



ISSN: 0067-2904

Efficient method to Recognition of Anemia Images based on Moment Invariants and Decision tree classifier

Nidaa Flaih Hassan^{1*}, Asma Ibrahim Hussein²

¹Computer Science Department, University of Technology, Baghdad, Iraq

²Ministry of Higher Education and Scientific Research, Baghdad, Iraq

Abstract

Anemia is one of the common types of blood diseases, it lead to lack of number of RBCs (Red Blood Cell) and amount hemoglobin level in the blood is lower than normal.

In this paper a new algorithm is presented to recognize Anemia in digital images based on moment variant. The algorithm is accomplished using the following phases: preprocessing, segmentation, feature extraction and classification (using Decision Tree), the extracted features that are used for classification are Moment Invariant and Geometric Feature.

The Best obtained classification rates was 84% is obtained when using Moment Invariants features and 74 % is obtained when using Geometric Feature. Results indicate that the proposed algorithm is very effective in detection distorted red blood cells and this helps the medical technician to decide the type of Anemia in Laboratory analyzes in the hospitals.

Keywords: Anemia, Feature Extraction, Moment invariant, Recognition.

طريقة كفاءة لتمييز صور مرض فقر الدم على اساس العزوم الثابتة والمصنف شجرة القرار

نداء فليح حسن^{1*}، أسماء إبراهيم حسين²

قسم علوم الحاسوب، الجامعة التكنولوجية، بغداد، العراق

وزارة التعليم العالي والبحث العلمي، بغداد، العراق

الخلاصة:

فقر الدم هو واحد من الأنواع الشائعة من أمراض الدم، وأنه يؤدي إلى نقص في عدد كريات الدم الحمراء (خلايا الدم الحمراء) وكمية مستوى الهيموجلوبين في الدم أقل من الطبيعي . في هذا البحث تم تقديم خوارزمية جديدة لتمييز مرض فقر الدم في الصور الرقمية وهذه الخوارزمية انجزت باستخدام المراحل التالية : معالجة الصورة ، تقطيع الصورة، استخراج الميزات والتصنيف باستخدام شجرة القرار، ويتم استخراج الميزات باستخدام نوعين من الميزات : العزوم الثابتة و الصفات الهندسية. تم الحصول على افضل نسبة للتصنيف 84% عند استخدام العزوم الثابتة ، و 74% تم الحصول عليها عند استخدام الميزات الهندسية، النتائج تشير الى ان الخوارزمية المقترحة فعالة جدا في الكشف عن تشويه خلايا الدم الحمراء وهذا يساعد أيضا فني طبي لاتخاذ قرار عن نوع فقر الدم في التحليلات المختبرية في المستشفيات.

*Email: nidaaalalousi_5@yahoo.com

1. Introduction

Pattern recognition is a field of research that examines the design of systems to identify patterns in the data. The longer term aim is to enable it to achieve near human level recognition for large number of categories under wide variety of conditions [1]. The Red Blood Cell (RBC) recognition system can be used for educational purposes in medical schools and assist in the development of workers in the field of Hematology. Anemic can be classified according to RBCs shape distortion, size of RBC and color or hemoglobin content, depended on the types of Anemia experienced by the patient [2].

Anemia is the most common disorder of the blood, "Anemia", the name is deriving from the ancient Greek word (*anemia*), which means "Lack of Blood" [3].

The current method used by the pathologists for identification of blood parameters is costly which are sometimes not affordable by the patient, also the generation of analysis report may require more time leading to loss of patient's life [4].

This paper proposed a new algorithm to recognize Anemia using moment invariant, the algorithm consists of the following phases : preprocessing, segmentation , feature extraction and classifier using Decision Tree (ID3) to classify the types of Anemic.

2. Related work

- **Bushra Q.Al-Abudi (2009)** proposed methods based on fractal geometry and invariants moment to recognize the printed or handwritten digits , The comparison proved that the fractal geometry possess high digit recognition capabilities and it gave 93% score of recognizing the printed digits during 6.6 s and 71% to recognize handwritten digit during 6.8 s, which is greater than the corresponding values 81% and 64% scores during 1.5s and 1.7 s, respectively given by the moments [5].
- **Fatin A. Dawood (2008)** proposed new method is presented to segment the microscopic images of White Blood Cells (WBC) into cell and non-cell regions depending on the image histogram, Experimental results show that the proposed unsupervised segmentation method have given reliable results comparable with thresholding method [6].
- **Wejdan A. Amer(2009)** algorithm to recognize distorted Machine and handwritten characters is proposed. It uses a feature point extraction-based recognition approach , experimental results on one database are presented, their results showing the robustness of the algorithm using small database [7].

