Spatio-temporal Level-Set Based Cell Segmentation in Time-Lapse Image Sequences

Fatima Boukari and Sokratis Makrogiannis

Mathematical Sciences Department, Delaware State University, Dover, DE 19901, USA
{fboukari11,smakrogiannis}@desu.edu

Abstract. Automated segmentation and tracking of cells in time-lapse imaging is a process of fundamental significance in several biomedical applications. In this work our interest is focused on cell segmentation over a set of fluorescence microscopy images with varying levels of difficulty with respect to cell density, resolution, contrast, and signal-to-noise ratio. We utilize a region-based approach to curve evolution based on the level-set formulation. We introduce and test the use of temporal linking for level-set initialization to improve the robustness and computational time of level-set convergence. We validate our segmentation approach against manually segmented images provided by the Cell Tracking Challenge consortium. Our method produces encouraging segmentation results with an average DICE score of 0.78 over a variety of simulated and real sequences and speeds up the convergence rate by an average factor of 10.2.

1 Introduction

The identification, quantification and characterization of cells using imaging techniques are being systematically integrated in cell biology studies [1]. Recent developments in time-lapse microscopy enable the observation and quantification of cell-cycle progression of individual cells [2]. The tasks of detecting and following individual particles in a time series of images are key elements in this process. Nevertheless, the large volume of data produced by fluorescence microscopy emphasizes the need for automated and robust techniques that can address the challenges presented by this imaging modality.

Cell tracking methodologies involve the tasks of preprocessing, cell segmentation and tracking [3], [4], [5], [6]. In this context, segmentation of cells is a particularly challenging task that has a direct impact on the overall quantification process. Image segmentation is a popular field in the domain of image analysis. More specifically, parametric [7] and nonparametric active contour models [8], [9] have been widely used in development of bioimaging and biomedical image analysis techniques.

An interesting aspect in cell analysis methods is the relation between image quality and segmentation accuracy. Several studies concluded that cell segmentation and tracking achieve good level of accuracy, when the signal-to-noise ratio is sufficiently high ($SNR \geq 4$) [3]; however, for low-quality images, the same methods may yield varying levels of performance depending on the level and type of distortion that corrupts the imaging data.

G. Bebis et al. (Eds.): ISVC 2014, Part II, LNCS 8888, pp. 41–50, 2014.

Our methodology builds upon the nonparametric active contour method by Chan and Vese [10], a region based approach to curve evolution that is largely robust to edge discontinuities and additive noise. We developed a system for cell segmentation applied to fluorescence microscope imaging. First, we assessed the image quality of our test data that was provided by the Cell Tracking Challenge [11]. Next, we studied the effect of temporal linking for initialization of level-sets between successive frames. To this end, we validated our approach over a set of sequences against manually segmented data. We also validated the segmentation accuracy of Chan-Vese segmentation without linking and perform comparisons with our approach. Furthermore, we compared the numbers of iterations that each numerical scheme needs to reach convergence. Our results indicate that our approach produces equivalent segmentation accuracies at a significantly smaller number of iterations than the original Chan-Vese method.

2 Methods

2.1 Preprocessing

Fluorescent microscopy imaging is often times subjected to a mixture of different types of noise. Therefore, the degradation model of this imaging process is highly non-linear. The main goal of a preprocessing step is to reduce the corruption caused by noise and to improve the image quality. For noise reduction we utilized a nonlinear diffusion filter proposed by Perona and Malik [12] which is based on the following PDE:

$$\frac{\partial u}{\partial s} = div(g(|\nabla \mathrm{u}|^2) \cdot \nabla \mathrm{u}) \tag{1}$$

where u is the image intensity, s the scale variable, and $g(.)$ a function that determines the amount of diffusion, also known as diffusivity function. This function controls the amount of diffusion according to the edge strength. Common selections of $g(.)$ are the sigmoid and exponential functions. Nonlinear diffusion is a robust method for image enhancement. In contrast to Gaussian blurring, it does not apply smoothing across the edges.

2.2 Active Contour without Edges Chan-Vese Model for Image Segmentation

In contrast to edge based methods like classical snakes [7] and early level-set methods [8], where an edge detector is used to stop the evolving curve, region based methods tend to be less sensitive to noise. The use of region-based statistics may prove advantageous for images characterized by edge discontinuity and higher level of noise.

