ISSN (e): 2250 – 3005 || Volume, 08 || Issue, 9|| Sepetember – 2018 ||
International Journal of Computational Engineering Research (IJCER)

Mathematical Model of Computer Aided Detection for Acute Leukemia

Najaat A. A. Abdullah[1], Prof. Mohammed A. M. Ibrahim[2], Prof. Adel S. M. Haider[3]

[1] *Department of Information Technology, Faculty of Engineering, Aden University, Yemen,*
[2] *Department of Information Technology, Faculty of Engineering, Taiz University, Yemen,*
[3] *Department of Information Technology, Faculty of Engineering, Aden University, Yemen,*
Corresponding Author: Najaat A. A. Abdullah

ABSTRACT: In the leukemia detection process, human errors, differences between observers based on their experiences and a huge variation in the time taken for diagnoses are some problems that need to overcome. Therefore, this paper proposes a computer-aided diagnostic mathematical model that assists in early detection of acute leukemia subtypes. The results showed that the conversion of images into L*a*b* color space produces uniformity in contrast compared to RGB color space. The resulted different features for each object indicate that the model can be used accurately to detect acute leukemia subtypes. To achieve this goal Lightness channel which gives a good contrast of the objects is selected to clearly extract the relevant object. The different features of selected objects suggest the ability to use the proposed model to distinguish the different acute leukemia subtypes. We hope the proposed model achieve a higher accuracy compared to available models in literature.
Keywords: Acute Leukemia, mathematical model, RGB color space, Lab color space, pre-processing, segmentation, feature extraction, theoretical test.

Date of Submission: 04-09-2018 Date of acceptance: 20-09-2018

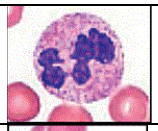
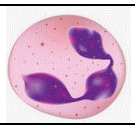
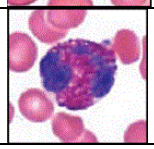
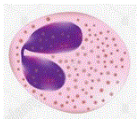
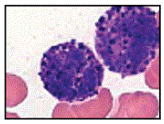
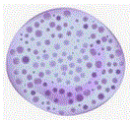
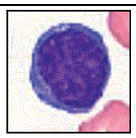
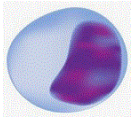
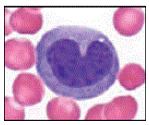
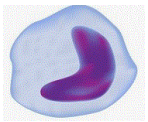
I. INTRODUCTION

Our body creates new cells continuously until growth is done, replace the dead cells and treat the damaged cells after injury by using certain genes. The damage to these genes leads to an abnormal behavior of cells called cancer. According to the World Health Organization, there were 8.8 million cancer deaths in 2015 compared to 8.2 million cases of 2012[1], which provides an overview of risk and cancer trends in the coming years.

Different types of cancers that effect the human body, including pancreatic cancer, lung cancer, colorectal, breast cancer, leukemia, etc. In leukemia cancer, tissues responsible for the production of blood cells including bone marrow and lymphatic system produces abnormal White Blood Cells (WBCs) or leukocytes. Theses abnormal leukocytes should die after a period of time but they don't and thus they become numerous in count[2]. There are five types of WBCs which classify based on granules into two main classes; granulocyte; that appears in the cytoplasm such as neutrophils of 40-70%, eosinophil of 5% and basophil of 0.5% and agranulocyte; that absent in the cytoplasm including lymphocytes of 20-50% and monocytes of 1-5%[3]. However, it easy to confuse these five types in the blood smear, thus we need to look to the morphology of nucleus. Table I shows the difference features of WBCs types according to granules staining and nucleus morphology[4].

The rapid development of the leukemia disease depends on how fast WBCs develop[5]. Therefore, leukemia also classify into two types: in **Acute Leukemia**; the abnormal blood cells (blasts) usually develop quickly and cannot do normal work. While, **Chronic Leukemia**; involves more mature blood cells that accumulate more slowly and can perform their normal function for a period of time. In diagnostic process, Leukemia is also named for the type of the WBCs that is affected. The rest of this paper is organized as follows. Section 2 discusses the literature review. Section 3 discusses subtypes of leukemia. Section 4 paper methodology. Section 5 provides a description of our proposed scheme in detail. Section 6 provides a brief analysis for our proposed model. Section 7 conclusion of our work. Section 8 future work.

Table I: Types of WBCs

Properties → / WBCs Subtypes ↓	Microscope appearance	Diagram	Diameter (µm)	Nucleus	Granules
Neutrophil			12 - 14	Multi-lobed	Fine, faintly pink
Eosinophil			12 - 17	Bi-lobed	Stain bright red, or red reddish - purple
Basophil			14 – 16	Bi-lobed or tri-lobed	Large blue
Lymphocytes			6 – 9 10 - 14	Deeply staining, eccentric	NK-cells and cytotoxic
Monocytes			Up to 20	Kidney shaped	None

https://www.histology.leeds.ac.uk/blood/blood_wbc.php.

II. LITERATURE REVIEW

The significant amount of research work has been carried out to improve the diagnostic process of the acute leukemia by using the machine learning and image processing as a computer aided system. The detailed description of the literature for acute leukemia detection on the blood smear images are listed in Table II. From the literature the highest accuracy of the classification of healthy leukocytes and leukemia leukocytes is 99.5 obtained in the [6, 7] studies which were gained by various extraction features by using classifiers like Decision tree, Artificial Neural Network (ANN) and Support Vector Machine (SVM), respectively. However, in the first study [6] study did not discriminate between lymphoid and myeloid subtypes, and the observed accuracy achieved by the second study [7] decreased to 97.1 and 98.5 % during the classification process of acute leukemia subtypes. From the extensive study of literature, it can be seen that there are rare studies related to discriminate of different acute leukemia subtypes. So there is a need to build a reliable computer aided classification system to improve diagnostic process of the acute leukemia with high accuracy and in short time.

