Automated Nucleus and Cytoplasm Segmentation of Overlapping Cervical Cells

Zhi Lu[1,*], Gustavo Carneiro[2], and Andrew P. Bradley[3,**]

[1] Department of Computer Science, City University of Hong Kong, China
[2] ACVT, The University of Adelaide, Australia
[3] School of Information Technology & Electrical Engineering,
The University of Queensland, Australia

Abstract. In this paper we describe an algorithm for accurately segmenting the individual cytoplasm and nuclei from a clump of overlapping cervical cells. Current methods cannot undertake such a complete segmentation due to the challenges involved in delineating cells with severe overlap and poor contrast. Our approach initially performs a scene segmentation to highlight both free-lying cells, cell clumps and their nuclei. Then cell segmentation is performed using a joint level set optimization on all detected nuclei and cytoplasm pairs. This optimisation is constrained by the length and area of each cell, a prior on cell shape, the amount of cell overlap and the expected gray values within the overlapping regions. We present quantitative nuclei detection and cell segmentation results on a database of synthetically overlapped cell images constructed from real images of free-lying cervical cells. We also perform a qualitative assessment of complete fields of view containing multiple cells and cell clumps.

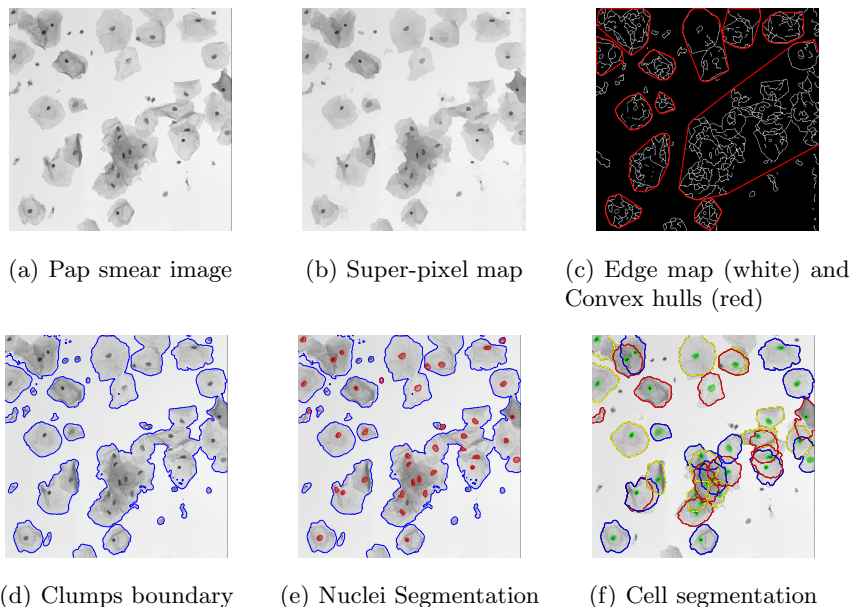**Keywords:** Overlapping cell segmentation, Pap smear image analysis.

1 Introduction

The Pap smear is a screening test used to detect pre-cancerous changes in a sample of cells from the uterine cervix and deposited onto a microscope slide for visual examination. The automated segmentation of overlapping cells in Pap smears remains one of the most challenging problems in image analysis. The main factors affecting the sensitivity of the Pap smear test are the number and type of cells sampled and the presence of mucus, blood and inflammatory cells [2], which affects both the intra- and inter-observer variability and leads to a large variation in false negative rates. [8]. These issues have motivated the development of both automated cell deposition and slide analysis techniques. The advantages of automated cell deposition techniques, such as mono-layer preparations, are that they remove a large portion of blood, mucus, and other

---

(a) Pap smear image

(b) Super-pixel map

(c) Edge map (white) and Convex hulls (red)

(d) Clumps boundary

(e) Nuclei Segmentation

(f) Cell segmentation

Fig. 1. (a) Typical Pap smear image; **(b)** Over-segmented super-pixel map is generated by Quick Shift; **(c)** super-pixel edge map and the convex hull of each clump; **(d)** Accurate clump boundary; **(e)** Nuclei detection and segmentation;**(f)** Overlapping cell segmentation

debris, reduce cell overlap and produce a mono-layer of easily focused cells. Automated slide analysis techniques attempt to improve both sensitivity and specificity by detecting, segmenting and then classifying all of the cells present on a slide [9,17,5,4,3].