Chan-Vese (CV) method [10] is a region based active contour model. Here, the segmentation of an image Ω is computed by minimizing the following energy functional:

$$\mathrm{F}(\mathrm{C}, \mathrm{c}_1, \mathrm{c}_2) = \mu \,.\, \mathrm{length}(\mathrm{C}) + \nu.\, \mathrm{area}(\mathrm{inside\ C}) + \lambda_1 \int_{\mathrm{inside\ (C)}} |\mathrm{u}_0 - \mathrm{c}_1|^2 \,\mathrm{dx\,dy} + \lambda_2 \int_{\mathrm{outside\ (C)}} |\mathrm{u}_0 - \mathrm{c}_2|^2 \mathrm{dxdy} \tag{2}$$

where C is the evolving curve; c_1 and c_2 are the average intensity levels inside and outside the contour C and $\mu > 0, \nu \geq 0, \lambda_1, \lambda_2 > 0$ are fixed parameters. The length of C and the area are regularizing terms.

The idea is to evolve the contour C from some initialization until the fitting energy is minimized:

$$\inf_{C,c_1,c_1} F(C, c_1, c_2). \tag{3}$$

This model is a special case of the Mumford-Shah functional [13] for segmentation in the case of piecewise constant approximation.

Level-set Formulation:
Chan and Vese used level-set functions to solve this optimization problem. In the level-set method, the contour is represented as the zero level-set of a Lipschitz function $\phi: \Omega \rightarrow IR$

$$C = \{(x,y) \epsilon \Omega \mid \phi(x,y) = 0\} \tag{4}$$

where ϕ is positive inside C and negative outside C:

$$\begin{aligned} \text{inside}(C) &= \{(x,y) \epsilon \Omega | \phi(x,y) > 0\} \\ \text{outside}(C) &= \{(x,y) \epsilon \Omega | \phi(x,y) < 0\}. \end{aligned}$$

Therefore, the energy functional in terms of the level-set $\phi(x,y)$ becomes:

$$\begin{aligned} F(\phi, c_1, c_2) = \mu \,.\, \text{length}\{\phi = 0\} + \nu.\, \text{area}\{\phi \geq 0\} + \lambda_1 \int_{\phi \geq 0} |u_0 - c_1|^2 \, dx \, dy \\ + \lambda_2 \int_{\phi < 0} |u_0 - c_2|^2 dx dy \end{aligned} \tag{5}$$

The Heaviside and Dirac functions are used to compute the length and area terms. In [10] the Euler-Lagrange equations and the gradient-descent method are used to derive the following evolution equation for the level-set function ϕ that will minimize the fitting energy. Minimization is done by solving the Euler-Lagrange equation for ϕ using time to parameterize the gradient descent:

$$\frac{\partial \phi(x,y)}{\partial t} = \delta(\phi(x,y)) \left[\mu \, \text{div} \left(\frac{\nabla \phi(x,y)}{|\nabla(\phi(x,y)|} \right) - \nu - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2 \right] \in \Omega \tag{6}$$

In practice, the Dirac delta function δ is implemented by a smooth approximation.

2.3 Temporally Linked Level-Set Segmentation

Our approach makes use of temporal connection between consecutive level-set results. That is, when segmenting an image, which is a part of a temporal sequence, we make use of the level-set results reached from minimization of the global energy associated with the contours of the segmented cells found in the previous time point $\phi_{n+1}(x, y; 0) = \phi_n(x, y; i_{final}), (x, y) \in \Omega$, where n is the frame number in the time-lapse sequence, and i_{final} is the number of iterations required to converge for frame n. These results are utilized to minimize the energy functional of the next image. If the segmentation in frame n is accurate, then this initialization will correspond to a point close to the global optimum of the energy functional in frame $n + 1$.

2.4 Algorithm Outline

Based on the previous description, our cell segmentation algorithm is completed in the following steps:

1. Given an input image
2. Apply preprocessing (nonlinear diffusion filter, intensity normalization)
3. Initialize the system for time $n = 0$ to be ϕ^0 using circular structure elements. Compute the signed distance of the initial level-set
4. Repeat steps 5-9 until a stationary solution is reached (i.e., when there is no change in ϕ, or $n > n_{max}$)
5. Compute the regularized Heaviside and Dirac functions
6. Compute the mean intensity of the image pixels inside and outside the level-set: $c_1(\phi^n)$ and $c_2(\phi^n)$ by discretizing the integrals into sums
7. Compute the normalized energy of image
8. Evolve the level-set function by solving the PDE in ϕ to obtain ϕ^{n+1} from (6) with $c_1(\phi^n)$ and $c_2(\phi^n)$
9. Reinitialize ϕ locally to the signed distance function to the curve
10. When the steady state is reached we obtain the final regions by thresholding ϕ
11. Apply morphological hole filling to the segmented regions

3 Experiments and Discussion

3.1 Datasets

The datasets consist of 2D fluorescent microscope time-lapse image sequences.