Table II: Studies carried out for the healthy leukocytes and acute leukemia classification

Authors, Year	Feature extraction method	Classifier used	Dataset used	No. of images	Accuracy (%)		
					Healthy or Leukemia	Leukemia subtypes	
						ALL	AML
Rawat, J. et al. 2015[8]	Shape and texture features	SVM	ALL-IDB	130	89.8	-	-
Negm, A. et al., 2017[9]	Geometry, color, texture and size features	Decision tree and ANN	-	642	99.519	-	-
Patel, N. and Mishra, A. 2015[2]	Color, geometry, texture and statistical features	SVM	ALL-IDB	27	93.57	-	-
Paswan, S. and Kumar, Y. 2017[10]	Color, texture, GLCM features and HD	KNN & SVM	ASH	100	61.11 & 83.33	-	-
Rawat, J. et al., 2017[7]	Color, geometric and texture features	SVM	ASH	240	99.5	97.1	98.5
Dumyan, S. and Gupta, A. 2017[11]	Shape, texture, statistics and moment invariants features	ANN	-	36	97.9	-	-

Kumar, P.and Vasuki, S., 2017[6]	GLCM, geometric and color features	SVM	-	70	90		
Priya, D. K. et al., 2015[12]	HD, LBP, shape and texture features	FFNN	ASH	80	98	-	-
Agaian, S. and Chronopoulos, A. T., 2014[13]	HD, color, shape and texture features	SVM	ASH	80	98	-	-
Sukanya, C. M. et al., 2016[14]	DRLBP and DRLTP	SVM	-	60	-	-	-
Kazemi, F. et al., 2016[15]	HD, shape, color and texture features	SVM	Shariati Hospital	330	98	-	87
Goutam, D. and Sailaja, S., 2015[16]	Texture features	SVM	ASH	90	98	-	-
GLCM is Gray Scale Co-occurrence Matrix, HD is Hausdroff dimension, LBP is Local Binary Pattern, DRLBP is Discriminative Robust Local Binary Pattern, DRLTP is Discriminative Robust Local Ternary Pattern, KNN is k-nearest neighbor, FFNN is feedforward Neural Network							

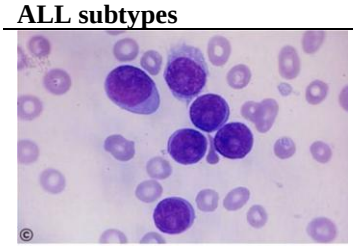
III. SUBTYPES OF LEUKEMIA

The routine examination of the blood smear under microscope classifies acute and chronic leukemia into four subtypes [5, 17, 18]; **Acute Myeloid Leukemia (AML)**, which starts in myeloid cells of bone marrow and quickly moves into the blood, is the most common type of leukemia diagnosed in both adults and children. **Acute Lymphocytic Leukemia (ALL)**, this type of cancer mostly spread in children under the age of 14 years. **Chronic Myeloid Leukemia (CML)**, is attributed to a defect in the Philadelphia chromosome that responsible of a genetic mutation in BCR ABL gene. This gene produces an abnormal protein called Tyrosine Kinase. **Chronic Lymphocytic Leukemia (CLL)**, as opposite of ALL is common in adults and characterized by the accumulation of small, mature-appearing lymphocytes in the blood, marrow and lymphoid tissues. However, the patient can feel good for several years without the need for any treatment.

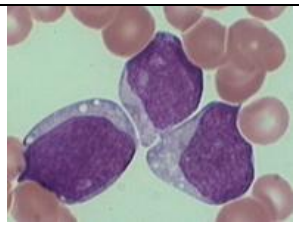
The French American British (FAB) classification system divides ALL and AML into eleven subtypes, for ALL L1 to L3 and for AML M0 to M7 based on the cell morphology[19]. Table III and Table IV show the FAB classification of ALL and AML subtypes [20-22]

Table III: FAB Classification

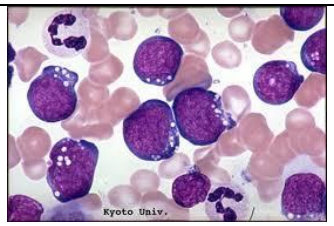
ALL subtypes



L1: Lymphoblastic leukemia with homogeneous structure



L2: Lymphoblastic leukemia with varied structure

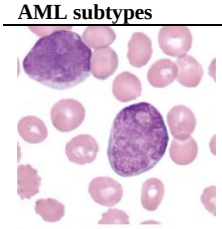


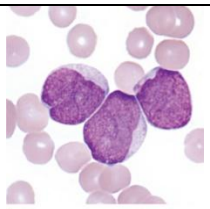
L3: Burkitt's leukemia

http://www.elsevier.es/en-revista-revista-medica-del-hospital-general-325-articulo-morphology-leukaemias-S0185106315000724

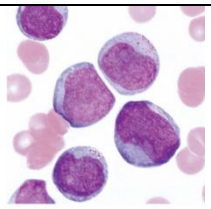
Table IV: FAB Classification

AML subtypes

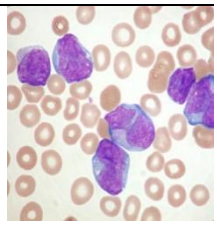




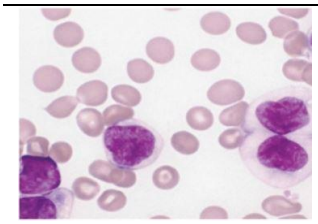
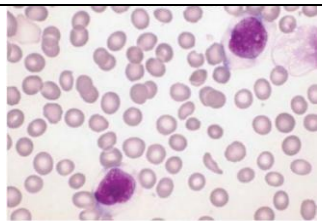
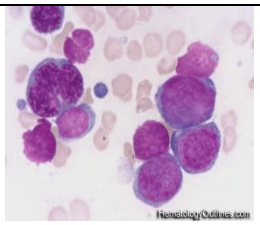
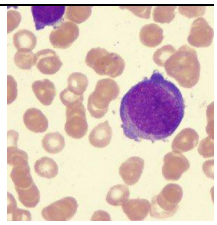
M1: Acute myeloblastic leukemia



M2: Acute myeloblastic



M3: Promyelocytic

M0: Acute myeloblastic leukemia with minimal differentiation	without maturation	leukemia with maturation	leukemia
			
M4: Acute myelomonocytic leukemia	**M5:** Acute monocytic leukemia	**M6:** Acute erythroid leukemia	**M7:** Acute megakaryocytic leukemia
https://www.wikidoc.org/index.php/Acute_myeloid_leukemia_pathophysiology			

However, the downside associated with diagnostic techniques used by hematologist in blood test make a difficult to differentiate among different subtypes. Thus this paper aims to generate a successful model that be used to assist hematologist to detect acute leukemia blood cells.