The classification of cervical cells is typically based on features, such as shape and area, extracted from the cytoplasm and nucleus of individual cells. This means that an accurate segmentation is critical to automated screening methods. Nevertheless, segmentation is complicated by the fact that they often form overlapping clumps on the glass slide, which means that cells in an upper layer can partially obscure cells lying underneath [10]. Current systems can segment the nucleus and cytoplasm of isolated cervical cells [5] (i.e., free-lying cells without any overlap with other cells); segment overlapping nuclei [3,10]; and segment nuclei and the whole region representing overlapping cytoplasm [4,2]. However, currently segmenting both the individual cytoplasm and nuclei from overlapping cells in Pap smear images is still a challenging problem due to the presence of overlapping cells and the poor contrast of cytoplasm [2] (see Fig. 1(a)).

In this paper, we propose a novel algorithm to segment both the cytoplasm and nucleus of each overlapping cell. The proposed methodology can be divided into two steps. First *scene segmentation*, comprising two sub-stages: 1) cell clump detection using unsupervised classification [14] and 2) nuclei detection using the

maximally stable extremal regions (MSER) algorithm [7]. Second, *cell segmentation*, which is the main contribution of this paper, assumes that each nucleus detected in step 1 represents an individual cell that forms a separate level set function. The optimization that generates the final overlapping cell segmentation minimizes an energy function that is constrained by the individual contour area and length [6], an ellipsoidal shape prior [13], and finally the area [18] and gray values within the overlapping regions. We quantitatively evaluate our methodology using a database of 18 synthetically constructed images consisting 60 real free-lying cell images. The results on this dataset show that we obtain a Jaccard index > 0.8 with a near zero false negative rate. We also show encouraging qualitative results using a small set of real Pap smear images. Furthermore, we demonstrate that the performance of our nuclei detection produces state of the art results [2].[1]

2   Literature Review

There are three primary approaches to cervical cell segmentation. Traditional methods segment nuclei from single or overlapping cells. For instance, Wu et al. [16] detect the boundary of nuclei by solving an optimal thresholding problem. Morphological analysis is also used to detect overlapping nuclei from cervical cell images [9]. Other approaches segment the nucleus and cytoplasm of single isolated cervical cells. Yang-Mao et al. [17] adapt the gradient vector flow (GVF) to cervical cell segmentation by estimating the orientation of the GVFs in the pixels near an edge. The GVF is also explored in the detection of nucleus and cytoplasm boundaries in a radiating field [5]. This approach produces competitive results, but can analyse only free-lying cells (and thus, a fraction of the cells present on a specimen). Finally, the segmentation of overlapping nuclei can be performed by identifying the nucleus and cytoplasm from candidate regions via a classification algorithm [4,2]. However, instead of accurate boundaries for each overlapping cell, these methods generate a contour of the whole cell clump, which reduces the number of features available for subsequent cellular analysis, i.e., each cell has its own nucleus, but not its own cytoplasm. In the above, the detection of overlapping nuclei is facilitated by their homogeneous texture, ellipsoidal shape and high-gradient boundaries. Unfortunately, none of these characteristics can be associated with the segmentation of overlapping cytoplasm. Furthermore, some of the important features extracted from nuclei are based on optical density and texture, which may be contaminated when nuclei overlap.

The segmentation of overlapping cells has been explored in other types of microscope images. For instance, Wahlby et al. [15] use a watershed segmentation algorithm and a statistical analysis to segment multiple CHO-cells stained with calcein. Furthermore, Quelhas et al. [11] explores a sliding band filter to segment nuclei and cytoplasm of overlapping cells using the Drosophila melanogaster Kc167 dataset. Although relevant to our paper, these works are applied to images

---

[1] Our code & data set are available at GitHub: `github.com/luzhi/miccai2013.git`

where the cells present smaller overlapping areas, when compared to the case of Pap smear images. Therefore, we believe that these approaches would need significant adaptation in order to segment cervical cell images.

3   Methods

3.1   Extended Depth of Field

A one-pass extended depth of field (EDF) algorithm [1], based on an overcomplete discrete wavelet transform, was applied to a "stack" of focal plane images to produce a single EDF image where all cellular objects are in focus. The advantage of this approach is that the scene segmentation, described below, need only be applied to a single (EDF) image rather than a set of images from different focal planes.

3.2   Scene Segmentation

Scene segmentation consists of two stages: 1) the segmentation of cell clumps, which facilitates 2) the detection and segmentation of nuclei.