3. Image Recognition

In many developed applications Image recognition techniques, especially the intelligent algorithms can be used the same attributes for a large group of training set images; these algorithms can classify and recognize a huge number of different groups of images. In the other word, these intelligent algorithms can discriminate any difference of images, if can be obtained enough number of training set images [8].

In the following sections, the main methods that will be used to extract features extraction on moment's invariant and Geometrical feature is presented, in addition one of the machines learning methods a decision tree that will be used for classification in the proposed algorithm is also presented .

3.1. Moment Invariants

Moments are considered the most widely used in the statistical theory which display an averages or variations for distributions, its works on the description of the contents of the image (or distributed) for the coordinates, thus it was designed to get geometric and detailed information about the image. Moments invariant are used in various image processing applications: identify airplanes, matching of the scene , analysis of shapes , normalization of images , discrimination of the letters written, determine the accuracy of the site , and recuperation of images [9].

Hu in 1961 suggest the idea of using moments in shape recognition The features of image remain fixed in the case underwent image any changes (scale, translation and rotation) when using a set of moment invariants, the following steps illustrate the computation of moment invariant :-

Step1: Compute the central moments, its define the moments that have the property of translation invariance and are denoted by μ_{pq} , it is defined in Eq.(1)[9] :

$$\mu_{pq} = \sum_{x=0}^{x=N-1} \sum_{y=0}^{y=M-1} (x - \bar{x})^p \cdot (y - \bar{y})^q f(x, y) \quad (1)$$

where \bar{x} and \bar{y} are the coordinates of the centered , and they are calculated using Eq. (2) and Eq. (3).

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad (2)$$

$$\bar{y} = \frac{m_{01}}{m_{00}} \quad (3)$$

The following Equations are computed to easily check that the central moments reach to up the order $p+q \leq 3$, Eq. (4) to Eq. (13) [9]:

$$\mu_{00} = m_{00} \quad (4)$$

$$\mu_{10} = 0 \quad (5)$$

$$\mu_{01} = 0 \quad (6)$$

$$\mu_{20} = m_{20} - \bar{x} m_{10} \quad (7)$$

$$\mu_{02} = m_{02} - \bar{y} m_{01} \quad (8)$$

$$\mu_{11} = m_{11} - \bar{y} m_{10} \quad (9)$$

$$\mu_{30} = m_{30} - 3\bar{x} m_{20} + 2\bar{x}^2 m_{10} \quad (10)$$

$$\mu_{12} = m_{12} - 2\bar{y} m_{11} - \bar{x} m_{02} + 2\bar{y}^2 m_{10} \quad (11)$$

$$\mu_{21} = m_{21} - 2\bar{x} m_{11} - \bar{y} m_{20} + 2\bar{y}^2 m_{01} \quad (12)$$

$$\mu_{03} = m_{03} - 3\bar{y} m_{02} + 2\bar{y}^2 m_{01} \quad (13)$$

Step2: normalized central moments using to get the Scale invariance η_{pq} , as Eq. (14) & Eq. (15) [9]:

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \quad (14)$$

Where

$$\gamma = \left\lceil \frac{(p+q)}{2} \right\rceil + 1 \quad (15)$$

Step3: From normalizing central moments for order three, that are invariant in the case underwent image any changes (scale, translation and rotation) compute seven moment are given as Eq. (16) to Eq. (22) [10]:

$$\phi_1 = \eta_{20} + \eta_{02} \quad (16)$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \quad (17)$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \quad (18)$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \quad (19)$$

$$\begin{aligned} \phi_5 = & (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12}) \\ & [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ & [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (20)$$

$$\begin{aligned} \phi_6 = & (\eta_{20} - \eta_{02}) [(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ & + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \end{aligned} \quad (21)$$

$$\begin{aligned} \phi_7 = & (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12}) \\ & [(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\ & + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \\ & [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (22)$$

3.2 Geometrical Features

Geometrical features is one of the most important image attributes that used for discriminating between shapes of cells because shape due to the fact that shape can influence the realization of human , geometrical shape features include:-

- A (Area): is a sum of pixels enclosed by cell boundary.
- P (Perimeter): is a sum of perimeter pixels.
- D (Diameter) = Area / (4×perimeter).
- SGF (Shape Geometric Factor) = larger diameter/smaller diameter.
- AP (Area Proportion) = Area of cell/area of central pallor.
- DV (Deviation): shape geometric [11].

In order to be able to discriminate types of Anemia in the images by using Geometrical features, a robust classifier should be used. The classification is performed by using Decision Tree (ID3) classifier.

3.3 Decision Tree (DT)

A decision tree is one of the machine learning (supervised learning) types that determines the class of a specific case. The tree has many nodes; each node defines either a class label or a specific test that divides the instances at the node depending on the results of the test and ability of producing rules from decision tree. This classifier has two stages [12]:

- Tree Building.
- Tree Pruning.