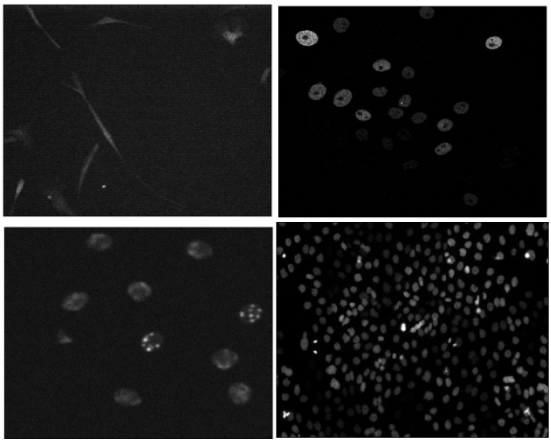

Fig. 1. Rat mesenchymal stem cells on a flat polyacrylamide substrate (2D) (*top left*); GFP-GOWT1 mouse stem cells (2D) (*top right*); (c) Simulated nuclei moving on a flat surface (2D) (*bottom left*); (d) HeLa cells stably expressing H2b-GFP (2D) (*bottom right*)

We used 12 time-lapse sequences, 6 videos consisted of real microscopy time-lapse sequences and 6 computer simulated videos with various cell densities and noise levels. We obtained the training and challenge data sets from the cell tracking challenge website [11].

Simulated videos: The six simulated videos displayed fluorescently labeled nuclei of the HL60 cell line migrating on a flat 2D surface (N2DH-SIM1, N2DH-SIM2, N2DH-SIM3, N2DH-SIM4, N2DH-SIM5, N2DH-SIM6). Fig. 1 shows some sample images. Some properties of the datasets are listed in Table 1.

Real videos: Consists of 3 datasets each containing 2 time-lapse sequences. An example is displayed in Fig. 1 (top right).Two videos sequences Fluo-C2DL-MSC1 and Fluo-C2DL-MSC2 rat mesenchymal stems cells, two video sequences N2DH-GOWT1_1 and N2DH-GOWT1_2, mouse embryonic stem cells and N2DL-HeLa1 and N2DL-HeLa2 expressing HeLa cells.

Table 1. Properties of the sequences used in our experiments from [11]

Data set name	Number of frames	Difficulty
C2DL-MSC	48	*High*: low signal-to-noise ratio, presence of cell stretching, which appear as discontinuous extensions of the cells.
N2DH-GOWT1	92	*Medium:* heterogeneous staining, prominent nucleoli, mitoses, cells entering and leaving the field of view.
N2DL-HeLa	92	*High:* High cell density and low resolution and intensity. The videos display frequent mitoses, both normal and abnormal.
N2DH-SIM	56-100	*Medium*: Different noise levels, cell density of the initial population and number of simulated mitotic events.

3.2 Image Quality and Segmentation Validation Measures

We used the ground truth data to estimate the average Signal-to-Noise Ratio (*SNR*) and Contrast-to-Noise Ratio (*CNR*) of each dataset (see Table 2). The SNR and CNR measures are defined as follows:

$$SNR = 20 \log_{10} \frac{\overline{u_C}}{\sigma_B} \tag{7}$$

$$CNR = \frac{|\overline{u_C} - \overline{u_B}|}{\sigma_B} \tag{8}$$

where $\overline{u_C}$ is the average image intensity over the cell regions, $\overline{u_B}$ is the average intensity over the background and σ_B is the standard deviation of the background pixels. In Table 2 we list the Average *SNR* (in dB) and average *CNR* that are means over all frames in each sequence.

