1. INTRODUCTION

The image processing techniques have been developed in the various areas such as pattern recognition, biometrics, image inpainting, medical image processing, image compression, information hiding, and multimedia security [1]. Medical image processing is a collection of techniques that help the clinician in the diagnosis of different diseases from medical images such as X-ray, magnetic resonance imaging, computed tomography, and ultrasound and microscopic images. Thus, various types of cancer can be detected based on medical image processing from medical images such as breast cancer, brain tumor, lung cancer, skin cancer, and blood cancer (leukemia). The advantage of creating a system based on medical image processing techniques is extracting the targeted diseases in higher accuracy, with reducing time consumption as well as decreasing cost, otherwise, the manual processing is taken a lot of time and it also needs a professional staff for detecting diseases [3]. In general, processing medical images include four main steps which are: Pre-processing, segmentation, feature extraction, and classification. This paper is mainly focused on the segmentation step in processing microscopic blood images which help the clinician in identifying various diseases in human’s blood such as blood cancer (leukemia), and anemia. Segmentation of white blood cells (WBCs) is the most important step in identifying leukemia. Leukemia is a type of cancer that affects blood, lymphocyte system, and bone marrow. Thus, the correct segmentation of WBCs has an impact on further steps, such as feature extraction and classification, to
obtain results more accurately [3]-[5]. In general, the main components of blood are four types as follows:

- **Red blood cells (RBCs):** Are responsible for delivering oxygen from the lungs to all the parts of the human’s body [6].
- **WBCs:** Are responsible for body’s immune system. The WBCs are larger in size and fewer in number compared to RBCs [6].
- **Platelets:** Are responsible for blood clotting [6].
- **Plasma:** Is responsible for carrying nutrients that are required by the cells [3]. It makes up about half of the blood volume.

Most blood cells are produced in the bone marrow, and there are cells known as stem cells (lymphoid and myeloid) that are responsible for producing different blood cells [7]. In general, there are different types of WBCs which are Basophil, eosinophil, neutrophil, lymphocyte, and monocyte, as illustrated in Fig. 1.

In this paper, a segmentation technique based on thresholding is proposed to segment WBCs from the microscopic blood images using public and well-known dataset acute lymphoblastic leukemia-image database (ALL-IDB). The proposed technique includes the following steps: Pre-processing step to enhance the image quality, segmentation step to separate WBCs from other blood components, and image cleaning step to remove the unwanted objects inside the segmented image. Consequently, the proposed segmentation technique is compared with other technique, which is known as color-k-means clustering, in terms of time consumption as well as performance of the segmentation technique based on probability random index (PRI), variance of information (VOI), and global consistency error (GCE) which are measurements that using for evaluating a segmentation techniques.

**1.1. Problem Statement**

The manual processing of medical images is time consuming and it requires a professional staff which mainly depends on personal skills, and sometimes it may produce inaccurate results.

**1.2. Objective of the Research**

The main objective of the research is to segment WBCs from microscopic blood images accurately and quickly because accurate segmentation of WBCs makes other processes (such as feature extraction) easier consequently producing an accurate result of classification. Thus, the overall process could help the clinician in identifying different diseases accurately in a faster way, which further helps them to provide treatment for the patients sooner.

The rest of the paper is organized as follows: In Section 2, the literature review is presented. Section 3 presents the proposed approach. The experimental results are illustrated in Section 4. Finally, our conclusions are given in Section 5.

### 2. LITERATURE REVIEW

Nowadays, processing medical images have a crucial role for early identification of diseases. Thus, segmenting WBCs from microscopic blood images are the most important step in identifying leukemia. This section focuses on reviewing the most important existing works on segmenting WBCs.

In 2009, Sadeghian et al. [8] proposed a segmentation technique to segment WBCs as well as their nuclei and a segmentation technique to separate the cytoplasm of the cell. The red, green, and blue (RGB) image is converted into gray image; then canny edge detection is applied followed by gradient vector flow to connect the boundary of the nucleus. Consequently, the hole filling technique is applied to get the nucleus. Furthermore, Zack algorithm is applied into the gray image to get the binary image to extract the
cytoplasm of the cell by subtracting the binary image from the gray image. Finally, their proposed work is obtained 92% accuracy for nucleus segmentation and 70% accuracy for cytoplasm segmentation.

