RANK FEATURES	RANK	INFO GAIN VALUE	RANK FEATURES
1	0.959	SOLID	17	0.446	FILLAREA
2	0.959	COMPACTNESS	18	0.426	HUE6
3	0.889	APEC	19	0.42	HUE7
4	0.882	ECCENTRICITY	20	0.395	ZERNIKE
5	0.875	ROUNDNESS	21	0.389	CONVEXAREA
6	0.823	CONCAVITY	22	0.38	HUE5
7	0.8	MINOR	23	0.35	HUE4
8	0.753	AREA MINOR	24	0.348	CN
9	0.647	CP	25	0.347	HUE3
10	0.55	MAJOR	26	0.303	FD2
11	0.55	RADIAN	27	0.268	FP1
12	0.526	AP	28	0.266	FP2
13	0.511	HUE2	29	0.252	AREA
14	0.511	AED	30	0.247	FD1
15	0.511	PED	31	0.236	PERIMETER
16	0.479	HUE1			

E. IDA and Thalassemia Cell Identification

Our main objective was to classify and compared nine different categories of abnormal and normal erythrocyte based on two conditions (24 and 31 features). Abnormal erythrocyte characterization in IDA and Thalassemia patients such as teardrop, target, elliptocyte, acanthocyte, hypochromic, echinocyte, microcytic, sickle cell and normal erythrocyte were identified. Here, we setup five classification approach namely radial basis function network, logistic regression, Naïve Bayes classifier, classification, regression trees and multilayer perceptron by using Weka version 3.7.5.

a) Logistic regression (LR)

Logistic regression is a methodology under supervised classification which determines the membership of each dataset. As our experiment involved nine classes of erythrocyte, multiclass logistic regression model with one against all algorithm was used [20].

b) Radial basis function network (RBF)

The radial basis function network is a sort of feed forward network [21]. It contents of input layer, output layer and single hidden layer. Let assume that $x = [x_i: i = 1,2,3, …. , d]$ and d dimensional input feature space, while target output, $t = [t_z = 1, …. , k]$. The output of the RBF is defined by Eq (6) as below,

$$y_k(x) = \sum_{j=1}^{m} w_{kj} \emptyset_j(x) + w_{k0} \tag{6}$$

Here, w_{k0} = the basis input, w_{kj} = weights of hidden node.

The basis function, $\emptyset_j(x)$ is explained in Eq. (7) as the following:

$$\emptyset_j(x) = exp\left(-\frac{\|x-\mu_j\|^2}{2\sigma^2_j}\right) \tag{7}$$

where μ_j is the centre of the radial basis function, x is input vector, σ_j is the width of $\emptyset_j$. In order to determine the basis function parameters, K-means clustering algorithm is applied. Let say, σ_j =0.1.

c) Multilayer perceptron (MLP)

This methods is one type of back propagation neural network algorithm with multiple layers [21]. This algorithm will update the weights of each input and output layer in order to minimize output error and give better accuracy. Eq. (8) is applied to calculate the training error,

$$J(w) = \frac{1}{2} \sum_{k=1}^{c} (t_k - z_k)^2 \tag{8}$$

where the network output z_k and target output is t_k. The number of node is depicted by C. Here, we have nine output nodes with single hidden layer.

d) Naïve bayes classifier (NB)

This classifier is rooted from Bayesian theoretic method for feature set independent assumption. This method will guess possession of specific feature set via posterior probability. Here, posterior probability is estimated and forecast for each class label based on maximum value [21].

e) Classification and regression tree (CART)

CART is commonly used in machine learning with different classification problem. This algorithm construct classification and regression tree via top down recursive and divide-and-conquer approach. The CART comprises of root node, internal node and leaf node. The bottom node is leaf node and top most node is root node. Class label is allocated by leaf node. Here, the splitting criterion for data partition is done by Gini index measure [19].

III. RESULT AND DISCUSSION

Based on both Table VI (31 features) and Table VII (24 features), it is perceived that sometimes, whole features might not contribute the highest prediction performance. Hence, we decided to diversify them into subsets via their information gain value. Each subset consists of incremental of five features separately that sum up to 7 subsets for 31 features and 5 subsets for 24 features. Then, all subset will go through the

five selected classifiers. This research includes 725 abnormal and 99 normal erythrocytes.

As tabulated in Table VI and Table VII, it shows variation performance of different subsets of morphological features ranked using five different classifiers. As in Table VI, the Accuracy value varies from 68.2 to 78.3 for Naïve Bayes classifiers, 73.0 to 80.5 for RBF, 75.8 to 83.3 for MLP, 75.6 to 83.5 for LR and 71.2 to 78.5 for CART respectively. While in Table VII, this Accuracy value is much lower within 46.9 to 58.6 (Naïve Bayes), 50.4 to 63.7 (RBF), 54.3 to 68.0 (MLP), 54.6 to 72.1 (LR) and 52.6 to 63.7 (CART). It is also noticeable that top fifteen features (from 31 features) with (Sensitivity:83.5%, PPV:83.3% and Accuracy:83.5%) via logistic regression give better performance compared to all features (from 24 features) (Sensitivity:72.2%, PPV:72.2% and Accuracy:72.1%) using the same classifier.