The building stage is to build the tree depending on local standard where repeated splitting the training dataset that belonging to each of the partitions has the same class label.

The pruning stage that is simplifies the tree, increases classification accuracy ID3 algorithm and solves over fitting problem for dataset in the tree [12].

- **Entropy:** Entropy is the measure for the goodness of a split and evaluates the amount of information in an features.

$$\text{Entropy}(T) = \sum_{n=1} (-P(n) \log_2 P(n)) \quad (23)$$

- P (n) is the probability of the attribute (T).

- **Information Gain:** Information gain is predictable reduction through measure of entropy produced by divided the dataset depended on features , It is considered measurement evaluates activity of the features in classifying the training dataset, where information gain, Gain(M, T) for attribute T, for a collection of examples M, shown in the following equation:

$$\text{Gain}(M, T) = \text{Entropy}(M) - \sum \left(\frac{|MV|}{|M|} \times \text{Entropy}(MV) \right) \quad (24)$$

Where: \sum is sum all values of the attribute T,

MV= subset of M for which attribute T has value V,

|MV|= set of element in MV,

|M|= set of element in M.

3. The Proposed Algorithm

In this proposed algorithm, Anemia in digital images is recognized based on Moment Invariant and Geometric Feature. The classifier ID3 is used to classify Red Blood Cell type, the optimal results of recognition is achieved using Moment Invariant method compared with that based on Geometric Features. The structure of the proposed algorithm is composed of two phases (Training phase and Testing phase) as shown in Figure-1, each phase has specific tasks which illustrated in the following sections.

