

RESEARCH ARTICLE

# PCSeg: Color model driven probabilistic multiphase level set based tool for plasma cell segmentation in multiple myeloma

Anubha Gupta<sup>1</sup>\*, Pramit Mallick<sup>2</sup>, Ojaswa Sharma<sup>3</sup>, Ritu Gupta<sup>4</sup>, Rahul Duggal<sup>1</sup>

**1** SBILab, Department of ECE, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), New Delhi, India, **2** Department of Computer Science, Courant Institute of Mathematical Sciences, New York University, New York City, New York, United States of America, **3** Department of CSE, Indraprastha Institute of Information Technology-Delhi (IIIT-Delhi), New Delhi, India, **4** Laboratory Oncology Unit, Dr. B. R.A. IRCH, All India Institute of Medical Sciences (AIIMS), New Delhi, India

\* These authors contributed equally to this work.

\* [anubha@iiitd.ac.in](mailto:anubha@iiitd.ac.in)



**OPEN ACCESS**

**Citation:** Gupta A, Mallick P, Sharma O, Gupta R, Duggal R (2018) PCSeg: Color model driven probabilistic multiphase level set based tool for plasma cell segmentation in multiple myeloma. PLoS ONE 13(12): e0207908. <https://doi.org/10.1371/journal.pone.0207908>

**Editor:** Yuanquan Wang, Beijing University of Technology, CHINA

**Received:** October 24, 2017

**Accepted:** November 8, 2018

**Published:** December 12, 2018

**Copyright:** © 2018 Gupta et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** The data set used in this paper has been uploaded to Harvard Dataverse and is available at: <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/XCX7ST>.

**Funding:** The work was funded by Department of Science and Technology, Govt. of India. Grant Number: EMR/2016/006183 and the Ministry of Communication and IT, Govt. of India. Grant Number: 1(7)/2014-ME&HI. The funders had no role in study design, data collection and analysis,

## Abstract

Plasma cell segmentation is the first stage of a computer assisted automated diagnostic tool for multiple myeloma (MM). Owing to large variability in biological cell types, a method for one cell type cannot be applied directly on the other cell types. In this paper, we present PCSeg Tool for plasma cell segmentation from microscopic medical images. These images were captured from bone marrow aspirate slides of patients with MM. PCSeg has a robust pipeline consisting of a pre-processing step, the proposed modified multiphase level set method followed by post-processing steps including the watershed and circular Hough transform to segment clusters of cells of interest and to remove unwanted cells. Our modified level set method utilizes prior information about the probability densities of regions of interest (ROIs) in the color spaces and provides a solution to the minimal-partition problem to segment ROIs in one of the level sets of a two-phase level set formulation. PCSeg tool is tested on a number of microscopic images and provides good segmentation results on single cells as well as efficient segmentation of plasma cell clusters.

## Introduction

Cell classification via image processing has recently gained interest from the point of view of building computer assisted diagnostic tools for hematological malignancies. The computer assisted image processing tools can evaluate morphological features that are not discernable with human eyes. If automated, these tools can be used to analyze large number of cells in an objective manner for reliable assessment of specific cell populations of interest. The process of ‘Cell Segmentation’ is a precursor to cell classification implemented via image processing and hence, is the first stage of any computer assisted diagnostic tool. Several methods for cell segmentation have been described in the literature and often multiple methods are combined to achieve reasonable results depending on the type of cell images. Broad categories of

decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

segmentation methods include intensity thresholding methods, region-based segmentation methods, machine learning based methods and active contour methods [1].

Intensity thresholding based segmentation is one of the simplest and fastest methods of image segmentation. Dorini et al. [2] used intensity thresholding to segment nuclei of mature lymphocytes. Sharif et al. [3] utilized information contained in YCBr color space along with intensity thresholding, morphological operations, and watershed segmentation to segment red blood cells from the microscopic images. The method of Dorini et al. [2] fails to delineate the regions of interest (ROI) and the method of Sharif et al. [3] does not accommodate spatial intensity variation in images as it depends on the structuring element chosen. Hence, both the methods do not yield robust results, especially, when cells are present in clusters.

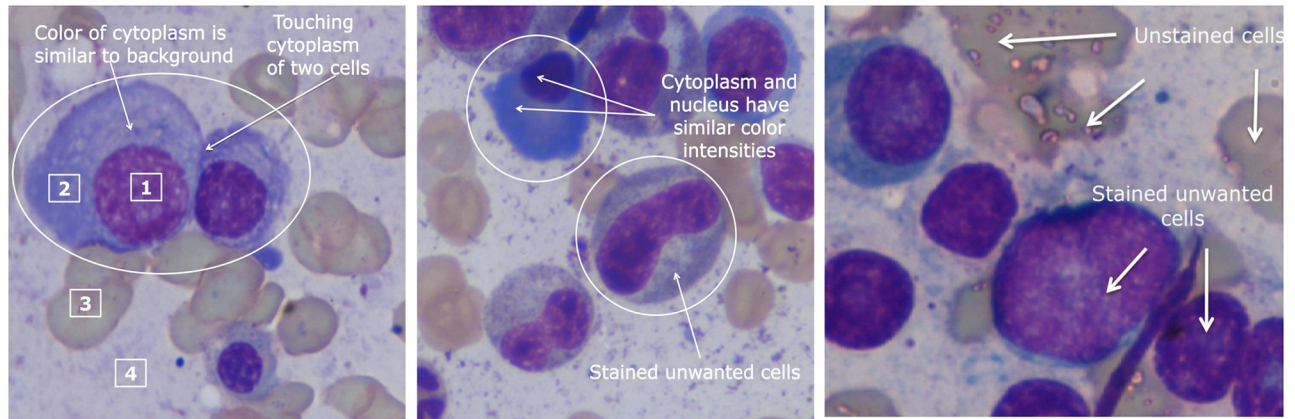
Region-based segmentation approaches look for connected components on the basis of properties such as texture and brightness. These approaches include seed based region growing and merging approaches [4–6]. In general, region growing methods are computationally expensive, are sensitive to noise, require correct identification of seeds, are local in nature without any global view, and at times have problem with the stopping criterion.

Machine learning based methods carry out segmentation via grouping of similar pixels (e.g. based on Euclidean distance on intensity) into clusters or by using other methods that learn pixel characteristics. Watershed, *k*-means clustering, and Support Vector Machines (SVM) are some of the most often used algorithms in segmentation [3, 7–10]. However, none of the above methods are able to segment cells of interest from cell clusters.

Active contour approach works on deformable curves that change their shapes according to the boundaries of targeted objects in an image using internal and external forces defining the motion of closed 2D contours [11–17]. Sadeghian et al. [12] carried out edge detection using Canny filter followed by geodesic snake contour method to segment leukocytes. However, these methods work on single cell windows extracted from the full microscopic image. Also, edge detection fails whenever intensity of nucleus and cytoplasm are similar.

With the focus of our study on building a robust automated pipeline for residual disease estimation in Multiple Myeloma (MM), a type of plasma cell (PC) cancer, the segmentation of the plasma cells was initiated as the first step. The pre-determined challenges specific to plasma cell segmentation are as follows (Fig 1): 1) Plasma cell segmentation requires segmentation of both nucleus and cytoplasm. At times, the color contrast of nucleus and cytoplasm and, more often, of the cytoplasm with the adjacent background is less due to overstaining or understaining. This poses difficulty in cell segmentation; 2) Plasma cells may be clustered together and hence, segmentation of the overlapping/touching cells is required. Generally, this is difficult because of different configurations as a) nuclei of different cells are touching, b) nuclei of one and cytoplasm of another cell are touching, or c) the cytoplasm of different cells are touching; and 3) Presence of more than one type of stained and unstained cells poses another challenge in extracting plasma cells of interest.

Although region growing and machine learning based methods have largely been used in cell segmentation, these methods are not effective in cluster segmentation [4–6, 8]. Contour based approaches such as snake models, level set models, and their variants are increasingly being used for segmentation in medical microscopic images [12–15, 17]. For example, Yang et al. [13] incorporated a color based gradient in the standard Gradient Vector Flow (GVF) model, a contour based approach to exploit the crucial information present in different histological components such as nucleus and cytoplasm of lymphocytes, follicle and mantle cells. Zamani and Safabakhsh [14] worked on a similar approach using GVF based on color gradients with the gradient flow initialized with the nuclei contours to identify nuclei using adaptive histogram thresholding to perform segmentation of lymphocytes. However, the accuracy of segmentation depends on the preliminary step of locating nuclei using histogram thresholding



**Fig 1. Challenges associated with plasma cell segmentation; numbers in boxes indicate the following image regions:** 1 nucleus of plasma cells, 2 cytoplasm of plasma cells, 3 unstained cells, and 4 background. Three challenges are highlighted via this Fig: 1) At times, the color difference between the cytoplasm with the adjacent background is less; 2) Plasma cells may be clustered together and hence, segmentation of the overlapping/touching cells is required; and 3) there may be more than one type of stained and unstained cells posing difficulty in extracting plasma cells of interest.

<https://doi.org/10.1371/journal.pone.0207908.g001>

that is generally not robust. Also, this approach fails to segment cell clusters. Yu et al. [15] used level set by Chan-Vese [16] to first segment only the nuclei of nerve cells and later used another level set to segment complete cells (nucleus and cytoplasm). Recently, Lu et al. [17] proposed a joint level set initialized with cell nuclei for pap smear cell segmentation. However, this approach fails in regions of low contrast between the nucleus and the cytoplasm.

From the above literature review, it appears that a contour based method may be able to provide a clear boundary of cells compared to morphological, thresholding, or clustering techniques. Since most of the above contour based methods are deterministic, a probabilistic level set formulation may be able to capture intra-subject and inter-subject related intensity variations within biological components such as within the cytoplasm. Region based and machine learning based methods have largely been used in cell segmentation but these methods are not observed to be effective in cluster segmentation. Contour based approaches such as snake models or level set models are the state-of-the-art medical image segmentation methods that are increasingly being used for segmentation in medical microscopic images [12–15, 17–19] as well as in other medical imaging applications, say, CT segmentation and brain MRI segmentation [20–23].

This motivates us to explore level set formulation within the probabilistic framework for plasma cell segmentation including cluster segmentation from microscopic images. In the present study, the existing methods as well as a recently described method by Saeedizadehet al. [24], using combination of thresholding, modified bottleneck algorithm, and watershed to segment plasma cells, was evaluated for segmentation of plasma cells in our set-up. The ultimate purpose of our work is to build an automated multiple myeloma residual disease detection tool for deployment in the hospital. Incorrect segmentation or partial segmentation of PCs will hinder the development of the subsequent classifier. Thus, we were motivated to explore the problem of PC segmentation afresh for robust results.

