RESEARCH

CrossMark

A novel white blood cells segmentation algorithm based on adaptive neutrosophic similarity score

A. I. Shahin[1,2], Yanhui Guo[3*], K. M. Amin[4] and Amr A. Sharawi[1]

Abstract

Background: White blood cells (WBCs) play a crucial role in the diagnosis of many diseases according to their numbers or morphology. The recent digital pathology equipments investigate and analyze the blood smear images automatically. The previous automated segmentation algorithms worked on healthy and non-healthy WBCs separately. Also, such algorithms had employed certain color components which leak adaptively with different datasets.

Methods: In this paper, a novel segmentation algorithm for WBCs in the blood smear images is proposed using multi-scale similarity measure based on the neutrosophic domain. We employ neutrosophic similarity score to measure the similarity between different color components of the blood smear image. Since we utilize different color components from different color spaces, we modify the neutrosphic score algorithm to be adaptive. Two different segmentation frameworks are proposed: one for the segmentation of nucleus, and the other for the cytoplasm of WBCs. Moreover, our proposed algorithm is applied to both healthy and non-healthy WBCs. in some cases, the single blood smear image gather between healthy and non-healthy WBCs which is considered in our proposed algorithm. Also, our segmentation algorithm is performed without any external morphological binary enhancement methods which may effect on the original shape of the WBC.

Results: Different public datasets with different resolutions were used in our experiments. We evaluate the system performance based on both qualitative and quantitative measurements. The quantitative results indicates high precision rates of the segmentation performance measurement A1 = 96.5% and A2 = 97.2% of the proposed method. The average segmentation performance results for different WBCs types reach to 97.6%.

Conclusion: In this paper, a method based on adaptive neutrosphic sets similarity score is proposed in order to detect WBCs from a blood smear microscopic image and segment its components (nucleus and the cytoplasm). The proposed segmentation algorithm can be utilized for fully-automated classification systems, such systems can be either for the healthy WBCs or even for non-healthy WBCs specially the leukemia cells.

Keywords: WBCs segmentation, Color based segmentation, Neutrosophic set, Adaptive neutrosophic similarity score

Background

The blood smear under microscope contains useful information for diagnosis of many diseases. The blood components are divided into three categories: red blood cells (RBCs), white blood cells (WBCs) and platelets 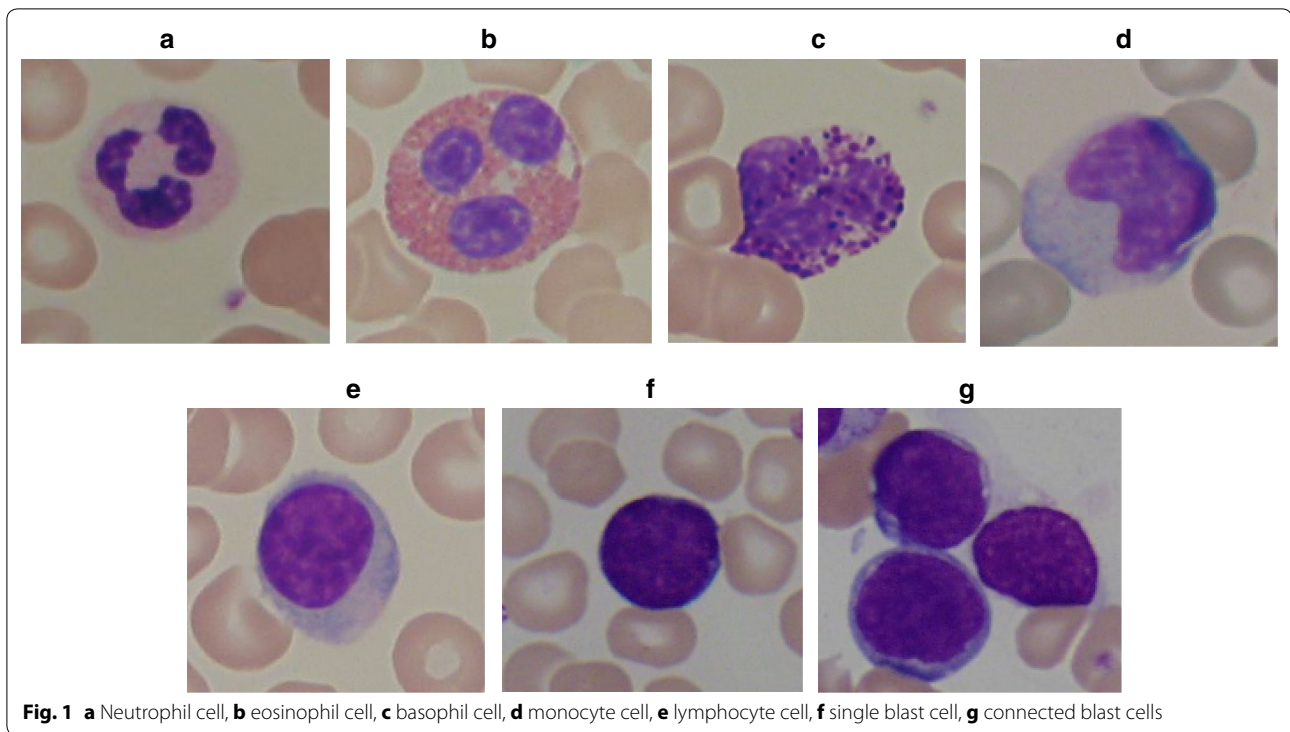[1]. WBCs are divided into 5 types by percentage: basophil (0–1%), eosinophil (1–5%), lymphocyte (20–45%), monocyte (2–10%) and neutrophil (50–70%) [1]. RBCs have no nuclei and each WBC type has its own shape of nucleus and cytoplasm [1]. The color appearance of each blood component is very essential in diagnosing and analysis of the blood smear microscopic image.

Each WBC consists from a nuclei and cytoplasm. Each of these WBCs have its own morphology and sometimes its own color as (shown in Fig. 1). Neutrophil as (shown in Fig. 1a) has a multi-lobed nuclei, and Eosinophil as

*Correspondence: yguo56@uis.edu
[3] Department of Computer Science, University of Illinois at Springfield, Springfield, IL, USA
Full list of author information is available at the end of the article

Shahin *et al. Health Inf Sci Syst* (2018) 6:1

Page 2 of 12

Fig. 1 **a** Neutrophil cell, **b** eosinophil cell, **c** basophil cell, **d** monocyte cell, **e** lymphocyte cell, **f** single blast cell, **g** connected blast cells

(shown in Fig. 1b) has a bi-lobed nuclei where the cytoplasm appears in red color. On the other hand, the nuclei and cytoplasm of Basophil are very difficult to be separated as the nature of the nuclei as (shown in Fig. 1c). Monocytes has a single nuclei with a weak cytoplasm color intensities as (shown in Fig. 1d). Lymphocyte as (shown in Fig. 1e) has the main focus of WBCs segmentation algorithms as it is responsible for the immune system in the body, and it has also two sizes which varies from 7 to 8 μm for small lymphocytes and 12–15 μm for large lymphocytes [2]. Each disease effect on each cell morphology and its characteristics. The leukemia disease affects mainly on lymphocyte cells which sometimes appear as single blast cell as (shown in Fig. 1f) or connected blast cells (shown in Fig. 1g) [3]. Each of previous cells characteristics should be taken into consideration to make an efficient automated segmentation algorithm.

