An Efficient Blood-Cell Segmentation for the Detection of Hematological Disorders

Pradeep Kumar Das, *Student Member, IEEE,* Sukadev Meher, *Member, IEEE,* Rutuparna Panda, *Member, IEEE,* and Ajith Abraham, *Senior Member, IEEE*

***Abstract*—The automatic segmentation of blood-cells for detecting hematological disorders is a crucial job. It has a vital role in diagnosis, treatment planning, and output evaluation. The existing methods suffer from the issues like noise, improper seed-point detection, and over-segmentation problems, which are solved here using an LoG-based modified highboosting operation, BO-FRS based seed-point detection, and hybrid ellipse-fitting, respectively. This paper proposes a novel hybrid ellipse-fitting based blood-cell segmentation approach, which may be used for detecting various hematological disorders. Our prime contributions are: (1) More accurate seed-point detection based on bounded opening followed by Fast Radial Symmetry (BO-FRS), (2) A novel least squares (LS)- based geometric ellipse-fitting approach, and (3) An improved segmentation performance by employing a hybridized version of geometric and algebraic ellipse-fitting techniques retaining the benefits of both the approaches. It is a computationally efficient approach since it hybridizes non-iterative- geometric and algebraic methods. Moreover, we propose to estimate the minor and major axes based on residue and residue offset factors. The residue offset parameter, proposed here, yields more accurate segmentation with proper ellipse-fitting. Our method is compared with the state-of-the-art methods. It outperforms the existing ellipse-fitting techniques in terms of dice similarity, Jaccard score, precision, and F1 score. It may be useful for other medical and cybernetics applications.**

***Index Terms*—Acute Lymphoblastic Leukemia, Acute Myeloid Leukemia, Ellipse-fitting, Hematological Disorder, Segmentation, Sickle Cell Anemia.**

I. Introduction

AUTOMATIC segmentation is a significant task in image processing applications [1]–[19]. It has an important role in the detection, classification, and diagnosis of hematological disorders: Sickle cell anemia (SCA), Acute Lymphoblastic Leukemia (ALL), and Acute Myeloid Leukemia (AML). SCA causes significant morphological variations of red blood-cells (RBCs) [1], [8]. However, AML and ALL affect white blood-cells (WBCs) [20], [21]. Hence, analysis of cell morphology has a vital role in the diagnosis of hematological disorders [1], [22], [23].

Manual counting and segmentation of cells are time-consuming and truely are a more challenging job [1], [8], [18]. The presence of noise, unwanted cells, overlapped cells, and poor contrast make this task more difficult. On the other hand, automatic segmentation enhances the robustness and thus makes diagnosis and treatment planning more accurate [1], [8]. Proper segmentation of overlapped or touched cells enhances detection and classification performance.

Recently, ellipse-fitting has become a preferred approach for the segmentation of overlapping objects/ cells [1], [4]–[13]. Gonzalez-Hidalgo et al. [8] concentrate on concave point identification, followed by an ellipse adjustment technique to effectively segment overlapped RBCs. On the other hand, Wang et al. [7] have suggested an efficient ellipse-fitting based approach to estimate fetal head circumference in ultrasound images automatically. In 2018, Panagiotakis and Argyros [13] have employed a region-based ellipse-fitting approach to segment the cells.

Azhar and Tjahjadi [9] have presented an ellipse-fitting and contour-movement based method to classify a human body-posture. Here, the ellipse-fitting approach is employed on the inverted-pendulum-human-body model to classify the body-posture. Chen and Epps [11] have suggested a convex-hull and ellipse-fitting based approach to extract the pupil boundary of the eye more accurately. They have employed a least-square-based ellipse-fitting approach to evaluate the pupil size efficiently. However, this ellipse-fitting technique depicts poor performance in some specific situations like a major part of the pupil is occluded.

Prasad et al. [5] have presented an elegant LS based ellipse-fitting technique. It is a non-iterative geometric ellipse fitting approach. It is stable and computationally efficient also. It has comparatively better selectivity to an elliptic curve than many current techniques, as it is a geometric based ellipse-fitting approach rather than algebraic based approach. The results presented in [5] signify that it delivers better performance than [27]–[29].

In 2015, Zafari et al. [4] have presented an efficient ellipse-fitting method to segregate overlapping objects for silhouette images. They [4] successfully extract seed points by applying bounded erosion [4], [24] and FRS [25]. Then edge and seed point information has been integrated to find out contour evidence. At last, an ellipse-fitting technique [26] is applied to estimate contour and segment overlapped objects as well [4]. They focus on the estimation of the proper eigenvector. Simplicity and computationally efficient are the salient features of the algebraic ellipse-fitting approach.

In 2020, Meng et al. [6] have presented a novel arc adjacency matrix-based ellipse-fitting technique. In this approach,

P. K. Das and S. Meher are with the Department of Electronics and Communication Engineering, National Institute of Technology, Rourkela 769008, India e-mail: (pdas391@gmail.com; smeher@nitrkl.ac.in).

R. Panda is with the Department of Electronics and Telecommunication Engineering, Veer Surendra Sai University of Technology, Burla 768018, India e-mail: (r_ppanda@yahoo.co.in).

A. Abraham is with the Machine Intelligence Research Labs (MIR Labs), Auburn, WA 98071 USA e-mail: (ajith.abraham@ieee.org).

diagram-based arc adjacency is estimated from the elliptic-arcs. Then, curvature and region constraints are applied to make it sparse. Finally, effective ellipse-fitting is achieved depending on twice the eigendecomposition of cumulative matrices based on the Jacobi approach.

The existing methods suffer from the problems of noise, intensity inhomogeneity, unwanted cells, and over-segmentation. Therefore, authors are motivated to propose a novel hybrid ellipse-fitting based blood-cell segmentation technique to mitigate the above issues and to yield more precise segmentation with proper ellipse-fitting. There is a strong need for a new theoretical investigation in this context. In this paper, we propose a new theoretical investigation on the estimation of major and minor axes of an ellipse based on residue and residue-offset.

This paper proposes a novel hybrid ellipse-fitting based blood-cell segmentation technique for the detection of hematological disorders. The proposed method emphasizes the elimination of unwanted cells (WBCs and platelets in SCA detection; RBCs and platelets in AML and ALL detection) using k-means clustering-based color segmentation [30], [31] and morphological operations. The work further emphasizes more accurate segmentation of overlapping cells. It consists of three crucial steps, including seed-point detection, contour evidence extraction, and ellipse-fitting. We propose an efficient geometric ellipse-fitting approach motivated by the ellipse-fitting method of Prasad et al. [5].

A. Contribution

The major contributions of the proposed method are the following:

1. Image quality improvement (edge enhancement, deblurring, and noise removal) using an LoG [33], [34] based modified highboosting operation;
2. Efficient seed-point detection by employing bounded opening followed by Fast Radial Symmetry (BO-FRS);
3. A novel geometric ellipse-fitting approach (where major and minor axes are accurately estimated based on residue and residue-offset values);
4. Combining this geometric ellipse fitting with the algebraic ellipse-fitting, suggested by Zafari et al. [4], to achieve better performance by retaining the advantages of both the approaches.

The paper is organized as follows. Section II highlights the proposed method for the segmentation of RBCs. The performance evaluation is presented in Section III. Finally, the paper is concluded in section IV.

II. Proposed Method

The schematic of the proposed method for the precise segmentation of blood-cells is illustrated in Fig. 1. It is composed of the following steps.