## Materials and methods

Microscopic images were captured from bone marrow aspirate slides of patients diagnosed with multiple myeloma as per the standard guidelines [25]. Slides were stained using Jenner-Giemsa stain. Images were captured at 1000x magnification using Nikon Eclipse-200

microscope equipped with a digital camera. Images were captured in raw BMP format with a size of 2560x1920 pixels. In all, our dataset consisted of 85 images. We trained our pipeline on 15 images. All the images were stain normalized, using the methodology proposed earlier [26], before being used for segmentation.

Written informed consent was obtained from all the subjects as per the guidelines of the Institute Ethics Committee (IEC) of All India Institute of Medical Sciences (AIIMS), New Delhi, India (Approval No. IEC/NP-145/2013 & RP-32/06.05.2013). Subsequently, a waiver for written informed consent for obtaining photomicrographs from the bone marrow aspirate slides was taken from IEC (approval No. OP-06/01.12.2017). One of the co-authors (RG) had access to the patient identifying information which was completely removed from the image data sets before sharing of data with the other co-authors for building up the PCSeg tool presented in this paper. The dataset is available at the public repository [37].

For the purpose of segmentation, an MM image can be divided into four regions of interest (ROI): (1) nucleus of PC, (2) cytoplasm of PC, (3) unstained cells, and (4) background (Fig 1). For efficient segmentation of all the four ROI, PCSeg Tool has been designed with the following four steps:

Step-1: Statistical modeling and computation of separability index of the four ROI in the images

Step-2: Removal of unstained cells

Step-3: Extraction of nucleus and cytoplasm of plasma cells using the proposed multiphase level set methodology

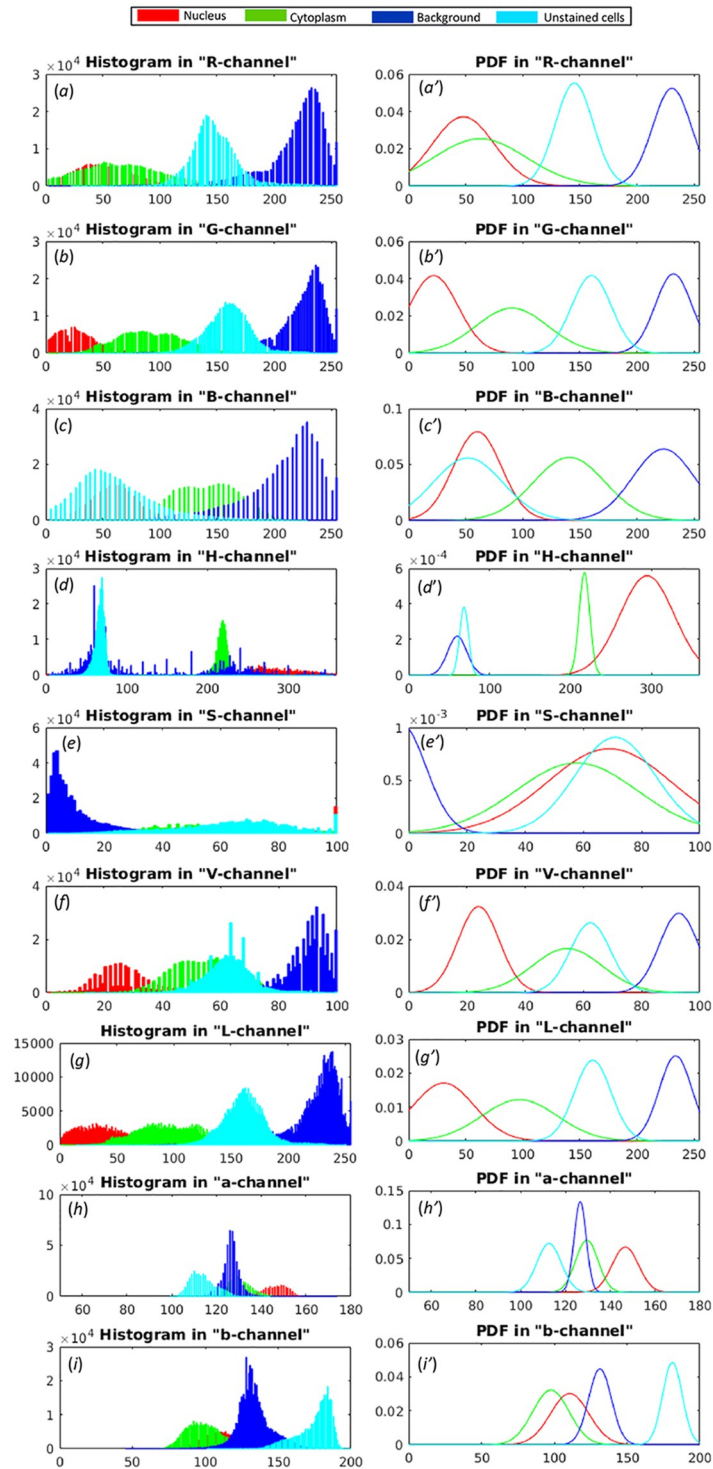
Step-4: Cluster cell segmentation using watershed and circular Hough transform

### Step-1: Statistical modeling and computation of separability index of regions of interest

First, the statistical characterization of the four ROI of MM images, i.e., nucleus of PC, cytoplasm of PC, unstained cells, and background (Fig 1) was done and the intensity profile of these regions was studied in different color spaces. A set of fifteen reference MM images, representative of the color histograms, were chosen and the histograms of RGB, HSV, and Lab color channels of the four ROI were marked in these reference images as shown in Fig 1. Since images were at a very high resolution of 2560 x 1920, sufficient numbers of pixels were available for computing the histogram. Histograms of RGB, HSV, and Lab color channels of the four ROI and their corresponding Gaussian probability density functions (PDFs) were fitted to the normalized histograms.

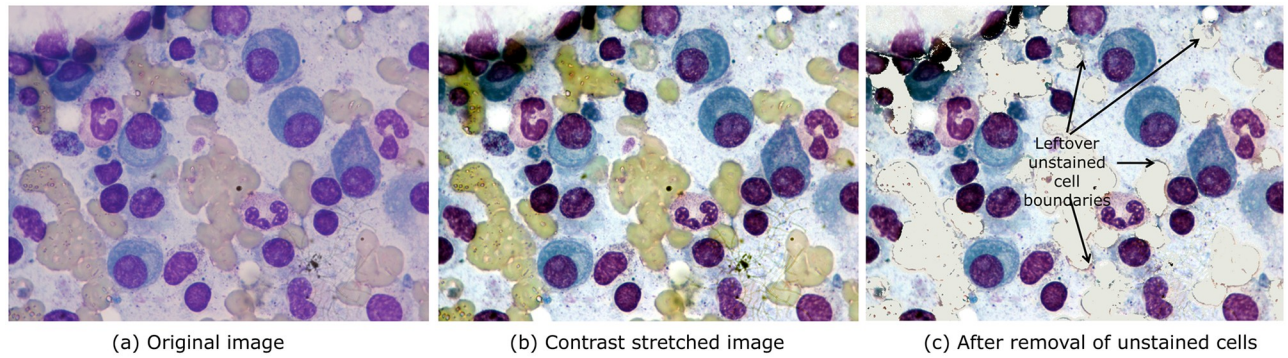
Fitted PDFs are drawn in Fig 2 in the intensity ranges of the original histograms. It is evident from these histograms that 1) nucleus and cytoplasm overlap in every color channel, although this overlap is considerably less in blue (B), hue (H), and value (V) channels; 2) nucleus and cytoplasm appear considerably separated from the background in red (R) and green (G) channels; and 3) the unstained cells do not overlap with nucleus and cytoplasm of plasma cells in the hue (H) channel. Thus, although it may be possible to remove unstained cells using the intensity profile in H-channel, nucleus and cytoplasm of plasma cells cannot be discerned using any single color channel. Rather, a combination of color channels would be required to separate these.

In order to quantify the separability of different image regions using probability distributions in RGB, HSV, and Lab color spaces, we used Bhattacharyya distance ( $D_B$ ) as a metric that



**Fig 2. Histogram of four image regions (nucleus, cytoplasm, background, and unstained cells) in:(a)-(c) RGB, (d)-(f) HSV, and (g)-(i): Lab color spaces; Corresponding fitted probability density functions (PDFs): (a')-(c') RGB, (d')-(f') HSV, and (g')-(i'): Lab color spaces.**

<https://doi.org/10.1371/journal.pone.0207908.g002>



**Fig 3. Contrast stretching followed by unstained cell removal, image patch size 2560 × 1920 this figure shows (a) an original image that is (b) contrast stretched such that 1% of lower and higher intensity values are saturated to 0 and 255, respectively. From the probability density functions of the four regions of interest (nucleus of PC, cytoplasm of PC, unstained cells, and background) of the resulting contrast stretched images, it is observed that nucleus and cytoplasm of plasma cells have maximum separability with unstained cells in H color channel. Unstained cells are removed by replacing intensity of pixels having values less than 120 in the H-channel with the background pixel intensity leading to (c).**

<https://doi.org/10.1371/journal.pone.0207908.g003>

quantifies separation between two PDFs  $p$  and  $q$  [27] as:

$$D_B(p, q) = -\log_e \left( \int \sqrt{p(x)q(x)} dx \right).$$

Prior to computing  $D_B$ , we applied contrast stretching on the RGB image such that 1% of lower and higher intensity values are saturated to 0 and 255, respectively. Next, we converted this contrast stretched image (Fig 3b) to HSV and Lab color spaces. The distance between all required combinations of two ROI in RGB, HSV, and Lab color spaces were computed (Table 1). These values are indicative of the separability between different regions and are used in the proposed modified multiphase level set method.