In 2015, Marzukia et al. [9] proposed a system to segment the nuclei of WBCs based on the active contour technique. The RGB image is converted to a gray image and then the active contour is applied to the resulted image to get the segmented nucleus. Continuously, the obtained image is converted to a binary image to find the roundness value to determine the grouped and individual WBCs. As they claimed, their proposed system can accurately extract the boundary of WBC nuclei.

In 2016, Sobhy et al. [11] proposed two segmentation techniques for segmenting WBCs. First, color correction is applied to extract the mean intensity from the histogram of each RGB channel of the original image. Moreover, the marker-control watershed technique is applied to separate overlapped cells. Furthermore, the seeded region growing technique is used to segment the nucleus of the cells. Finally, the performance of their approach is evaluated using relative ultimate measurement accuracy and misclassification error to measure the accuracy and it achieves 96% for WBCs segmentation and 94% for nucleus segmentation.

In 2015, Madhloom et al. [10] proposed an approach to segment WBCs and their nucleus. First, it segments the WBCs based on the color transformation as well as mathematical morphology. Moreover, the marker-control watershed technique is applied to separate overlapped cells. Furthermore, the seeded region growing technique is used to segment the nucleus of the cells. Finally, the performance of their approach is evaluated using relative ultimate measurement accuracy and misclassification error to measure the accuracy and it achieves 96% for WBCs segmentation and 94% for nucleus segmentation.

In 2017, Gowda and Kumar [12] proposed a system to segment WBCs based on k-means clustering and Gram-Schmidt orthogonalization. The proposed system is started by pre-processing step which includes: (1) Converting image from RGB to gray image, (2) median filter is implemented to remove noise, and (3) image normalization is applied for contrast stretching. Consequently, k-means clustering segmentation technique is applied to segment WBC and its subtypes such as: Basophil, eosinophil, neutrophil, lymphocyte, and monocyte. Furthermore, Gram-Schmidt orthogonalization scheme is applied to segment the nucleus from the cell.

Finally, in 2017, Nain et al. [13] proposed a system to differentiate individual and overlapped WBCs based on the watershed segmentation scheme to segment the cells. The edge map extraction technique, which is a circle fitted on each cell, is used to identify both individual and overlapped cells. To evaluate the performance of the proposed work, two evaluation measurements which are detection rate (DR) and false alarm rate (FAR) are used. As they claimed, the proposed system is achieved 97.10% and 7.80%, respectively, for DR and FAR, the lower value of FAR and higher value of DR means better segmentation of WBCs.

3. PROPOSED APPROACH

A microscopic image of the blood sample contains three components, as discussed before, which are RBCs, WBCs, and background (platelets and plasma). The proposed approach focuses on segmenting WBCs from other components (which are RBCs and background). The segmentation is done based on the thresholding technique. This section concentrates to explain the steps of the proposed work, as illustrated in Fig. 2.

3.1. Pre-processing

The first step of the proposed work is pre-processing, which is an important step to provide a better quality of the input image and make the next step (segmentation step) easier. In this step, the three channels of the input image, which are RGB are separated, and then the median filter (to remove noise) and histogram equalization (to contrast enhancement) are applied on the only green channel of the input image to enhance the image quality, Fig. 3.

The median filter is a non-linear filter, which is used for removing noise. While, the histogram equalization is a technique that uses image’s histogram to improve low or high contrast of the image which makes better distribution of intensities of the image’s histogram [14].

3.2. Segmentation

To segment the WBCs from the other components of the microscopic image, i.e., RBCs and background, the thresholding segmentation technique is applied to the image.
Mohammed and Abdulla: White Blood Cells Segmentation

Fig. 2. Steps of the proposed approach.
produce a binary image from a gray image (the value of 0, which represents background and it is black and the value of one which represents object inside the image and it is white color). This process has the advantage of reducing complex information and simplifies the classification and recognition processes [16]. The threshold value of this technique can be selected manually or automatically based on the information from the features of the image [17]. There are types of thresholding which are:

3.2.1. Global thresholding (single thresholding)
This technique use only one threshold (T) to segment the whole image into objects and background; this technique is more appropriate for those images that have bimodal histogram. The global thresholding can be defined as equation (S1) [16], [17]: Suppose a sample image of \( f(x, y) \) has the following diagram (Fig. 4):

Then, the binary image \( g(x, y) \) of the \( f(x, y) \) can be defined as following equation:

\[
g(x, y) = \begin{cases} 
1 & \text{if } f(x, y) > T \\
0 & \text{otherwise if } f(x, y) \leq T 
\end{cases}
\]

(1)

Where T is the threshold value.