A. Preprocessing

The quality of the captured image suffers due to the presence of blur and unwanted noise, which may result a faulty

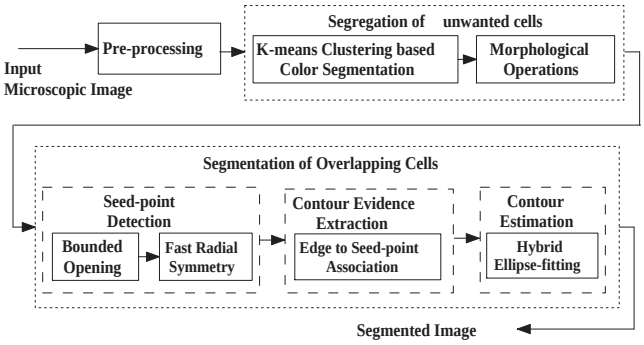

Fig. 1. Schematic of the proposed method.

diagnosis [1]. Thus, pre-processing is required to improve the image quality. Laplacian operator can be employed for edge enhancement and deblurring. However, it is sensitive to noise. Hence, to minimize the noise-effect, a 2D Gaussian filter is employed prior to the Laplacian operator; making it a Laplacian-of-Gaussian (LoG) filter [33]. It is also known as 'Mexico hat'. Rotational symmetry is one of the salient features of LoG [34]. Here, we employ an LoG based modified highboosting operation to produce a noise eliminated deblurred image. This operation is represented as follows.

$$F_I(x,y) = C_I(x,y) + kL_G(x,y) \tag{1}$$

where $C_I(x,y)$ depicts input color image while $F_I(x,y)$ represents the pre-processed image. k is a weight factor $(k>1)$. The weight factor k should be properly selected to boost the image quality efficiently. Here, we take $k=1.5$. $L_G(x,y)$ denotes LoG on $C_I(x,y)$, which is represented as:

$$L_G(x,y) = \frac{-1}{\pi\sigma^4}\left(1-\frac{x^2+y^2}{2\sigma^2}\right)\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right). \tag{2}$$

where σ symbolizes standard deviation, which should be properly selected to improve the image quality. Here, we employ an LoG filter of 7×7 with $\sigma=1$.

B. Elimination of Unwanted Blood-cells

The image may have unwanted cells (WBCs and platelets in SCA detection; RBCs and platelets in AML and ALL detection), which may cause a false diagnosis [1]. Our proposed method emphasizes the elimation of unwanted cells using k-means clustering-based color segmentation [30], [31] and morphological operations to improve segmentation accuracy and reliability. K-means clustering is an unsupervised machine learning technique [30]–[32]. It is employed to extract WBCs and platelets from an image, as highlighted in Pseudo Code 1. The detection of WBCs is a binary-class color segmentation problem, where each class is represented by three color-channels. Hence, the number of clusters (n) in K-means clustering is selected as 6. Moreover, n=6 is the optimal number of clusters at which WBCs are detected efficiently.

For the detection of ALL and AML, morphological operations are employed on (W) to eliminate the platelets like small particles and get the desired WBCs. However, in SCA dataset to get the desired RBCs, we have first to convert (F_I) and

Pseudo Code 1 Detection of WBCs and platelets.

Input: An RGB color image (F_I)
Output: An RGB color image having only WBCs and platelets
Begin
1: Convert the input color image to L*a*b color space.
1: Initialize the number of clusters: n=6.
3: K-means clustering is employed to classify color in 'a*b*' space.
4: Level each pixel based on its cluster index.
5: Find out the resultant RGB color-image (W).
End

(W) into HSV color space. Since the saturation component (S) of RBCs and WBCs differ significantly, we subtract the S component of (W) from that of (F_I) to eliminate the detected WBCs and platelets. Then, apply morphological operations to completely remove WBCs, and platelets and get the S component of desired RBCs. Finally, the resultant RGB color image containing only RBCs are extracted successfully.

C. Segmentation of overlapped cells

This paper presents a novel approach to efficiently segment cells, especially when a few cells overlap over each other as demonstrated in Pseudo Code 2. It consists of three crucial steps: seed-point detection, contour-evidence extraction [4], and contour-estimation. In the following, we highlight these crucial steps for the segmentation of overlapped cells.

1) Seed-point Detection: Seed-point detection is a crucial task in the segmentation of overlapping cells. It influences the performance of the final segmentation outputs. The main objective of seed-point detection is to identify each particle in an image by seed point. It also gives information about the total number of particles (cells) available in the image. Hence, proper seed point detection boosts the performance of contour evidence extraction as well as contour estimation.

Zafari et al. [4] have suggested an effective seed-points detection approach based on bounded erosion (BE) followed by FRS [25]. It can efficiently detect seed-points in silhouette images. However, the minor axis length gradually decreases due to BE. Hence, it is unable to properly detect the seed-points of cells, which have a relatively large aspect ratio [1]. To mitigate this issue, we employ the bounded opening (BO) in place of BE, which can efficiently separate convex-shape cells. BO is simply an iterative opening process having a specific number of iteration. Here, the number of opening operations in BO is selected as four since it yields an optimal tread off between the seed-point detection performance and the number of iterations. FRS [25] is employed to detect seed regions having high rotational symmetry. Moreover, the proposed BO-FRS gives priority to both convexity and rotational symmetry of cells.

A gray-scale image G_I is represented as the union of convex cells C_j. Mathematically, G_I can be expressed as:

$$G_I = \bigcup_{j=1}^{n} C_j. \tag{3}$$

The main motive of employing BO is to enhance the separability of each convex cell C_j. In the kth iteration of BO, every connected object C_j of $G_I(k-1)$ is processed to produce resultant cell R_j . R_j is represented as:

$$R_j = C_j^k \circ D\left(0,1\right). \tag{4}$$

Here, $D(0,1)$ represents closed disc structuring object having unity radius and $\circ$ symbolizes morphological opening [24].

The resultant image in the kth iteration of BO is obtained as a union of all R_j.

$$G_I^{k+1} = \bigcup_{j} R_j \tag{5}$$

FRS is a computationally efficient feature extraction approach [25]. It concentrates on local radial symmetry (LRS) to extract seed-points [4]. While LRS is employed with $N \times N$ neighborhood over the image having M pixels, it yields the complexity of order $O(MN)$ [25]. FRS is usually employed on the gradient of an image. Here, each edge-pixel of an image delivers a vote to achieve the best LRS within a certain distance from this pixel [4]. We have to select an appropriate range of distance value d. Let d vary in $[R_{MIN},\, R_{MAX}]$ for each pixel of gradient image I.

FRS emphasizes the estimation of magnitude projection image M_d and orientation projection image O_d. For this purpose, it first estimates positively-affected A_{+ve} and negatively-affected A_{-ve} pixels.

$$A_{+ve}\left(x,y\right) = \left(x,y\right) + round\left(\left(\frac{I\left(x,y\right)}{\left\|I\left(x,y\right)\right\|}\right) \times d\right) \tag{6}$$

$$A_{-ve}\left(x,y\right) = \left(x,y\right) - round\left(\left(\frac{I\left(x,y\right)}{\left\|I\left(x,y\right)\right\|}\right) \times d\right) \tag{7}$$

$$O_d\left(A_{+ve}\left(x,y\right)\right) \rightarrow O_d\left(A_{+ve}\left(x,y\right)\right) + 1 \tag{8}$$

$$O_d\left(A_{-ve}\left(x,y\right)\right) \rightarrow O_d\left(A_{-ve}\left(x,y\right)\right) - 1 \tag{9}$$

$$M_d\left(A_{+ve}\left(x,y\right)\right) = M_d\left(A_{+ve}\left(x,y\right)\right) + \left\|I\left(x,y\right)\right\| \tag{10}$$

$$M_d\left(A_{-ve}\left(x,y\right)\right) = M_d\left(A_{-ve}\left(x,y\right)\right) - \left\|I\left(x,y\right)\right\| \tag{11}$$

The contribution of radial symmetry S_d is estimated as:

$$S_d = \frac{M_d\left(x,y\right)}{k_d}\left(\frac{\left|\tilde{O}_d\left(x,y\right)\right|}{k_d}\right)^{\lambda} * H_d. \tag{12}$$

Where, $*$ symbolizes convolution operation and H_d denotes 2D Gaussian. Here, λ is the radial strictness and k_d is a scaling-factor responsible for the normalization of M_d and O_d. $\tilde{O}_d$ is represented as:

$$\tilde{O}_d\left(x,y\right) = \begin{cases} O_d\left(x,y\right), & \text{if} \quad O_d\left(x,y\right) < k_d \\ k_d, & \text{otherwise.} \end{cases} \tag{13}$$

The full FRS-transform is computed as the mean of the symmetry contributions over all the radii $d \in [R_{MIN}, R_{MAX}]$.

$$S = \frac{1}{|N|} \sum_{d \in \left[\ R_{MIN}, R_{MAX}\ \right]} S_d \tag{14}$$

Finally, the seed-points are predicted based on average location, centroid, of the identified symmetry regions, in S.