### Step-2: Removal of unstained cells

It is observed that both nucleus and cytoplasm of plasma cells have maximum separability with unstained cells in H color channel with large values of  $D_B$  distance (Table 1). Thus, we identified unstained cells using the H-color channel. Since both background and unstained

**Table 1. Bhattacharyya distance calculated between different image regions using the ground truth data.**

Color Channels	Bhattacharyya distance between two regions of interest					
	PC Nucleus and PC Cytoplasm	PC Nucleus and Background	PC Nucleus and Unstained Cells	PC Cytoplasm and Background Cells	PC Cytoplasm and Unstained Cells	Background and Unstained Cells
R	0.06	8.27	2.38	3.86	1.06	3.31
G	0.85	<b>15.89</b>	6.27	<b>4.21</b>	1.02	2.38
B	1.17	5.74	0.04	1.00	1.03	4.44
H	<b>1.76</b>	1.10	<b>11.19</b>	1.26	<b>73.13</b>	1.09
S	0.02	0.96	0.03	0.7	0.08	1.31
V	1.25	14.48	3.75	2.17	0.15	2.85
L	0.68	7.67	3.30	2.52	0.57	3.05
a	0.87	2.09	4.74	0.18	1.54	1.97
b	0.14	0.45	4.38	1.28	6.70	<b>4.89</b>

<https://doi.org/10.1371/journal.pone.0207908.t001>

cells are unwanted regions for the purpose of plasma cell segmentation, we replaced unstained cell pixels with the background pixels. This is carried out by replacing intensity of pixels having values less than 120 in the H-channel with the background pixel intensity (Fig 2). This replacement of unstained cells' intensity with the background intensity ensured that no additional region is created for the subsequently used multiphase level set algorithm (Fig 3(c)). Although unstained cells were removed, some outliers were still left (Fig 3(c)) that were subsequently removed in Step-3 of the proposed algorithm.

### Step-3: Stained cell extraction using the proposed modified multiphase level set method

We modified the multiphase level set formulation by utilizing statistical information of the four ROI in the image, i.e., nucleus, cytoplasm, unstained cells, and the background, wherein the four ROI were modeled via four-phases of two level sets. Each of the ROI was assigned one phase and the corresponding label. We assigned labels  $\Omega_{11}$  and  $\Omega_{10}$  to the first level set  $\phi_1$  for the nucleus and cytoplasm, respectively. Likewise, we assigned labels  $\Omega_{01}$  and  $\Omega_{00}$  to the second level set  $\phi_2$  for the background and the remaining unstained cells, respectively. Next, four probability maps of the entire image were created corresponding to each of the four phases of the level set (one each for their respective regions of interest, namely, nucleus, cytoplasm, unstained cells, and the background) as below:

$$p(U_0|\Omega_{ij}) = \frac{\sum_{c=1}^9 w_{c,\Omega_{ij}} * p(U_{0,c}|\Omega_{ij})}{\sum_{c=1}^9 w_{c,\Omega_{ij}}}, \tag{1}$$

where  $U_0$  is the contrast stretched image and  $c$  corresponds to color channels with  $c = 1, 2, 3, 4, \dots, 8, 9$  for channels R, G, B, H, S, V, L, a, and b, respectively.  $U_{0,c}$  is the  $c^{th}$  color channel of the image  $U_0$ ,  $\Omega_{ij}$  is one of the four ROI with  $i, j \in \{0, 1\}$ ,  $w_{c,\Omega_{ij}}$  are the weights,  $p(U_{0,c}|\Omega_{ij})$  is the conditional probability (over the ROI  $\Omega_{ij}$ ) of the color channel image, and  $p(U_0|\Omega_{ij})$  is the conditional probability of the original image over all four phases constructed using the weighted probability in all color channels.

Bhattacharyya distance  $D_B$  in Table 1 was used to determine weights because it provides an appropriate metric for discerning ROI. Weights for a channel were assigned based on the ability of discerning the desired ROI from all other ROI in that channel. Since Bhattacharyya distance would be higher for larger separation, it can be used as the weight, provided this distance is larger than some minimum threshold. For example, nucleus is discernible from both cytoplasm and background in blue channel with  $D_B > 1$  for each ROI (Table 1). Hence, the maximum of the two distances (distance between nucleus and cytoplasm, and distance between nucleus and background) is chosen as the weight for nucleus in blue channel. On the other hand, nucleus cannot be separated from cytoplasm in the red channel, although it is widely separated from the background in this color channel. This implies that nucleus cannot be extracted from all other ROI in red channel and hence, a zero weight is chosen for nucleus in this channel. Likewise, weights were chosen in stepwise manner for all the ROI, as detailed below.

For determining weights,  $w_{c,\Omega_{11}}$ , in (1) for nucleus, a channel was chosen (R, G, B, H, S, V, L, a, or b) and if the Bhattacharyya distance  $D_B$  (Table 1) between both 1) nucleus and cytoplasm, and 2) nucleus and background was greater than 1, the maximum distance of the above two was chosen as the weight for nucleus in that channel. Else a value of zero was assigned to

the weight in that channel (2). The process was repeated for all the channels.

$$w_{c,\Omega_{11}} = \begin{cases} \max[D_B(p(U_{0,c}|\Omega_{11}), p(U_{0,c}|\Omega_{10})), D_B(p(U_{0,c}|\Omega_{11}), p(U_{0,c}|\Omega_{01}))] \\ \quad \text{if } D_B(p(U_{0,c}|\Omega_{11}), p(U_{0,c}|\Omega_{10})) > 1 \\ \quad \text{and } D_B(D_B(p(U_{0,c}|\Omega_{11}), p(U_{0,c}|\Omega_{10})), p(U_{0,c}|\Omega_{01})) > 1 \\ 0 \quad \text{otherwise} \end{cases} \quad (2)$$

Similarly, weights were assigned for cytoplasm based on the Bhattacharya distance  $D_B$  between both 1) cytoplasm and nucleus, and 2) cytoplasm and background in each of the channels (3).

$$w_{c,\Omega_{10}} = \begin{cases} \max[D_B(p(U_{0,c}|\Omega_{10}), p(U_{0,c}|\Omega_{11})), D_B(p(U_{0,c}|\Omega_{10}), p(U_{0,c}|\Omega_{01}))] \\ \quad \text{if } D_B(p(U_{0,c}|\Omega_{10}), p(U_{0,c}|\Omega_{11})) > 1 \\ \quad \text{and } D_B(D_B(p(U_{0,c}|\Omega_{10}), p(U_{0,c}|\Omega_{11})), p(U_{0,c}|\Omega_{01})) > 1 \\ 0 \quad \text{otherwise} \end{cases} \quad (3)$$

Weights were assigned for background based on the Bhattacharya distance  $D_B$  between both 1) background and nucleus, and 2) background and cytoplasm. Weights were assigned for unstained cells based on the Bhattacharya distance  $D_B$  between both 1) unstained cells and nucleus, and 2) unstained cells and cytoplasm. Since plasma cells are required to be clearly delineated from the background and unstained cells, a greater threshold of 3 was considered for background (4) and unstained cells (5).

$$w_{c,\Omega_{01}} = \begin{cases} \max[D_B(p(U_{0,c}|\Omega_{01}), p(U_{0,c}|\Omega_{11})), D_B(p(U_{0,c}|\Omega_{01}), p(U_{0,c}|\Omega_{10}))] \\ \quad \text{if } D_B(p(U_{0,c}|\Omega_{01}), p(U_{0,c}|\Omega_{11})) > 3 \\ \quad \text{and } D_B(D_B(p(U_{0,c}|\Omega_{01}), p(U_{0,c}|\Omega_{11})), p(U_{0,c}|\Omega_{10})) > 3 \\ 0 \quad \text{otherwise} \end{cases} \quad (4)$$

$$w_{c,\Omega_{00}} = \begin{cases} \max[D_B(p(U_{0,c}|\Omega_{00}), p(U_{0,c}|\Omega_{11})), D_B(p(U_{0,c}|\Omega_{00}), p(U_{0,c}|\Omega_{10}))] \\ \quad \text{if } D_B(p(U_{0,c}|\Omega_{00}), p(U_{0,c}|\Omega_{11})) > 3 \\ \quad \text{and } D_B(D_B(p(U_{0,c}|\Omega_{00}), p(U_{0,c}|\Omega_{11})), p(U_{0,c}|\Omega_{10})) > 3 \\ 0 \quad \text{otherwise} \end{cases} \quad (5)$$

Final weights ( $w_{c,\Omega_y}$ ) obtained using the above scheme are summarized in Table 2.

Next, we defined an energy functional  $E_p(\phi_1, \phi_2)$  for the multiphase level set formulation, where  $\phi_1$  and  $\phi_2$  are the two level set functions that capture the curves of cell boundaries. The



Table 2. Weights of each ROI for level set equations.

Color channels (c)	Weights ( $w_{c,\Omega_{ij}}$ )			
	Region $\Omega_{11}$ (Nucleus)	Region $\Omega_{10}$ (Cytoplasm)	Region $\Omega_{01}$ (Background)	Region $\Omega_{00}$ (Unstained cells)
R	0	0	8.27	0
G	0	0	15.89	0
B	5.74	1.17	0	0
H	1.76	1.76	0	73.13
S	0	0	0	0
V	14.48	2.17	0	0
L	0	0	0	0
a	0	0	0	0
b	0	0	0	6.70

<https://doi.org/10.1371/journal.pone.0207908.t002>

energy functional  $E_p(\phi_1, \phi_2)$  utilizes the above constructed probability maps and was added to the overall functional required to be minimized for the derivation of level set equations.

$$\begin{aligned}
 E_p(\phi_1, \phi_2) = & - \int_{\Omega} \log_e(p(U_0(\mathbf{x})|\Omega_{11}))H(\phi_1(\mathbf{x}))H(\phi_2(\mathbf{x}))d\mathbf{x} \\
 & - \int_{\Omega} \log_e(p(U_0(\mathbf{x})|\Omega_{10}))H(\phi_1(\mathbf{x}))[1 - H(\phi_2(\mathbf{x}))]d\mathbf{x} \\
 & - \int_{\Omega} \log_e(p(U_0(\mathbf{x})|\Omega_{01}))[1 - H(\phi_1(\mathbf{x}))]H(\phi_2(\mathbf{x}))d\mathbf{x} \\
 & - \int_{\Omega} \log_e(p(U_0(\mathbf{x})|\Omega_{00}))[1 - H(\phi_1(\mathbf{x}))][1 - H(\phi_2(\mathbf{x}))]d\mathbf{x},
 \end{aligned} \tag{6}$$

where  $H$  is the Heaviside function,  $\Omega \subset \mathbb{R}^2$  is an open and bounded domain,  $U_0 : \Omega \mapsto \mathbb{R}$  is the given bounded function representing the initial image, and  $C$  as the closed subset in  $\Omega$  made up of finite set of smooth curves.