We also observed a fluctuation trend in these tables with different subset with different classifier. This can be concluded that each of feature have their unique value that led to better or worst result. Lesser feature is worth to get better performance and accuracy, in which approximately subset of ten, fifteen and twenty are worth than 31 features ranked. However, more features need to get more accurate result based on 24 features ranked. Hence, it is notable that each feature will get more precise and optimized result with the introduction of information gain ranking.

TABLE VI. PERFORMANCE STATISTICS OF 31 FEATURES FOR ERYTHROCYTE CHARACTERIZATION (A) SENSITIVITY (B) PPV AND (C) ACCURACY

(A)

Set of ranked features	Sensitivity				
	NB	RBF	MLP	LR	CART
FIVE	68.2	73.0	75.9	75.5	71.2
TEN	71.8	76.5	75.7	81.3	73.2
FIFTEEN	75.9	75.3	79.9	83.5	75.1
TWENTY	78.1	80.5	77.5	82.3	77.7
TWENTY FIVE	78.3	80.3	81.1	80.3	78.5
THIRTY	77.1	76.5	83.3	78.3	78.3
THIRTY ONE	77.1	76.3	80.3	81.5	78.3

(B)

Set of ranked features	PPV				
	NB	RBF	MLP	LR	CART
FIVE	69.3	71.8	75.2	75.3	70.1
TEN	72.9	76.5	74.5	81.0	73.1
FIFTEEN	76.4	75.1	79.6	83.3	74.9
TWENTY	78.0	80.6	77.1	82.1	77.5
TWENTY FIVE	78.1	80.4	80.8	80.1	78.5
THIRTY	77.1	76.2	83.2	78.6	78.2
THIRTY ONE	77.2	76.2	80.0	81.5	78.2

(C)

Set of ranked features	Accuracy				
	NB	RBF	MLP	LR	CART
FIVE	68.2	73.0	76.0	75.6	71.2
TEN	71.8	76.5	75.8	81.3	73.2
FIFTEEN	76.0	75.4	79.9	83.5	75.2
TWENTY	78.1	80.5	77.5	82.3	77.7
TWENTY FIVE	78.3	80.3	81.1	80.3	78.5
THIRTY	77.1	76.5	83.3	78.7	78.3
THIRTY ONE	77.1	76.3	80.3	81.5	78.3

TABLE VII.   PERFORMANCE STATISTICS OF 24 FEATURES FOR ERYTHROCYTE CHARACTERIZATION (A) SENSITIVITY (B) PPV AND (C) ACCURACY

(A)

Set of ranked features	Sensitivity				
	NB	RBF	MLP	LR	CART
FIVE	46.9	50.4	54.3	54.6	52.6
TEN	53.9	56.9	59.8	62.3	56.5
FIFTEEN	57.6	63.7	63.1	67.2	62.5
TWENTY	58.6	63.1	65.7	69.3	62.1
TWENTY FOUR	57.2	61.4	68.0	72.2	63.7

(B)

Set of ranked features	PPV				
	NB	RBF	MLP	LR	CART
FIVE	44.0	49.5	53.3	52.6	50.2
TEN	52.5	56.7	58.6	61.7	54.9
FIFTEEN	56.2	63.6	62.7	67.2	62.1
TWENTY	57.6	63.3	65.6	69.2	61.5
TWENTY FOUR	56.4	60.9	67.9	72.2	63.3

(C)

Set of ranked features	Accuracy				
	NB	RBF	MLP	LR	CART
FIVE	46.9	50.4	54.3	54.6	52.6
TEN	53.9	56.9	59.8	62.3	56.5
FIFTEEN	57.6	63.7	63.1	67.2	62.5
TWENTY	58.6	63.1	65.7	69.3	62.1
TWENTY FOUR	57.2	61.36	68.0	72.1	63.7

IV.   CONCLUSION

As a conclusion, we have characterized and classified nine types of erythrocyte for identification of IDA and thalassemia condition. A set of morphological features were extracted from segmented blood smear images and ranked via information gain value into two conditions: (i) 24 features and (ii) 31 features. The experimental amongst the five classifiers has shown that logistic regression had given the optimal value of accuracy, sensitivity and positive predictive value (83.5%, 83.5% and 83.3%) respectively. Thus, we proposed these geometrical features to characterize each of the shape.

It also detected that a fewer feature in subset (for 31 feature) provides better performance in most classifier. This result proven that each of features has different significant value depend on the segmented image and it is important to ranked this feature correctly. This research could be extended by using image dataset from other blood based disease for future work.

REFERENCES

[1] Y. M. Alomari, S. N. H. Sheikh Abdullah, R. Zaharatul Azma, and K. Omar, "Automatic Detection and Quantification of WBCs and RBCs Using Iterative Structured Circle Detection Algorithm.," Comput. Math. Methods Med., vol. 2014, p. 979302, 2014.