IV. METHODOLOGY

In this **section**, the mathematical equations of image processing, which is programmed using MATLAB software, will be discussed. The proposed system for automated acute leukemia detection from blood smear images is shown in Figure 1.

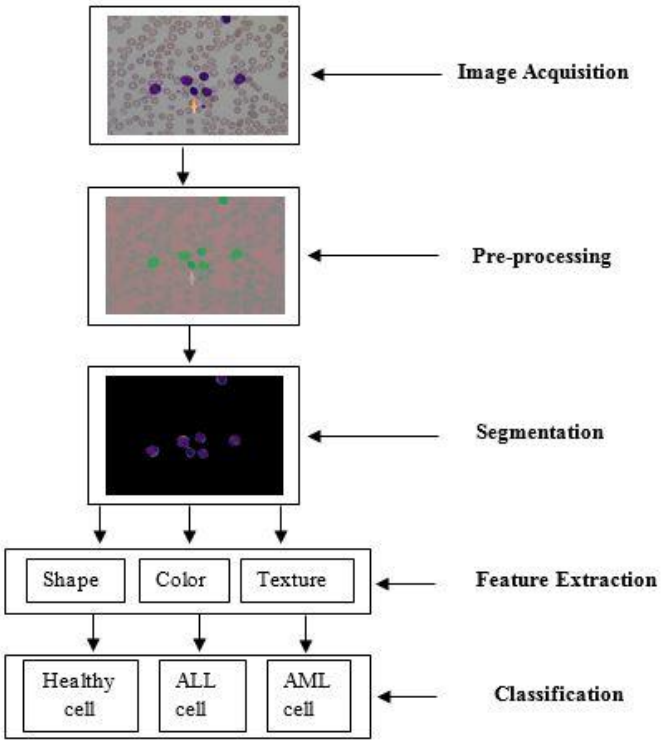


Figure 1. Flowchart of the steps will be achieved in the proposed mathematical model

V. DESCRIPTION OF OUR PROPOSED SCHEME

1.1. Image Acquisition

In this step, two data sets of acute leukemia images; Acute Lymphocytic Leukemia – Image Database (ALL-IDB)[23] and the American Society of Hematology (ASH) [24], captured with an optical laboratory microscope will be taken. This provides a wide range of high quality images for the following steps.

1.2. Pre-processing

The acquired image mostly contains some sort of noise and bluer regions. So before further processing, a median filter will be suggested to enhance the quality of the input image [25]. The non-uniformity in the contrast between cells and background prevents a right segmentation thus to facilitate the interpretation of the image and extract the object information, the RGB color space of the input image should be converted to the L*a*b* color space by using the basic equations [26].

Assuming that CC is denoted to each RGB color channel, which is normalized by divided by $2^{\text{bit depth}}$ (255). The following boundary conditions of $C_{R,G,B}$ is applied to the normalized CC to get the XYZ three dimensional space values as follows.

$$C_{R,G,B} = \begin{cases} \dfrac{CC}{12.92} & if\ CC \leq 0.04045 \\ \left(\dfrac{CC + 0.055}{1.055}\right)^{2.4} & otherwise \end{cases}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.608 & 0.174 & 0.201 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.117 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The transformation to CIELab is done by using a reference white point which is ($X_{ref.}$, $Y_{ref.}$ and $Z_{ref.}$)

$$L = (116 * f(Y/Y_{ref.})) - 16$$

$$a = 500 * [f(Z/Z_{ref.}) - f(Y/Y_{ref.})]$$

$$b = 200 * [f(Y/Y_{ref.}) - f(X/X_{ref.})]$$

where $f(x) = \begin{cases} x^{1/3} & if\ x > 0.008856 \\ (7.787 * x) + (16/116) & otherwise \end{cases}$

1.3. Segmentation

Although the b channel gives more localized blast cells without the background interfering, the L channel is selected in the theoretical test due to the higher contrast to identify the specific region of acute leukemia blast cells.

To detect the blast cell regions from L color channel subspace the Otsu's method is used [27] to calculate the optimum threshold that has the lowest weighted within-class variance. The value of each point in the selected channel is compared to the optimum threshold in order to get the binary image based on the following equation.

$$\forall (i,j) \in I_{output}, I_{output}(i,j) = \begin{cases} 1 & I(i,j) \geq T \\ 0 & I(i,j) < T \end{cases}$$

To extract the correct features of acute leukemia cells, the single cells crop from the binary image while the overlapped cells will be separated by using a watershed segmentation which is defined as follows[28].

$$Wshed(f) = D \cap \left(\bigcup_{i \in I} CB(m_i)\right)^c$$

where D is the topographical distance between two points p and q and can be given by:

$$T_f(p,q) = inf_\gamma \int_\gamma \|\nabla f(\gamma(s))\| ds$$

$CB(m_i)$ is the catchment basin of the minimum and can be given by:

$$CB(m_i) = \{x \in D | \forall_j \in I \backslash \{i\} : f(m_i) + T_f(x, m_i) < f(m_j) + T_f(x, m_j)\}$$

The resulted blast cell consists of two parts cytoplasm and nucleus. Therefore, we need to segregate the nucleus from the cytoplasm.

1.4. Feature Extraction

The aim of this step is to define a large set of data into feature set with a reduced dimension. The extracted features based on shape, texture and color from regions selected in the segmentation process; nucleus and whole cell,[29] provides valuable information for the classification of acute leukemia cells. In the following we will describe the three important features will be used in this study.