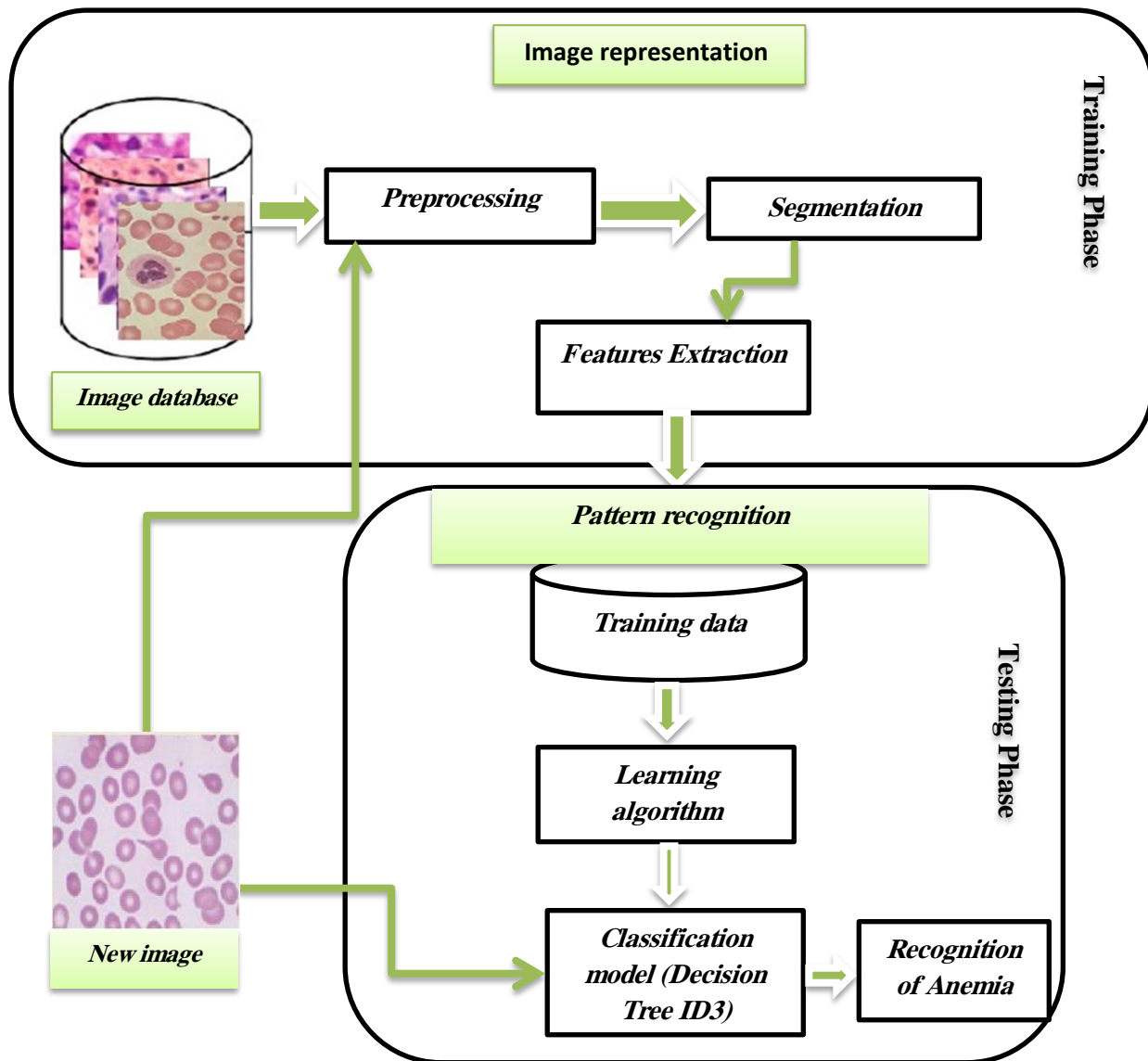


Figure 1- Structure of Proposed algorithm

1. Pre-processing

This stage consists of the following three steps:

- **Collection of Anemia Dataset:** due to the lack of available standard Anemia images, Anemia Dataset is collected from number of hospitals (Al- karama Teaching Hospital, Al-Yarmouk Teaching Hospital and Ibn Al-Balady Hospital). Images are captured in the JPEG format with the maximum resolution. Total number of images is (455), where (280) images are selected for training phase and (175) images are selected for testing phase.
- **Image Conversion:** the color (RGB) images are converted to the gray scale.
- **Noise Removal:** Removal of noise is done by using median filter so as to improve image quality, while preserving the integrity of edges and detail information.

2. Image Segmentation

In this stage, asset of sub divisions each of an (RBC) Red Blood Cell image is achieved to produced number of regions, each region represent one cell from smear of blood, Otsu's thresholding method is used to segmented image by divided the pixels into two clusters (object and background) based on threshold , Histogram is based on clustering , to find the threshold value by the following steps which including : compute the total of histogram, then compute sum histogram for each background and object , find the weight and mean for each class , compute the Variance Between two Class , finally the threshold value consider maximum value of variance between two class. Figure-3 shows the segmented cell.

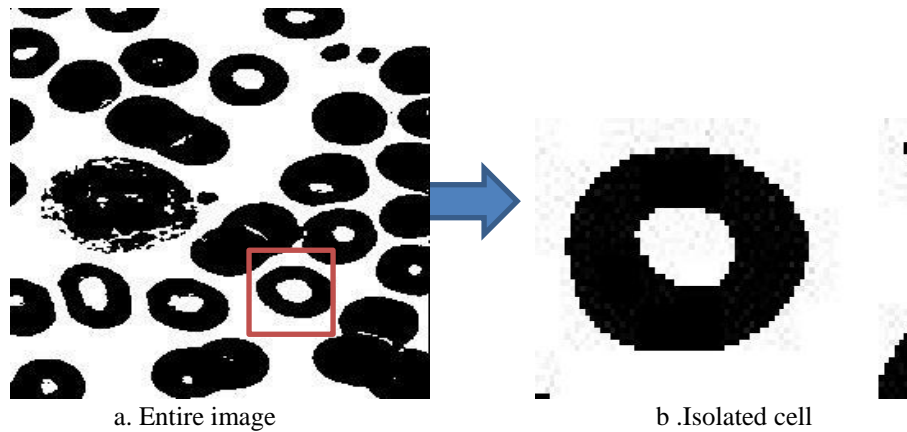


Figure 3- Segmentation of cells in RBC image

3. Feature Extraction

After the images are segmented in pervious stage, in this stage to recognize segmented objects or regions from other objects, so each object consider as pattern with its features, the features are extracted to reduce the original data set and to distinguish one input pattern from another, in this algorithm two types of feature are considered:-

A .Moment Invariants: The moment Invariants feature (as described in section (2.1)) are applied on black and white images, Figure-4 shows the conversion of gray image to black and white image where each cell is represented by 1 for the white pixel and 0 for the black pixel, Moment Invariants Algorithm is summarized by the Algorithm (1.3).

Algorithm 1.3: Moment Invariants
Input : binary image with edge detection
Output : moment invariants each cell in image RBC
Begin
Step 1: Read a black and white image
Step 2: For i=0 to image. width – 1
For j=0 to image. height – 1
Calculate the Central Moments apply equation (1- 3).
Step 3: Compute the Scale Invariance applies equation (4- 13).
Step 4: Calculate the Seven Moment Invariants (based on μ_{pq}) by apply equation (16 - 22).
End