[2] D. Albashish, A. Sahran, S., M. A., Alweshah, and A. A., "A hierarchical classifier for multiclass prostate histopathology image gleason grading.," Pertanika J. Sci. Technol., vol. 2, no. 2, pp. 323–346, 2018.

[3] A. Rz et al., "A Comparative Study of Red Blood Cell Parameters of Alpha and Beta Thalassaemia Patients Diagnosed in a University Hospital in," vol. 3, no. 1, pp. 23–27, 2018.

[4] Y. Hirimutugoda and G. Wijayarathna, "Image Analysis System for Detection of Red Cell Disorders Using Artificial Neural Networks," Sri Lanka J. Bio-Medical Informatics, vol. 1, no. 1, pp. 35–42, 2010.

[5] J. A. Alkrimi, L. E. George, A. Suliman, A. R. Ahmad, and K. Al-jashamy, "Isolation and Classification of Red Blood Cells in Anemic Microscopic Images," Int. J. Medical, Heal. Pharm. Biomed. Eng., vol. 8, no. 10, pp. 686–689, 2014.

[6] S. Chandrasiri and P. Samarasinghe, "Morphology Based Automatic Disease Analysis through Evaluation of Red Blood Cells," 2014 5th Int. Conf. Intell. Syst. Model. Simul., pp. 318–323, 2014.

[7] E. Suryani and K. N. Wahyudiani, "Berdasarkan Morfologi Sel Darah Merah," Fakt. thalasemia pada anak, vol. 2, no. 1, pp. 15–28, 2015.

[8] S. M. M. Kahaki, M. J. Nordin, W. Ismail, S. J. Zahra, and R. Hassan, "Blood cancer cell classification based on geometric mean transform and dissimilarity metrics," Pertanika J. Sci. Technol., vol. 25, no. S6, pp. 223–234, 2017.

[9] V. Acharya and P. Kumar, "Identification and red blood cell classification using computer aided system to diagnose blood disorders," 2017 Int. Conf. Adv. Comput. Commun. Informatics, ICACCI 2017, vol. 2017–Janua, pp. 2098–2104, 2017.

[10] S. J. Zahra, R. Sulaiman, and A. S. Prabuwono, "SCIENCE & TECHNOLOGY Invariant Feature Descriptor based on Harmonic Image Transform for Plant Leaf Retrieval," vol. 25, pp. 107–114, 2017.

[11] N. F. Hassan and A. I. Hussein, "Efficient method to Recognition of Anemia Images based on Moment Invariants and Decision tree classifier ارقلا ةرقشجر لاصملاو فلا ف ةناب ةلاعزموسا ساسا لع دلادم رقف ضرم صور لث يم زيك ةوفةطرق،" vol. 57, no. 3, pp. 2360–2370, 2016.

[12] D. Das, M. Ghosh, C. Chakraborty, M. Pal, and A. K. Maity, "Invariant moment based feature analysis for abnormal erythrocyte recognition," Int. Conf. Syst. Med. Biol. ICSMB 2010 - Proc., no. December, pp. 242–247, 2010.

[13] X. Gual-Arnau, S. Herold-García, and A. Simó, "Erythrocyte shape classification using integral-geometry-based methods," Med. Biol. Eng. Comput., vol. 53, no. 7, pp. 623–633, 2015.

[14] R. C. Gonzales and R. E. Woods, Digital Image Processing, 2nd editio. Prentice Hall, 2002.

[15] M. M. R. Krishnan et al., "Automated characterization of sub-epithelial connective tissue cells of normal oral mucosa: Bayesian approach," TechSym 2010 - Proc. 2010 IEEE Students' Technol. Symp., no. April, pp. 44–48, 2010.

[16] M. Gonzalez-Hidalgo, F. a Guerrero-Pena, S. Herold-Garcia, A. Jaume-I-Capo, and P. D. Marrero-Fernandez, "Red Blood Cell Cluster Separation from Digital Images for use in Sickle Cell Disease.," IEEE J. Biomed. Heal. informatics, vol. 2194, no. c, pp. 1–11, 2014.

[17] H. Berge, D. Taylor, S. Krishnan, and T. S. Douglas, "Improved Red Blood Cell Counting in Thin Blood Smears," in 978-1-4244-4128-0/11/$25.00 © IEEE, 2011, pp. 204–207.

[18] F. Statistics, A. M. Goon, and M. K. Gupta, "BOOK-REVIEWS," vol. 2, pp. 0–1.

[19] M. a Karaolis, J. a Moutiris, D. Hadjipanayi, and C. S. Pattichis, "Assessment of the risk factors of coronary heart events based on data mining with decision trees.," IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 3, pp. 559–566, 2010.

[20] D. K. Das, C. Chakraborty, B. Mitra, A. K. Maiti, and A. K. Ray, "Quantitative microscopy approach for shape-based erythrocytes characterization in anaemia," J. Microsc., vol. 249, no. 2, pp. 136–149, 2013.

[21] R. .Duda, Pattern Classification, 2nd edition. wiley, 2001.