We also defined another energy functional, namely, distance energy functional  $E_d(\phi_1, \phi_2)$  that measures the intensity difference of a given pixel from each of the region’s mean color value. To this end, we first defined and computed the distance images  $U_{d,\Omega_{ij}}$  for each of the regions in (7) as:

$$U_{d,\Omega_{ij}}(\mathbf{x}) = \left[ \frac{\sum_{c=1}^9 w_{c,\Omega_{ij}}(U_{0,c}(\mathbf{x}) - \mu_{c,\Omega_{ij}})^2}{\sum_{c=1}^9 w_{c,\Omega_{ij}}} \right], \tag{7}$$

where  $\mu_{c,\Omega_{ij}}$  is the mean color value of the distribution in region  $\Omega_{ij}$  and  $w_{c,\Omega_{ij}}$  is the weight of color channel  $c$  in region  $\Omega_{ij}$  as tabulated in Table 2. Accordingly, we defined the distance

energy functional  $E_d(\phi_1, \phi_2)$  in (8) as:

$$\begin{aligned}
 E_p(\phi_1, \phi_2) = & \int_{\Omega} U_{d,\Omega_{11}}(\mathbf{x})H(\phi_1(\mathbf{x}))H(\phi_2(\mathbf{x}))d\mathbf{x} \\
 & + \int_{\Omega} U_{d,\Omega_{10}}(\mathbf{x})H(\phi_1(\mathbf{x}))[1 - H(\phi_2(\mathbf{x}))]d\mathbf{x} \\
 & + \int_{\Omega} U_{d,\Omega_{01}}(\mathbf{x})[1 - H(\phi_1(\mathbf{x}))]H(\phi_2(\mathbf{x}))d\mathbf{x} \\
 & + \int_{\Omega} U_{d,\Omega_{00}}(\mathbf{x})[1 - H(\phi_1(\mathbf{x}))][1 - H(\phi_2(\mathbf{x}))]d\mathbf{x}.
 \end{aligned}
 \tag{8}$$

Adding the regularization terms of length and area, the proposed modified multiphase level set energy functional is defined in (9) as:

$$\begin{aligned}
 E(\phi_1, \phi_2) = & \eta_1 E_p(\phi_1, \phi_2) + \eta_2 E_d(\phi_1, \phi_2) + \alpha_1 \int_{\Omega} H(\phi_1(\mathbf{x}))d\mathbf{x} \\
 & + \alpha_2 \int_{\Omega} H(\phi_2(\mathbf{x}))d\mathbf{x} + \beta_1 \int_{\Omega} |\nabla H(\phi_1(\mathbf{x}))|d\mathbf{x} \\
 & + \beta_2 \int_{\Omega} |\nabla H(\phi_2(\mathbf{x}))|d\mathbf{x},
 \end{aligned}
 \tag{9}$$

where  $\alpha$  is a constant regularizer that controls the area inside the contour  $C$ ,  $\beta$  controls the length of the contour and,  $\eta_1$  and  $\eta_2$  are the constants that control relative weighting of the two energy functionals. The level set  $(\phi_1, \phi_2)$  is periodically re-initialized to the signed distance function [28].

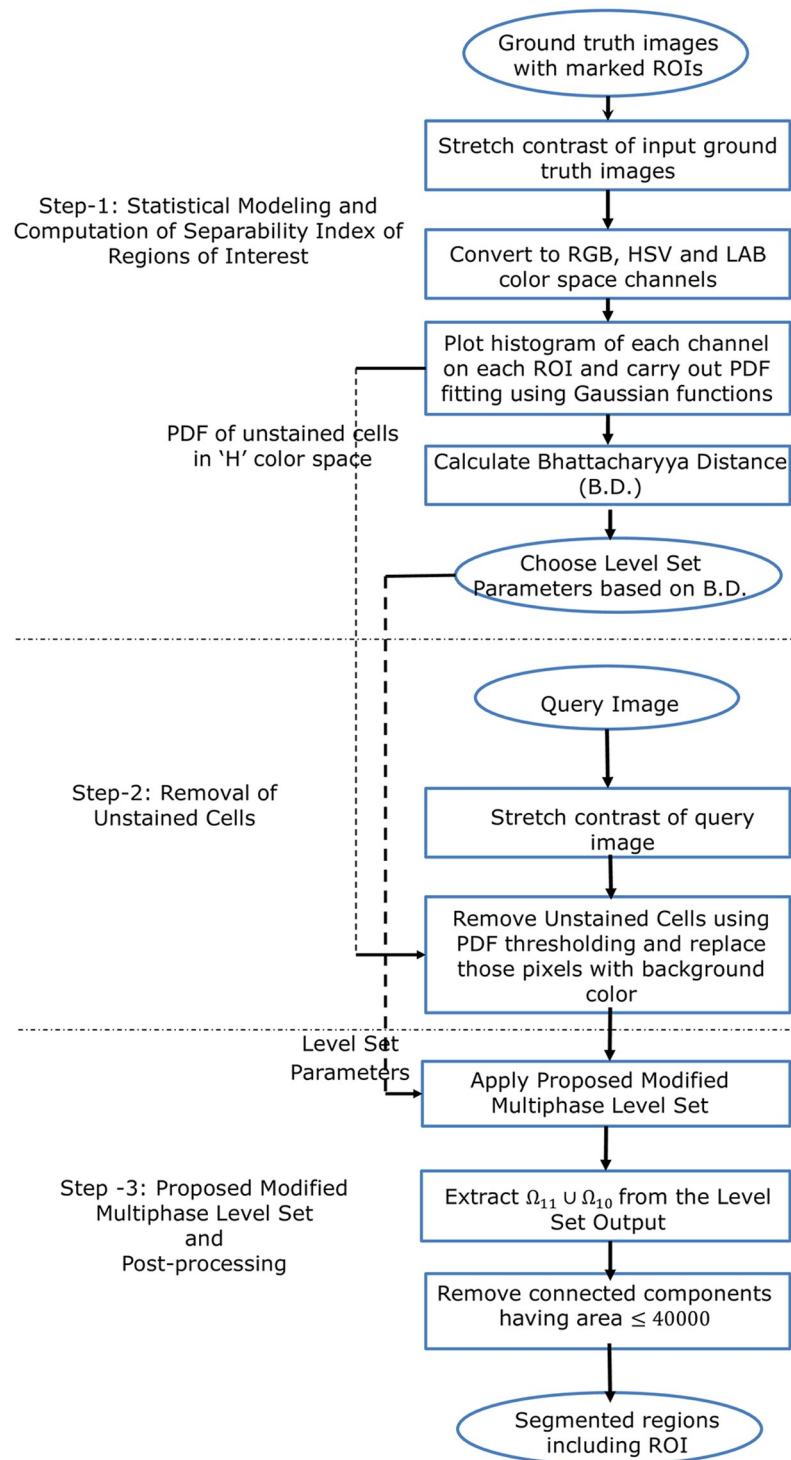
Fig 4 presents Steps 1 to 3 of the proposed method including the modified multiphase level set formulation. The extracted mask of  $\Omega_{11} \cup \Omega_{10}$  from the level set output provides the segmented plasma cells. Fig 5 presents the segmentation results after processing the input image with Steps-1 to 3.

On evaluation of the output of the multiphase level set step in Fig 5, it was noted that some unwanted stained cells, such as lymphocytes, were segmented to final output. On careful observation, we noted that the cytoplasm of stained PCs covered large cell area compared to unwanted stained cells and therefore, unwanted cells could be rejected using a threshold on cytoplasm cell area. In addition, some small disconnected components that were noisy patches owing to faulty manual staining were also captured by the multiphase level set. These small noisy disconnected components that are too small to form any ROI were rejected at the output of level set by thresholding on the size of the component.

Although all the four ROI were captured by the multiphase level set, plasma cell clusters were not segmented as observed from Fig 5. To address this problem, Step-4 was added to the tool as detailed below.

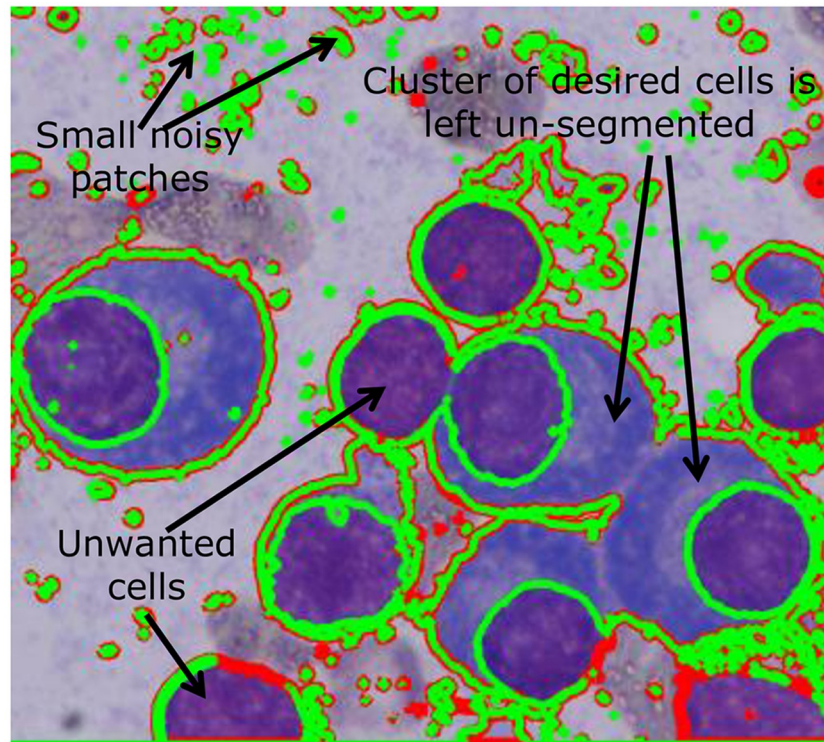
### Step-4: Cluster cell segmentation using watershed and circle Hough transform

Since stained plasma cells are approximately circular in shape, a combination of watershed and circular Hough transform (CHT) was applied to segment PC clusters. First, it is necessary to segment nuclei as the cases of touching nuclei will lead to improper cell segmentation. The nucleus of a plasma cell has a few distinct features as: 1) the nucleus is dark colored, 2) it is



**Fig 4. Schematic diagram of Steps 1 to 3 of the proposed method of PCSeg tool:** Regions of interest (ROI) are: Nuclei of Plasma Cells ( $\Omega_{11}$ ), Cytoplasm of Plasma Cells ( $\Omega_{10}$ ), unstained cells ( $\Omega_{01}$ ), and background ( $\Omega_{00}$ ).

<https://doi.org/10.1371/journal.pone.0207908.g004>



**Fig 5. Output of the modified multiphase level set method after Step-3 on an image patch.** It was noted that some unwanted stained cells (e.g. lymphocytes) were segmented to final output. In addition, some small disconnected components that were noisy patches owing to faulty manual staining were also captured by the multiphase level set. Also plasma cell clusters were not segmented.