In recent years, the digital pathology automated systems is exponentially growing, such systems help the pathologist to save effort and time. The blood smear images analysis has the main scope of the pathologist and a lot of researches in this field for microscopic images analysis have been proposed [4–17]. Since many beneficial explorations have been carried out for WBC segmentations, but majority of these methods have some defects to different extent, such as complexity of arithmetic, difficulty to ensure parameters, and so on.

The main contributions of our work in this paper are as follows; providing a fully-automated segmentation system which is able to count WBCs and measure each WBC structure that is an important step to classify the WBCs disorders later, detection and cropping the WBCs ROI in pathology images automatically, accurately segmentation of each WBC to a nuclei and a cytoplasm, segmentation of both the healthy WBCs (Neutrophil, Eosinophil, Basophil, Monocyte, small-large lymphocyte) and non-healthy cells (connected blast cells, single blast cell), applying the neutrosophic sets to a new domain of images, modifying the neutrosophic sets similarity score measure to be adaptive with multi-scale and multi-criteria environment. Moreover, our proposed method is not based on morphological enhancement operations which keeps the original structure of the WBCs. Finally, the performance of the segmentation algorithm is very promising to work on fully-automated classification system.

The rest of the paper is organized as follows: in the next sections, we present the related work on neutrosophic sets and the previous methods that have been proposed for WBCs segmentation, then the proposed method based on adaptive neutrosophic sets similarity score. Finally, the experimental results and discussion followed by the conclusion sections are presented.

Shahin *et al. Health Inf Sci Syst (2018) 6:1*

Page 3 of 12

Related work

According to the color nature of microscopic blood smear images, the segmentation on color channels is the most efficient technique to separate the components inside pathological microscopic images. On the other hand, all utilized color models cannot be used alone for segmentation and a complementary algorithm must be provided with them like thresholding, clustering, supervised learning, region growing or active contour model. In this section, we discuss both the previous utilized color spaces with their complementary algorithms in the literature. Then, we will discuss the neutrosphic sets similarity score algorithm since we will employ it as a complementary algorithm with threshloding.

For the previous utilized color spaces in the literature, the green channel in Red–Green–Blue (RGB) color space has been used for WBCs segmentation as it contains the contrast information between the leukocyte nucleus and other regions [9]. However, the RGB color space is not efficient on images with different illumination conditions, since, color transfer is an important procedure to overcome such conditions. CMYk color space has been employed in segmentation procedure [12, 14]. However, it also suffers from the same problems of RGB color space as it does not separate the luminance from the color information. In [10, 16–18], the HSV color space has been used as the advantage of separating the hue and saturation values. In [13], the CIE-Lab color space has been used as it is the most color model simulating the human visual system. In [12], three color spaces (RGB, HSI and CMYk) were employed to increase the performance of the segmentation algorithm.

We summarize the previous proposed WBCs segmentation systems in (Table 1). In [8], the authors combined the RGB with HSV color spaces to extract the nucleus, then, the gradient vector flow algorithm was applied to extract the WBC boundary. The system was very complex and works only with healthy WBCs. In [9], the author used the Gram–Schmidt Thresholding for RGB color space. However, the used dataset was very limited which cannot prove the concept.). In [10], the authors used the HSV with RGB color spaces to obtain the nuclei using thresholding, then to obtain the whole cell boundary using gradient vector flow (GVF). However, the thresholding values applied to red and blue channels are empirically defined. In [11], the authors presented a segmentation model consists from two classifiers (support vector machine and artificial neural network) and the watershed algorithm. However, it depends on RGB color space and does not work well under different light conditions, and the system has a very high processing time and needs a training procedure. In [12], the authors utilized the k-means clustering algorithm where three color spaces (RGB, HSI and CMYk. In [13], the authors employed the CIE-LAB color space with fuzzy- Means clustering algorithm, however the performance of the system was evaluated through only qualitative measurement. In [14], the author used the mean-shift algorithm which requires high processing time. In [15], the authors used yellow and black color components extracted from CMYk color model and the spatial kernel fuzzy c-means (SKFCM) was employed to segment WBCs in the image. In [16], the author used the fuzzy decision tree with HSV color model to segment WBCs. In [17], the author worked on the 5-WBCs in HSV domain. The thresholding was applied firstly to obtain the nuclei, then the active contour was applied to get the cell boundary. However, the Active contour method does not work if there is connected RBCs on the cell boundary.

From the literature, it is clearly defined that we cannot depend on a single color space or even certain color components. Also, in the literature, a lot of researches like

Table 1 Survey of recent WBCs segmentation systems

Reference	Cell type	Color space	Segmentation technique	Segmentation accuracy
[8]	5 WBCs	RGB, HSV	Thresholding, GVF	93%
[9]	5 WBCs	RGB	Gram–schmidt thresholding	93%
[10]	5 WBCs	HSV + RGB	Otsu's thresholding	96.5%
[11]	CLL Cell	RGB	Watershed + supervised training	90%
[12]	5 WBCs	RGB, HSI and CMYk	K-means-clustering	94.6%
[13]	5 WBCs	CIE-LAB	Fuzzy-K-means-clustering	Q
[14]	ALL Cell + 5 WBCs	RGB	Mean-shift +watershed	95%
[15]	NLCell	CMYk	Spatial-kernel fuzzy-K-means-clustering	98%
[16]	5 WBCs	HSV	Thresholding	85%
[17]	5 WBCs	HSV	Thresholding + active contour	92%

WBCs white blood cells, *NL* neoplastic lymphoid, *CLL* chronic lymphoid leukemia, *ALL* acute lymphoid leukemia, *RGB* red–green–blue, *HSI* hue-saturation-intensity, *HSV* hue-saturation-value, *CMYk* cyan-magneta–yellow–black, *GVF* gradient vector flow, *Q* qualitative

Shahin *et al. Health Inf Sci Syst (2018) 6:1*

Page 4 of 12

[5, 18, 19] in WBCs segmentation utilized the morphological operations as a main step of WBCs segmentation to refine the segmentation result. However, such operations affect the shape of WBCs and morphological structure [20]. All these previous reasons create the real need to build a robust WBCs segmentation algorithm and be adaptive. In this paper, the proposed segmentation algorithm is adaptive with different datasets, applied to healthy and non-healthy cells, and the proposed system performance are evaluated using qualitative and quantitative measurements.