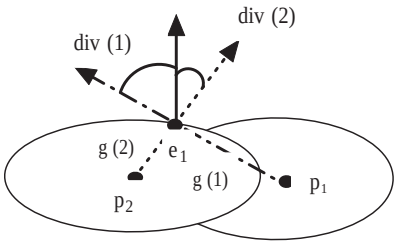

Fig. 2. Example of edge-to-seed-point association.

2) Contour Evidence Extraction: In this section, we emphasize the detection of contour-evidence based on the edge-to-seed point association technique [35]. This technique can successfully assign each edge-pixel to the corresponding seed-point based on the Euclidean distance and the divergence factor. For this purpose, we first extract the edge-points based on the Sobel edge detection technique [22], which is focused on high spatial frequency to identify the appropriate edge-points. For example, let the set of detected seed-points be $P = \{p_1, p_2, ..., p_n\}$. Each edge-point e_j in $E = \{e_1, e_2, ..., e_r\}$ is associated to the corresponding seed-points depending on a relevance factor $rel(e_j, p_k)$ given as:

$$rel\,(e_j, p_k) = \frac{1-\omega}{1+dist\,(e_j, p_k)} + \omega\frac{div\,(e_j, p_k)+1}{2}. \quad (15)$$

where ω is a constant taking on values [0, 1]. Here, $dist\,(e_j, p_k)$ is the normalized Euclidean distance between e_j and p_k, and $div\,(e_j, p_k)$ denotes the divergence function between e_j and p_k. Both the terms lie within $(0, 1]$. The edge-pixel e_j is associated with a seed-point p_k only when relevance value between them is maximum.

It is considered that each pixel on the line joining the edge-point-to-seed-point $l(e_j, p_k)$ should lie in the image foreground I_f :

$$g\,(x) = \begin{cases} |e_j - p_k|\,, & \text{if} \quad l\,(e_j, p_k) \subset I_f \\ \infty, & \text{otherwise.} \end{cases} \quad (16)$$

The deviation between the direction of $l(e_j, p_k)$ and the gradient direction at e_j is represented by $div\,(e_j, p_k)$. It is expressed as:

$$div\,(e_j, p_k) = \frac{l\,(e_j, p_k)\,g\,(e_j)}{\|l\,(e_j, p_k)\|\,\|g\,(e_j)\|}. \quad (17)$$

The search space is adaptively determined through a circular region around every edge-point. Hence, less number of seed-points are processed, and erroneous edge-points are eliminated successfully. Moreover, it makes the edge-to-seed-point association more robust and computationally more efficient.

Fig. 2 depicts the edge-to-seed-point association. Here, the dashed arrow illustrates the line from seed-point to edge-point. The solid arrow indicates the gradient direction at e_j.

3) Contour Estimation: The prime objective of this section is to segment overlapping cells using contour estimation properly. Here, we propose a novel geometrical ellipse-fitting approach inspired by the ellipse-fitting technique of Prasad et al. [5]. It is a non-iterative, LS-based ellipse-fitting technique. Moreover, the major and minor axes are updated more precisely based on residue and residue-offset factors. The geometric equation of an ellipse is represented as:

$$\begin{array}{l} \frac{((x-x_0)\cos\theta_0-(y-y_0)\sin\theta_0)^2}{m^2}+ \\ \frac{((x-x_0)\sin\theta_0+(y-y_0)\cos\theta_0)^2}{n^2} = 1 \end{array} \quad (18)$$

where, m: Semi-major axis length
n: Semi-minor axis length
θ_0: Angle of orientation (angle between x-axis and major axis)
$C\,(x_0, y_0)$: Center

Eq. (18) represents an ellipse, provided the following conditions are satisfied.

$$\begin{array}{ll} C1: & m, n \in R^+ \\ C2: & m \geq n \\ C3: & \theta_0 \in [0, \pi) \\ C4: & x_0, y_0 \in R \end{array} \quad (19)$$

Eq. (18) can be rewritten as:

$$\alpha(x-x_0)^2 + \beta(y-y_0)^2 + \gamma\,(x-x_0)\,(y-y_0) = (mn)^2 \quad (20)$$

where,

$$\begin{array}{l} \alpha = 0.5\left(\left(m^2+n^2\right) - \left(m^2-n^2\right)\cos\left(2\theta_0\right)\right) \\ \beta = 0.5\left(\left(m^2+n^2\right) + \left(m^2-n^2\right)\cos\left(2\theta_0\right)\right) \\ \gamma = \left(m^2-n^2\right)\sin\left(2\theta_0\right) \end{array} \quad (21)$$

The mathematical model of ellipse-fitting is demonstrated as follows.

Suppose $\bar{E}$ is a five-dimensional vector, which consists of ellipse parameters: m, n, θ_0, x_0, y_0, the vector utilizing the sequence of pixels $P_k{}'\,(x_k{}', x_k{}') : k = 1 : N$. The vector $\bar{Y}$ contains y-coordinates of the pixels.

$$\bar{E} = [m, n, \theta_0, x_0, y_0]^T \quad (22)$$

$$\bar{Y} = \left[\ -{y_1}^2 \quad -{y_2}^2 \quad .... \quad -{y_N}^2\ \right]^T, \quad \bar{Y} \ \in Z^N \quad (23)$$

$$\bar{H} = \left[\ h_1 \quad h_2 \quad h_3 \quad h_4 \quad h_5\ \right]^T, \quad \bar{H} \in R^5 \quad (24)$$

$$G : \bar{E} \rightarrow \bar{H}; \quad X : \bar{H} \rightarrow \bar{Y} \quad (25)$$

$\bar{H}$ is a five-dimensional vector, which contains new real-valued variables h_1 to h_5. It is employed to split the nonlinear mapping $F : \bar{E} \rightarrow \bar{Y}$ into two maps G and X as given by (25). The variables h_1 to h_5 are estimated in such a manner that it yields linear mapping between $\bar{H}$ and $\bar{Y}$, represented by:

$$\bar{Y} = X\bar{H}. \quad (26)$$

$G : \bar{E} \rightarrow \bar{H}$ is a nonlinear map whereas the mapping between $\bar{H}$ and $\bar{Y}$ is linear; and thus maintaining the mapping of $\bar{E}$ to $\bar{Y}$ is nonlinear. However, the parameters of $\bar{H}$ and $\bar{E}$ make G as a one-to-one mapping. Thus, m, n, θ_0, x_0, y_0 can be evaluated distinctively from h_1 to h_5.