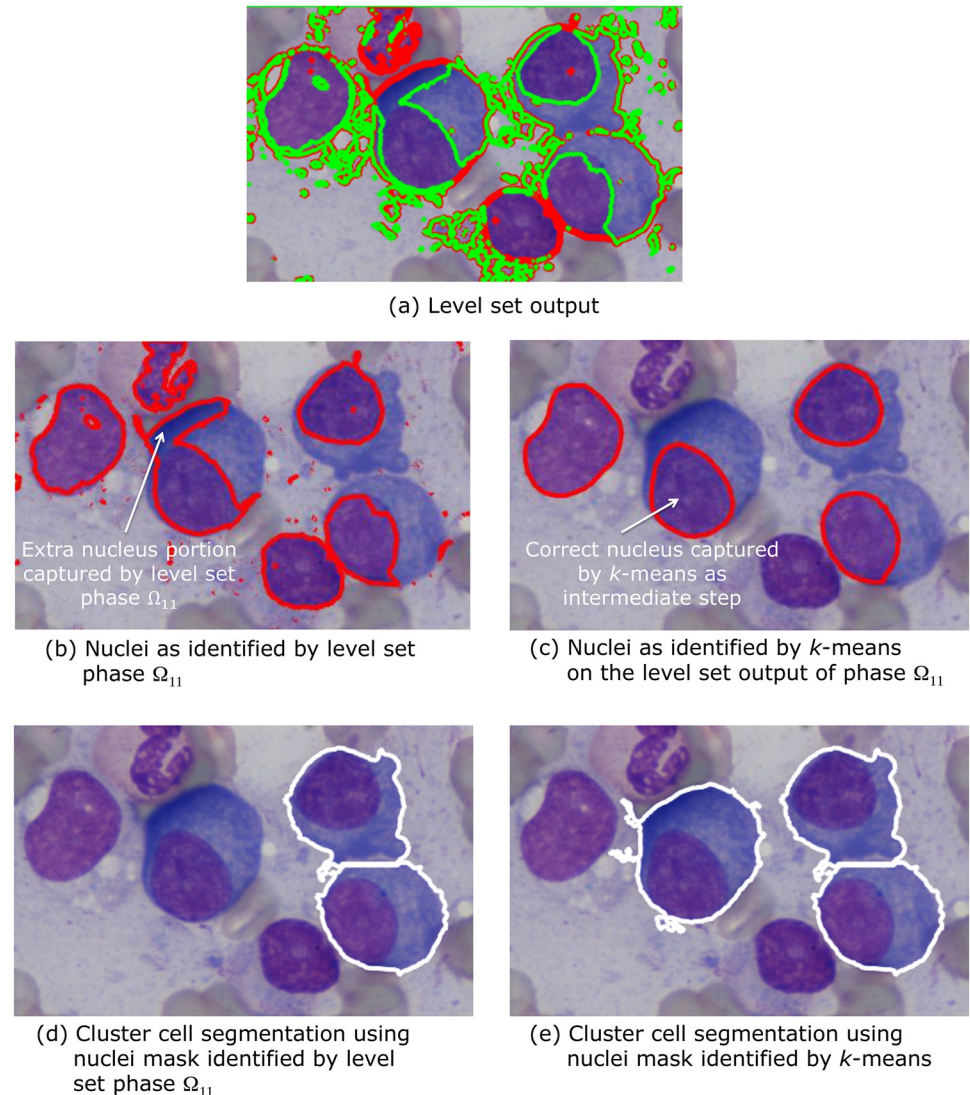
<https://doi.org/10.1371/journal.pone.0207908.g005>

differently colored than the background, and 3) it is always encapsulated within the cytoplasm. Thus, the center of the nucleus could serve as an ideal seed for the watershed algorithm and the nuclei mask obtained as an output of  $\Omega_{11}$  phase of the level set could be used to compute the distance transform required by the watershed algorithm.

However, due to the variability in cytoplasm staining, the  $\Omega_{11}$  phase was observed to capture nuclei regions more liberally in some images as shown in Fig 6b. This led to unnecessary rejections of some PCs as shown in Fig 6d. Since it is vital for any medical imaging work to segment as many correct cells as possible,  $k$ -means (instead of  $\Omega_{11}$  phase of the level set) was used to extract the nuclei mask (Fig 6c) and distance transform was applied on this mask. This basin was used by the watershed algorithm to segment the nuclei from clusters. From the watershed output, only those segmented nuclei regions were retained that were circular in nature, i.e., segmented regions that contained a center point of CHT. The non-circular regions identified as nuclei were discarded.

Following this, the distance transform of the mask of the stained portions ( $\Omega_{11} \cup \Omega_{10}$ ) was obtained. The centers of the segmented nuclei obtained from  $k$ -means above were used to impose a minima on this basin and subsequently used by the watershed algorithm to segment full plasma cells from clusters. Again, we retained only those segmented regions as cells that were circular in nature, i.e., segmented regions that contained a center point of CHT.

We have named the developed tool as PCSeg Tool-1 for the complete pipeline with  $k$ -means based nuclei mask for cluster cell segmentation in Step-4 and named the developed tool as PCSeg Tool-2 for the complete pipeline with  $\Omega_{11}$  phase based nuclei mask for cluster cell



**Fig 6.** Output of cluster cell segmentation on an image patch Subfigure (a) shows the level set output wherein red boundary shows nucleus of PC being captured by levelset and green boundary shows cytoplasm of PC being captured. However, the cluster of cells are not segmented. (b) shows nuclei identified by levelset phase  $\Omega_{11}$  and the problem therein of some extra mask of cytoplasm in nucleus. (d) shows cluster cell segmentation using nuclei mask identified by level set  $\Omega_{11}$  in (b). One cell is falsely rejected. (c) shows nuclei identified by  $k$ -means on the levelset phase  $\Omega_{11}$  and (e) shows correct cluster cell segmentation using the mask of (c).

<https://doi.org/10.1371/journal.pone.0207908.g006>

segmentation in Step-4. The complete process pipeline of cluster cell segmentation with PCSeg Tool-1 is shown in Fig 7.

We noted that while most PCs were segmented, some of the stained lymphocytes and unwanted regions were also retained. In general, the amount of cytoplasm in lymphocytes is considerably less in comparison to PCs. Thus, for each segmented region, the ratio of the nucleus area to the total cell area was used to detect and discard these unwanted cells. As stated earlier, it was observed that the use of nuclei segmented from the phase  $\Omega_{11}$  led to inadvertent rejection of some PCs (Fig 6d), while the use of nuclei segmented from  $k$ -means helped us in retaining such cells of interest (Fig 6e). Hence, PCSeg Tool-1 based segmentation pipeline

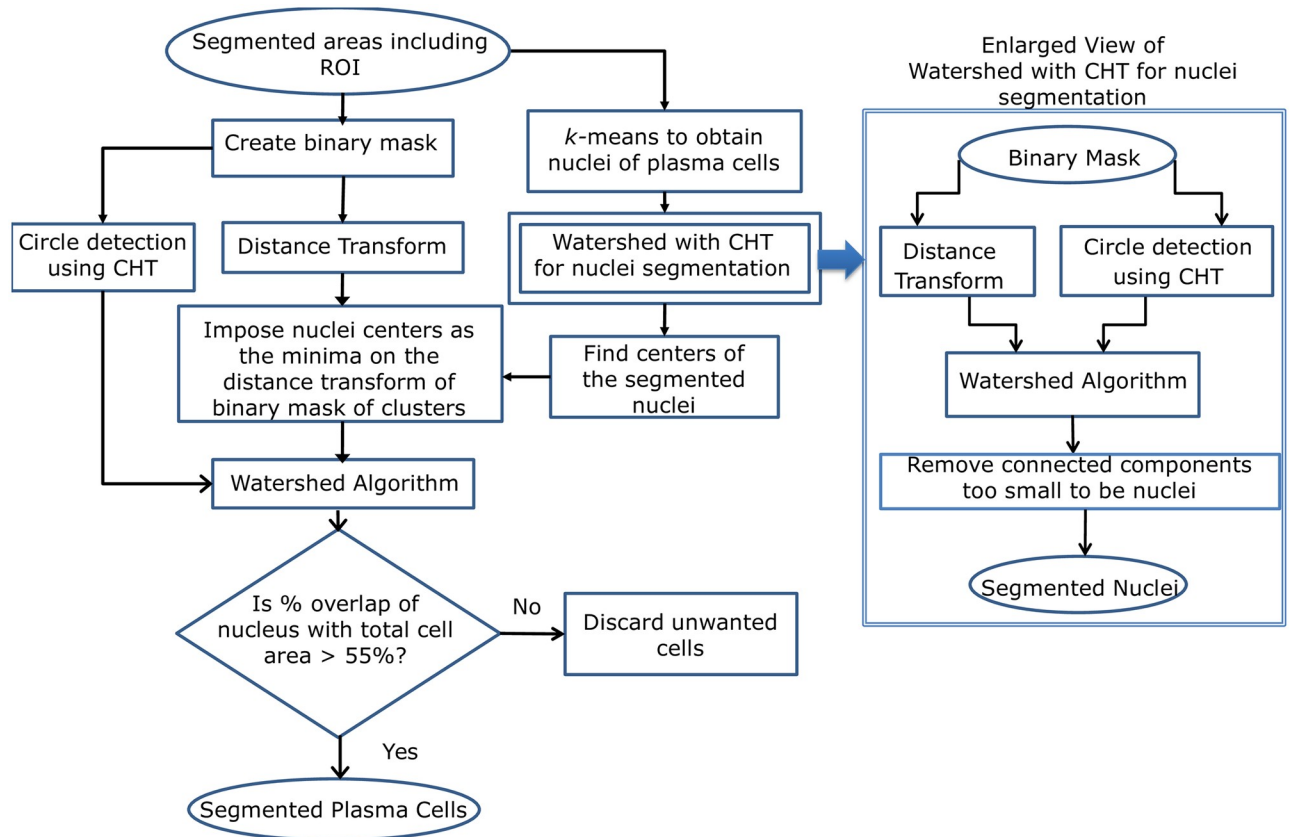


Fig 7. Schematic diagram of cluster cell segmentation (Step 4 of PCSeg tool) using watershed and circle Hough transform (CHT).

<https://doi.org/10.1371/journal.pone.0207908.g007>

appeared more useful. Fig 6e shows the correctly segmented cluster, while the complete process pipeline is shown in Fig 7.