Neutrosophy [21] is one of the most interesting philosophy theory that was introduced by Florentin Smarandache in 1980. This theory studies the origin, nature, and scope of neutralities. Neutrosophic sets have been used in many applications such as image enhancement [22], image edge detection [23], image segmentation [24] and handwritten recognition [25]. Neutrosophic set similarity measure (NSSM) gives rich information about the neutrosophic sets interval (NSI) and the degree of similarity between each of them [26, 27]. Neutrosophic similarity score (NSS) is a novel measurement defined in [28] which has been used specifically in many computer vision applications like image thresholding [28], image segmentation [29] and image classification [30].

To define the NSS, the first step is to represent the gray-scale intensity image in the neutrosophic set domain where the neutrosophic pixel intensity P_{NS} (T,I,F) is described using three membership values: True (T), Indeterminacy (I) and False (F) memberships. For each neutrosophic set, there are a set of alternatives A = {A_1, A_2 … A_m} at a specific criteria C = {C_1, C_2 … C_g}. Then, the three membership values can be defined as [28]:

$$T_{c_g}(x,y) = \frac{g(x,y) - g_{min}}{g_{max} - g_{min}} \tag{1}$$

$$I_{c_g}(x,y) = 1 - \frac{Gd(x,y) - Gd_{min}}{Gd_{max} - Gd_{min}} \tag{2}$$

$$F_{c_g}(x,y) = 1 - T_{c_g}(x,y) \tag{3}$$

where g(x,y) and Gd(x,y) are the intensity value and gradient value at the position of (x,y) on the image.

Then, the similarity score is derived to extract the degree of similarity according to the ideal object as the following equation [28]:

where A^* represent the ideal alternative and $T_{C_j}(x,y), F_{C_j}(x,y)$ and $I_{C_j}(x,y)$ represent each pixel value in the neutrosophic domain for each criteria.

Methods

There are two common methods in WBCs segmentation. The first method is based on one step segmentation as in [12, 14] where the entire image is processed to segment the WBCs directly. The second method is based on two-step segmentation as in [8] where the WBCs region of interest (ROI) is firstly detected, then each WBC is cropped, and the segmentation procedure is finally performed for each WBC structure. Our proposed system (shown in Fig. 2) is based on a two-step WBCs segmentation algorithm. This technique reduces the error rate and the processing time [31]. The segmentation is also performed well whether the cell is single blast cell or connected blast cell. The proposed method considers the color information of WBCs, the similarity between different color components is measured based on modified neutrosophic set similarity score.
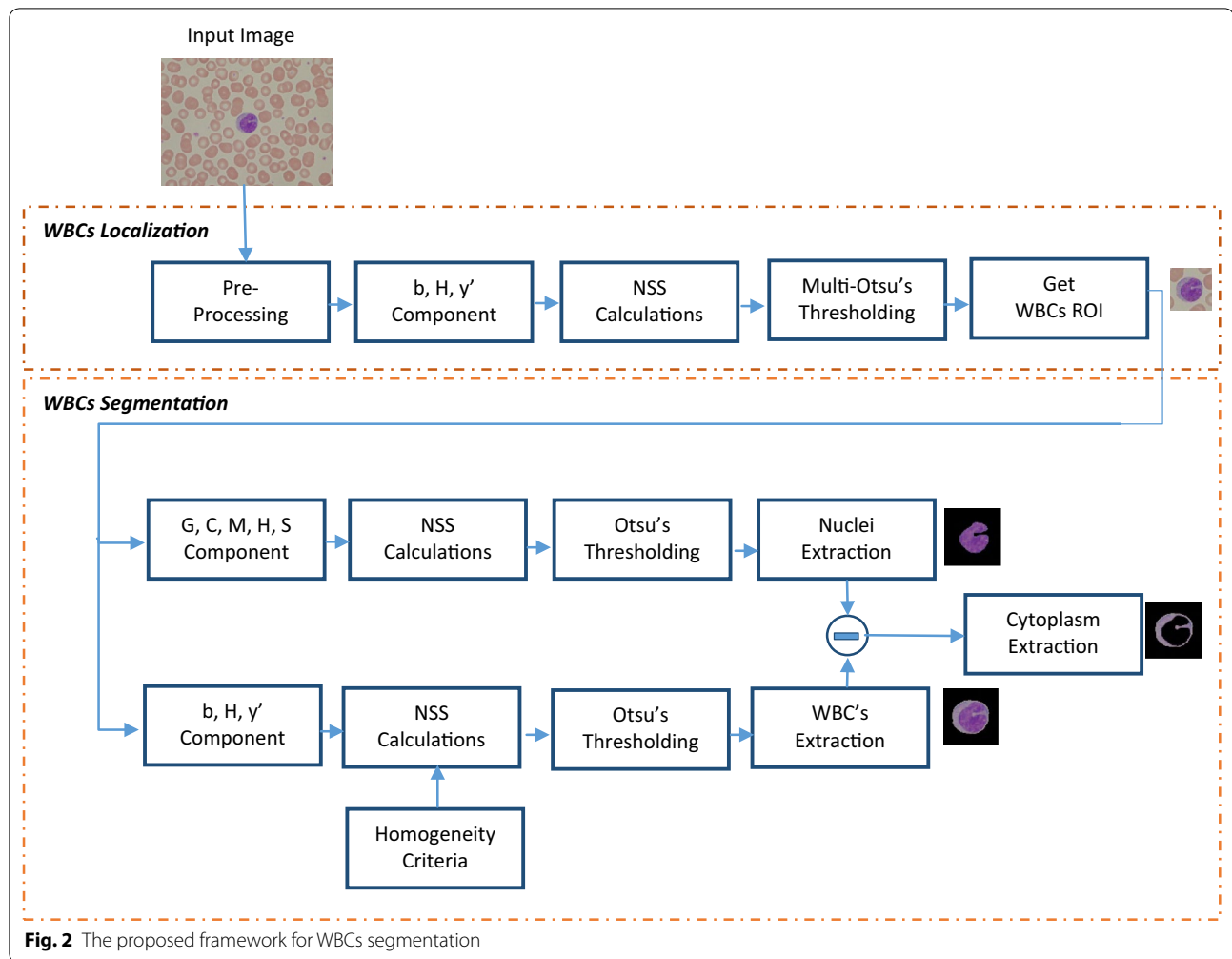
WBCs localization

The target of this stage is localizing the WBCs in the pathological images and obtaining the regions of interest (ROIs) that contain WBCs. In our proposed method, the connected blasted WBCs are taken into consideration, on the other hand, WBCs at corners are neglected. The initial detection of the WBCs is processed to eliminate the false regions after applying the smoothing procedure. We increase the cropping area to have more accurate segmentation for the cytoplasm area [31] as the WBC's cytoplasm sometimes have a low color intensities values according to staining artifacts.