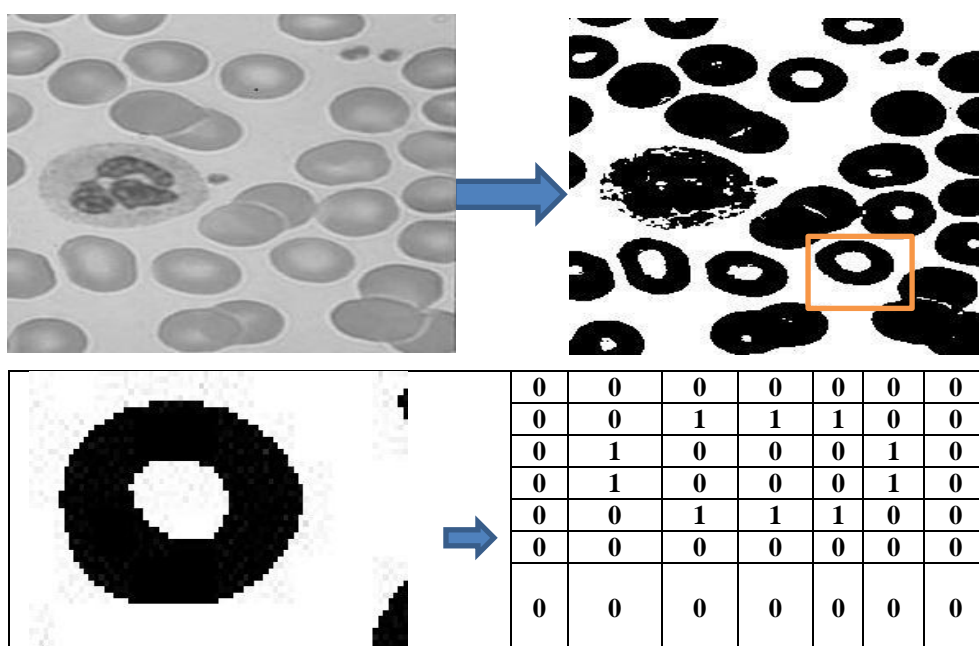


Figure 4- Conversion of gray image to black and white image

The following steps illustrate as an example for moment invariant computation of Figure-4 according to the equations mentioned in section (2.1).

Step 1: Calculate the moment of order shown in the Table-1.

Table 1-Moment order

Moment Order	m_{00}	m_{01}	m_{10}	m_{11}	m_{02}	m_{20}	m_{12}	m_{21}	m_{30}	m_{03}
Value	10	25	40	75	110	77	275	231	256	450

Step 2: Calculated the coordinates center of mass using Eq. (2) , Eq. (3) .

$$\bar{x} = \frac{m_{10}}{m_{00}} = \frac{25}{10} = 2.5 , \bar{y} = \frac{m_{01}}{m_{00}} = \frac{40}{10} = 4$$

Step 3: Calculate the central moment of the order according equations from Eq. (4) to Eq. (13), shown in the Table-2.

Table 2- Central Moment

Central Moment	μ_{00}	μ_{01}	μ_{10}	μ_{11}	μ_{02}	μ_{20}	μ_{12}	μ_{21}	μ_{03}	μ_{30}
Value	10	0	0	25	50	14.5	200	48	410	625

Step 4: Calculate the normalized central moment for the order (η_{20} , η_{02} , η_{11} , η_{30} , η_{03} , η_{21} , η_{12}) as Eq. (14) and Eq. (15) as in the Table-3.

Table 3- Normalized Central Moment

Normalized Central Moment	η_{11}	η_{20}	η_{02}	η_{03}	η_{30}	η_{12}	η_{21}
Value γ	2	2	2	2.5	2.5	2.5	2.5
Value of η	0.25	0.12	0.5	0.12965	1.9769	0.6324	0.1517

Step 5: calculate the seven moment invariant ($\phi_1, \phi_2, \phi_3, \phi_4, \phi_5, \phi_6, \phi_7$) according Eq.(16) to Eq. (22) ,The Table-4 show the result of the seven Moment .

Table 4- The Seven Moment Invariant

Moment invariant	ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7
Value of ϕ	0.65	0.3796	0.21346	2.2325	1.31412	729863	3528833.4

B. Geometrical Features: in this stage six Geometric Features were calculated (as mentioned in the section (2.2)) , Table-5 illustrates the results of applying Geometric Features on each cell of image in Figure-4.

Table 5- Geometric features

Cell Number	Image class	AP	DV	SGF	D	A	P
1	normal	0.02	1	1.31	6.4	33	10
2	iron	0.01	1	1.09	5.3	2	45
3	Target cell	0.08	1	1.18	6.8	4.5	32
4	normal	0.09	1	1.24	6.4	0.5	41
5	normal	0.04	1	1.16	6	2	43
6	iron	0	0	1.11	5.9	2.1	22
7	Sickle Cell	0	0	1.08	6	3	100
8	Target cell	0.05	1	1.20	6.9	0.6	345
9	teardrop	0.13	1	1.04	6.9	2	24
10	iron	0.11	1	1.14	6.7	2.2	24
11	iron	0.04	1	1.10	7	2.8	67
12	hypo	0.09	1	1.25	6.6	1.11	77
13	normal	0.12	1	1.33	6.6	0.02	109
14	iron	0.12	1	1.13	6.7	7.1	453
15	Sickle Cell	0.18	1	1.28	6.8	3.2	56
16	iron	0.06	1	1.20	6.8	2	43
17	Sickle Cell	0.04	1	1.07	5.9	1.2	28

4. Decision Tree

In this stage built decision tree depended on training Anemia images where to test a new image in forecasting phase which is not exist in the training phase to given decision to classify new image, Table-6 shows range value of each feature extracted from the sample of images.