1.4.1. Shape Features: that provides information on image geometric features; such as area, perimeter, diameter, solidity, etc., from the cell and its nucleus. Then the Elongation, Eccentricity, Rectangularity, Convexity and Compactness [2] will be calculated as follows:

- Elongation:

$$Elongation = 1 - \frac{minor\ axis}{major\ axis}$$

- Rectangularity

$$\text{Rectangularity} = \frac{area}{major\ axis * minor\ axis}$$

- Eccentricity

$$\text{Eccentricity} = \frac{\sqrt{major\ axis^2 - monor\ axis^2}}{major\ axis}$$

- Convexity

$$\text{Convexity} = \frac{perimeter_{convex}}{perimeter}$$

- Compactness

$$\text{Compactness} = \frac{4 * \pi * area}{perimeter^2}$$

1.4.2. Texture Features: which study the nucleus chromatin structure and divides into two types; histogram-based features and GLCM features. In histogram-based features: a histogram plot displays the distribution of pixels among grayscale values. The common features such as mean, standard deviation, etc., and their equations are listed in the Appendix.

GLCM features: will be used to determine the spatial dependency, the relationship between pairs of pixels at a time, of grayscale values in the image by relying on the second order statistics [30-32] as:

$$G_{dx,dy}(i,j) = \sum_{x=1}^{n}\sum_{y=1}^{m}\begin{cases}1, if\ I(x,y) = i\ and\ I(x+dx, y+dy) = j\\0, \qquad\qquad otherwise\end{cases}$$

where i and j are the gray level values, x and$y$ are the spatial position in the image and the offset (dx, dy) depends on the direction θ and distance D. The sum of the squared elements in the GLCM (energy), the measurement of the intensity between two neighboring pixels over the whole image (contrast), a correlation between two neighboring pixels over the whole image (correlation), and the measurement of the proximity of the distribution of elements in the GLCM to GLCM diagonal (homogeneity) will be handed by MATLAB function (graycomatrix) while, another set of texture features will be calculated. The equations for each of these features are listed in the Appendix.

1.4.3. Color Features: the mean values of the gray image will be produced for each color channel such as; mean, standard deviation, skewnees, kurtosis, energy and entropy.

1.5. Classification

In the final stage, the calculated feature values will be used to categorize leukemia patients into ALL and AML subtypes. To do that, first the entire specific data will be divided into training and testing data sets. Second, the feature vector of training dataset will be assigned into the class label of ALL and AML subtypes. Finally, the values of the testing dataset will be checked with the previous calculated values in the training step. Based on the values of the input image the classifier classifies the test image into ALL and AML subtypes with their discrimination from healthy cells. To evaluate the efficiency of the classifier a confusion metric shown in Table V will be introduced.

Table V: Confusion Metrics

		True Class	
		Abnormal	**Normal**
Predicate Class	**Abnormal**	True Positive (TP)	False Positive (FP)
	Normal	False Negative (FN)	True Negative (TN)

where TP is the number of the correct samples that will be categorized as a cancerous cell. TN is the number of the correct samples that will be categorized as a noncancerous cell. FP is the number of the wrong samples that will be categorized as a cancerous cell. FN is the number of the wrong samples that will be categorized as noncancerous cells [33]. The confusion metrics contain information about the true and predicate values of the testing dataset obtained from a classifier. Based on the outcomes of the confusion metric accuracy, sensitivity and specificity will be measured as follows [34].

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

A sensitivity, which used as a measure of the ability of the classifier to correctly identify the blast cells and can be calculated as:

$$Sensitivity = \frac{TP}{TP + FN}$$

While specificity, used as a measure of the ability of the classifier to correctly identify the healthy cells and can be calculated as:

$$Specificity = \frac{TN}{TN + FP}$$

VI. TEST AND ANALYSIS

To test our proposed model the experiments are conducted mathematically by 16*16*3 matrix on RGB image. The result of the proposed color space conversion is shown in Table VI and in Figure 2.