### Experimental set-up

All experiments were performed on a Ubuntu 14.04 system with an Intel<sup>®</sup> Xeon(R) CPU E5-2630 v2 @ 2.60GHz 12 processor and a GeForce GTX 980/PCIe/SSE2 graphics card supporting CUDA. Level set results depend on the initialization of  $\phi$ 's and therefore, initial contour was set to small circles covering the entire image to ensure faster convergence. The initialization and the energy functionals were calculated on the CPU. The level set propagation and the re-initialization of  $\phi$ 's was implemented on GPU using MEX compiled files containing CUDA code. While the implementation was memory bound, the code achieved upto 75x speed up as compared to the MATLAB R2015b implementation. The rest of the pipeline was implemented in MATLAB. The GPU versus CPU computational speed results are shown in Fig 8.

### Evaluation metrics

A total of 85 images were considered, where parameters were fine tuned on 15 randomly chosen images. For our experiments, the level set parameters  $\alpha_1 = \alpha_2 = 0$ ,  $\beta_1 = 2$ ,  $\beta_2 = 1$ ,  $\eta_1 = 1$  and  $\eta_2 = 3$  provided the best results. Parameters  $\alpha_1$  and  $\alpha_2$  are related to the compactness and the area of level set phase captured at the end. Since our images contained both isolated single cells as well as cluster of cells, putting constraints on the area provided poor results. Hence, we

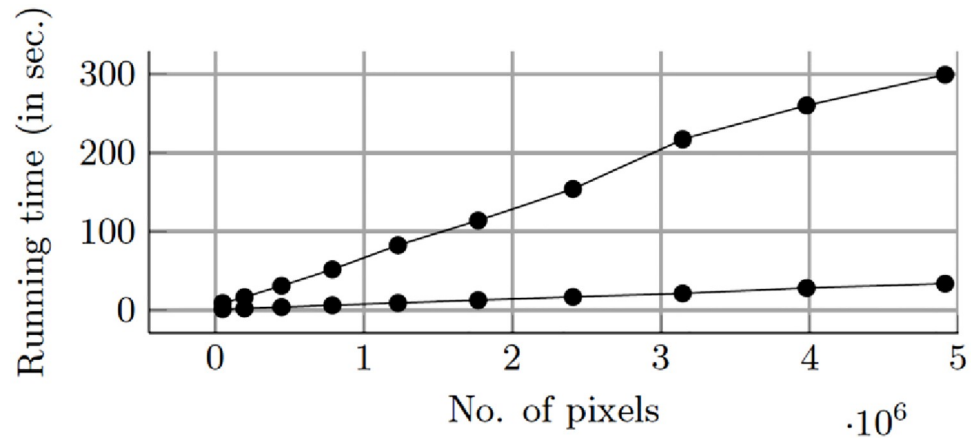


Fig 8. Running times on CPU and GPU.

<https://doi.org/10.1371/journal.pone.0207908.g008>

chose these parameters to be zero. Parameters  $\beta_1$  and  $\beta_2$  control the tightness of the boundary. Since Level set 1 captures the desired plasma cell, we wanted its boundary to be tighter compared to that of Level set 2. Hence, a higher  $\beta$  value was chosen for Level set 1 compared to Level set 2, i.e.,  $\beta_1$  was chosen to be greater than  $\beta_2$ . Since we used stain normalized images, the mean color vector based energy functional captured cells of interest neatly, while the probability based energy functional term took care of the slight color variations owing to subject variability and/or other variability. Hence,  $\eta_2 \geq \eta_1$  provided us best results.

The image dataset evaluated contained 260 single cells and 45 clusters in total. These 45 clusters had a total of 102 cells, with each cluster having two or more cells. For a quantitative assessment of the proposed pipeline and the method, we used TPR (True Positive Rate) or Recall rate defined as

$$TPR = \frac{TP}{TP + FN},$$

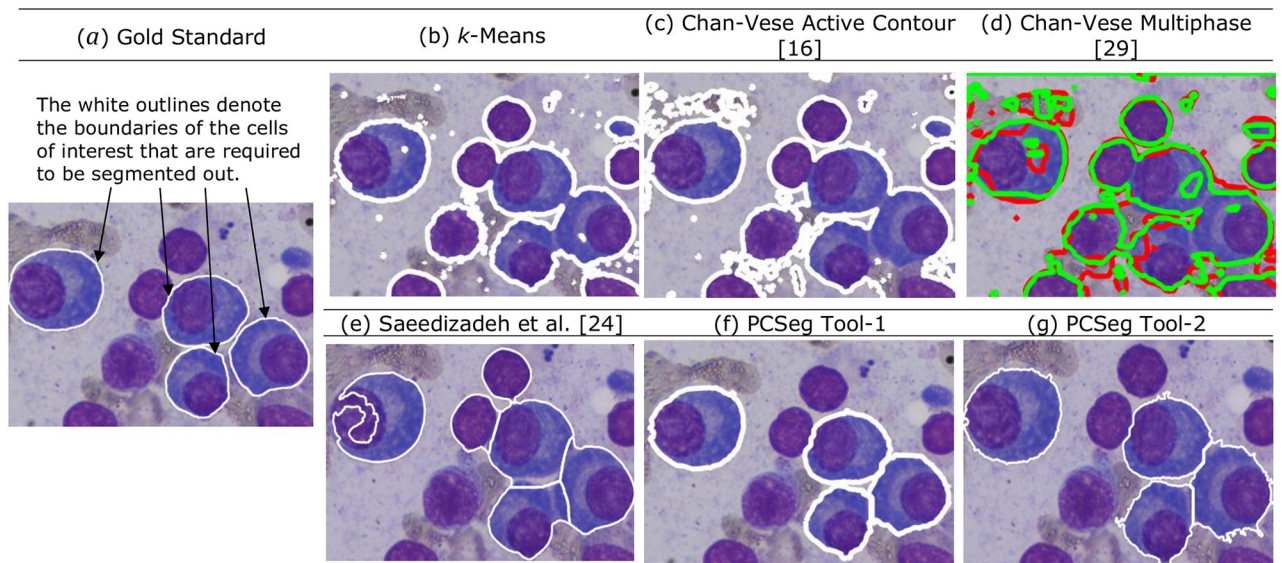
PPV (Positive Predictive Value) or Precision defined as

$$PPV = \frac{TP}{TP + FP},$$

and F<sub>1</sub>-Score defined as

$$F_1 - \text{Score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}},$$

where TP, TN, FP, and FN stand for true positive (PC detected and segmented as PC), true negative (non-PC rejected), false positive (non-PC segmented as PC) and false negative (PC rejected as non-PC), respectively. In TP or true positives, we only considered those plasma cells that were completely segmented from the image. Any plasma cell that was over-segmented or segmented out with partial portion was discarded. This is to note that F<sub>1</sub>-Score is same as the Dice coefficient that is used as a standard metric to assess the performance of segmentation.



**Fig 9. Qualitative comparison of MM cell segmentation using different methods:** (a) Gold standard (showing cells of interest with white outlines), (b) *k*-means, (c) Chan-Vese active contour method [16], (d) Chan-Vese multiphase method [29], (e) Saeedizadeh et al. [24] method, (f) PCSeg Tool-1, and (g) PCSeg Tool-2. All white outlines in (b)-(g) denote the outlines of regions segmented out. These regions are required to be compared with the regions contained in the Gold Standard shown in (a).

<https://doi.org/10.1371/journal.pone.0207908.g009>

Since the rate of detection of false positives is also crucial in medical applications, we evaluated False Discovery Rate (FDR) in our samples as

$$FDR = \frac{FP}{TP + FP}.$$

## Results

We compared the results with the traditional levelset, multiphase levelset and *k*-means in Fig 9. From Fig 9, we notice that the output of these methods do not yield correctly segmented cells (including both nucleus and cytoplasm). Hence, quantification of results will not yield any accuracy with these methods that is worth comparison. Since the work by Saeedizadeh et al. [24] specifically addresses the problem of plasma cell segmentation, its pipeline is tuned to this cell type. Hence, of the existing methods including levelset, we could quantify results of this method only and thus, used the method by Saeedizadeh et al. [24] as the state-of-the-art work as of today for comparison on the given problem statement.

Quantitative results obtained on 260 numbers of isolated single PC and 45 clusters (with 102 PC) are tabulated in Table 3 and the statistical quantities on TPR, PPV,  $F_1$ -score, and FDR are tabulated in Table 4. The PCSeg Tool-1 correctly segmented 83.5% of single and isolated plasma cells and 93.3% of PC clusters (71.1% complete clusters and 22.2% partial clusters). PCSeg Tool-2 performed slightly inferior and segmented 55.8% of single plasma cells and 64.5% of PC clusters (48.9% complete clusters and 15.6% partial clusters). Further, the number of false positives detected with PCSeg Tool-1 were about 90 cells leading to an FDR of 23.44% (Table 4). Compared to this, PCSeg Tool-2 detected 102 false positives and thus, had a higher FDR of 34.11% (Table 4). Thus, the PCSeg Tool-1 performed better than PCSeg Tool-2 in terms of both TPR and FDR.

Further, we compared both the proposed methods, i.e., PCSeg Tool-1 (PCSeg Tool with *k*-means as nuclei mask for cluster cell segmentation) and PCSeg Tool-2 (PCSeg Tool with  $\Omega_{11}$



**Table 3. Quantitative results on plasma cell segmentation.**

Method	Performance on 260 single/isolated cells	Performance on clusters, Out of (45 clusters, 102 cells in 45 clusters)		*completely rejected clusters	False positives
	Correctly segmented cells	(completely segmented clusters, completely segmented cells from clusters)	(partially segmented clusters, completely segmented cells from these clusters)		
PCSeg Tool-1	217	(32,64)	(10,13)	3	90
PCSeg Tool-2	145	(22,43)	(7,9)	16	102
Saeedizadeh et al. [24]	161	(26,46)	(3,5)	16	192

\*Some clusters were completely rejected out of the total of 45 clusters.

<https://doi.org/10.1371/journal.pone.0207908.t003>

phase of level set as nuclei mask for cluster cell segmentation) with existing cell segmentation methods relevant to our problem. These existing methods included *k*-means method, Chan-Vese active contour, Chan-Vese multiphase methods, and a recently described method by Saeedizadeh et al. [24]. *k*-means is a widely used method, while the proposed multiphase level set method in PCSeg Tool-1 and PCSeg Tool-2 is a refinement over standard active contour methods of Chan-Vese active contour [16] and Chan-Vese multiphase [29] methods. All these methods, i.e., *k*-means and standard active contour methods, have been used earlier on segmentation of cells other than plasma cells. The method of Saeedizadeh et al. [24] addresses specifically the plasma cell segmentation. Hence, all the above four methods were chosen for qualitative comparison with PCSeg Tool-1 and PCSeg Tool-2.