Figure-5 shows the decision tree based on select the best attribute for branch using Entropy and Information gain in the Eq. (23) and Eq.(24) , The algorithm (4.1) of ID3 learning Algorithm.

Algorithm (4.1): ID3 learning Algorithm	
Input: A data set, S	
Output: A decision tree	
<ul style="list-style-type: none"> ▪ Begin: If all the instances have the same value for the target attribute then return a decision tree that is simply this value (not really a tree – more of a stub). ▪ Else <ol style="list-style-type: none"> 1. Calculate Information Gain value (see equation (23)) for all attributes, select an attribute with the highest value, and generate a node for this attribute. 2. Make a branches from this node for every value of this attribute 3. Assign all possible values of this attribute to branches. 4. Follow each branch by partitioning the dataset to be only instances whereby the value of the branch is present and then go back to 1. 	

Table 6- Categorical feature values

Features	Feature-values	Range of values
AP	AP1	[0.23 , 0.363]
	AP2	[0.364 , 0.496]
	AP3	[0.497 , 0.628]
DV	DV1	[0.778 , 0.833]
	DV2	[0.834 , 0.888]
	DV3	[0.889 , 0.942]
SGF	SGF1	[0.359 , 0.529]
	SGF2	[0.53 , 0.699]
	SGF3	[0.7 , 0.87]
D	D1	[0.871 , 1.757]
	D2	[1.758 , 2.643]
	D3	[2.644 , 3.53]
A	A1	[0.587 , 1.144]
	A2	[1.145 , 1.701]
	A3	[1.702 , 2.258]
p	P1	[0.431 , 1.22]
	P2	[0.44 , 1.701]
	P3	[1.302 , 2.678]