Table VI: The 16*16*3 matrix of RGB image and L*b*a* color space

Red	Green	Blue
141 151 140 149 144 143 140 160 152 156 158 154 124 115 154 152 154 154 114 141 147 162 158 155 161 154 155 154 146 155 144 130 153 153 152 160 147 161 161 158 161 159 159 158 154 157 161 131 156 163 167 164 160 158 158 116 142 157 132 155 159 158 153 157 153 159 163 162 155 157 153 135 139 146 106 155 160 161 153 157 152 156 156 158 158 156 149 160 170 137 142 165 154 153 151 130 156 157 159 157 160 156 153 140 137 103 123 162 155 154 150 121 157 153 154 161 161 156 161 81 74 84 151 165 167 103 135 140 155 159 157 160 157 160 152 86 84 134 152 161 159 112 144 151 155 161 157 161 158 156 146 120 145 144 143 156 151 152 161 154 154 153 159 162 158 158 157 143 158 153 154 157 154 153 159 153 150 153 160 161 133 143 156 158 158 152 151 155 156 152 152 152 151 159 155 161 123 117 156 152 151 149 111 146 157 150 152 155 153 155 150 150 156 154 155 148 147 156 139 158 159 152 151 154 155 149 157 151 157 162 156 145 139 162 161 159 153 153 154 152 133 150 159 156 151 154 153 148 117 154 155 154 148 148 150 153	162 140 111 130 144 161 157 159 141 149 152 144 105 93 144 143 146 146 61 116 170 157 153 147 159 146 144 144 134 151 147 101 143 146 140 156 170 156 158 151 157 153 153 153 149 153 161 105 149 165 171 165 155 151 154 71 115 151 111 146 159 155 144 153 143 154 162 160 145 151 143 111 112 131 79 147 160 163 145 152 141 148 149 152 151 147 135 157 174 118 130 166 147 147 162 106 150 151 154 150 155 149 142 124 121 76 103 159 149 150 139 84 152 143 145 159 158 146 123 40 29 46 160 166 173 74 122 117 150 157 152 156 151 155 143 50 41 115 141 161 158 85 136 144 150 162 153 160 152 147 132 98 131 164 163 152 144 145 162 151 148 147 157 162 152 154 150 162 152 144 145 153 151 150 160 149 141 145 159 161 107 126 151 156 156 142 141 150 156 148 147 147 143 159 151 163 89 77 151 143 142 136 58 131 157 143 147 153 149 152 141 141 153 149 151 135 131 151 113 156 163 148 146 153 154 140 156 143 154 166 154 129 115 166 162 161 152 151 152 150 118 146 162 154 143 151 149 139 78 150 152 151 163 144 147 153	156 143 138 143 158 157 155 154 138 146 149 139 129 130 138 140 143 150 132 149 162 155 151 146 157 146 144 144 135 149 158 140 143 150 157 161 165 155 157 151 154 152 150 151 147 150 153 135 149 163 166 161 153 151 155 130 146 151 131 146 155 154 143 147 144 154 161 158 147 153 144 142 145 140 131 148 156 159 145 145 142 149 150 153 153 148 138 153 165 131 147 161 146 145 141 138 151 152 155 151 155 150 144 142 150 133 142 155 147 149 140 131 152 145 146 158 157 146 136 123 122 123 146 160 165 133 145 135 151 159 154 156 152 156 152 126 124 138 140 157 153 127 149 143 152 162 158 160 153 149 141 140 142 159 160 151 143 141 156 147 150 149 158 162 152 156 152 159 153 145 144 151 150 148 156 146 144 148 141 140 141 151 156 155 143 146 150 153 145 145 145 145 145 140 153 161 141 137 151 146 145 150 130 145 153 141 145 150 151 154 145 145 158 160 153 148 154 159 148 159 158 145 143 149 153 144 158 147 155 164 155 148 149 162 160 158 150 148 149 148 144 152 161 154 145 153 150 146 135 156 149 149 141 141 146 149
L*	a*	b*
169 152 132 146 171 169 166 167 153 159 162 155 122 113 155 154 156 157 98 137 176 166 162 158 167 157 156 155 146 160 173 124 154 157 154 166 176 165 167 161 166 163 163 162 159 162 168 126 159 172 177 172 164 161 163 103 136 161 129 157 167 164 155 162 155 164 170 168 157 161 155 131 134 145 103 158 168 170 156 161 153 159 160 162 162 158 148 166 179 134 144 173 157 157 153 126 160 161 164 160 165 160 154 140 138 101 123 167 159 159 151 111 162 155 156 167 167 157 139 75 68 78 153 173 178 100 137 135 160 166 162 165 161 165 155 81 77 132 153 168 166 108 149 154 160 170 163 168 162 158 146 119 145 171 171 161 154 155 169 160 158 157 166 170 162 163 161 170 162 155 156 162 160 159 167 158 153 156 168 169 128 143 161 165 165 154 153 160 164 157 157 157 154 167 161 170 116 107 161 155 154 150 96 146 165 153 157 161 159 161 153 153 163 160 161 149 147 162 134 165 169 157 156 161 162 152 165 154 163 172 163 145 136 172 169 168 160 160 161 159 135 156 169 163 154 160 159 151 107 160 161 160 153 153 156 161	127 134 147 140 127 128 129 128 133 131 131 133 142 145 132 132 132 133 164 146 125 130 130 132 129 132 134 133 135 130 125 149 133 132 137 131 126 130 129 132 130 131 131 130 130 127 146 132 127 125 127 130 132 130 159 146 131 142 133 127 129 133 129 133 131 128 129 134 131 133 145 147 137 150 132 127 126 132 130 134 132 132 131 132 133 136 129 125 140 137 127 132 131 133 145 131 131 131 132 131 132 134 139 141 151 144 129 131 130 134 154 131 134 133 129 129 133 139 161 163 159 135 127 124 152 138 143 131 129 131 130 131 130 134 158 162 141 134 127 128 149 134 132 131 127 130 129 131 133 137 146 137 127 128 130 132 131 127 129 131 131 129 128 131 130 132 128 131 133 133 130 129 127 130 133 133 129 128 147 141 131 129 129 133 134 131 128 130 130 130 133 128 130 127 153 157 131 133 133 137 164 138 127 131 130 129 130 130 133 133 130 132 130 137 140 132 147 129 125 130 130 128 128 133 129 133 130 126 129 139 145 125 127 127 128 129 129 129 140 131 126 129 133 130 130 134 156 131 129 129 130 130 129 127	131 129 119 125 131 131 130 131 132 131 131 133 119 112 133 131 131 127 101 115 132 130 130 130 129 130 130 130 130 130 132 113 130 127 121 126 130 130 129 129 130 130 131 130 130 130 132 117 129 129 130 130 130 129 128 106 117 129 121 130 130 129 130 132 129 129 129 129 129 128 129 116 116 126 105 129 130 130 130 133 130 129 129 129 128 129 129 131 132 125 121 130 130 130 130 115 129 129 128 129 129 129 129 121 115 102 111 131 130 129 130 110 129 129 129 129 129 130 125 92 88 94 127 131 131 101 118 123 128 127 128 129 129 130 125 94 92 119 131 130 131 110 122 130 128 128 128 128 129 129 126 109 125 131 130 129 130 132 131 131 128 128 128 128 129 129 128 130 129 129 130 130 129 130 130 130 128 128 127 129 115 118 129 128 129 129 127 129 130 130 130 130 129 127 128 129 107 104 129 128 128 123 101 123 130 131 130 130 128 128 128 128 126 123 128 124 119 125 114 127 130 130 131 130 129 128 127 127 128 128 128 121 114 129 129 129 129 130 130 129 117 126 128 128 129 128 128 126 105 126 130 130 130 130 130 130