Preprocessing

Many preprocessing techniques for WBCs segmentation have been proposed. Some techniques employed color correction [12, 14] and others used the traditional enhancement techniques [29]. In our proposed method, the smoothing procedure is applied to each channel of the input image. An averaging filter with a disk element whose radius r = 5 pixels and a square averaging kernel of size E = 2*r + 1. The filter size is minimized to overcome the interference between cytoplasm and the background in the pathological image, and also prevent blurring effect. This smoothing stage is important in the

$$S_{C_j}(P(x,y),A^*) = \frac{\left[T_{C_j}(x,y)T_{C_j}(A^*) + I_{C_j}(x,y)I_{C_j}(A^*) + F_{C_j}(x,y)F_{C_j}(A^*) \right]}{\sqrt{T_{C_j}^2(x,y) + I_{C_j}^2(x,y) + F_{C_j}^2(x,y))}\sqrt{\left(T_{C_j}^2(A^*) + I_{C_j}^2(A^*) + F_{C_j}^2(A^*) \right)}} \tag{4}$$

Shahin *et al. Health Inf Sci Syst* (2018) 6:1

Page 5 of 12

Fig. 2 The proposed framework for WBCs segmentation

NSS calculations according to their sensitivity to noisy pixels as reported by [28].

Color space conversions

In (Fig. 3), we present a blood smear image as example (shown in Fig. 3a), different color components (shown in Fig. 3b–g), and the proposed calculated NSS image (shown in Fig. 3i). It is important to realize that the WBCs appearance in this example cannot be generalized for all blood smear images, however, the calculated NSS image can be generalized for all blood smear images. This is specifically our main contribution in this paper. The successive color components which had been used in the literature for WBCs segmentation are evoked in our proposed system. In this paper, we firstly extract the whole WBC boundary to detect and segment the WBC ROI, then the nuclei region is segmented. The color components that have been applied in WBCs segmentation are as follow: the green (G) color component [9, 7], the blue (b) component in CIE-LAB color space [13], the hue (h) and saturation (s) [31], the Cyan (C) and magenta (M) color components [12], the yellow (Y) color component [14]. In this paper, we employ G, C, M, H and S components for WBC nuclei's segmentation. This makes the system more robust and adaptive since, it does not depend on a single color component or a specific color space.

NSS calculations

The previous NSS measure [28–30] depended on specific criteria which are intensity, homogeneity and local mean intensity criteria. In order to make the proposed method more robust, we propose new criteria using intensity values of several color components. The utilized

Shahin *et al. Health Inf Sci Syst* (2018) 6:1

Page 6 of 12
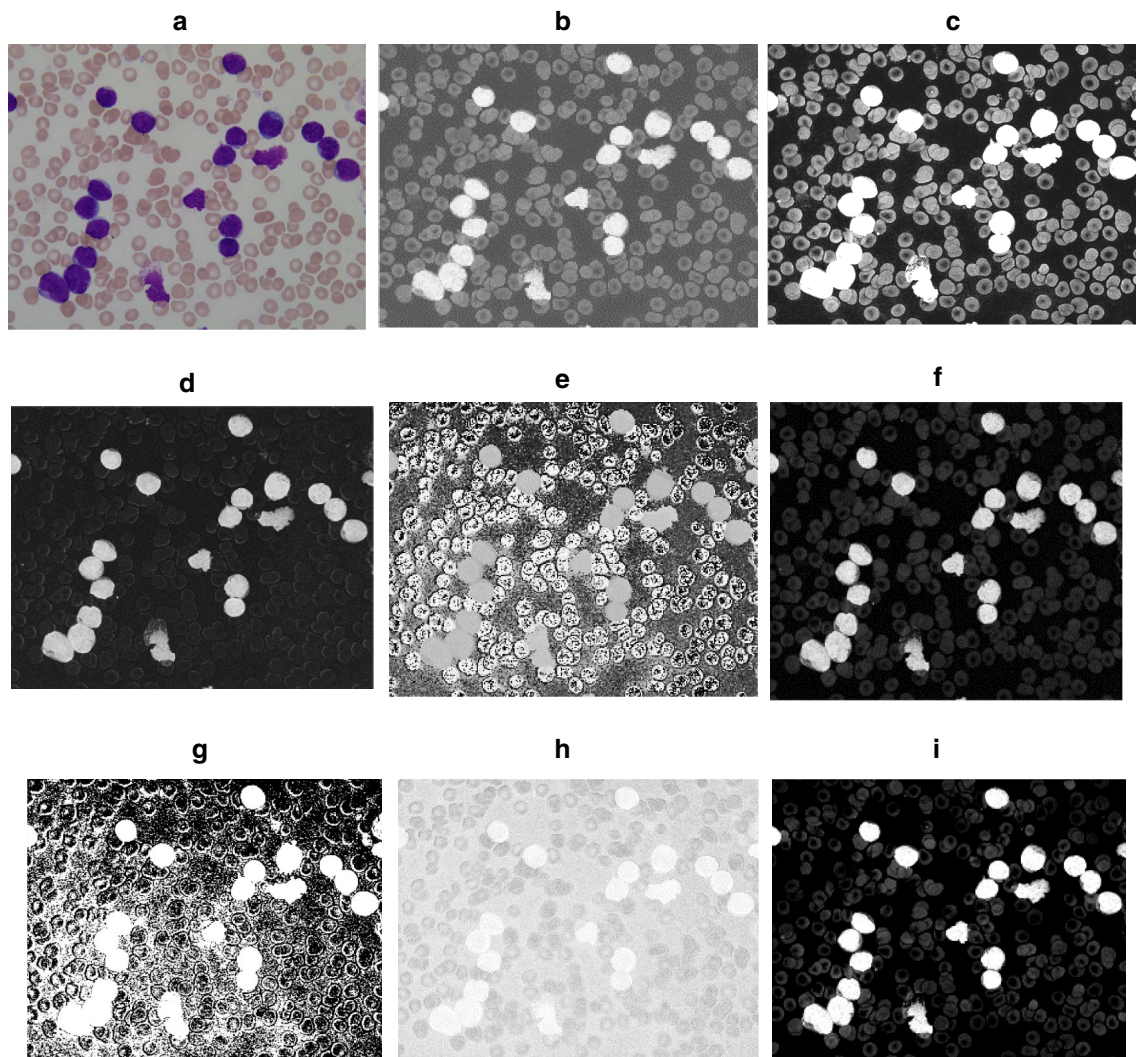
Fig. 3 **a** Original image, **b** the G color component, **c** the M color component, **d** the C color component, **e** the H color component, **f** the S color component, **g** the b color component, **h** Y' color component and **i** the NSS between previous color components

color components have a WBC with more bright pixels relative to the other pixels as (shown in Fig. 3). For this similarity measure, we use the criteria b color component in CIE-LAB color space, H color component in the HSV color space and the negative of y color components in the CMY color space.