Table 2. Average Signal-to-Noise Ratio (*SNR*) and Contrast-to-Noise Ratio (*CNR*) of each dataset

Dataset sequence name	Average *SNR*	Average *CNR*
C2DL-MSC1	0.21 ± 2.55	4.5 ± 2.02
C2DL-MSC2	5.52 ± 0.27	2.49 ± 0.16
N2DH-GOWT1	23.45 ± 0.6	12.62± 0.76
N2DH-GOWT2	20.07 ± 1.01	8.32 ± 0.9
N2DL-HeLa1	19.4 ± 1.93	19.23± 7.67
N2DL-HeLa2	15.22 ±1.25	5.4 ± 1.08
N2DH-SIM1	9.7 ± 1.16	7.96 ± 0.95
N2DH-SIM2	9.88 ± 1.17	8.34 ± 0.94
N2DH-SIM3	5.9 ± 0.67	4.2 ± 0.47
N2DH-SIM4	5.8 ± 0.56	4.09 ± 0.49
N2DH-SIM5	5.79 ± 0.62	4.24 ± 0.6
N2DH-SIM6	9.54 ± 0.9	7.9 ± 0.8

In the preprocessing step we first applied nonlinear diffusion on each 2D frame as well as contrast enhancement and intensity normalization, to address increased level of noise. In nonlinear diffusion we used a large number of diffusion iterations, greater than or equal to 50 because even if we increase the number of diffusion iterations the edges will still be preserved.

The ground truth, consisting of manually annotated videos (segmentation) was obtained from [11] along with a short description and links to the raw datasets. For segmentation evaluation performance we used the ground truth image. The main purpose of the segmentation performance measurement is to evaluate how well the segmented cells match the cell regions of the ground truth image.

We quantify the accuracy of the segmentation performance by computing the DICE coefficient. This is defined as:

$$\text{DICE coefficient} = 2 * \frac{(R_S \cap R_G)}{|R_S| + |R_G|} \in [0,1] \tag{9}$$

where R_G is the set of all pixels that belong to cell regions in the ground truth image, R_s is the set of all binary regions delineated by our segmentation technique. The DICE coefficient measures the relative similarity between two binary images over their cardinalities. It is frequently used for image segmentation validation. The value one indicates perfect matching.

3.3 Comparison of CV Method with Temporal Linking Method

In this experiment we segmented the test sequences using CV method and our approach. To estimate the segmentation performance we computed DICE coefficients between each method and the reference manual segmentations. We computed means and the standard deviations of DICE scores over all frames for each sequence as displayed in Fig.2. Overall the accuracy in DICE scores derived from the linking method appears to be similar in value with that obtained from CV method. However, there is a clear difference in the values of the standard deviation. That is, the standard

deviations obtained from our temporal-linking method (0.006-0.1) are significantly smaller than those derived from the CV method (0.02-0.4), indicating better convergence and stability.

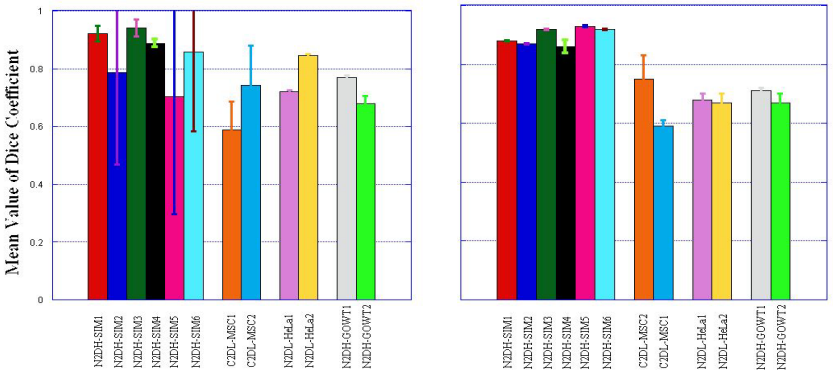

Fig. 2. Segmentation accuracy produced by the original CV algorithm (*left*) and the temporally linked level-sets (*right*)

Better insight into the improvement from the temporal-linking method can be obtained by looking at frame-by-frame segmentation and convergence. Consider the SIM5 dataset, for example. In Fig. 3 we show the results of energy minimization derived using the CV and temporal-linking methods. Because of the non-convexity of the energy functional (allowing therefore many local minima), the CV method reached on frame 8 a local minimum of energy where, however, the moving contour got trapped (Fig. 3 – left side). In contrast, the temporal-linking method led to a global minimum of the energy, yielding the actual cell regions (Fig.3 – right side). Fig. 4 provides better insight into the convergence of both methods. This figure plots the DICE score for each frame. In particular, notice the abrupt decrease of the DICE scores to values very close to zero on certain frames when using the CV method, whereas DICE scores produced by the temporal-linking method appear robust indicating improved robustness.