### Discussion

The *k*-means method provided many false positives, missed many PCs, and could not segment clusters of PCs (Fig 9). Similar was the case with standard active contour methods of Chan-Vese active contour and Chan-Vese multiphase methods. Since *k*-means, Chan-Vese active contour [16], and Chan-Vese multiphase [29] methods performed poorly on PC segmentation, quantitative results have been presented on only the rest of the three methods (Tables 3 and 4).

As compared to Saeedizadeh et al. [24], PCSeg Tool-1, and PCSeg Tool-2 performed far more superior to *k*-means and standard active contour methods (Tables 3 and 4). While PCSeg Tool-2 did not outperform [24] in the number correct PC cell segmentation, it did better with FDR compared to [24] which led to a very high number of 192 false positives with an FDR of 47.52% (Table 4). The method by [24] performed second best by segmenting 62% of single plasma cells and 64.5% of PC clusters (57.8% complete clusters and 6.7% partial clusters). However, one of the 4 plasma cells present in Fig 9 has been incorrectly segmented by [24]. On the other hand, PCSeg Tool-1 and PCSeg Tool-2 captured all 4 cells. This result shows that both the variants of PCSeg Tool (1 and 2) performed better in capturing the cells of interest.

**Table 4. Statistical results on a total of 364 (260+102) plasma cells.**

Method	True Positive Rate or Recall (%)	Positive Predictive Value or Precision (%)	False Discovery Rate (%)	F <sub>1</sub> -Score (%)
PCSeg Tool-1	81.66	76.56	23.44	79.03
PCSeg Tool-2	54.72	65.88	34.12	59.78
Saeedizadeh et al. [24]	58.88	52.47	47.53	55.49

<https://doi.org/10.1371/journal.pone.0207908.t004>

This is an expected result because the methods specifically designed for plasma cell segmentation, i.e., PCSeg Tools 1 and 2 and [24] take into account the problems inherent to plasma cell segmentation. Further, these results establish that the method applicable on one cell type cannot be ported for segmentation of another cell type directly, i.e., no single method or segmentation pipeline can be applied to all the different cell types.

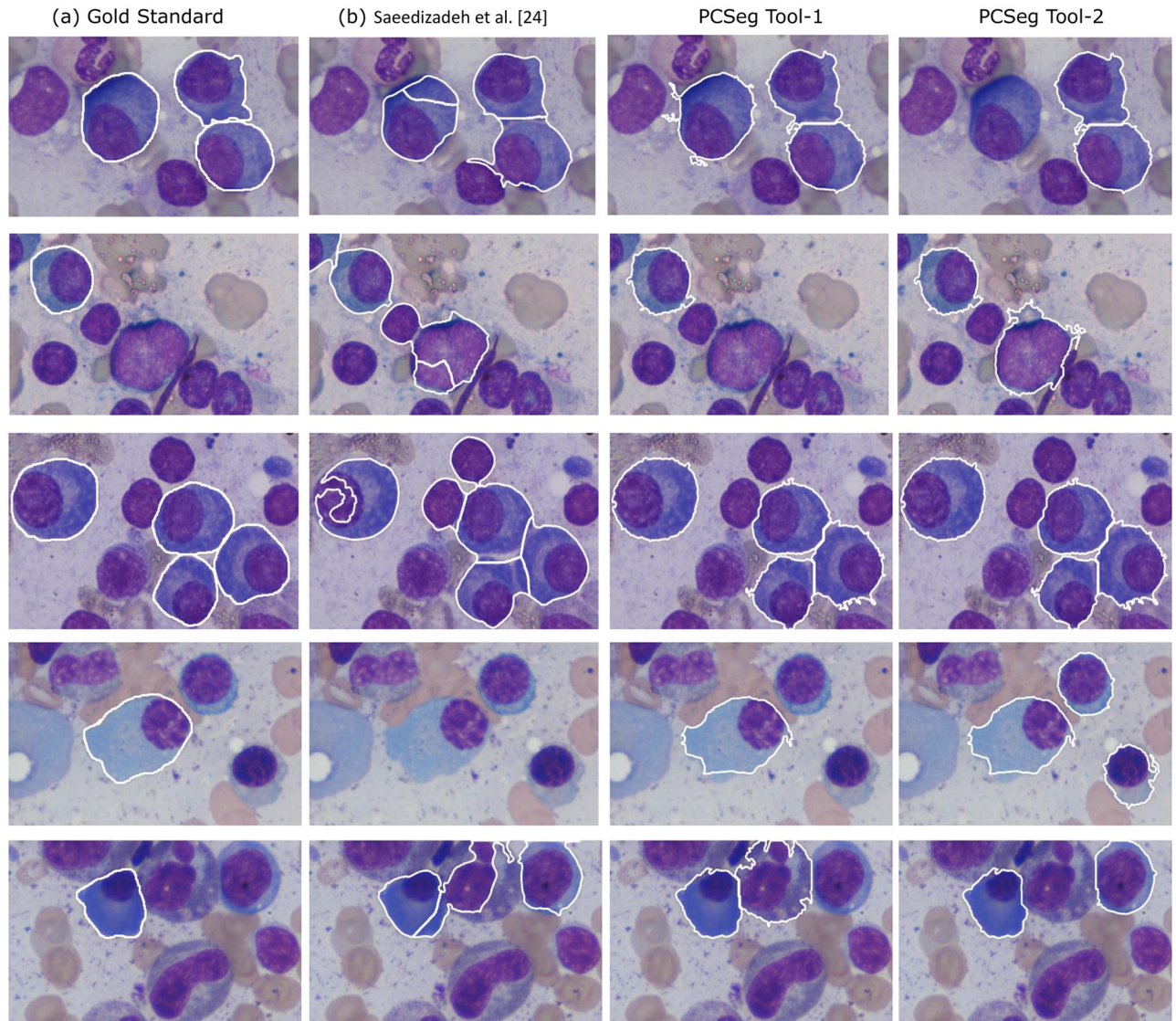
We also compared the performance of the above three segmentation tools with reference to Precision, Recall, and F1-score. While recall rate quantifies a tool's performance with respect to false negatives (how many plasma cells were missed out in segmentation), precision rate informs us about the performance of the tool with respect to false positives, i.e., other cells segmented out as plasma cells. Thus, recall informs what ratio of correct plasma cells could be segmented, while precision tells how many of the cells segmented as plasma cells were erroneous. Thus, both precision and recall have a significance and both these are imbedded in the  $F_1$ -score that should be as high as possible and informs about the overall performance of the segmentation tool. It is noted that recall of PCSeg Tool-1 is good as 81.66% compared to 54.72% of PCSeg Tool-2 and 58.88% of [24]. Precision of PCSeg Tool-1 is also good as 76.56% compared to 65.88% of PCSeg Tool-2 and 52.47% of [24]. Thus, as expected, PCSeg Tool-1 provides best F1-score of 79% compared to 59.78% of PCSeg Tool-2 and 55.49% of [24]. Thus, although PCSeg Tool-2 could not segment as many PCs as [24], it yielded more correct segmented cells compared to [24].

Overall, PCSeg Tool-1 performed best and provided better TPR, PPV, and smaller FDR. It was not only able to reduce the number of false negatives; it was also able to reduce false positives. For use in real clinical treatment, false negative rate of PCs should be low as well as false positives should be low because less than required number of chemotherapy sessions due to poor recall (or missing of large number of true PCs) or more than required number of chemotherapy sessions on aged people owing to high false discovery can prove to be fatal.

Fig 10 presents qualitative (visual) comparison of these methods on some more images. The good performance of PCSeg Tool-1 with  $k$ -means identified nuclei mask for cluster cell segmentation in the modified level set formulation implies that perhaps the proposed probability based and mean color vector based energy functionals in the multiphase level set, added with the robustness of the  $k$ -means on the level set output for nuclei mask for subsequent cluster cell segmentation, is playing a key role in the robust segmentation of PCs. Thus, we name the PCSeg Tool-1 as the final PCSeg Tool product for use with plasma cell segmentation.

## Conclusions and future work

In this paper, we designed, described, and implemented PCSeg tool for the segmentation of plasma cells from microscopic images. This tool has a robust pipeline consisting of modified multiphase level set method that utilizes statistical information about the probability densities of regions of interest (ROI) and the mean color vector of ROI in the color spaces in the multiphase level set. The level set stage removed the background and most of the unwanted cells. Only the stained single cells or the clusters of cells were retained after its application. Next, we tried two variations of PCSeg Tool: Tool-1 that utilized  $k$ -means based nuclei mask in the cluster cell segmentation and Tool-2 that utilized one of the phases of the level set for nuclei mask in cluster cell segmentation, where cluster segmentation was carried out with watershed and circular Hough transform and unwanted cells are completely removed in the post-processing stage of PCSeg Tool. PCSeg Tool-1 provided best results with better recall, precision, and F1-score. Further, the implemented PCSeg Tool-1 provided good results on segmentation of single isolated plasma cells as well as segmentation of plasma cells from cell clusters.



**Fig 10. Qualitative comparison of MM cell segmentation using different methods over five images:** (a) Gold standard (showing cells of interest with white outlines), (b) Saeedizadeh et al. [24] method, (c) PCSeg Tool-1, and (d) PCSeg Tool-2. White outlines in all figures (b)-(d) denote the outlines of regions segmented out. These regions are required to be compared with the regions contained within the white boundaries in Gold Standard shown in (a).

<https://doi.org/10.1371/journal.pone.0207908.g010>

Recently, cell segmentation with deep learning (DL) has started picking up pace. However, as of now there are only a few papers with DL on cell segmentation [30–35]. Most of these methods have dealt with nucleus segmentation and so far, there is no paper on plasma cell segmentation using deep learning. Plasma cell segmentation is a more challenging problem compared to nucleus segmentation because (i) it requires both nucleus and cytoplasm segmentation, (ii) the color contrast of cytoplasm is sometimes very near to background, and (iii) cluster segmentation is also a problem because it can include a cluster of touching nuclei, touching nucleus with cytoplasm, touching cytoplasm of different cells, etc. Hence, recently proposed DL methods cannot be directly ported on this dataset. Solving the problem of plasma cell segmentation using DL is a challenging research problem that we plan to attempt in the near

future. In addition, recent optimization/regularization based methods used in other domains of medical image segmentation similar to low rank and sparse decomposition method of [36] can also be explored in cell segmentation.

## Acknowledgments

Authors gratefully acknowledge the research funding support (Grant Number: 1(7)/2014-ME&HI) from the Ministry of Communication and IT, Govt. of India and funding support (Grant: EMR/2016/006183) from the Department of Science and Technology, Govt. of India for this research work. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We would also like to thank the support of Dr. Hossein and Mrs. Saeedizadeh [24] for providing their code for the experiments.