For an ideal alternative A = [0 0 1], the NSS measure MS_{C_j} under the b, H and negative y criteria is defined as:-

The previous NSS measure [28–30] neglected the weights of each criteria during the similarity calculations. To make the similarity measure adaptive with different criteria, we modify the NSS to be adaptive with the variation of intensities with the diversity of blood smear images. The weights coefficients w_{k1} can be defined as;

$$MS_{C_j}\left(P(x,y),A^*\right) = \frac{w_{k1}\left[T_{C_j}(x,y)\,T_{C_j}(A^*) + I_{C_j}(x,y)I_{C_j}(A^*) + F_{C_j}(x,y)F_{C_j}(A^*)\right]}{\sqrt{w_{k1}(T_{C_j}^2(x,y) + I_{C_j}^2(x,y) + F_{C_j}^2(x,y))}\sqrt{w_{k1}\left(T_{C_j}^2(A^*) + I_{C_j}^2(A^*) + F_{C_j}^2(A^*)\right)}} \tag{5}$$

Shahin *et al. Health Inf Sci Syst* (2018) 6:1

Page 7 of 12

$$W_{k1} = \frac{1}{t \times u} \sum_{i=0}^{t-1} \sum_{j=0}^{u-1} P(i,j) \tag{6}$$

The weights coefficients w_{k1} value are derived from the mean intensity values of each criterion, t and u represent the image height and width respectively, and P (i, j) represent the pixel intensity value at position (i,j). The similarity result between the proposed criteria is (shown in Fig. 3i).

$$MS_{C_j}(P(x,y),A^*) = \frac{w_{k2}\left[T_{C_j}(x,y)T_{C_j}(A^*) + I_{C_j}(x,y)I_{C_j}(A^*) + F_{C_j}(x,y)F_{C_j}(A^*)\right]}{\sqrt{w_{k2}(T_{C_j}^2(x,y) + I_{C_j}^2(x,y) + F_{C_j}^2(x,y))}\sqrt{w_{k2}\left(T_{C_j}^2(A^*) + I_{C_j}^2(A^*) + F_{C_j}^2(A^*)\right)}} \tag{7}$$

After NSS calculation, a multilevel thresholding using Otsu's method [32] is used to get the binary masks. In (Fig. 3i), the pixel intensities can be classified into three levels. The first darkest intensities describe the background, the second describe the RBCs and the third with more bright pixels describe the WBCs and platelets. The thresholding result is very clear where the binary mask contains only the WBCs with platelets. As the platelets have a very small area relative to the WBCs, we use a binary area filter to remove the platelets.

WBCs cropping

For each blood smear image contains a certain number of WBCs, using the masks resulted from the thresholding procedure, we perform a cropping to the original image as (shown in Fig. 2). We extend 64 pixels to the original boundary of each WBC region to ensure that the existence of weak cytoplasm in the cropped image as proposed in [8].

Nucleus extraction

According to the deviation of color information of each WBC type on healthy cells and non-healthy cells, the nucleus segmentation should not be depend on only specific color component. Therefore, we choose G, C, M, H and S color components. For each color component, we define the neutrosophic sets using Eqs. (1), (2) and (3). For an ideal alternative A = [0 0 1], the NSS measure MS_{C_j} under the criteria G, C, M, H and S is defined as:

To make the similarity measure adaptive with different criteria, we employ the modified NSS to be adaptive with the variation of intensities with the diversity of blood smear images. The weights coefficients w_{k2} can be defined as Eq. (6). After the NSS calculation, the Otsu's thresholding is applied to the NSS image to extract the nuclei mask as (shown in Fig. 2).

WBCs extraction

In this stage, the whole WBC boundary is segmented. The accurate segmentation for the WBC boundary is highly affected by the segmentation accuracy specifically the cytoplasm region. The b, H and negative y criteria are combined with the homogeneity criteria of the green channel in the similarity measure. For an ideal alternative A = [0 0 1], the neutrosophic similarity measure GS_{C_j} for the green channel under the homogeneity criteria can be defined as:-

$$GS_{C_j}(P(x,y),A^*) = \frac{w_{k3}\left[T_{C_j}(x,y)T_{C_j}(A^*) + I_{C_j}(x,y)I_{C_j}(A^*) + F_{C_j}(x,y)F_{C_j}(A^*)\right]}{\sqrt{w_{k3}(T_{C_j}^2(x,y) + I_{C_j}^2(x,y) + F_{C_j}^2(x,y))}\sqrt{w_{k3}\left(T_{C_j}^2(A^*) + I_{C_j}^2(A^*) + F_{C_j}^2(A^*)\right)}} \tag{8}$$

WBCs segmentation

The second stage of segmentation is extracting the WBC region of interest (ROI). Each ROI contains a WBC will be accurately segmented to nuclei and the whole WBC. Finally, the nuclei are subtracted from the whole WBC to extract the cytoplasm region.

where the weights coefficients $w_{k3} = [0\ 1\ 0]$ to extract the homogeneity criteria.

The neutrosophic subsets T, I and F for each criteria are extracted. For an ideal alternative A = [0 0 1], the similarity FS_{C_j} under b, H, negative y and homogeneity of the green channel criteria can be defined as:

$$FS_{C_j}(P(x,y),A^*) = \frac{w_{k4}\left[T_{C_j}(x,y)T_{C_j}(A^*) + I_{C_j}(x,y)I_{C_j}(A^*) + F_{C_j}(x,y)F_{C_j}(A^*)\right]}{\sqrt{w_{k4}\left(T_{C_j}^2(x,y) + I_{C_j}^2(x,y) + F_{C_j}^2(x,y)\right)}\sqrt{w_{k4}\left(T_{C_j}^2(A^*) + I_{C_j}^2(A^*) + F_{C_j}^2(A^*)\right)}} \tag{9}$$

Shahin *et al. Health Inf Sci Syst* (2018) 6:1

Page 8 of 12

To make the similarity measure adaptive with different criteria, we employ the adaptive NSS. The weights coefficients w_{k4} can be defined as Eq. (6).

After NSS calculation, the Otsu's thresholding is applied which easily determine the WBC area as (shown in Fig. 2). Finally, The WBCs' cytoplasm mask are extracted by subtracting the nuclei's region from the whole WBC region as (shown in Fig. 2).

Experimental results
Dataset
The experiments were performed using different public pathology image datasets. These images have different resolutions, different contrasts, different illuminations, and were extracted from different sources. All these datasets have been widely used in many researches before.