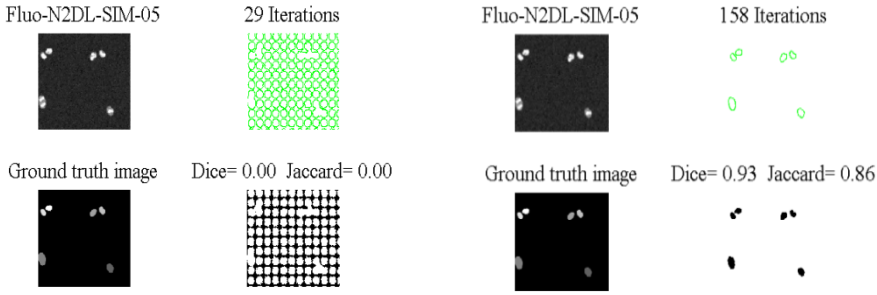

Fig. 3. CV level-set without initialization converges to a local minimum for the frame number 8 of the N2DL-SIM-05 sequence (*left*).Temporal linking improves segmentation accuracy (*right*)

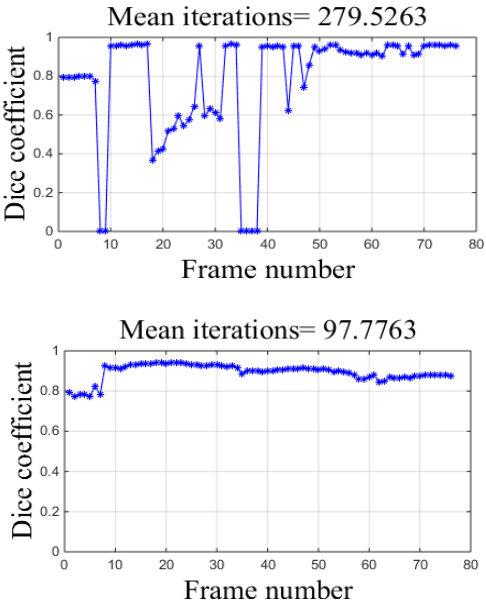

Fig. 4. Comparison of the DICE coefficient of each frame (*total frame=77*) of the Fluo-N2DH SIM5 data using the traditional CV segmentation (*top*) and the temporal-linking method (*bottom*). We note the greater DICE scores and faster convergence of the proposed method (*see top of each figure*).

The previous observation can be explained in terms of energy minimization. Depending on several factors, the energy can sometimes converge to a local energy minimum without reaching the desired global energy minimum. The contour may sometimes stop its evolution before reaching the desired boundaries. In particular cases the splitting of the cells is not delineated correctly; this case is very often encountered in cell segmentation. In some cases, especially in MSC data, stretched cells can leave a trace which can be detected by single or few pixel objects which may be mis-delineated as a cell. Because our energy function is non-convex, allowing therefore many local minima, conventional CV initialization may lead to premature convergence to a local minimum and segmentation errors.

Another consideration in our comparisons was the number of iterations until convergence. Fig. 5 shows the results of the segmentation of a sample frame from the N2DL-Hela 2 sequence. We observe a significant reduction in the number of iterations (from 5000 to 173), because the initial level-set is very close to the actual regions to segment. However, the DICE score produced by the temporally linked level-set is lower than the one produced by the CV method. As indicated in Table 1, this dataset has a high level of difficulty due to the high density and especially the frequent mitoses, low resolution and low fluorescence intensity. In this case, the temporally linked level-set converges significantly faster than CV. The local minimum is reached in a few iterations compared to the proposed technique that reaches a more accurate solution nonetheless.