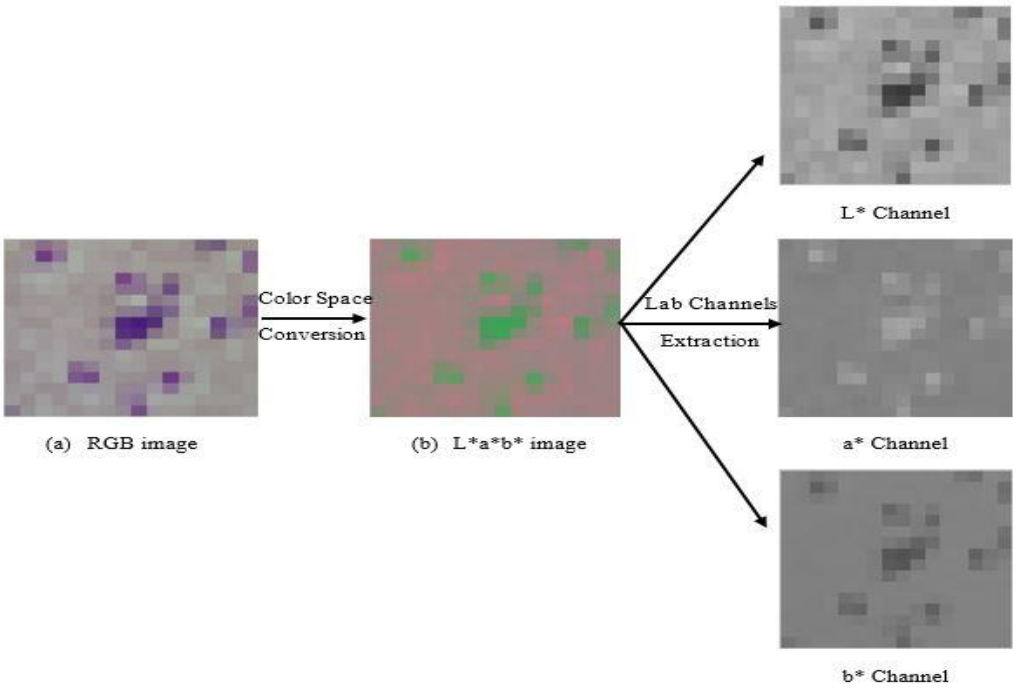


Figure 2. The Lab color space channels

As can be seen the objects are more contrasted in L* channel which provide a good chance to recognize the object region. To highlights the desired objects of the binary image resulting from the segmentation process using Otsu's method, the complementation is done in Table VII and in Figure 3.

Table VII: Otsu's method and its complementation

1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1
1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1
1	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1
1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0
1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	0
1	1	1	1	1	1	1	0	0	0	1	1	1	0	0	0
1	1	1	1	1	1	1	0	0	0	1	1	1	0	1	1
1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
1	1	1	1	0	0	1	1	1	1	0	1	1	1	1	1
1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1

0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0
0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0
0	0	0	0	0	0	0	1	1	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	1
0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	1
0	0	0	0	0	0	0	1	1	1	0	0	0	1	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0

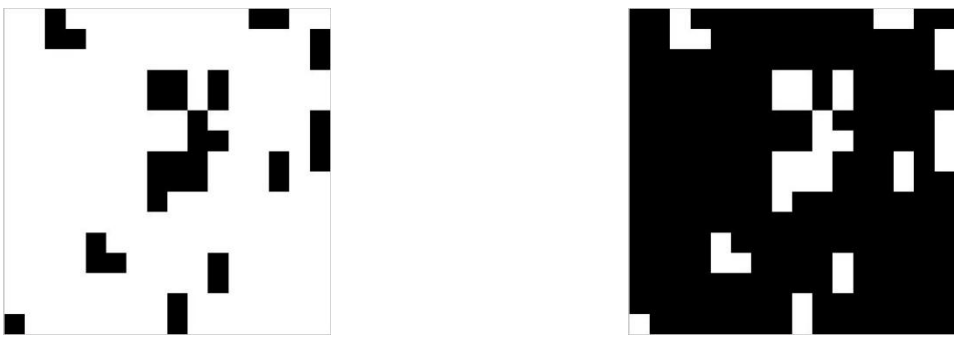


Figure 3. Segmentation process of binary image in left and complementation of binary image in right

These results suggest that the proposed model provides an easy way that will help to reflect the object in the input image and facilitate its feature extraction. In order to crop each object individually in the binary image and find out its features the connected objects are labeled after removing the boundary objects. The resulted cropped objects are shown in Figure 4.

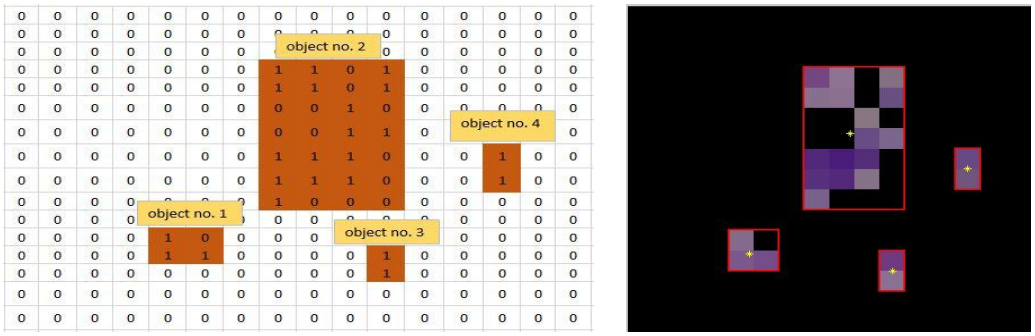


Figure 4. Labeling process of connected components, matrix in left and input image in right

In our proposed model the shape, color and texture features of the extracted objects are calculated in Table VIII.

Table VIII: Feature extractions from the image

Feature Extracted	Object no. 1	Object no. 2	Object no. 3	Object no. 4
Shape Features				
Size	3	16	2	2
Perimeter	3.4142	23.3134	2	2
Diameter	1.95441	4.513517	1.595769	1.595769
Roundness	3.234098	0.369929	6.283185	6.283185
GLCM Features				
Energy	0.3888889	0.208984375	0.333333333	0.333333333
Contrast	3	3.28125	4.333333333	4.333333333
Homogeneity	0.750	0.622395833	0.527777778	0.527777778
Entropy	1.011404265	1.934941269	1.098612289	1.098612289
Color Feature				
Red Channel				
Mean	0.162527233	1.1225	0.163398693	0.140522876
Standard Deviation	0.230333408	1.42298981	0.233244523	0.19899036
Skewness	0.720604663	0.596676053	0.765236979	0.71543576
Kurtosis	-1.461442769	-1.444437345	-1.336341669	-1.47645753
Entropy	1.002718265	1.747389797	0.867563228	0.867563228
Energy	0.481481481	0.37125	0.5	0.5
Green Channel				
Mean	0.119854248	0.821875	0.111764706	0.103921569
Standard Deviation	0.170526871	1.120395961	0.169880585	0.14749391
Skewness	0.792842461	0.884306187	0.099784505	0.729693862
Kurtosis	-1.253016201	-0.907658972	-0.426357066	-1.43622791
Entropy	1.002718265	1.712732438	0.867553228	0.867563228
Energy	0.481431481	0.3725	0.5	0.5
Blue Channel				
Mean	0.182570806	1.329375	0.181699346	0.169934641
Standard Deviation	0.258229404	1.633036737	0.257768365	0.240419842
Skewness	0.707970853	0.426645669	0.726918996	0.709645388
Kurtosis	-1.497571742	-792889054	-1.444050126	-1.492821167
Entropy	0.848685558	1.630336516	0.867563228	0.867563228
Energy	0.50617284	0.37625	0.5	0.5

According to the results, each object can distinguish by different features. That is our proposed model is a promising strategy that can be used to characterize acute leukemia subtypes.