## Author Contributions

**Conceptualization:** Anubha Gupta, Ritu Gupta.

**Data curation:** Ritu Gupta.

**Formal analysis:** Anubha Gupta.

**Funding acquisition:** Anubha Gupta, Ritu Gupta.

**Investigation:** Anubha Gupta, Ritu Gupta.

**Methodology:** Anubha Gupta, Pramit Mallick, Ojaswa Sharma.

**Project administration:** Anubha Gupta, Ritu Gupta.

**Resources:** Anubha Gupta.

**Software:** Pramit Mallick, Rahul Duggal.

**Supervision:** Anubha Gupta, Ritu Gupta.

**Validation:** Anubha Gupta, Pramit Mallick, Ritu Gupta.

**Visualization:** Anubha Gupta, Ritu Gupta.

**Writing – original draft:** Anubha Gupta, Pramit Mallick.

**Writing – review & editing:** Anubha Gupta, Ritu Gupta.

## References

1. Meijering E. "Cell segmentation: 50 years down the road [life sciences]." *Signal Processing Magazine, IEEE*. 2012; 29(5), pp. 140–145. <https://doi.org/10.1109/MSP.2012.2204190>
2. Dorini LB, Minetto R, Leite NJ. "White blood cell segmentation using morphological operators and scale-space analysis." In: *Computer Graphics and Image Processing, 2007. SIBGRAPI 2007. XX Brazilian Symposium on. IEEE; 2007. pp. 294–304.*
3. Sharif JM, Miswan MF, Ngadi MA, Salam M, Jamil M. "Red blood cell segmentation using masking and watershed algorithm: A preliminary study." In: *Biomedical Engineering (ICoBE), 2012 International Conference on. IEEE; 2012. pp. 258–262.*
4. Gao W, Tang Y, Li X. "Segmentation of microscopic images for counting leukocytes." In: *Bioinformatics and Biomedical Engineering, 2008. 2nd International Conference on. IEEE; 2008. pp. 2609–2612.*
5. Anoraganingrum D, Kroner S, Gottfried B. "Cell segmentation with adaptive region growing." *ICIAP Venedig, Italy. 1999; pp. 27–29.*
6. Stoklasa R, Balek L, Krejci P, Matula P. "Automated cell segmentation in phase-contrast images based on classification and region growing." In: *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on. IEEE; 2015. pp. 1447–1451.*

7. Yan M, Cai J, Gao J, Luo L. "K-means cluster algorithm based on color image enhancement for cell segmentation." In: Biomedical Engineering and Informatics (BMEI), 2012 5th International Conference on. IEEE; 2012. pp. 295-299.
8. Mohapatra S, Patra D. "Automated cell nucleus segmentation and acute leukemia detection in blood microscopic images." In: Systems in Medicine and Biology (ICSMB), 2010 International Conference on. IEEE; 2010. pp. 49-54.
9. Orozco-Monteagudo M, Mihai C, Sahli H, Taboada-Crispi A. "Combined Hierarchical Watershed Segmentation and SVM Classification for Pap Smear Cell Nucleus Extraction." *Computacion y Sistemas*. 2012; 16(2), pp.133-145.
10. Yin Z, Bise R, Chen M, Kanade T. "Cell segmentation in microscopy imagery using a bag of local Bayesian classifiers." In: Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on. IEEE; 2010. pp. 125-128.
11. Kass M, Witkin A, Terzopoulos D. "Snakes: Active contour models." *International Journal of Computer Vision*. 1988; 1(4), pp.321-331. <https://doi.org/10.1007/BF00133570>
12. Sadeghian F, Seman Z, Ramli AR, Kahar BHA, Saripan MI. "A framework for white blood cell segmentation in microscopic blood images using digital image processing." *Biological procedures online*. 2009; 11(1), pp. 196. <https://doi.org/10.1007/s12575-009-9011-2> PMID: 19517206
13. Yang L, Meer P, Foran DJ. "Unsupervised segmentation based on robust estimation and color active contour models." *Information Technology in Biomedicine, IEEE Transactions on*. 2005; 9(3), pp. 475-486. <https://doi.org/10.1109/TITB.2005.847515>
14. Zamani F, Safabakhsh R. "An unsupervised GVF snake approach for white blood cell segmentation based on nucleus." In: *Signal Processing, 2006 8th International Conference on*. vol. 2. IEEE; 2006.
15. Yu W, Lee HK, Hariharan S, Bu W, Ahmed S. "Level set segmentation of cellular images based on topological dependence." In: *Advances in Visual Computing*. Springer; 2008. pp. 540-551.
16. Chan TF, Vese LA. "Active contours without edges." *Image processing, IEEE transactions on*. 2001; 10(2), pp. 266-277. <https://doi.org/10.1109/83.902291>
17. Lu Z, Carneiro G, Bradley AP. "Automated nucleus and cytoplasm segmentation of overlapping cervical cells." In: *Medical Image Computing and Computer-Assisted Intervention, MICCAI 2013*. Springer; 2013. pp. 452-460.
18. Bilkova Z et al. "Cell Segmentation Using Level Set Methods with a New Variance Term." In: *International Conference Image Analysis and Recognition*. Springer; 2016. pp. 183-190.
19. Zhang P, Wang F, Teodoro G, Liang Y, Brat D, Kong J. "Automated level set segmentation of histopathologic cells with sparse shape prior support and dynamic occlusion constraint." In: *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*. IEEE; 2017. p. 718-722.
20. Swierczynski P, Papiez BW, Schnabel JA, Macdonald C. "A level-set approach to joint image segmentation and registration with application to CT lung imaging." *Computerized Medical Imaging and Graphics*. 2018; 65(4), pp. 58-68. <https://doi.org/10.1016/j.compmedimag.2017.06.003> PMID: 28705410
21. Wang J, Cheng Y, Guo C, Wang Y, Tamura S. "Shape intensity prior level set combining probabilistic atlas and probability map constrains for automatic liver segmentation from abdominal CT images." *International journal of computer assisted radiology and surgery*. 2016; 11(5), pp. 817-826. <https://doi.org/10.1007/s11548-015-1332-9> PMID: 26646416
22. Patel A, van Ginneken B, Meijer FJ, van Dijk EJ, Prokop M, Manniesing R. "Robust cranial cavity segmentation in CT and CT perfusion images of trauma and suspected stroke patients." *Medical image analysis*. 2017; 36, pp.216-228. <https://doi.org/10.1016/j.media.2016.12.002> PMID: 28011374
23. Chen YT. "A novel approach to segmentation and measurement of medical image using level set methods. *Magnetic resonance imaging*." 2017; 39, pp.175-193. <https://doi.org/10.1016/j.mri.2017.02.008> PMID: 28219649
24. Saeedizadeh Z, Mehri Dehnavi A, Talebi A, Rabbani H, Sarrafzadeh O, Vard A., "Automatic recognition of myeloma cells in microscopic images using bottleneck algorithm, modified watershed and SVM classifier." *Journal of microscopy*. 2015; 261(1), pp. 46-56. <https://doi.org/10.1111/jmi.12314> PMID: 26457371
25. Rajkumar SV, Dimopoulos MA, Palumbo A, Blade J, Merlini G, Mateos MV, et al. "International Myeloma Working Group updated criteria for the diagnosis of multiple myeloma." *The lancet oncology*. 2014; 15(12): e538-e548. [https://doi.org/10.1016/S1470-2045\(14\)70442-5](https://doi.org/10.1016/S1470-2045(14)70442-5) PMID: 25439696
26. Gupta R, Mallick P, Duggal R, Gupta A, Sharma O. "Stain Color Normalization and Segmentation of Plasma Cells in Microscopic Images as a Prelude to Development of Computer Assisted Automated Disease Diagnostic Tool in Multiple Myeloma." *Clinical Lymphoma Myeloma and Leukemia*. 2017; 17(1, Supplement):e99. <https://doi.org/10.1016/j.clml.2017.03.178>.

27. Bhattachayya A. "On a measure of divergence between two statistical population defined by their population distributions." *Bulletin Calcutta Mathematical Society*. 1943; 35, pp. 99–109.
28. Chan TF, Sandberg BY, Vese LA. "Active contours without edges for vector-valued images." *Journal of Visual Communication and Image Representation*. 2000; 11(2), pp. 130–141. <https://doi.org/10.1006/jvci.1999.0442>
29. Vese LA, Chan TF. "A multiphase level set framework for image segmentation using the Mumford and Shah model." *International journal of computer vision*. 2002; 50(3), pp. 271–293. <https://doi.org/10.1023/A:1020874308076>
30. Duggal R, Gupta A, Gupta R, Wadhwa M, Ahuja C. "Overlapping cell nuclei segmentation in microscopic images using deep belief networks." In: *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing*. ACM; 2016, p.82.
31. Song Y, Tan EL, Jiang X, Cheng JZ, Ni D, Chen S, et al. "Accurate cervical cell segmentation from overlapping clumps in pap smear images. *IEEE transactions on medical imaging*. 2017; 36(1), pp. 288–300. <https://doi.org/10.1109/TMI.2016.2606380> PMID: 27623573
32. Xing F, Xie Y, Yang L. "An automatic learning-based framework for robust nucleus segmentation." *IEEE transactions on medical imaging*. 2016; 35(2), pp. 550–566. <https://doi.org/10.1109/TMI.2015.2481436> PMID: 26415167
33. Akram SU, Kannala J, Eklund L, Heikkila J. "Cell segmentation proposal network for microscopy image analysis." In: *Deep Learning and Data Labeling for Medical Applications*. Springer; 2016. pp. 21–29.
34. Zhang L, Sonka M, Lu L, Summers RM, Yao J. "Combining fully convolutional networks and graph-based approach for automated segmentation of cervical cell nuclei." In: *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*. IEEE; 2017. pp. 406–409.
35. Hatipoglu N, Bilgin G. "Cell segmentation in histopathological images with deep learning algorithms by utilizing spatial relationships." *Medical and biological engineering and computing*. 2017; 55(10), pp.1829–1848. <https://doi.org/10.1007/s11517-017-1630-1> PMID: 28247185
36. Shi C, Cheng Y, Wang J, Wang Y, Mori K, Tamura S. "Low-rank and sparse decomposition based shape model and probabilistic atlas for automatic pathological organ segmentation." *Medical image analysis*. 2017; 38, pp.30–49. <https://doi.org/10.1016/j.media.2017.02.008> PMID: 28279915
37. Gupta A. "MiMM\_SBI Lab Dataset: Microscopic Images of Multiple Myeloma", <https://doi.org/10.7910/DVN/XCX7ST>, Harvard Dataverse, V1.