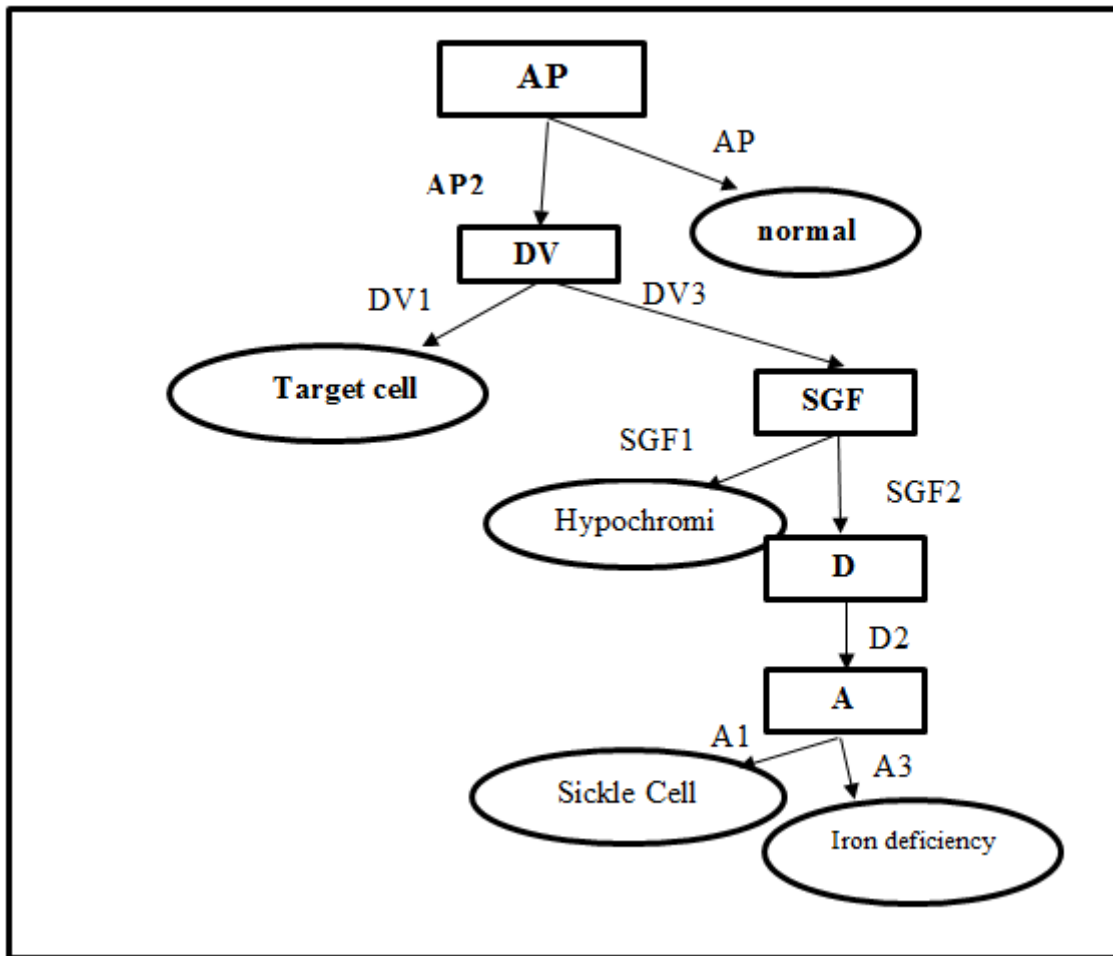


Figure 5- Decision Tree

5. Discussion of Results

Experimental evaluation of the proposed algorithm is performed on two stages (training and testing); these experimental results are compared for two features: Moment Invariant and Geometric Feature.

The training images in this algorithm consists of 7 classes (teardrop, target cell, Ovalocytes, sickle cell, Normal, hypochromic and iron deficiency), each class consist of different number of images. Total number of images is (455), where (280) images are selected for training phase and (175) images are selected for testing phase.

Accuracy measure is the most widely used measurement method to evaluate the classification images, thus to compute the accuracy of the decision tree the following equation is used:

$$\text{Accuracy} = \frac{\text{Number of Correctly Classified Images}}{\text{Total Number of Images}} \times 100\% \tag{25}$$

Table-7 shows the accuracy rate of features extracted by using moment invariant, and Table (8) shows the accuracy rate of features extracted by using Geometric feature.

The optimal results of recognition are achieved using Moment Invariant with rate (84%) compared with rate (74%) when using Geometric Feature.

Figure-6 shows the percentage distribution of the classification accuracy for the types of anemia ,the iron deficiency has higher accuracy than others by using Moment Invariant and Figure-7 shows ratio accuracy classifier the types of anemia the iron deficiency has higher accuracy by using Geometric Features.

Table 7- Accuracy Rate Based on Moment Invariant

Type of Classification	Correct classifiers	Wrong classifiers	No Decision	Accuracy
Teardrop	70	20	8	86%
Target cell	8	2	0	80%
Ovalocytes	25	5	6	83%
Sickle cell	9	1	0	90%
Normal	45	11	5	80%
Hypochromic	80	10	3	87%
Iron deficiency	100	47	0	86%
Average				84%

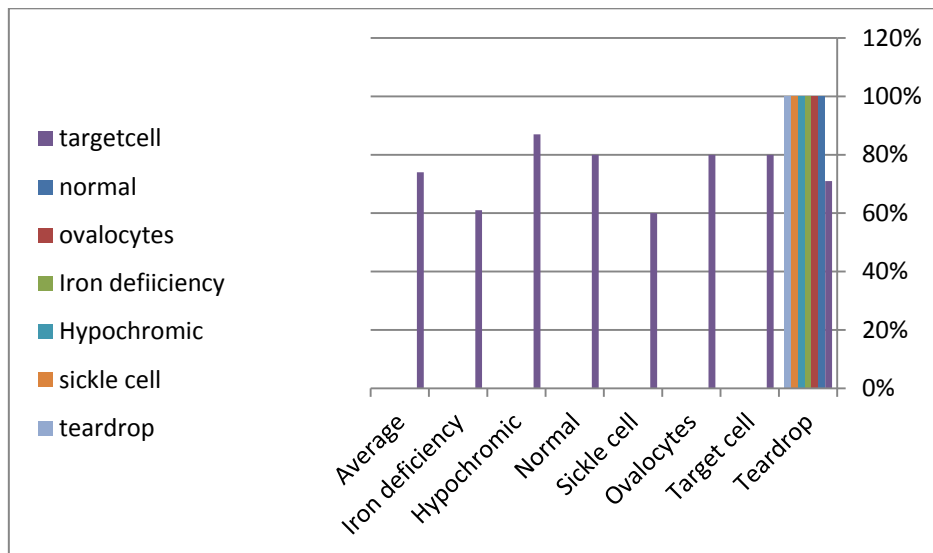


Figure 6- Result of Accuracy Classification Based on Moment Invariant

Table 8- Accuracy Rate Based on Geometric Feature

Type of Classification	Correct Classifiers	Wrong classifiers	No Decision	Accuracy
Teardrop	80	13	5	71%
Target cell	8	2	0	80%
Ovalocytes	30	3	3	80%
Sickle cell	6	4	0	60%
Normal	45	11	5	80%
Hypochromic	80	10	3	87%
Iron deficiency	90	57	0	61%
Average				74%