In BS_DB3, a low cost system is consisting from CCD camera is applied to the microscope. The microscope magnification is adjusted at 100 × objective lens. The image's resolution was 640 × 480 pixels. The dataset has been previously used in [9, 19, 33, 34]. The total no. of WBCs presented in BS_DB3 dataset as described in [9] are 365 WBCs which consist from 271 neutrophil, 40 eosinophil, 33 lymphocyte, 19 monocyte and 2 basophil.

In ALL_DB1 and ALL_DB2, images in the dataset have been captured with an optical laboratory microscope coupled with a digital camera. All images are in JPG format with 24 bits color depth. The images were acquired with a resolution of 2592 × 1944 pixels. The microscope magnification range is from 300 to 500 x objective lens. Despite that the ALL_IDB1 and the ALL_IDB2 datasets are specified in Acute-Leukemia disease, these datasets contain other healthy WBCs which will be helpful in our experiments. The dataset has been widely used in [35–38].

The ALL_IDB1 contains 108 images. 59 healthy images and 49 non healthy image. Non healthy lymphocyte cell presented in ALL_IDB1 are 510 [36]. The other healthy WBCs presented in ALL_IDB1 are 48 neutrophil, 6 eosinophil, 115 lymphocyte, 66 monocyte and 1 basophil.

The ALL_IDB2 contains 260 images (130 healthy images and 130 non healthy) with (257 × 257) resolution. The total no. of WBCs in ALL_IDB2 are 260 which consist from 25 neutrophil, 2 eosinophil, 56 healthy lymphocyte, 30 small lymphocyte cell, 130 non healthy lymphocyte cell, 16 monocyte and 1 basophil.

Our experiments have used totally 1371 WBCs count as (shown in Table 2) which include 344 neutrophil, 48 eosinophil, 138 healthy lymphocyte and 101 monocyte cell. There are also total 640 non-healthy lymphocyte cells, which are 319 connected blast cell and 321 single non-healthy cells. The basophil cells count are 4, this low count is according to its low percentage in the blood.

Results
This section presents experimental outcomes of the proposed segmentation technique based adaptive NSS approach for two datasets: ALL_IDB and BS_DB3 of WBCs and also provides a comparison of the present method with four recent reported methods for dataset BS_DB3. One problem we encountered while measure the performance of our proposed system was that many authors tested their system with only a few sample images, or with their own datasets which is not publicly available. On the other hand, in the literature, few examples performed a combination of available datasets for the proposed system [14]. Thus, we could not directly compare our findings for the combined dataset with the results obtained by previous systems. We present the system performance based on qualitative [4, 6, 37] and quantitative [9, 12, 39] segmentation results.
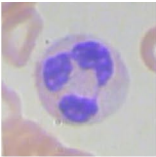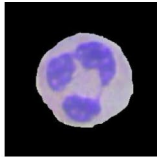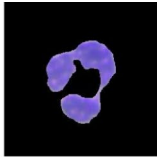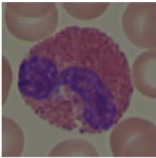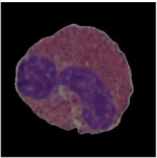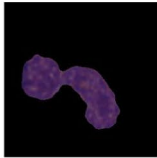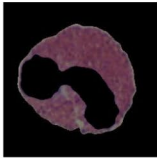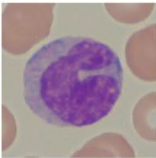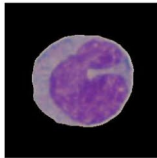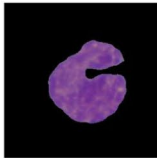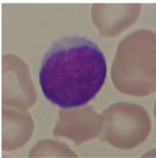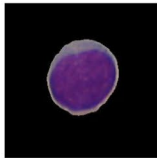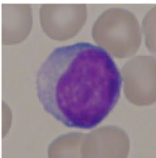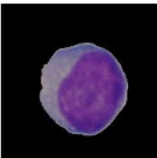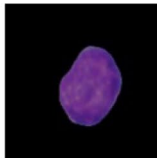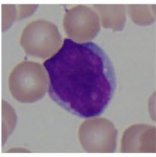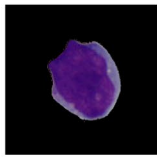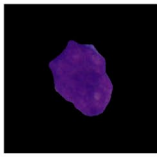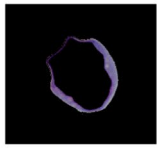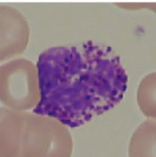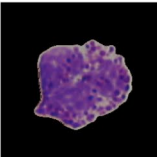
For the qualitative segmentation results visualization, in (Fig. 4), eight samples of different WBCs are selected to visualize the segmentation result of our proposed algorithm. The first column displays the original image with original size, the second displays the segmented WBCs, the third displays the segmented WBC's nuclei, and the last one displays the WBC's cytoplasm region. The important note in the (Fig. 4g) is that the basophil has not a cytoplasm due to the inherent nature of the basophiles and their cytoplasm and nuclei appear together (they are not separated) [39]. The segmentation results of either nuclei or even the cytoplasm appear to be more robust and adaptive with lighting conditions or image resolution as (shown in Fig. 4a–c).

In (Fig. 4e), the proposed algorithm is successfully segment the WBCs which are connected with the RBCs. In the previous alg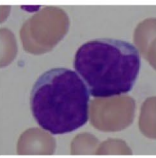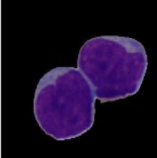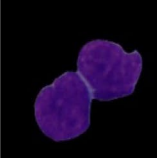orithms, solving that issue require boundary tracing. Our proposed method is also adaptive with healthy cases as (shown in Fig. 4d), non-healthy lymphocyte cells as (shown in Fig. 4f) and blasted cells as (shown in Fig. 4h).

Table 2 WBCs presented in used dataset

WBC Type	Neutrophil	Eosinophil	Healthy lymphocyte	Non healthy lymphocyte		Monocyte	Basophil	Total WBCs
				Single blast cell	Connected blast cells			
Count	344	48	234	321	319	101	4	1371

Fig. 4 Segmentation result of WBCs from different datasets **a** Neutrophil, **b** Eosinophil, **c** Monocyte, **d** Healthy Lymphocyte, **e** small lymphocyte, **f** Non Healthy lymphocyte- single cell, **g** Basophil, **h** Non Healthy lymphocyte- connected blasted cells

Shahin *et al. Health Inf Sci Syst (2018) 6:1*

Page 10 of 12

For the quantitative segmentation results, we use two different metrics. The first describes the total WBC segmentation accuracy performance [12] which refers to the quantitative counting of WBCs. These metrics are defined as:

$$A_1 = 100 \times \frac{The\ number\ of\ correctly\ detected\ WBCs}{Total\ number\ of\ detected\ WBCs} \quad (10)$$

$$A_2 = 100 \times \frac{The\ number\ of\ detected\ WBCs}{Total\ number\ of\ WBCs\ that\ existed\ in\ all\ dataset\ images} \quad (11)$$

For both nucleus and the cytoplasm of WBC segmentation area, we use the segmentation performance metric (SPM) [41] which indicates the quality of segmentation relative to the expert accuracy (ground truth) which is determined by pathologist. SPM is defined as:-

$$SPM = 100 \times \frac{A_{program} \cap A_{expert}}{\max(A_{program}, A_{expert})} \quad (12)$$

where $A_{program}$ is the segmented area by the proposed algorithm and A_{expert} is the segmented area by an expert (ground truth). If these two areas are the same, *SPM* is 100%.