VII. CONCLUSION

Leukemia as a fatal disease caused due to exceed in the number of WBCs cells must be diagnosed early. So, this paper suggests an effective mathematical model which will be used to detect the acute leukemia subtypes from the blood smear images. The proposed model was tested mathematically. The result showed that the conversion of images into L*a*b* color space produce uniformity compared to RGB color space. The

proposed model suggested different features for each object indicating that the model can be used accurately to detect acute leukemia subtypes.

VIII. FUTURE WORK

The proposed model will apply to upgrade quality and accuracy of segmentation and classification tasks in order to gain more confidence diagnostic process. The associated problem of limited leukemia classification subtypes gives the proposed model an opportunity to seek a critical task to classify the various acute leukemia subtypes.

REFERENCES

[1]. World Health Organization. (2018). World Cancer Day 2018 Available: http://www.who.int/cancer/en/
[2]. N. Patel and A. Mishra, "Automated leukaemia detection using microscopic images," Procedia Computer Science, vol. 58, pp. 635-642, 2015.
[3]. A. Tareef, Y. Song, D. Feng, M. Chen, and W. Cai, "Automated multi-stage segmentation of white blood cells via optimizing color processing," in Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on, 2017, pp. 565-568.
[4]. The Histology Guide. White blood cells. Available: https://www.histology.leeds.ac.uk/blood/blood_wbc.php
[5]. T. J. Kipps, F. K. Stevenson, C. J. Wu, C. M. Croce, G. Packham, W. G. Wierda, et al., "Chronic lymphocytic leukaemia," Nature Reviews Disease Primers, vol. 3, p. 16096, 2017.
[6]. S. Kumar, "Automated diagnosis of acute lymphocytic leukemia and acute myeloid leukemia using multi-SV," Journal of Biomedical Imaging and Bioengineering, vol. 1, 2017.
[7]. J. Rawat, A. Singh, H. Bhadauria, J. Virmani, and J. S. Devgun, "Computer assisted classification framework for prediction of acute lymphoblastic and acute myeloblastic leukemia," Biocybernetics and Biomedical Engineering, vol. 37, pp. 637-654, 2017.
[8]. J. Rawat, A. Singh, H. Bhadauria, and J. Virmani, "Computer aided diagnostic system for detection of leukemia using microscopic images," Procedia Computer Science, vol. 70, pp. 748-756, 2015.
[9]. A. S. Negm, O. A. Hassan, and A. H. Kandil, "A decision support system for Acute Leukaemia classification based on digital microscopic images," Alexandria Engineering Journal, 2017.
[10]. S. Paswan and Y. K. Rathore, "Detection and Classification of Blood Cancer from Microscopic Cell Images Using SVM KNN and NN Classifier," International Journal of Advance Research, Ideas and Innovations in Technology, vol. 3, pp. 315-324, 2017.
[11]. S. Dumyan and A. Gupta, "An Enhanced Technique for Lymphoblastic Cancer Detection Using Artificial Neural Network," International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE), vol. 6, pp. pp: 38-42, 2017.
[12]. D. K. Priya, S. Krithiga, P. Pavithra, and J. R. Kumar, "Detection of leukemia in blood microscopic images using fuzzy logic," Int. J. Engg. Res. & Sci. & Tech, vol. 240, 2015.
[13]. S. Agaian, M. Madhukar, and A. T. Chronopoulos, "Automated screening system for acute myelogenous leukemia detection in blood microscopic images," IEEE Systems journal, vol. 8, pp. 995-1004, 2014.
[14]. C. Sukanya and V. Paul, "AML Detection in Blood Microscopic Images Using DRLBP and DRLTP Feature Extraction," International Journal of Engineering Science, vol. 6942, 2016.
[15]. F. Kazemi, T. A. Najafabadi, and B. N. Araabi, "Automatic recognition of acute myelogenous leukemia in blood microscopic images using K-means clustering and support vector machine," Journal of medical signals and sensors, vol. 6, p. 183, 2016.
[16]. D. Goutam and S. Sailaja, "Classification of acute myelogenous leukemia in blood microscopic images using supervised classifier," in Engineering and Technology (ICETECH), 2015 IEEE International Conference on, 2015, pp. 1-5.
[17]. A. Hochhaus, R. A. Larson, F. Guilhot, J. P. Radich, S. Branford, T. P. Hughes, et al., "Long-term outcomes of imatinib treatment for chronic myeloid leukemia," New England Journal of Medicine, vol. 376, pp. 917-927, 2017.
[18]. J. N. Saultz and R. Garzon, "Acute myeloid leukemia: a concise review," Journal of clinical medicine, vol. 5, p. 33, 2016.
[19]. THE FREE DICTIONARY. (2018). Fab classification, acute leukemias. Available: https://medical-dictionary.thefreedictionary.com/FAB+classification%2c+acute+leukemias
[20]. W. Ladines-Castro, G. Barragán-Ibañez, M. Luna-Pérez, A. Santoyo-Sánchez, J. Collazo-Jaloma, E. Mendoza-García, et al., "Morphology of leukaemias," Revista Médica del Hospital General de México, vol. 79, pp. 107-113, 2016.
[21]. S. Bashasha, A. Kordofani, I. Osman, O. Musa, and H. Altayb, "Prevalence of the different FAB sub type of Acute Myeloid Leukemia related to hematological parameters in Sudanese," J Hematol Blood Disord, vol. 3, p. 102, 2017.
[22]. WikiDoc. (2017). Acute myeloid leukemia pathophysiology. Available: https://www.wikidoc.org/index.php/Acute_myeloid_leukemia_pathophysiology
[23]. R. D. Labati, V. Piuri, and F. Scotti, "All-IDB: The acute lymphoblastic leukemia image database for image processing," in Image processing (ICIP), 2011 18th IEEE international conference on, 2011, pp. 2045-2048.
[24]. ASH Image Bank. (2018). Acute myeloid leukemia. Available: https://imagebank.hematology.org/
[25]. N. Kumar and M. Nachamai, "Noise Removal and Filtering Techniques Used in Medical Images," Indian Journal of Computer Science and Engineering, vol. 3, pp. 146-153, 2012.
[26]. V. S. Rathore, M. S. Kumar, and A. Verma, "Colour based image segmentation using l* a* b* colour space based on genetic algorithm," International Journal of Emerging Technology and Advanced Engineering, vol. 2, pp. 156-62, 2012.
[27]. S. N. M. Safuan, R. Tomari, W. N. W. Zakaria, and N. Othman, "White blood cell counting analysis of blood smear images using various segmentation strategies," in AIP Conference Proceedings, 2017, p. 020018.
[28]. J. B. Roerdink and A. Meijster, "The watershed transform: Definitions, algorithms and parallelization strategies," Fundamenta informaticae, vol. 41, pp. 187-228, 2000.
[29]. N. Aggarwal and R. Agrawal, "First and second order statistics features for classification of magnetic resonance brain images," Journal of Signal and Information Processing, vol. 3, p. 146, 2012.
[30]. R. M. Haralick and K. Shanmugam, "Textural features for image classification," IEEE Transactions on systems, man, and cybernetics, pp. 610-621, 1973.
[31]. L.-K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," IEEE Transactions on geoscience and remote sensing, vol. 37, pp. 780-795, 1999.
[32]. D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," Canadian Journal of remote sensing, vol. 28, pp. 45-62, 2002.