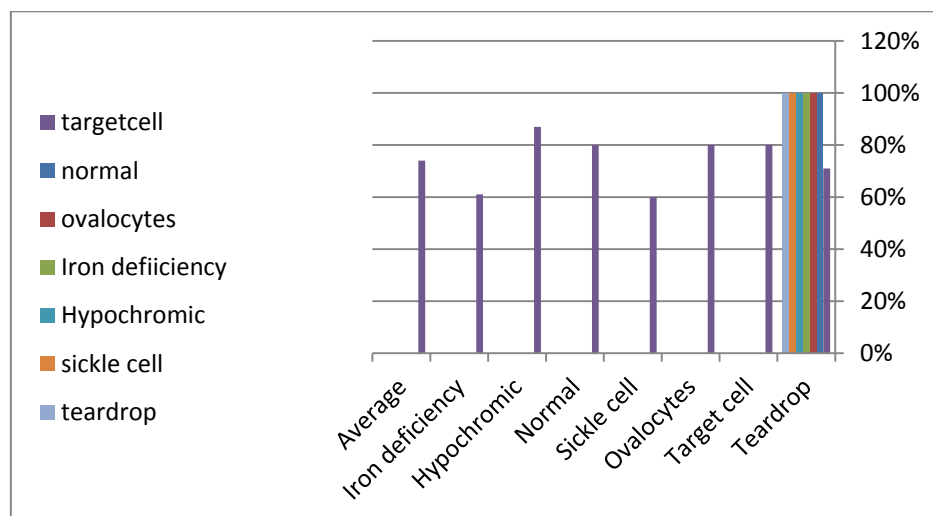


Figure 7- Result of Accuracy Classification Based on Geometric Feature

6. Conclusion

In this paper, recognition of Anemia images algorithm is presented, this algorithm includes four stages: preprocessing, segmentation, feature extraction and decision Tree; the feature are extracted using Moment Invariants and Geometrical Features, these features values are classified into classes by using one of the machine learning algorithm classifier which is Decision Tree (ID3) algorithm. Accuracy results indicate that discrimination power of Moment Invariants is better than Geometrical Features since it produced (84%) accuracy rate compared with (74%) accuracy rate when Geometric Feature is used.

References

1. Seng, W. C., Mirisae, S.H. **2016**. A New Method for Fruits Recognition System. IEEE Xplore, University of Malaya.
2. Ali, J., Ahmad, A.R., George, L.E., Der, C.S. and Aziz, S. **2013**. Red Blood Cell Recognition using Geometrical Features, *International Journal of Computer Science*, 10(1).
3. Barpanda, S. S. **2013**. Use of Image Processing Techniques to Automatically Diagnose Sickle-Cell Anemia Present in Red Blood Cells Smear. Department of Electrical Engineering National Institute of Technology Rourkela, ODISHA, INDIA.
4. Chandrasiri, S. and Samarasinghe, P. **2014**. Morphology Based Automatic Disease Analysis Through Evaluation of Red Blood Cells. Department of Information Technology/ Sri Lanka Institute of Information Technology, 2014 Fifth International Conference on Intelligent Systems, Modelling and Simulation.
5. Al-Abudi, B. Q. **2009**. Digit Recognition Using Fractal and Moment Invariants. *Iraq Journal of Science*, 50, pp:110-119.
6. Dawood, F.A. **2008**. Unsupervised Segmentation of Microscopic White Blood Cells Images Using Histogram Equalization Technique, *Iraq Journal of Science*, 49, pp:204-211.
7. Amer, W. A. **2009**. Efficient and Fast Distorted Character Recognition Algorithm, *Iraq Journal of Science*, 50, pp:110-119.
8. Umbaugh, S. **2011**. *Digital Image Processing and Analysis*. Second Edition, Taylor & Francis Group, LLC.
9. Hosny, K.M. **2007**. Exact and fast computation of geometric moments for gray level images, *Applied Mathematics and Computation*, pp:1214-1222.
10. George, D. N. **2013**. Tumor Type Recognition Using Artificial Neural Networks, M.Sc. Thesis, Iraqi Commission for Computers and Informatics Informatics Institute for Postgraduate Studies.
11. Taherisadr, M., Nasirzonouzi, M., Baradaran, B. and Mehdizade, A. **2013**. New Approach to Red Blood Cell Classification Using Morphological Image Processing, *Shiraz E-Medical Journal*, 14(1).
12. Mazien, H. **2015**. Proposed system for the diagnosis of skin diseases using Multiwavelet Transform and Decision Tree, M.Sc. Thesis, University of Technology, Department of Computer Science.