Pattern Recognition 58 (2016) 1–11

Contents lists available at ScienceDirect

Pattern Recognition

journal homepage: www.elsevier.com/locate/pr

# Segmentation of cell nuclei in fluorescence microscopy images: An integrated framework using level set segmentation and touching-cell splitting

Amin Gharipour *, Alan Wee-Chung Liew

*School of Information and Communication Technology, Gold Coast Campus, Griffith University, QLD4222, Australia*

ARTICLE INFO

*Article history:*
Received 21 July 2015
Received in revised form
30 January 2016
Accepted 28 March 2016
Available online 6 April 2016

*Keywords:*
Fluorescence microscopy images
Level set segmentation
Touching-cell splitting

ABSTRACT

Accurate segmentation of cells in fluorescence microscopy images plays a key role in high-throughput applications such as quantification of protein expression and the study of cell function. In this paper, an integrated framework consisting of a new level sets based segmentation algorithm and a touching-cell splitting method is proposed. For cell nuclei segmentation, a new region-based active contour model in a variational level set formulation is developed where our new level set energy functional minimizes the Bayesian classification risk. For touching-cell splitting, the touching cells are first distinguished from non-touching cells, and then a strategy based on the splitting area identification is proposed to obtain splitting point-pairs. To form the appropriate splitting line, the image properties from different information channels are used to define the surface manifold of the image patch around the selected splitting point-pairs and geodesic distance is used to measure the length of the shortest path on the manifold connecting the two splitting points. The performance of the proposed framework is evaluated using a large number of fluorescence microscopy images from four datasets with different cell types. A quantitative comparison is also performed with several existing segmentation approaches.



## 1. Introduction

Numerous areas of analysing and quantifying fluorescence microscopy images rely on quantitative cell nucleus image analysis [1–3]. Specifically, the basis for all automatic image analysis required in high-throughput applications is cell image segmentation. Semi-automatic and manual segmentation methods are tedious, need intensive labor, and suffer from inter-and intra-reader variability. Therefore automatic methods with the ability to deal with different cell types and image artifacts are required.

Specifically, this paper has two main goals. The first goal is motivated by the fact that in cell nuclei image segmentation, pixels in an image patch possess nearly the same intensity. Therefore, the spatial relationship of the pixels in an image patch can be utilized as an important characteristic that improves the performance of level set segmentation methods [4–7]. Consequently, the first goal is to develop a segmentation algorithm based on the image patch information [8–11]. The segmentation algorithms often fail to separate the individual cells which form clumps. Since quantitative cell nucleus image analysis is dependent on the characteristics of each individual cell, the overlapping of cells could have an adverse effect on the performance of the quantitative high-throughput automated image analysis. Thus, the second goal is to develop a splitting algorithm for touching cells.

The contributions of this paper are as follows. 1) A new approach is introduced for cell nucleus segmentation in fluorescence microscopy images. First, a region-based active contour model in a variational level set formulation, which is based on the image patch information, is used to segment the image. Compared to previous approaches, we define a novel local energy functional based on the Bayesian classification risk [12–14] for an image patch. In addition, a weighting scheme is used to enable the pixels in each image patch to have anisotropic weights. 2) A three-step touching-cell splitting algorithm is utilized for splitting. In the first step, morphological features, and the distance between the most likely radial-symmetry point and the geometrical center, are utilized to distinguish touching-cell clumps from non-touching cells. After touching-cell identification, splitting areas are identified, and pairs of splitting points are selected for each clump. Once the splitting points are recognized, the image patch around splitting points is defined as their joint neighborhood. The image properties from different information channels are used to define the surface manifold of the image patch, and geodesic distance is used to

* Corresponding author.
*E-mail addresses:* amin.gharipour@griffithuni.edu.au (A. Gharipour), a.liew@griffith.edu.au (A. Liew).

http://dx.doi.org/10.1016/j.patcog.2016.03.030
0031-3203/

measure the length of the shortest path on the manifold connecting the two splitting points. Finally, Dijkstra's algorithm [15] is used to find the weighted shortest path between splitting points.

The remainder of this paper is organized as follows. Section 2 introduces the proposed approaches. In Section 3, experimental results are presented and analysed using different cell types. The results of the proposed methods are also compared with previous approaches. This paper is summarized in Section 4.

## 2. Proposed approach

### 2.1. Background

Recently, many methods have been proposed for the segmentation of cell nuclei in fluorescence microscopy images. One of the most common approaches for cell segmentation is intensity thresholding which suffers from intensity inhomogeneity. Watershed-based methods are commonly utilized for clustered nuclei separation. A cell segmentation method based on the watershed algorithm and the cluster splitting technique has been introduced in [1]. A method based on the improved distance transform and statistical model-based merging has been proposed in [2]. Marker-controlled watershed techniques are proposed by Cheng and Rajapakse [3] as well as Jung and Kim [16] for cluster splitting, where the H-minima transform is used to find the optimal number of markers.