3.2.2. Local thresholding (multiple thresholding)
In this thresholding technique, the image segments based on multiple threshold values and the image partitions into multiple region of interests (ROIs) and backgrounds [15]. Moreover, it segments an image by partitioning an image into \((n \times n)\) sub-images and then selects a threshold \( T_{ij} \) for every sub-image [16], as illustrated in Fig. 5.

This technique is more appropriate for images that are contained disparate illuminations, this technique can be defined as follows [16]:

\[
g(x, y) = \begin{cases} 
0 & \text{if } f(x, y) < T(x, y) \\
1 & \text{otherwise if } f(x, y) \geq T(x, y) 
\end{cases}
\]

(2)

Where \( f(x, y) \) is an input image and \( g(x, y) \) is a binary image produces depending on multiple threshold value \( T(x, y) \).

In our proposed approach, we segment the microscopic images into ROI (which are WBCs) and background based on global threshold value with \( T = 50 \) as clarified in the diagram of the work (Fig. 2.). We apply this technique in the resulted image obtained after the pre-processing step is done (i.e., image d from Fig. 3). Moreover, we obtain a binary image that contains WBCs which are larger in size in the image and some remaining as unwanted objects (RBCs and platelets), as illustrated in Fig. 6.
3.3. Image Cleaning
The resulted image obtained in the previous step, as illustrated in Fig. 6, includes some unwanted objects such as RBCs or platelets. To overcome this drawback, the morphological opening operator is applied to the resulted image of the segmentation step.

The opening operator is a morphological operation that uses to eliminating imperfections in the images that impact on texture and shape of images. Thus, morphological operations have a crucial role in image processing especially image segmentation process because they treat with shape extraction within the image and they describe the structure of the image [18]. The most significant morphological operations are dilation and erosion. There are also two other operations which are opening and closing which work depending on the combination of dilation and erosion [18], [19]. The opening operator is a combination of both erosion and dilation. It first erodes an image depends on the structuring element and then dilated it by the same structural element. Opening smoothens the boundary of an object and deletes small unwanted objects inside the image. It can be defined as the following equation [18], [19]:

\[
A \circ B = (A-B)+B
\]

Where A: Is an image
B: Is a structuring element.

Moreover, some of the WBCs are located in the edge of the image; thus, the border cleaning scheme is used to remove those cells because they have a negative impact on the accuracy rate in further steps such as feature extraction and classification. Thus, to make only complete WBCs remains in the image, the border cleaning is applied to the resulted image of the opening operator, as illustrated in Fig. 7. The border cleaning is a morphological technique that used to remove an object that touches the edge of the image [11].

3.4. WBCs Cropping
To crop each WBCs as an individual image, we used the bounding box technique. This technique can be defined as the smallest rectangle that surrounded an object of interest, and it extracts the minimum area of the box and it can be represented as follows [20]:

\[
\text{Area} = \text{Major axis length} \times \text{Minor axis length}
\]

Where the major axis represents the longest line that can be drawn between two points in the object, and the minor axis is the longest line between two points of the object that can be drawn perpendicularly to the major axis, as illustrated in Fig. 8. [20].

We applied the bounding box technique on the original image (input image) based on the location of WBCs in segmented and cleaned images obtained in Fig. 7, which further can be used for more processing such as feature extraction and classification.
The result of the bounding box is represented in Fig. 9. And Fig. 10 illustrate cropped WBCs as individual images.

4. EXPERIMENTAL RESULTS

The main purpose of the proposed approach is to segment the WBCs from microscopic blood images. In this section, to evaluate the performance of the proposed approach, experiments are conducted on the dataset (explained in 4.1).