For evaluating the parameters h_1 to h_5 and constructing the map X, we characterize residual distance as:

$$\sum_{k=1}^{N} {r_k}^2 = \left\|\bar{Y} - X\bar{H}\right\|^2. \quad (27)$$

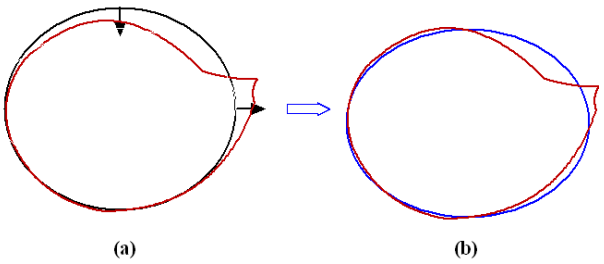

Fig. 3. The red-line represents the contour of an object. (a) The black-line represents the ellipse-fitting without considering the terms R_n and $R_n^{'}$ in Eq. 37 and 38. (b) The blue-line represents the ellipse-fitting by considering R_n and $R_n^{'}$.

where, $\|.\|$ symbolizes the Euclidean norm of a vector. This residual distance is minimized to obtain the unique optimal-solution of $\bar{H}$. $\bar{H}$ can be estimated as:

$$\bar{H} = \left(X^T X\right)^{-1} X^T \bar{Y}. \tag{28}$$

The ellipse formulation, given in (20), may be rewritten as:

$$\begin{aligned} -y^2 = \tfrac{\alpha}{\beta}\left(x^2\right) + \tfrac{\gamma}{2\beta}\left(2xy\right) - 2x\left(\tfrac{\alpha x_o + \frac{\gamma}{2} y_0}{\beta}\right) - \\ 2y\left(\tfrac{\beta y_0 + \frac{\gamma}{2} x_0}{\beta}\right) - \left(\tfrac{m^2 n^2 - \left(\alpha x_0^2 + \beta y_0^2 + \gamma x_0 y_0\right)}{\beta}\right) \end{aligned} \tag{29}$$

The parameters h_1 to h_5, and the map X can be evaluated by solving (29) as follows.

$$h_1 = \alpha/\beta \tag{30}$$

$$h_2 = \gamma/2\beta \tag{31}$$

$$h_3 = \frac{2\alpha x_0 + \gamma y_0}{2\beta} = h_1 x_0 + h_2 y_0 \tag{32}$$

$$h_4 = \frac{2\beta y_0 + \gamma x_0}{2\beta} = y_0 + h_2 x_0 \tag{33}$$

$$\begin{aligned} h_5 &= \tfrac{m^2 n^2 - \left(\alpha x_0^2 + \beta y_0^2 + \gamma x_0 y_0\right)}{\beta} \\ &= \tfrac{m^2 n^2}{\beta} - \left(h_1 x_0^2 + y_0^2 + 2h_2 x_0 y_0\right) \end{aligned} \tag{34}$$

$$X = \begin{bmatrix} x^2 & 2xy & -2x & -2y & -1 \end{bmatrix} \tag{35}$$

The normalized residue is estimated as:

$$R_n = \frac{\left\|X\bar{H} - \bar{Y}\right\|}{\left\|\bar{Y}\right\|}. \tag{36}$$

where $\|.\|$ depicts the Euclidean norm of a vector. Perfect ellipse-fitting can be achieved at a minimal normalized residue. The normalized residue varies between 0 and 1.

The parameters of the ellipse: m, n, θ_0, x_0, y_0 can be evaluated from h_1 to h_5 as shown in Eq. (37)-Eq.(41), respectively.

$$\theta_0 = -0.5\tan^{-1}\left(\frac{2h_2}{1-h_1}\right) \tag{39}$$

$$x_0 = \frac{h_3 - h_2 h_4}{h_1 - h_2^2} \tag{40}$$

Pseudo Code 2 Segmentation of Overlapping Cells

Input: The resultant color image after removal of unwanted cells
Output: Final segmented color image
Begin
1: Initialization of parameters
Set number of opening operations in BO as 4.
Set radial range:
$[R_{MIN}, R_{MAX}]$= [12, 21] for SCA and ALL datasets
= [14, 31] for AML dataset.
2: Detect the seed-points by employing BO-FRS.
3: Extract the contour evidence using edge-to-seed-point association.
4: Construct the matrix X and $\bar{Y}$ using Eq. (35) and (23), respectively.
5: Compute $\bar{H}$ utilizing Eq. (28).
6: Evaluate the normalized residue (R_n) using Eq. (36).
7: Estimate m, n, θ_0, x_0, y_0 employing Eq. (37)-(41), respectively.
8: To make the ellipse-fitting more efficient, employ the suggested algebraic ellipse-fitting particularly when $R_n > 0.5$, $m > \rho$, $n > \rho$.
End

$$y_0 = \frac{h_1 h_4 - h_2 h_3}{h_1 - h_2^2} \tag{41}$$

where, the term $R_n^{'}$, the *residueoffset* in (37) and (38), is estimated as:

$$R_n^{'} = mean\left(sign\left(X\bar{H} - \bar{Y}\right)\right) R_n \tag{42}$$

The terms R_n and $R_n^{'}$ are responsible for fine-tunning and precise fine-tunning (the ellipse-fitting), respectively. The term R_n modifies the minor and major axes to provide better ellipse-fitting. Furthermore, $R_n^{'}$ optimizes the major and minor axes depending upon the $mean\left(sign\left(X\bar{H} - \bar{Y}\right)\right)$, as shown in Fig. 3. In this example, $mean\left(sign\left(X\bar{H} - \bar{Y}\right)\right)$ is negative, which makes $R_n^{'}$ negative. Hence, it optimizes the ellipse-fitting by increasing major-axis length and decreasing the minor-axis length based on $R_n^{'}$. Similarly, positive $mean\left(sign\left(X\bar{H} - \bar{Y}\right)\right)$ enhances the minor-axis (n) and reduces the major-axis (m); thus yielding a precise fine-tunning of ellipse-fitting.

Usually, the above proposed ellipse-fitting method demonstrates excellent performance. However, sometimes it may display relatively poor performance particularly, when $R_n > 0.45$, $m > \rho$, $n > \rho$. Specifically, in such situations, we employ the LS-based non-iterative algebraic ellipse-fitting method proposed by Zafari et al. [4] to improve the performance and make the ellipse-fitting more accurate. We take $\rho = 30$ for SCA and ALL datasets, whereas $\rho = 60$ for AML dataset. Hence, the proposed method combines the benefits of both geometric and algebraic ellipse-fitting techniques.

III. Performance Evaluation

In this section, we present the performance of proposed method as well as various existing methods for the segmentation of cells. Next, we discuss the datasets and the performance measures used. Finally, the performance analysis of the proposed method is presented.

A. Dataset

The proposed method is validated using the following publicly available datasets: SCA dataset (ErythrocytesIDB [36]), ALL dataset (ALLIDB1 [37]), and AML dataset [38].

$$m = \left| \sqrt{\frac{2\left(\left(h_2+h_5\right)\left(h_2^2-h_1\right)^2+\left(h_2h_3-h_1h_4\right)^2+h_1\left(h_2h_4-h_3\right)^2\right)}{\left(h_2^2-h_1\right)^2\left(1+h_1+1.5R_n-\sqrt{\left(1-h_1-R_n'\right)^2+4\left(h_2-0.5R_n'\right)^2}\right)}} \right| \tag{37}$$

$$n = \left| \sqrt{\frac{2\left(\left(h_2+h_5\right)\left(h_2^2-h_1\right)^2+\left(h_2h_3-h_1h_4\right)^2+h_1\left(h_2h_4-h_3\right)^2\right)}{\left(h_2^2-h_1\right)^2\left(1+h_1+1.5R_n+\sqrt{\left(1-h_1-R_n'\right)^2+4\left(h_2-0.5R_n'\right)^2}\right)}} \right| \tag{38}$$

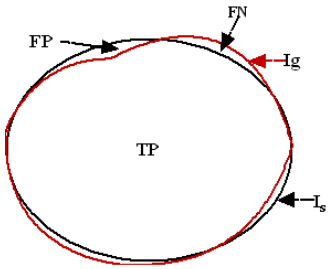

Fig. 4. Confusion Matrix Venn Diagram.