Segmentation of Cell Clumps. The segmentation of cell clumps facilitates the detection of nuclei by constraining the search space to clumps only. For instance, in the image of Fig. 1(a), the goal is to segment the six free-lying cells and the set of overlapping cells in different clumps. This segmentation process involves three stages. First, we run the quick shift algorithm [14] in order to find local maxima of a density function that takes into account gray value similarities and spatial proximity. The outcome of this step is a map of super-pixels (Fig. 1(b)), which are labeled with gray values in the range $[0, 1]$, representing the mode of the respective super-pixel. The second stage consists of running an edge detector on this super-pixel map, resulting in a clean edge map that detects the most prominent super-pixel edges and removes most of the background information (Fig. 1(c)). In order to find candidate cell clumps, the third stage utilises an unsupervised binary classifier, where the classes are "background" and "cell clump". The initial assignment is provided by building a convex hull around the connected components of the edge map computed in stage 2 (Fig. 1(c)). Hence, pixels inside the convex hull initially belong to the cell clump class and pixels outside belong to the background. Using maximum likelihood estimation, we learn a Gaussian mixture model (GMM) for each class, based on the gray value of each pixel (Fig. 1(d)).

Detection and Segmentation of Nuclei. This is a critical step of our algorithm because each nucleus represents one cell. Nuclei can be characterized by relatively low gray values, homogeneous texture, and well defined, almost circular borders. If we assume that the nuclei do not overlap, then we can use the Maximally Stable Extremal Regions (MSER) algorithm [7] using the cell clumps as the input. The MSER algorithm uses pixel gray value and proximity to detect

stable connected components, which are characterized by blobs that represent the candidate nuclei. We filter out some of these candidates if the eccentricity of the blob detected is larger than a threshold (i.e., keeping only the most circular blobs).

3.3   Joint Level Set Segmentation of Overlapping Cells

The joint level set optimisation presented in this section is the main contribution of this paper. The segmentation of overlapping cells uses the set of nuclei described in Sec. 3.2 as the initial guess for each level set function. Consider that $\phi : \Omega \to \mathbb{R}$ denotes a level set function (LSF) (Ω represents the image domain), and that N nuclei have been detected, then the set of LSF's are available is denoted by $\{\phi_i\}_{i=1}^{N}$. The energy functional to be minimized is defined as:

$$\mathcal{E}(\{\phi_i\}_{i=1}^{N}) = \sum_{i=1}^{N} \mathcal{E}_u(\phi_i) + \sum_{i=1}^{N} \sum_{j \in \mathcal{N}(i)} \mathcal{E}_b(\phi_i, \phi_j), \tag{1}$$

where $\mathcal{E}_u(.)$ denotes the unary energy functional defined for each LSF independently, $\mathcal{E}_b(.,.)$ represents the binary function defined over pairs of LSF's, and $\mathcal{N}(i)$ represents the level set functions ϕ_j such that their zero level set intersects the zero level set of ϕ_i. The unary functional is defined by:

$$\mathcal{E}_u(\phi_i) = \mu \mathcal{R}(\phi_i) + \lambda \mathcal{L}(\phi_i) + \alpha \mathcal{A}(\phi_i) + \rho \mathcal{P}_p(\phi_i), \tag{2}$$

where $\mu, \lambda > 0$, $\alpha, \rho \in \mathbb{R}$, $\mathcal{R}(\phi)$ is a regularization term [6] that maintains the signed distance property $|\bigtriangledown \phi_i| = 1$ (guaranteeing that the LSF is smooth), $\mathcal{L}(\phi_i) = \int_{\Omega} g\delta(\phi_i)|\bigtriangledown \phi_i|d\mathbf{x}$ measures the length of the zero level set (with $\delta(.)$ denoting the Dirac delta function, $g = \frac{1}{1+|\nabla G_\sigma * I|}$, and G_σ the Gaussian kernel with standard deviation σ), $\mathcal{A}(\phi_i) = \int_{\Omega} gH(-\phi_i)d\mathbf{x}$ measures the area of $\phi_i < 0$ (H(.) is the Heaviside function), and the shape prior term is defined by [13]:

$$\mathcal{P}_p(\phi_i) = \int_{\Omega} gH(-p(\phi_i))d\mathbf{x}, \tag{3}$$

where $p(\phi_i)$ returns an ellipsoidal shaped LSF estimated from the covariance matrix of the coordinates $\mathbf{x}$ for which $\phi_i(\mathbf{x}) < 0$. The binary functional in (1) is defined as:

$$\begin{aligned} \mathcal{E}_b(\phi_i, \phi_j) = & \zeta f_a \left(\frac{\int_{\Omega} gH(-\phi_i)H(-\phi_j)d\mathbf{x}}{\int_{\Omega} gH(-\phi_i)d\mathbf{x}} \right) + \\ & \omega f_g \left(\frac{\int_{\Omega} vH(-\phi_i)d\mathbf{x}}{\int_{\Omega} gH(-\phi_i)d\mathbf{x}} - \frac{\int_{\Omega} vH(-\phi_i)H(-\phi_j)d\mathbf{x}}{\int_{\Omega} gH(-\phi_i)H(-\phi_j)d\mathbf{x}} \right), \end{aligned} \tag{4}$$

where the first term in (4) denotes the ratio between the areas of $(\phi_i < 0) \bigcap (\phi_j < 0)$ and $\phi_i < 0$ (with $f_a(y)$ defined as y for $y > \tau_a$ and 0 otherwise, where τ_a is the maximum overlapping area accepted by the optimisation) [18], and the

second term represents the difference between the average gray value in $\phi_i < 0$ subtracted by the average gray value in the intersection $(\phi_i < 0) \bigcap (\phi_j < 0)$ (with $f_g(y)$ defined as y for $y < 0$ and 0 otherwise, and v returns the gray value of position $\mathbf{x}$). The minimization of the energy functional in (1) follows that defined by $\frac{\partial \phi}{\partial t} = -\frac{\partial \mathcal{E}(\{\phi_i\}_{i=1}^N)}{\partial \phi}$. The derivation of these functions is presented in the supplemental material.[2]

4   Material and Experimental Setting

Our dataset consists of four Pap smear images with extended depth of field (EDF) [1]. Four non-overlapping fields of view (FOVs) were captured from a single specimen consisting of Papanicolaou-stained cervical cells, prepared using the AutoCyte PREP technology. The specimen was approximately $20\mu m$ "thick" in the focal-dimension. The numerical aperture of the microscope's $\times 40$ objective was 0.75, which gives a depth of field of approximately $1\mu m$. Therefore, for each FOV, a stack of twenty focal plane images were acquired with a separation of $1\mu m$. In these images, all 135 nuclei were manually annotated, with the 13 free-lying cells being fully annotated with nucleus and cytoplasm delineations.

We assess our methodology both qualitatively and quantitatively. The qualitative experiment consists of a visual inspection of the result of our algorithm using the four EDF images. For the quantitative assessment, we generated 18 synthetic Pap smear images (of size 512×512) containing 2 to 5 cells with different degrees of overlap. In order to generate these synthetic images, we divide the aforementioned set of 13 manually annotated cells, into: 1) a training set of 5 cells that is used to build 3 synthetic training images, and 2) a test set of 8 cells to form 15 test images[3]. In order to build these synthetic images, we take samples from the background of the original EDF images, and place them in the 512×512 image, using mirror transformation to smooth the transitions between the background patch borders. Then, we pick one of the cells from the training/test set, apply a random rigid transform and random linear brightness transform, and place them on the synthetic image, using a random value (from 0 to 1) for the alpha channel to simulate the partial transparency effect observed in real Pap smear images. Note that the training images are used to assess the best combination of the level set parameters in (1), while the testing images are used to verify the generalization ability of the method with respect to the level set parameters found on the training images. Finally, this quantitative performance is assessed with the average Jaccard index (JI) computed over the "good" cell segmentations [12], where the cell segmentation has a JI above a threshold of {0.5, 0.6, 0.7, 0.8}. We also report the object based false negative rate (FNR) obtained as the proportion of cells having a JI below this threshold. In addition, pixel based evaluation through true positive rate (TPR) and false positive rate (FPR) for both training and test set are also shown.

---

[2] Appendix at `www.cs.cityu.edu.hk/~luzhi/publications/app_MICCAI13.pdf`
[3] Early experiments showed results stabilise when at least 13 test images are used.

We also compare our nuclei detection methodology from Pap smear images with that in [2] using the following measurements. First, we compute the precision and recall of nuclei detection by considering the detection region A and annotation B, and noting that a correct detection is defined by $(A \bigcap B)/A > 0.6$ and $(A \bigcap B)/B > 0.6$. Second, we compute the pixel-based precision and recall values of the correct detections above, in addition to the Dice coefficient.