[33]. L. Wang, K. Zhang, X. Liu, E. Long, J. Jiang, Y. An, et al., "Comparative analysis of image classification methods for automatic diagnosis of ophthalmic images," Scientific reports, vol. 7, p. 41545, 2017.
[34]. K. Kavitha, A. Gopinath, and M. Gopi, "Applying improved svm classifier for leukemia cancer classification using FCBF," in Advances in Computing, Communications and Informatics (ICACCI), 2017 International Conference on, 2017, pp. 61-66.

APPENDIX: TEXTURE FEATURE EQUATIONS

Features	Equations
First Order Statistics	
Mean	$\mu = \sum_{i=1}^{N_g} iH(i)$ where; N_g is the number of the non-zero histogram bins and $H(i)$ is the first order histogram.
Standard Deviation	$\sigma = \sqrt{\sum_{i=1}^{N_g} (i-\mu)^2 H(i)}$
Skewness	$s = \sigma^{-3} \sum_{i=1}^{N_g} (i-\mu)^3 H(i)$
Kurtosis	$k = \left(\sigma^{-4} \sum_{i=1}^{N_g} (i-\mu)^4 H(i)\right) - 3$
Entropy	$-\sum_{i=1}^{N_g} H(i) \log_2 H(i)$
Energy	$\sum_{i=1}^{N_g} H(i)^2$
Second Order Statistics	
Energy	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} [g(i,j)^2]$ where $g(i,j)$ denote to the co-occurrence matrix
Contrast	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \lvert i-j\rvert^2 g(i,j)$
Correlation	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{(i-\mu_x)(j-\mu_y)g(i,j)}{\sigma_x \sigma_y}$ where μ is the mean of $g(i,j)$, $\mu_x(i)$ is the mean of the row i, $\mu_y(j)$ is the mean of column j, $\sigma_x(i)$ is the standard deviation of row i, $\sigma_y(j)$ is the standard deviation of column j
Homogeneity	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{g(i,j)}{1+\lvert i-j\rvert}$
Entropy	$-\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} g(i,j) \log_2[g(i,j)$
Sum of square variance	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} (i-\mu)^2 g(i,j)$
Inverse difference moment	$\sum_{i=1}^{N_g}\sum_{j=1}^{N_g} \frac{g(i,j)}{1+\lvert i-j\rvert^2}$
Sum average	$\sum_{i=2}^{2N_g} [i g_{x+y}(i)]$ $g_{x+y}(k) = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} g(i,j)$, *where* $i+j=k$ *and* $k = 2,3,\ldots 2N_g$
Sum variance	$\sum_{i=2}^{2N_g} (i - sum\_entropy)^2 g_{x+y}(i)$
Sum entropy	$-\sum_{i=2}^{2N_g} g_{x+y}(i) \log_2[g_{x+y}(i)]$

Difference variance	$\sum_{i=0}^{N_g-1} i^2 g_{x-y}(i)$ $g_{x-y}(k) = \sum_{i=1}^{n_g}\sum_{j=1}^{n_g} g(i,j)$, *where* $i-j=k$ *and* $k = 0,1,\ldots,N_g-1$
Difference entropy	$-\sum_{i=0}^{N_g-1} g_{x-y}(i)\log_2[g_{x-y}(i)]$
Information measure of correlation I	$\frac{H-HXY1}{\max(HX,HY)}$ $H = -\sum_{i=1}^{n_g}\sum_{j=1}^{n_g} g(i,j)\log_z[g(i,j)]$ $HXY1 = -\sum_{i=1}^{n_g}\sum_{j=1}^{n_g} g(i,j)\log_z\left(g_x(i)g_y(j)\right)$ $HXY2 = -\sum_{i=1}^{n_g}\sum_{j=1}^{n_g} g_x(i)g_y(j)\log_z\left(g_x(i)g_y(j)\right)$
Information measure of correlation II	$IMC_{II} = \sqrt{1-e^{-2(HXY2-H)}}$

Najaat A. A. Abdullah "Mathematical Model of Computer Aided Detection for Acute Leukemia "International Journal of Computational Engineering Research (IJCER), vol. 08, no. 09, 2018, pp.39-50