B. Performance Measures

We emphasize various performance measures: precision [1], recall [1], F1 score [1], Jaccard coefficient (J_c) [4], and Dice similarity coefficient (DSC) [39] for quantitative performance analysis.

$$Precision = \frac{TP}{TP+FP} \tag{43}$$

$$Recall = \frac{TP}{TP+FN} \tag{44}$$

$$F1 \quad Score \quad = \frac{2 \times (Recall \times \ Precision)}{(Recall + \ Precision)} \tag{45}$$

$$J_c = \frac{|I_s \cap I_g|}{|I_s \cup I_g|}. \tag{46}$$

$$DSC = \frac{2\,|I_s \cap I_g|}{|I_s| + |I_g|}. \tag{47}$$

where I_s represents segmented image and I_g symbolizes ground truth image. True Positive (TP) characterizes the number of properly segmented pixels. False Negative illustrates the number of undetected pixels. False Positive (FP) is the number of falsely identified pixels, which truly are not so, as shown in Fig. 4. The ideal value of DSC, (J_c), and (A_s) is 1.

C. Results

This section presents the performance of the proposed segmentation method for accurate detection of hematological disorders. We emphasize the performance evaluation of the elimination of unwanted cells, seed-point detection, and segmentation of overlapping cells. Moreover, the performance of the proposed ellipse-fitting approach is compared with other existing methods including ellipse-fitting methods of Zafari et al. [4], Prasad et al. [5], and Meng et al. [6]. Moreover, we apply the same pipeline structure as presented in the proposed method to make the comparison more fair and effective.

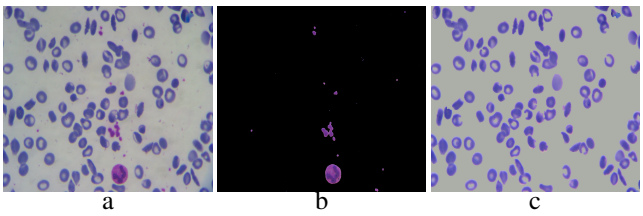

Fig. 5. Segmentation of unwanted cell in sickle-cell dataset: (a) Original image; (b) Extracted unwanted cells; (c) Resultant image having only RBCs

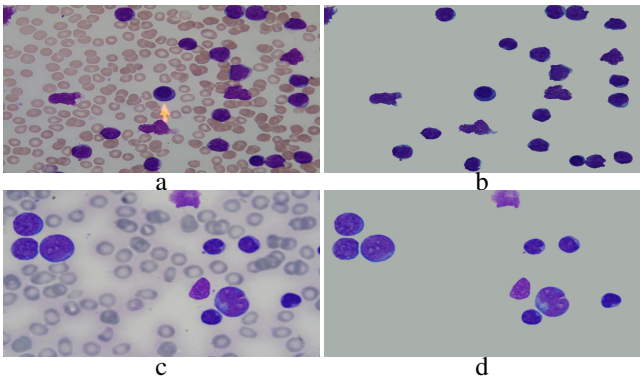

Fig. 6. Segmentation of unwanted cell in ALL and AML datasets: (a) and (c) Original images; (b) and (d) Resultant images containing only WBCs.

1) Elimination of WBCs and Platelets: In this section, we illustrate the qualitative performance of the elimination of unwanted cells. Fig. 5(a) depicts the input image, which contains RBCs along with WBCs and platelets. These unwanted cells are extracted successfully by employing the k-means clustering-based color segmentation technique, as shown in Fig. 5(b). Fig. 5(c) illustrates the desired resultant image, which contains only RBCs. From the figure, we visualize that WBCs and platelets are eliminated perfectly, which will make the segmentation more accurate. Moreover, from Fig. 6 (b) and Fig. 6 (d), we observe that WBCs are extracted successfully from ALL and AML dataset images, respectively.

2) Seed-point Detection: The prime objective of seed-point detection is to identify the location of the cells properly. Fig. 7 demonstrates the qualitative performance of seed-point detection. BO-FRS can detect seed-points more precisely than BE-FRS. From these figures, we observe that BE-FRS illustrates comparatively poor-performance due to false detection of seed-points, particularly the cells having a large aspect ratio. Table I illustrates that BOFRS demonstrates superior performance in terms of pecision, recall, and F1 score as compair to BEFRS.

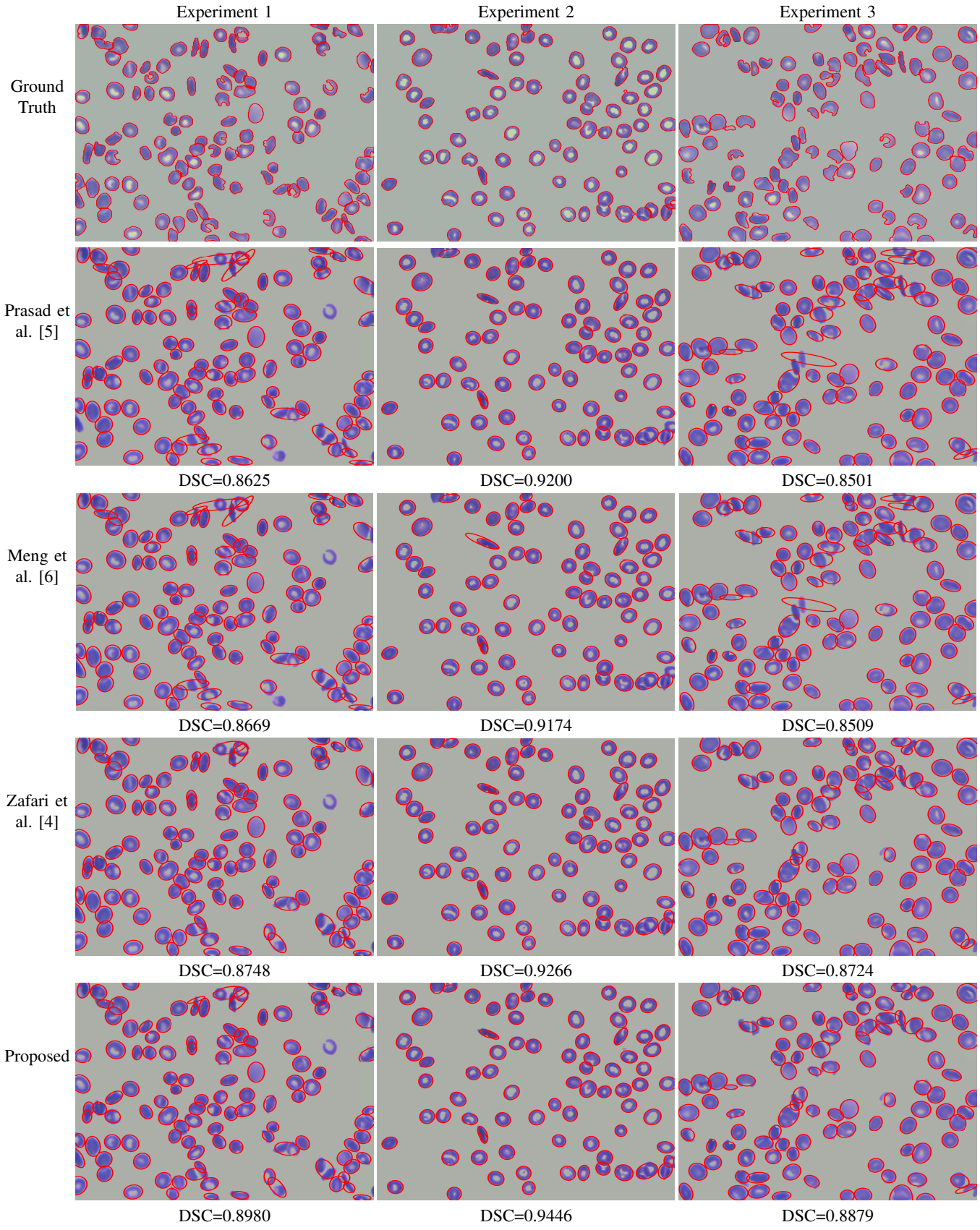

Fig. 8. Segmentation of cells in SCA dataset [36].