5   Experimental Results

We first show the influence of the parameters λ, α, ρ in the unary functional (2), and ζ and ω in the binary functional (4). The parameter μ is fixed at 0.2 per time step [6], where time step is 5. The JI and corresponding FNR for each parameter combination and the best parameters combination for training are shown in Table 1. This table shows the results on the test set using the best parameter combination obtained on the training set. Moreover, in Table 2 we show the pixel-based TPR and FPR on the training and test sets for the "good" segmentations. In Fig. 2, we show examples of the synthetic and real Pap smear image segmentations.

Table 1. Qualitative evaluation to show training process and test results in terms of the JI for "good" segmentations

| Training set				
α λ ρ ζ ω	JI> 0.5	JI> 0.6	JI> 0.7	JI> 0.8
-5 4 0.15 3 4.1	0.91 (FNR=0)	0.91 (FNR=0)	0.91 (FNR=0)	0.93 (FNR=0.08)
-5 0 0 0 0	0.87 (FNR=0)	0.87 (FNR=0)	0.87 (FNR=0)	0.92 (FNR=0.3)
0 4 0 0 0	0.93 (FNR=0.67)	0.93 (FNR=0.67)	0.93 (FNR=0.67)	0.93 (FNR=0.67)
0 0 0.15 0 0	0.77 (FNR=0.25)	0.83 (FNR=0.42)	0.86 (FNR=0.5)	0.93 (FNR=0.67)
0 0 0 3 0	0.91 (FNR=0)	0.91 (FNR=0)	0.91 (FNR=0)	0.92 (FNR=0.08)
0 0 0 0 4.1	0.91 (FNR=0)	0.91 (FNR=0)	0.91 (FNR=0)	0.92 (FNR=0.08)
Test set				
-5 4 0.15 3 4.1	0.83 (FNR=0.02)	0.85(FNR=0.09)	0.88(FNR=0.21)	0.91(FNR=0.34)

We compare our nuclei detection with the approach by Aksoy et al. [2], but this comparison is not ideal as they are run on different data sets[4]. Specifically, Aksoy et al.'s approach is tested on Hacettepe data set (which has 139 nuclei), while ours are the Pap smear images described above (with 135 nuclei). In terms of object-based nuclei detection, we achieve a precision of .69 and recall .90, while [2] has precision of .74 and recall .93. Furthermore, the pixel-based result for our method consists of a precision of .97($\pm$.04), recall .88($\pm$.08) and Dice .92 ($\pm$.04); while [2] has precision of .91($\pm$.08), recall .88($\pm$.07) and Dice .89 ($\pm$.04).

Finally, our algorithm shows an average running time of 56 seconds per cell through the test set using an unoptimized Matlab code on a PC with 2.7GHz Intel Xeon processor and 128 GB RAM.

---

[4] As the database in [2] is not publicly available.

Table 2. Pixel-based TPR/FPR on training & test sets for "good" segmentations

α λ ρ ζ ω	JI> 0.5	JI> 0.6	JI> 0.7	JI> 0.8
Training set (TPR/FPR)				
-5 4 0.15 3 4.1	0.92/0.0005	0.92/0.0005	0.92/0.0005	0.94/0.0005
-5 0 0 0 0	0.94/0.0038	0.94/0.0038	0.94/0.0038	0.94/0.0013
0 4 0 0 0	0.93/0.0001	0.93/0.0001	0.93/0.0001	0.93/0.0001
0 0 0.15 0 0	0.77/0.0001	0.84/0.0001	0.87/0.0001	0.93/0.0001
0 0 0 3 0	0.93/0.0011	0.93/0.0011	0.93/0.0011	0.94/0.0011
0 0 0 0 4.1	0.93/0.0010	0.93/0.0010	0.93/0.0010	0.94/0.0009
Test set (TPR/FPR)				
-5 4 0.15 3 4.1	0.88/0.0032	0.89/0.0025	0.92/0.0023	0.93/0.0017

(a) Synthetic images   (b) Real Pap smear images
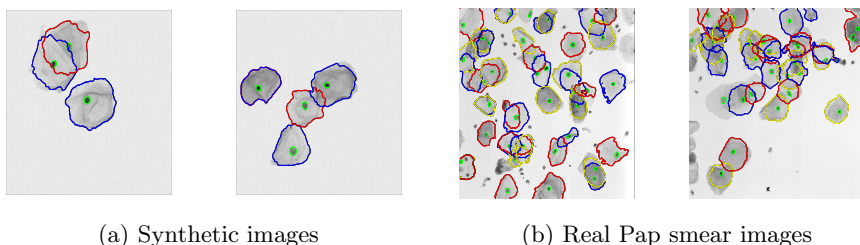
Fig. 2. Qualitative assessment of results