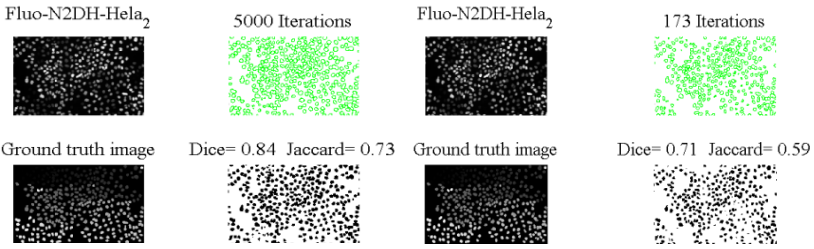

Fig. 5. Sample results obtained by (a) CV method, (b) temporally linked level-set method on the same frame (*number 16*) of N2DL-Hela2

Building upon the previous observation, we tested our hypothesis of faster convergence on our complete set of sequences. We report the average number of iterations used to achieve segmentation of the cells for each dataset over all the frames in Table 3. This table demonstrates the major reduction of the computational time for minimizing the energy with the use of the temporal linking method that is equal to a factor of 10.2. This reduction is achieved by exploiting the previously computed level-set and using it to find an initial solution for the current frame.

Table 3. The mean number of iterations required to achieve segmentation of each sequence by both methods: CV and Temporal Linking

Dataset name	Number of iterations CV	Number of iterations Temporal linking
N2DH-SIM1	261	106
N2DH-SIM2	1088	88
N2DH-SIM3	412	96
N2DH-SIM4	456	43
N2DH-SIM5	279	97
N2DH-SIM6	228	72
C2DL-MSC1	240	104
C2DL-MSC2	343	224
N2DL-HeLa1	1039	67
N2DL-HeLa2	4470	82
N2DH-GOWT1	322	64
N2DH-GOWT2	2469	141
Mean	**998**	**98**

4 Conclusion

We have introduced a level-set based cell segmentation method with temporal linking. We validated our approach on datasets from the online Tracking Challenge [11]. In contrast to standard CV method the proposed method converges faster with similar detection and localization accuracy. Further, temporal linking produces more robust results as it avoids trapping in undesirable local minima. However, parameter retuning may be required for sequences with significantly different quality levels, cell types, and image acquisition configurations. Future goals are the integration of motion tracking techniques, further validation, and extension to 3-D sequences.

Acknowledgments. We acknowledge the support by the Center for Research and Education in Optical Sciences and Applications (CREOSA) of Delaware State University funded by NSF CREST-8763.

References

1. Eils, R., Athale, C.: Computational imaging in cell biology. J. Cell. Biol. 161(3), 477–481 (2003)
2. Stephens, D.J., Allan, V.J.: Light microscopy techniques for live cell imaging. Science 300(5616), 82–86 (2003)
3. Meijering, E., Dzyubachyk, O., Smal, I.: Methods for cell and particle tracking. Methods Enzymo. 504, 183–200 (2012)
4. Dzyubachyk, O., van Cappellen, W.A., Essers, J., Niessen, W.J., Meijering, E.H.W.: Advanced Level-Set-Based Cell Tracking in Time-Lapse Fluorescence Microscopy. IEEE Trans. Med. Imaging 29(3), 852–867 (2010)
5. Yang, X., Li, H., Zhou, X.: Nuclei Segmentation Using Marker-Controlled Watershed, Tracking Using Mean-Shift, and Kalman Filter in Time-Lapse Microscopy. IEEE Transactions on Circuits and Systems I: Regular Papers 53(11), 2405–2414 (2006)
6. Maška, M., et al.: A benchmark for comparison of cell tracking algorithms. Bioinformatics 30(11), 1609–1617 (2014)
7. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. International Journal of Computer Vision 1(4), 321–331 (1988)
8. Malladi, R., Sethian, J.A., Vemuri, B.C.: Shape Modeling with Front Propagation: A Level-set Approach. IEEE Transactions on Pattern Analysis and Machine Intelligence 17, 158–175 (1995)
9. Cremers, D., Rousson, M., Deriche, R.: A Review of Statistical Approaches to Level-set Segmentation: Integrating Color, Texture, Motion and Shape. International Journal of Computer Vision 72(2), 195–215 (2006)
10. Chan, T.F., Vese, L.A.: Active Contours Without Edges. IEEE Transactions on Image Processing 10(2), 266–277 (2001)
11. Cell Tracking Challenge (2013), `http://www.grand-challenge.org/`
12. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 12(7), 629–639 (1990)
13. Mumford, D., Shah, J.: Optimal approximations by piecewise smooth functions and associated variational problems. Comm. on Pure and Applied Mathematics 42(5), 577–685 (1989)