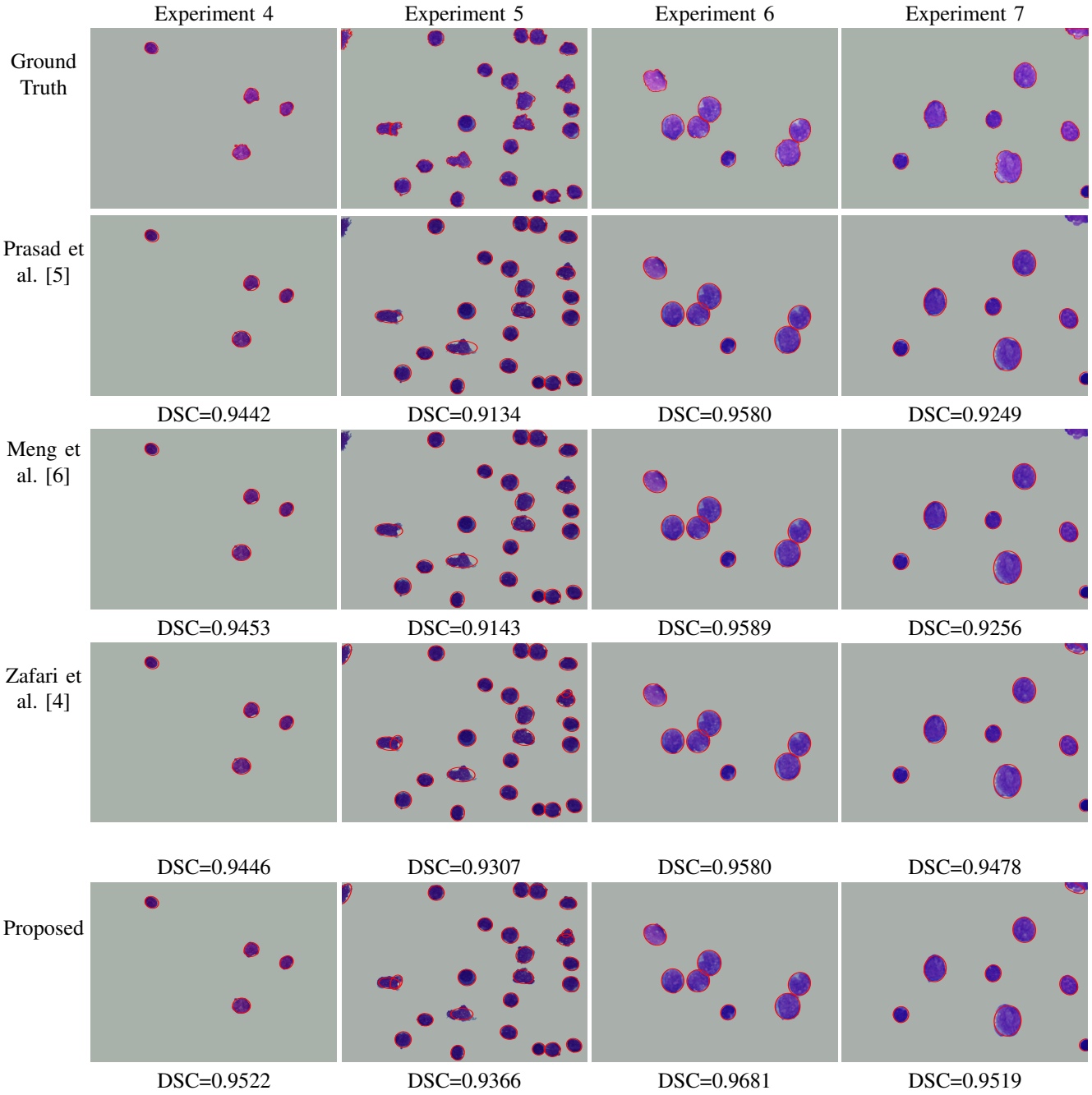

Fig. 9. Segmentation of cells in Leukemia datasets [37], [38].

TABLE I
PERFORMANCE COMPARISON OF SEED-POINT DETECTION

Dataset	Method	Precision (%)	Recall (%)	F1 Score (%)
Sickle cell [36]	BEFRS	95.80	95.00	95.40
	BOFRS	**96.64**	**95.83**	**96.23**
ALL [37]	BEFRS	98.63	98.18	98.40
	BOFRS	**99.09**	**98.64**	**98.86**
AML [38]	BEFRS	99.00	98.5	98.75
	BOFRS	**99.33**	**99.00**	**99.16**

3) Segmentation of Overlapping Cells: In this section, we present both the qualitative and quantitative performance of the segmentation of cells using ellipse-fitting approaches. The proposed hybrid ellipse-fitting technique is compared with var-

TABLE II
PERFORMANCE COMPARISON USING SCA DATASET [36]

Ex.	Ellipse-fitting Method	Precision (%)	F1 Score (%)	J_c	DSC
1	Prasad et. al [5]	79.36	86.25	0.7582	0.8625
	Meng et al. [6]	79.69	86.69	0.7651	0.8669
	Zafari et al. [4]	80.47	87.48	0.7775	0.8748
	Proposed	**85.97**	**89.81**	**0.8150**	**0.8597**
2	Prasad et. al [5]	86.63	92.00	0.8519	0.9200
	Meng et al. [6]	86.19	91.74	0.8474	0.9174
	Zafari et al. [4]	86.71	92.66	0.8632	0.9266
	Proposed	**91.67**	**94.46**	**0.8950**	**0.9446**
3	Prasad et. al [5]	75.90	85.01	0.7393	0.8501
	Meng et al. [6]	76.07	85.09	0.7405	0.8509
	Zafari et al. [4]	78.86	87.24	0.7737	0.8724
	Proposed	**83.82**	**88.79**	**0.7984**	**0.8879**

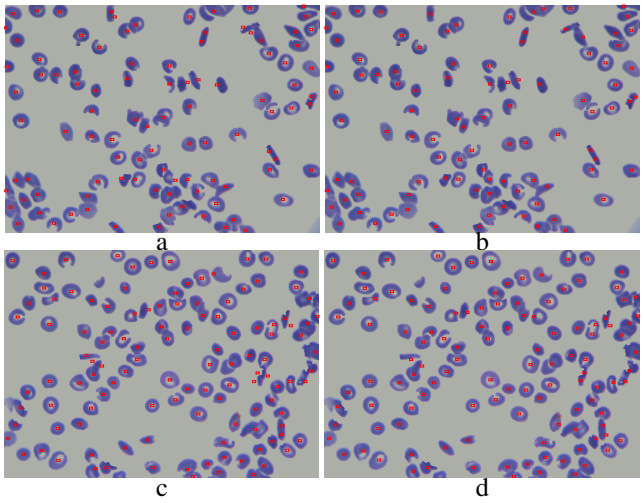

Fig. 7. Seed point detection: (a) and (c) BE-FRS; (b) and (d) BOFRS.