6   Discussion and Conclusion

In the results, we can see that on the test set our methodology produces a JI = 0.83 with a nearly zero FNR, which in general can be considered to be an acceptable result, if the "good" cell segmentations has the minimal JI > 0.5. Indeed, the qualitative results in Fig. 2 demonstrate that for all the isolated cells and most of the overlapping cells, our approach successfully segments the individual cytoplasm and nuclei. According to the results in Table 1, it is possible to conclude that both binary terms have a strong positive influence on the results. The main failures observed are caused by the weak transitions between the cytoplasm and background that make the segmentation border move either outside or inside the cytoplasm. Moreover, our nuclei detection produces quantitative results on par with the state of the art [2].

The methodology proposed in this paper provides reasonably accurate results on the challenging problem of segmenting both nuclei and cytoplasm from overlapping cervical cells. Nevertheless, there are a few points that need to be addressed to improve the effectiveness of our approach. The most important issue is that the system still misses the boundary of the cytoplasm on some overlapping cells in regions of poor contrast. Also, to obtain more robust results the system should be extended to process the original stack of multi-focal plane images (see Sec. 4) as per manual analysis.

References

1. Bradley, A., Bamford, P.: A one-pass extended depth of field algorithm based on the over-complete discrete wavelet transform. In: IVCNZ 2004, pp. 279–284 (2004)
2. Gençtav, A., Aksoy, S., Önde, S.: Unsupervised segmentation and classification of cervical cell images. Pat. Recognition 45, 4151–4168 (2012)
3. Jung, C., Kim, C., Chae, S., Oh, S.: Unsupervised segmentation of overlapped nuclei using bayesian classification. IEEE TBE 57(12), 2825–2832 (2010)
4. Kale, A., Aksoy, S.: Segmentation of cervical cell images. In: ICPR (2010)
5. Li, K., Lu, Z., Liu, W., Yin, J.: Cytoplasm and nucleus segmentation in cervical smear images using Radiating GVF Snake. Pat. Recognition 45, 1255–1264 (2012)
6. Li, C., Xu, C., Gui, C., Fox, M.: Distance regularized level set evolution and its application to image segmentation. IEEE TIP 19(12), 3243–3254 (2010)
7. Matas, J., et al.: Robust wide baseline stereo from maximally stable extremal regions. In: Proc. BMVC, pp. 384–396 (2002)
8. Noorani, H.: Assessment of techniques for cervical cancer screening. CCOHTA 1997: 2E, Canadian Coordinating Office for Health Technology Assessment (1997)
9. Plissiti, M., et al.: Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering. IEEE TITB 15, 233–241 (2011)
10. Plissiti, M., Nikou, C.: Overlapping cell nuclei segmentation using a spatially adaptive active physical model. IEEE TIP 21(11), 4568–4580 (2012)
11. Quelhas, P., et al.: Cell nuclei and cytoplasm joint segmentation using the sliding band filter. IEEE TMI 29(8), 1463–1473 (2010)
12. Radau, P.: Evaluation framework for algorithms segmenting short axis cardiac MRI. The MIDAS J. - Cardiac MR Left Ventricle Segmentation Challenge (2009)
13. Rousson, M., Paragios, N.: Shape priors for level set representations. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002, Part II. LNCS, vol. 2351, pp. 78–92. Springer, Heidelberg (2002)
14. Vedaldi, A., Soatto, S.: Quick Shift and Kernel Methods for Mode Seeking. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) ECCV 2008, Part IV. LNCS, vol. 5305, pp. 705–718. Springer, Heidelberg (2008)
15. Wahlby, C., et al.: Algorithms for cytoplasm segmentation of fluorescence labelled cells. Analytical Cellular Pathology 24(3), 101–111 (2002)
16. Wu, H.-S., Gil, J., Barba, J.: Optimal segmentation of cell images. IEE Proc. on Vision, Image and Signal Processing. 145, 50–56 (1998)
17. Yang-Mao, S.-F., et al.: Edge enhancement nucleus and cytoplast contour detector of cervical smear images. IEEE TSMC, Part B: Cybernetics 38, 353–366 (2008)
18. Zimmer, C., Olivo-Marin, J.-C.: Coupled parametric active contours. IEEE TPAMI 27(1), 1838–1842 (2005)