4.1. Dataset

The input images are taken from a public and commonly used dataset known as ALL-IDB, which consists of two groups of images:

- The first one is ALL-IDB1 which was designed for testing the segmentation techniques, and it contains 108 microscopic blood images, including (49 abnormal images of leukemia patients and 59 normal images) [21]

- The second one is called ALL-IDB2 which was designed for testing the classifier techniques and it consists of 260 cropped WBC images (50% abnormal cells and 50% normal cells) which are taken from ALL-IDB1 [21].

As this study is mainly focused on the segmentation of WBCs, thus the first group of the dataset is used (i.e., ALL-IDB1). Because as mentioned above, this group of the dataset is dedicated for testing segmentation techniques.

4.2. Results

As mentioned in Section 3, in the thresholding segmentation, only the green channel of the input microscopic image was used for the segmentation process. Because the green channel gives better results in terms of segmentation compared to the other two channels (red and blue), since it holds more contrast information regarding to WBC [22]. In the segmentation step, the thresholding technique used for segmenting WBCs with \( T = 50 \), in our experiments first different threshold values are tested on all three channels of normal and abnormal images and the experimental results demonstrate that the best result is given by the green channel with \( T = 50 \). As illustrated in Fig. 11, the WBCs have higher contrast in the green channel of the image compared to two other channels; thus, it makes the segmentation process more accurate.
To evaluate the performance of the proposed approach, we compare our proposed approach with the commonly used segmentation technique which is known as color-k-means clustering, which was used by many existing works such as Mishra et al., Kumar and Vasuki, Ferdosi et al., Sarrafzadeh and Dehnavi [23]-[26]. We also applied the color-k-means clustering on the same dataset (ALL-IDB1) to evaluate both segmentation technique, the comparison was done in terms of time consumption as well as the performance of the segmentation techniques based on the PRI, GCE, and VOI. The PRI is a nonparametric technique used to measure the performance of the segmentation technique and it can be calculated as follows [22], [27]:

\[
PRI \left(S_{\text{test}}, G_k\right) = \frac{1}{N} \sum_{i < j} C_{i,j} P_{i,j} + \left(1 - C_{i,j}\right) \left(1 - P_{i,j}\right)
\]

Moreover, the GCE is a region-based segmentation consistency, which measures to quantify the consistency between image segmentation of differing granularities, and it can be calculated as follows [28]:

\[
GCE(S_1, S_2) = \frac{1}{n} \min\left\{ \sum_i X_i(S_1, S_2), \sum_i X_i(S_1, S_2) \right\}
\]

While the VOI is calculated the distance between two segmented areas which provide information about the participation of pixels in different clusters and measures the knowledge lost or gained in two clusters it can be calculated as follows [28]:

\[
VOI = H(X) = H(Y) - 2I(X,Y)
\]

The experimental results are as follows:

4.2.1. Execution time

The time consumption has been calculated for both segmentation techniques which implemented on the ALL-IDB1, as illustrated in Table 1. The systems were implemented using MATLAB on Lenovo computer with 8 GB of RAM and 64-bit Operating System\Windows 10.

Table 1 demonstrates that the proposed thresholding-based technique needs less time compared to the color-k-means technique for segmenting the WBCs from microscopic images which applies to 49 abnormal images and 59 normal images.

<table>
<thead>
<tr>
<th>Segmentation techniques</th>
<th>Execution time/\text{ms}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proposed</td>
<td>70.8144</td>
</tr>
<tr>
<td>Color-k-means clustering</td>
<td>204.7188</td>
</tr>
</tbody>
</table>

4.2.2. Performance of the segmentation techniques

Three well-known segmentation measurements such as PRI, GCE, and VOI are calculated for both proposed technique and color-k-means clustering and Table 2 illustrates the performance of each technique. The lower value of PRI, GCE, and VOI means the better/higher performance of the segmentation technique [27].

Table 2 demonstrates that the proposed technique has a lower value of PRI, GCE, and VOI. Thus, it is better for segmenting WBCs from microscopic images.

5. CONCLUSION

The progression of using techniques of image processing in medical fields provides better accuracy for identifying different diseases from medical images. This paper focuses on the segmentation step, which is the most important step in medical image processing for segmenting WBCs from microscopic blood images, depending on the threshold-based technique. Experimental results prove that the proposed segmentation technique can extract the WBCs from images better than color-k-means clustering in terms of segmentation evaluation as well as time consuming.

REFERENCES