TABLE III
PERFORMANCE COMPARISON USING ALL DATASET [37]

Ex.	Ellipse-fitting Method	Precision (%)	F1 Score (%)	J_c	DSC
4	Prasad et. al [5]	90.55	94.42	0.8943	0.9442
	Meng et al. [6]	90.87	94.53	0.8963	0.9453
	Zafari et al. [4]	90.54	94.46	0.8950	0.9446
	Proposed	**95.67**	**95.22**	**0.9088**	**0.9522**
5	Prasad et. al [5]	88.40	91.34	0.8406	0.9134
	Meng et al. [6]	88.61	91.43	0.8421	0.9143
	Zafari et al. [4]	88.44	93.07	0.8704	0.9307
	Proposed	**94.02**	**93.66**	**0.8808**	**0.9366**

ious existing methods including the ellipse-fitting techniques suggested by Zafari et al. [4], Prasad et al. [5], and Meng et al. [6].

Fig. 8 demonstrates the qualitative performance of the segmentation of cells in the SCA dataset [36]. From these figures, we visualize that the proposed ellipse-fitting technique yields the best performance among all with proper detection of cells. However, Zafari et al. [4] depict the second-best performance.

In Fig. 9, Experiment 4 and 5 represents the qualitative performance of segmentation of cells in ALL dataset [37], whereas Experiment 6 and 7 demonstrate the performance in AML dataset [38] . From the figure, we observe that the proposed ellipse-fitting approach yields outstanding performance in all these experiments.

In Experiment 4, the ellipse-fitting approach presented by Meng et al. [6] yields the second-best performance, whereas the ellipse-fitting technique suggested by Zafari et al. [4] ranked second in Experiment 5. Experiments 6 and 7 illustrate all these ellipse-fitting technique yields excellent performance. Moreover, the proposed method outperforms with proper segmentation of cells.

The quantitative performance of segmentation of cells in SCA [36], ALL [37], and AML [38] datasets are shown in Table II, III, and, IV, respectively. From these tables, we notice that the proposed method yields superior performance with the best DSC, Jaccard score, precision, and F1 score. In all the experiments except Experiment 4 and 6, the ellipse-fitting technique suggested by Zafari et al. [4] ranked second. On the other hand, in Experiment 4 and 6, the ellipse-fitting approach presented by Meng et al. [6] demonstrates the second-best performance. Moreover, in all these seven experiments, the proposed hybrid ellipse-fitting technique outperforms with the best DSC, Jaccard score, precision, and F1 score since it retains the benefits of both the geometric and algebraic methods. Furthermore, the efficient estimation of minor and major axes based on residue and residue-offset helps it to achieve the best performance.

4) Computation Time: The simulation experiment is carried out on a MATLAB R2019b platform operating on a hardware computing device comprising an Intel Core i5 processor clocked at 3.40 GHz and a 12GB RAM. Table V highlights the computation time required for the ellipse-fitting. From the table, we notice that the algorithms of Zafari et al. [4], Prasad et al. [5], and the proposed hybrid ellipse-fitting approach yields similar performance regarding computational time. However, the proposed method is faster than the ellipse-fitting method of Meng et al. [6]. Moreover, the proposed hybrid-ellipse fitting is preferred over others since it yields outstanding performance with comparable computation time.

TABLE IV
PERFORMANCE COMPARISON USING AML DATASET [38]

Ex.	Ellipse-fitting Method	Precision (%)	F1 Score (%)	J_c	DSC
6	Prasad et. al [5]	92.40	95.80	0.9194	0.9580
	Meng et al. [6]	92.61	95.89	0.9210	0.9589
	Zafari et al. [4]	92.39	95.80	0.9194	0.9580
	Proposed	**95.84**	**96.81**	**0.9382**	**0.9681**
7	Prasad et. al [5]	93.09	92.49	0.8603	0.9249
	Meng et al. [6]	93.30	92.56	0.8615	0.9256
	Zafari et al. [4]	92.67	94.78	0.9008	0.9478
	Proposed	**96.15**	**95.19**	**0.9082**	**0.9519**

TABLE V
COMPUTATIONAL TIME FOR ELLIPSE-FITTING (S)

Dataset	Prasad et. al [5]	Meng et al. [6]	Zafari et al. [4]	Proposed
SCA [36] (Ex. 1-3)	3.59 s	5.84 s	3.64 s	3.68 s
ALL [37] (Ex. 4-5)	1.02 s	2.93 s	1.06 s	1.11 s
AML [38] (Ex. 6-7)	1.09 s	3.12 s	1.10 s	1.14 s

IV. CONCLUSION

Unlike the general segmentation of blood-cells algorithms, which are mostly based on watershed segmentation, the segmentation algorithm discussed in this paper is based on BO-FRS based seed-point detection and the hybrid ellipse-fitting based contour estimation. The proposed methods have the ability to extract the seed-points more accurately and to segment the overlapping cells more precisely, even if from the low-contrast inhomogeneous visual features. Therefore, the method is suitable for solving complex blood cell segmentation problems. The LS-based geometric ellipse fitting approach adopted in this article helps in more localization, therefore better accuracy. Usually, in algebraic ellipse-fitting methods, the fitting parameters are biased, and the fitting errors are unwillingly weighted, which leads to inaccurate boundaries, leading to problems in the detection of hematological disorders. On the other hand, our proposed method performs

better than the existing methods because of the combined benefits of geometric and algebraic ellipse-fitting methods. Moreover, it detects the cells more accurately due to proper ellipse-fitting. It is a computationally efficient technique since it hybridizes non-iterative-geometric and algebraic methods. Furthermore, the noise problem is solved by employing an LoG based modified highboosting operation. Hence, it is less sensitive to noise. Unlike other methods, our scheme never faces over-segmentation problem. Finally, we believe that it can also be employed for the segmentation of cells in other medical applications like- MRI, CT, ultrasound, and X-ray images, cybernetics applications, the segmentation of overlapping objects, etc.

REFERENCES

[1] P. K. Das, S. Meher, R. Panda, and A. Abraham, "A Review of Automated Methods for the Detection of Sickle Cell Disease," *IEEE Reviews in Biomedical Engineering*, vol. 13, pp. 309–324, 2020.

[2] P. K. Mishra, S. Agrawal, R. Panda, and A. Abraham, "A Novel Type-2 Fuzzy C-Means Clustering for Brain MR Image Segmentation," *IEEE Trans. Cybern.*, doi: 10.1109/TCYB.2020.2994235.

[3] S. Agrawal, R. Panda, and A. Abraham, "A novel diagonal class entropy-based multilevel image thresholding using coral reef optimization," *IEEE Tran. Syst., Man, and Cybern.: Syst.*, vol. 99, pp. 1–9, 2018.

[4] S. Zafari, T. Eerola, J. Sampo, H. Kalviainen, and H. Haario, "Segmentation of overlapping elliptical objects in silhouette images," *IEEE Trans. Image Process.*, vol. 24, no. 12, pp. 5942–5952, Dec. 2015.

[5] D. K. Prasad, M. K. Leung, and C. Quek, "Ellifit: An unconstrained, non-iterative, least squares based geometric ellipse fitting method," *Pattern Recognition*, vol. 46, pp. 1449–1465, 2013.

[6] C. Meng, Z. Li, X. Bai, and F. Zhou, "Arc Adjacency Matrix-Based Fast Ellipse Detection," *IEEE Trans. Image Process.*, vol. 29, pp. 4406–4420, 2020.