We evaluate the proposed segmentation technique for all WBCs presented in ALL_IDB datasets and BS_DB3 dataset based on A1, A2 and SPM values. The experimental results reflect a high segmentation performance accuracy of the proposed method. The basophil's SPM

value of cytoplasm is not defined as the cause that have been mentioned before. The A1 and A2 of the proposed method are highest with 96.5 and 97.2% for all WBCs.

The performances of the proposed method were evaluated on each type of WBCs. The qualitative segmentation results of the proposed method were more robust and adaptive with different lighting conditions in the blood smear image as (shown in Table 3).

In (Table 3), the SPM values of non-healthy lymphocyte are not defined as the previous techniques did not took in consideration the presence of non-healthy lymphocyte or the connected blasted cells in the blood smear image. The average SPM values of the nucleus with different WBCs were 98.3%. The nucleus SPM values were better than the SPM values of the cytoplasm with 97.3%. The SPM values of non-healthy lymphocyte are very promising with accuracy 99.1% to use it in the classification process.

As mentioned before, we cannot present the comparison results for the combined dataset, as there were no reported research results available in the literature. To compare our results with the previous techniques in the literature as in [40, 41], we compare our results with each dataset results in the literature separately.

For ALL_IDB dataset, the authors evaluated their segmentation technique based on A1 and A2 values [12]. Our proposed system reflects higher A1 (95.4%) and A2

Table 3 SPM Results for different WBCs using our proposed method

	Basophil	Eosinophil	Healthy lymphocyte	Non- healthy lymphocyte		Monocyte	Neutrophil	Overall
				Single cells	Connected blast cells			
Nucleus (%)	97.1	98.2	99.3	98.7	97.2	97.8	98.4	98.1
Cytoplasm	–	98.6%	99.1%	99.4%	95.3%	95.2%	94.3%	97%
Average (%)	97.1	98.4	99.2	99.1	96.3	96.5	96.4	97.6

Table 4 SPM results comparison

	Basophil (%)	Eosinophil (%)	Healthy lymphocyte (%)	Monocyte (%)	Neutrophil (%)	Overall (%)
Mad. et al. [42].	64.3	44.8	58.9	57.4	57.0	55.9
Moh. et al. [33]	80.4	69.3	83.8	86.3	80.3	79.7
Rez. et al. [43]	75.7	88.9	79.6	89.6	82.3	83.2
Moh. et al. [9]	78.6	90.1	78.3	83.0	85.8	85.4
Proposed method	96.2	98.6	98.8	97.2	94.2	97

Shahin *et al. Health Inf Sci Syst* (2018) 6:1

Page 11 of 12

(96.2%) values than the previous system with A1 (94.6%) and A2 (95.1%) values [12].

For BS_DB3 dataset, the previous results in the literature were evaluated through SPM values [9, 33, 42, 43] as (shown in Table 4). Our proposed method achieves the highest SPM values through different WBCs. The overall accuracy value of our proposed method achieves higher accuracy than Mohamed et al. [33], Madhloom.et al. [42], Rezatofighi et al. [43], and Mohamed et al. [9]. The low performance of the previous methods may be occurred according to employing a single color component and the heavy usage of morphological binary operations.

Conclusion

Here, we propose an innovative WBC segmentation technique. The proposed technique is based on adaptive neutrosophic set similarity measure between different color components. The utilizing of multi-color components makes the proposed system more robust and adaptive. We also propose two novel frameworks for WBCs segmentation with its both structure; the nuclei or even the cytoplasm. The results for all presented datasets indicates a high precision rates of the quantitative segmentation performance A1 = 96.5% and A2 = 97.2% of the proposed method. The average SPM results for different WBCs types reach to 97.6%. The proposed method achieves high overall accuracy in nucleus segmentation specifically in non-healthy lymphocyte cells. The proposed method overcomes the problem on WBCs connected to the RBCs. Moreover, the proposed method is adaptive with different resolution or light conditions. All of the above mentioned results recommended that the proposed method is accurate and effective for WBC segmentation, and its performance is very promising.

In the future, we suggest to make a complete CAD system for WBCs identification based on the proposed segmentation system and we also suggest to separate the connected blast cells, identify the staining artifacts, accelerate and optimize the proposed algorithm as it work under multi-criteria which consume more time.

Author details
[1] Department of Biomedical Engineering, Cairo University, Cairo, Egypt.
[2] Department of Biomedical Engineering, HTI, Ramadan city, Egypt. [3] Department of Computer Science, University of Illinois at Springfield, Springfield, IL, USA. [4] Department of Information Technology, Menoufia University, Menoufia, Egypt.