Deformable models, which are able to capture a wide spectrum of different shapes, can be considered as a major category of cell segmentation techniques. There are two main types of deformable models: parametric models [17], which use an explicit representation of objects [18–20], and implicit models [21]. A region based parametric energy functional is introduced in [18]. Their proposed model utilizes a coupling term for multiple contours as well as a penalty term to prevent merging. In [19], the authors used texture-adaptive weights to overcome the internal pseudo-edges and low-contrast cell boundaries problems of Zimmer and Olivo-Marin's model [18]. Butenuth and Heipke proposed a method for image segmentation which was based on the parametric active contour model and a graph-based approach [20]. Although parametric active contour models can be considered as one of the common approaches for image segmentation, these models depend on the parameterization and are not able to deal with topological changes.

Implicit models [6,7,22–42] using level sets have been widely used in cell segmentation with promising results. The implicit models are able to handle topological changes, which are typically not possible in parametric models. A two-step level set method has been developed for cell segmentation in [29]. An energy functional which is based on the multiple active contours integration as well as a combination of gradient based and region-based terms has been introduced in [30]. In [13], Voronoi tessellation has been utilized to determine regions corresponding to single cells and then for each of these regions, region based energy functional is used. In [14], the flux tensor is used for initialization of their proposed level set method which was based on Bayesian energy functional. The cell and the background of cell images are separated using a gradient-based level set approach and then topology preserving level sets are employed to perform cluster splitting in [16]. A level set method based on Bayesian energy functional which uses a non-PDE-based minimization is developed in [39]. For cell image segmentation, the level set methods and the graph partitioning approaches have been combined within in a variational framework in [38,40]. Sequential integrations of fuzzy clustering and implicit models have been proposed in [43–45]. A multi-phase graph partitioning active contour approach which uses regional density functions has been developed in [38]. A two-step level set method for histopathological images has been proposed in [40] where first, a geodesic active contour model is initialized using a hierarchical normalized cuts scheme and then a level set functional is used. An approach has been proposed in [46] for segmenting cell nuclei s based on active contours using level sets and convex energy functionals.

### 2.2. Local level set method based on the Bayesian risk and weighted image patch (LLBWIP)

#### 2.2.1. Modeling

Assuming that an image is formed by two regions, cell and background pixels, the following two hypotheses can be used to characterize the image segmentation.

- A null hypothesis $H_1$, in which the cell is absent.
- An alternative hypothesis $H_2$, in which the cell is present.

The segmentation method is utilized to decide which hypothesis is correct. Therefore, one of two decisions can be made:

- $D_1$: the classifier declares that the cell is absent.
- $D_2$: the cell is present, and thus should be chosen by segmentation procedure.

The following four conditional probabilities are defined for the combinations of decisions in the hypothesis test:

(1) $P(D_1|H_1)$ is the probability of declaration that the cell is absent when it is actually absent.
(2) $P(D_2|H_1)$ is the probability of declaration that the cell is present when it is absent.
(3) $P(D_1|H_2)$ is the probability of declaration that the cell is absent when it is present.
(4) $P(D_2|H_2)$ is the probability of declaration that the cell is present when it is actually present.

Using the statistical terminology, the first probability, which is the probability of rejecting the null hypothesis $H_1$ when it is actually true, is called *type I* risk. On the other hand, $P(D_1|H_2)$, which can be considered the probability of accepting $H_1$ when $H_1$ is actually false, is called *type II* risk. The consequence of each combination of hypothesis and decision is quantified with an associated loss. The losses of $P(D_1|H_1)$, $P(D_2|H_1)$, $P(D_1|H_2)$, and $P(D_2|H_2)$ can be denoted as $L(1,1)$, $L(2,1)$, $L(1,2)$, and $L(2,2)$, respectively. $L(1,1)$ and $L(2,2)$ can be viewed as the losses arising from the correct decision while $L(2,1)$ and $L(1,2)$ denote the losses that arise from the incorrect decision. As the general rule, we set $L(1,1)=L(2,2)=0$ and $L(2,1)=L(1,2)=1$ [12]. Now the Bayesian risk for segmenting an image into cell and background can be written as follows:

$$r = P(H_1)P(D_2|H_1)+P(H_2)P(D_1|H_2) \quad (1)$$

Assuming $\Omega=\Omega_{i=1}^{2}$ denotes the image domain, where $\Omega_1$ and $\Omega_2$ denote cell and background pixels, respectively. $I:\Omega\rightarrow\mathbb{R}^{+}$ denotes a given image, for each point $\mathbf{x}$ in the image domain $\Omega$, the image patch centered on $\mathbf{x}$ can be represented as:

$$P_{\mathbf{x}} = (I(\mathbf{y}), \mathbf{y}\in N_{\mathbf{x}}) \quad (2)$$

where $N_{\mathbf{x}}$ can be considered as a $q\times q$ neighbourhood of point $\mathbf{x}$. Now the image patch $P_{\mathbf{x}}$ with domain $N_{\mathbf{x}}$ can be partitioned by $\Omega_{i=1}^{2}$ into the following disjoint regions:

$R_1=\{\Omega_1\cap N_{\mathbf{x}}\}$ and $R_2=\{\Omega_2\cap N_{\mathbf{x}}\}$, where $N_{\mathbf{x}}=\cup_{i=1}^{2} R_i, R_i\cap R_j=\varnothing\,\forall i\neq j$.

$D_1$ and $D_2$ can now be redefined as follows; all pixels of the image patch $P_{\mathbf{x}}$ that lead the segmentation procedure to choose