[7] J. Li, Y. Wang, B. Lei, J. Z. Cheng, J. Qin, T. Wang, S. Li, and D. Ni, "Automatic fetal head circumference measurement in ultrasound using random forest and fast ellipse fitting," *IEEE Journal on Biomedical and Health Informatics*, vol. 22, no. 1, pp. 215–223, 2018.

[8] M. Gonzalez-Hidalgo, F. A. Guerrero-Pena, S. Herold-Garca, A. Jaume-i-Capo, and P. D. Marrero-Fernandez, "Red Blood Cell Cluster Separation from Digital Images for use in Sickle Cell Disease," *IEEE Journal on Biomedical and Health Informatics*, vol. 19, no. 4, pp. 1514-1525, 2015.

[9] F. Azhar and T. Tjahjadi, "Significant Body Point Labeling and Tracking," *IEEE Trans. Cybern.*, vol. 44, no. 9, pp. 1673–1685, Sept. 2014.

[10] M. Winter, W. Mankowski, E. Wait, E. C. De La Hoz, A. Aguinaldo, and A. R. Cohen, "Separating touching cells using pixel replicated elliptical shape models," *IEEE Trans. Med. Imag.*, vol. 38, pp. 883–893, 2019.

[11] S. Chen and J. Epps, "Efficient and Robust Pupil Size and Blink Estimation From Near-Field Video Sequences for HumanMachine Interaction," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2356–2367, Dec. 2014.

[12] S. Baidya, A. M. Hassan, W. Al-Shaikhli, B. A. P. Betancourt, J. F. Douglas, and E. J. Garboczi, "Analysis of Different Computational Techniques for Calculating the Polarizability Tensors of Stem Cells with Realistic Three-Dimensional Morphologies," *IEEE Trans. on Biomed. Eng.*, vol. 66, no. 7, pp. 1816–1831, 2019.

[13] C. Panagiotakis and A. A. Argyros, "Cell Segmentation Via Region-Based Ellipse Fitting," *2018 25th IEEE International Conference on Image Processing (ICIP), Athens, 2018*, pp. 2426–2430, doi: 10.1109/ICIP.2018.8451852.

[14] B. Sheng, P. Li, S. Mo, H. Li, X. Hou, Q. Wu, J. Qin , R. Fang, and D. D. Feng, "Retinal Vessel Segmentation Using Minimum Spanning Superpixel Tree Detector," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2707–2719, July 2019.

[15] A. Tareef, Y. Song, H. Huang, D. Feng, M. Chen, Y. Wang, and W. Cai, "Multi-pass fast watershed for accurate segmentation of overlapping cervical cells," *IEEE Trans. Med. Imag.*, vol. 37, pp. 2044–2059, 2018.

[16] X. Bai, Y. Zhang, H. Liu and Z. Chen, "Similarity Measure-Based Possibilistic FCM With Label Information for Brain MRI Segmentation," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2618–2630, July 2019.

[17] H. Xu, C. Lu, R. Berendt, N. Jha, and M. Mandal, "Automatic nuclear segmentation using multiscale radial line scanning with dynamic programming," *IEEE Trans. on Biomed. Eng.*, vol. 64, no. 10, pp. 2475–2485, 2017.

[18] O. C. Linares, B. Hamann, J. B. Neto, "Segmenting Cellular Retinal Images by Optimizing Super-Pixels, Multi-Level Modularity, and Cell Boundary Representation," *IEEE Trans Image Process.*, vol. 29, pp. 809–818, 2019.

[19] C. L. Chowdhary and D. P. Acharjya, "Segmentation and feature extraction in medical imaging: a systematic review," *Computer Science*, vol. 167 pp. 26–36, 2020.

[20] O. Sarrafzadeh, H. Rabbani, A. M. Dehnavi, "A. Talebi, Detecting different sub-types of acute myelogenous leukemia using dictionary learning and sparse representation," *2015 IEEE Int. Conf. Image Processing (ICIP)*, pp. 3339–3343, 2015.

[21] S. Mishra, B. Majhi, and P. K. Sa, "Texture feature based classification on microscopic blood smear for acute lymphoblastic leukemia detection," *Biomedical Signal Processing and Control*, vol. 47, pp. 303–311, 2019.

[22] P. Rakshit and K. Bhowmik, "Detection of Abnormal Finding in Human RBC in Diagnosing Sickle Cell Anaemia Using Image Processing," *Procedia Technology*, vol. 10, pp. 28–36, 2013.

[23] M. A. Fadhel, A. J. Humaidi, and S. R. Oleiwi, "Image processing based diagnosis of sickle cell anemia in erythrocytes," *IEEE Annual Conf. on New Treand in Information and Communication Technology Applications*, pp. 203–207, 2017.

[24] R. M. Haralick, S. Zhuang, C. Lin, and J. Lee, "The digital morphological sampling theorem," *IEEE Trans. Acoust., Speech, and Signal Process.*, vol. 37, no. 12, pp. 2067–2090, 1989.

[25] G. Loy and A. Zelinsky, "Fast radial symmetry for detecting points of interest," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 959–973, 2003.

[26] M. Pilu, A. W. Fitzgibbon, and R. B. Fisher, "Ellipse-specific direct least-square fitting," *In Proc. Int. Conf. Image Process. (ICIP)*, vol. 3, pp. 599–602, 1966.

[27] Z. Y. Liu and H. Qiao, "Multiple ellipses detection in noisy environments: A hierarchical approach," *Pattern Recognition*, vol. 42, pp. 2421–2433, 2009.

[28] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, pp. 476–480, 1999.

[29] D. Chaudhuri, "A simple least squares method for fitting of ellipses and circles depends on border points of a two-tone image and their 3-D extensions," *Pattern Recognition Lett.*, vol. 31, pp. 818–829, 2010.

[30] A. Z. Chitade, and S. K. Katiyar, "Colour based image segmentation using k-means clustering," *Int. Jour. Eng. Sci. Techno.*, vol. 2, no. 10, pp. 5319–5325, 2010.

[31] H. Yadav, P. Bansal, and R. K. Sunkaria, "Color dependent K-means clustering for color image segmentation of colored medical images," *In 2015 1st International Conference on Next Generation Computing Technologies (NGCT), IEEE*, pp. 858-862, 2015.

[32] L. Duan, S. Cui, Y. Qiao, and B. Yuan, "Clustering Based on Supervised Learning of Exemplar Discriminative Information," *IEEE Tran. Syst., Man, and Cybern.: Syst.*, Dec. 2018. DOI-10.1109/TSMC.2018.2870549

[33] X. Jingbo, L. Bo, L. Haijun, and L. Jianxin, "A new method for realizing LOG filter in image edge detection," *Proc. of 6th International Forum on Strategic Technology IEEE*, pp. 733–737, 2011.

[34] H. Xu, C. Lu, R. Berendt, N. Jha, and M. Mandal, "Automatic nuclear detection based on generalized Laplacian of Gaussian filters," *IEEE Journal on Biomedical and Health Informatics*, vol. 21, no. 3, pp. 826–837, 2017.

[35] C. Park, J. Z. Huang, J. X. Ji, and Y. Ding, "Segmentation, inference and classification of partially overlapping nanoparticles," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 669–681, 2013.

[36] ErythrocytesIDB database. ULR http://erythrocytesidb.uib.es/.

[37] ALL-IDB1 dataset for ALL classification. ULR http://crema.di.unimi.it/fscotti/all/.

[38] Medical image and signal processing research center. URL https://misp.mui.ac.ir/en/database

[39] A. Khadidos, V. Sanchez, and C. T. Li, "Weighted level set evolution based on local edge features for medical image segmentation," *IEEE Trans. Image Process.*, vol. 26, no. 4, pp. 1979–1991, 2017.