References

1. Lichtman MA. Williams manual of hematology. New York: McGraw-Hill Higher Education; 2016.
2. Mohan H. Textbook of pathology. New Delhi: Jaypee Brothers; 2005.
3. Barbara BJ. Diagnosis from the blood smear. N Engl J Med. 2005;353(5):498–507.
4. Sadeghian F, Seman Z, Ramli AR, Abdul Kahar BH, Saripan M-I. A framework for white blood cell segmentation in microscopic blood images using digital image processing. Biol Proced Online. 2009;11:196–206. https://doi.org/10.1007/s12575-009-9011-2.
5. Mohammed EA, Mohamed MMA, Far BH, Naugler C. Peripheral blood smear image analysis: a comprehensive review. J Pathol Inf. 2014;5:9. https://doi.org/10.4103/2153-3539.129442.
6. Ghane N, Vard A, Talebi A, Nematollahy P. Segmentation of white blood cells from microscopic images using a novel combination of K-means clustering and modified watershed algorithm. J Med Signals Sens. 2017;7(2):92–101.
7. Prinyakupt J, Pluempitiwiriyawej C. Segmentation of white blood cells and comparison of cell morphology by linear and naïve Bayes classifiers. BioMed Eng Online. 2015;14:63. https://doi.org/10.1186/s12938-015-0037-1.
8. Ramesh N, Dangott B, Salama ME, Tasdizen T. Isolation and two-step classification of normal white blood cells in peripheral blood smears. J Pathol Inf. 2012;3(1):13.
9. Mohamed MMA, Far B. A fast technique for white blood cells nuclei automatic segmentation based on gram-schmidt orthogonalization. In Tools with Artificial Intelligence (ICTAI), 2012 IEEE 24th International Conference, 2012.
10. Huang DC, Hung KD, Chan YK. A computer assisted method for leukocyte nucleus segmentation and recognition in blood smear images. J Syst Softw. 2012;85(9):2104–18.
11. Mohammed EA, Mohamed MM, Naugler C, Far BH. Toward leveraging big value from data: chronic lymphocytic leukemia cell classification. Netw Model Anal Health Inf Bioinform. 2017;6(1):6.
12. Zhang C, Xiao X, Li X, Chen Y, Zhen W, Chang J, Zheng C, Liu Z. White blood cell segmentation by color-space-based K-means clustering. Sensors. 2014;14:16128–47.
13. Sarrafzadeh O, Dehnavi AM. Nucleus and cytoplasm segmentation in microscopic images using K-means clustering and region growing. Adv Biomed Res. 2015;4:174. https://doi.org/10.4103/2277-9175.163998.
14. Liu Z, Liu J, Xiao X, Yuan H, Li X, Chang J, Zheng C. Segmentation of white blood cells through nucleus mark watershed operations and mean shift clustering. Sensors. 2015;15:22561–86.
15. Alférez S, Merino A, Bigorra L, Mujica L, Ruiz M, Rodellar J. Automatic recognition of atypical lymphoid cells from peripheral blood by digital image analysis. Am J Clin Pathol. 2015;143(2):168–76.
16. Fatichah C, Purwitasari D, Hariadi V, Effendy F. Overlapping white blood cell segmentation and counting on microscopic blood cell images. Int J Smart Sens Intell Syst. 2014;7(3):71–86.
17. Mathur A, Tripathi AS, Kuse M. Scalable system for classification of white blood cells from Leishman stained blood stain images. J Pathol Inf. 2013;1:15.
18. Nazlibilek S, Karacor D, Ercan T, Kalender O, Sazli MH, Ege Y. Automatic segmentation, counting, size determination and classification of white blood cells. Measurement. 2014;55:58–65.
19. Mohammed EA, Far BH, Mohamed MMA, Naugler C. Automatic working area localization in blood smear microscopic images using machine learning algorithms. In IEEE International Conference on Bioinformatics and Biomedicine, Shanghai, 2013.
20. Jones KW. Evaluation of cell morphology and introduction to platelet and white blood cell morphology. Clin Hematol Fundam Hemost. 2009;93:116.
21. Smarandache F. Neutrosophic set, a generalization of the intuitionistic fuzzy sets. Int J Pure Appl Math. 2005;24:287–97.
22. Mohamed EA. New Approach for Enhancing Image Retrieval using Neutrosophic Sets. Int J Comput Appl. 2014;95(8):0975–8887.
23. Guo Y, Şengürb A. A novel image edge detection algorithm based on neutrosophic. Comput Electr Eng. 2014;40(8):3–25.
24. Yu B, Niu Z, Wang Z. Mean shift based clustering of neutrosophic domain for unsupervised constructions detection. Optik. 2013;124:4697–706.

Shahin *et al. Health Inf Sci Syst (2018) 6:1*

Page 12 of 12

25. Leng WY, Shamsuddin SM. Writer identification for Chinese handwriting. Int J Adv Soft Comput Appl. 2010;2(2):142–73.
26. Ye J. Multicriteria decision-making method using the correlation coefficient under single-valued neutrosophic environment. Int J Gen Syst. 2013;42(4):386–94. https://doi.org/10.1080/03081079.2012.761609.
27. Hanafy IM, Salama AA, Mahfouz K. Correlation of neutrosophic Data. Int Refereed J Eng Sci (IRJES). 2012;1(2):39–43.
28. Guo Y, Şengürb A, Yec J. A novel image thresholding algorithm based on neutrosophic similarity score. Measurement. 2014;58:175–86.
29. Guo Y, Şengürb A, Tian JW. A novel breast ultrasound image segmentation algorithm based on neutrosophic similarity score and level set. Comput Methods Programs Biomed. 2016;123:43–53. https://doi.org/10.1016/j.cmpb.2015.09.007.
30. Amin KM, Shahin A, Guo Y. A novel breast tumor classification algorithm using neutrosophic score features. Measurement. 2016;81:210–20.
31. Ghosh P, Bhattacharjee D, Nasipuri M. Blood smear analyzer for white blood cell counting: a hybrid microscopic image analyzing technique. Appl Soft Comput. 2016;46:629–38. https://doi.org/10.1016/j.asoc.2015.12.038.
32. Otsu N. A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern. 1979;9(1):62–6. https://doi.org/10.1109/tsmc.1979.4310076.
33. Mohamed M, Far B, Guaily A. An efficient technique for white blood cells nuclei automatic segmentation. In 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 220–225, 2012.
34. Mohamed, M, Far B. An enhanced threshold based technique for white blood cells nuclei automatic segmentation. In: e-Health Networking, Applications and Services (Healthcom), 2012 IEEE 14th International Conference; 2012. pp. 202–207. ‬‬
35. Amin MM, Kermani S, Talebi A, Oghli MG. Recognition of acute lymphoblastic leukemia cells in microscopic images using K-means clustering and support vector machine classifier. J Med Signals Sens. 2015;5(1):49.
36. Labati RD, Piuri V, Scotti F. All-IDB: The acute lymphoblastic leukemia image database for image processing. In: 2011 18th IEEE International Conference on Image Processing; 2011. https://doi.org/10.1109/icip.2011.6115881.
37. Putzu L, Di Ruberto C. White blood cells identification and counting from microscopic blood image. In: Proceedings of World Academy of Science, Engineering and Technology; 2013, 73:363.
38. Putzu L, Caocci G, Di Ruberto C. Leucocyte classification for leukaemia detection using image processing techniques. Artif Intell Med. 2014;62(3):179–91.
39. Siuly S, Kabir E, Wang H, Zhang Y. Detection of motor imagery EEG signals employing Naïve Bayes based learning process. Measurement. 2016;86:148–58.
40. Siuly S, Li Y. Discriminating the brain activities for brain–computer interface applications through the optimal allocation-based approach. Neural Comput Appl. 2015;26(4):799–811.
41. Rezatofighi SH, Soltanian-Zadeh H. Automatic recognition of five types of white blood cells in peripheral blood. Comput Med Imaging Graph. 2011;35(4):333–43.
42. Madhloom HT, Kareem SA, Ariffin H, Zaidan AA, Alanazi HO, Zaidan BB. An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold. J Appl Sci. 2010;10(11):959–66.
43. Rezatofighi SH, Soltanian-Zadeh H, Sharifian R, Zoroofi RA. A new approach to white blood cell nucleus segmentation based on gram-schmidt orthogonalization. In: International Conference on Digital Image Processing, 2009.