Two-Step Watershed Segmentation of Epithelial Cells

Kevin Keraudren[1], Martin Spitaler[2], Vania M.M. Braga[2], Daniel Rueckert[1], Luis Pizarro[1]

[1]Department of Computing, Imperial College London, UK
[2]National Heart and Lung Institute, Imperial College London, UK

***Abstract*— Epithelial cells are present in many different organs and are essential for physiology. Understanding the molecular mechanisms via which cell-cell contacts are stabilised has important implications for a variety of diseases and cancers. To elaborate on such studies, we develop an automatic approach to cell segmentation based on a simple but effective combination of well-established image filters, morphological operations and watershed segmentation. Aiming at preserving the localisation of the different cell structures, we are able to extract nuclei, cell walls and cell-cell contacts with high accuracy. These turn out to be important for masking the image readouts of cadherin receptors and actin reorganisation to distinguish between junctional and cytoplasmic cell phenotypes, which makes the proposed approach well-suited for image-based high-throughput RNAi screens. Although we focus on epithelial cells, our approach has general applicability to other cell-based screens in confocal microscopy.**

***Index Terms*— Cell segmentation, epithelial cells, high-content, high-throughput screening, biological image analysis**

I. Introduction

Image-based RNA interference (RNAi) screening has become a widely used technique to knock-down targeted genes to analyse their effect on specific cellular functions. The identification of phenotypes corresponding to distinct phases of the cell cycle in cells of different health status is the first step to a better understanding of cellular function in disease and in response to treatment. Large-scale high-throughput, high-content screens produce thousands of images and time-lapse videos capturing particular instants of these cellular processes. Recent studies have emphasised the need of robust and efficient computational tools to analyse this overwhelming amount of data, which is the real bottleneck in terms of data interpretation [1], [2], [3].

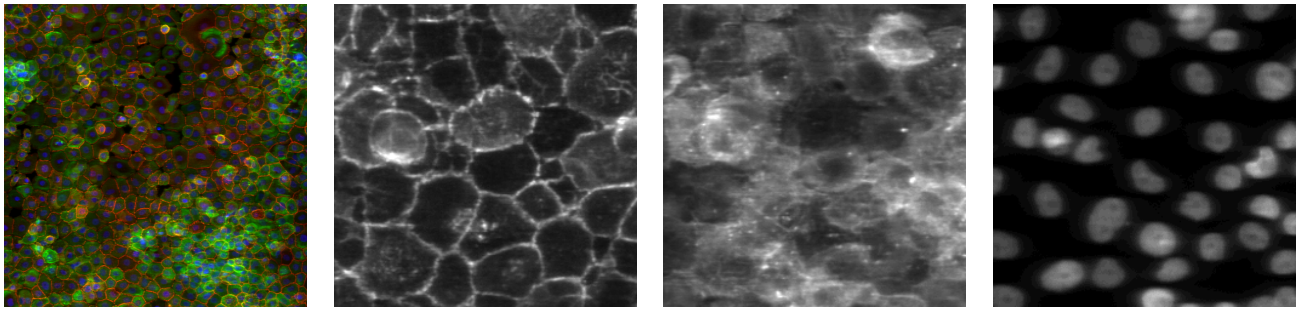
We are interested in studying the biological processes that regulate cell-cell adhesion signalling. Cell-cell adhesion is essential for the life of many different cell types and for their organisation into higher ordered structures (tissues and organs). Epithelial cells attach strongly to neighbouring cells via specialised adhesive molecules. Sheets of tightly packed epithelial cells play a key role in absorption, secretion and as protective layers. Perturbation of their ability to interact tightly is found in different pathologies such as detachment of malignant tumour cells or pathogen infection [4], [5]. One of the master regulators of epithelial adhesion are members of the cadherin family of receptors. Establishment of E-cadherin cell-cell contacts is a multifactorial event; signalling and structural actin-binding proteins contribute to receptor clustering and stabilisation at adhesive sites. Clearly, disruption of actin cytoskeleton severely prevents stable cell-cell adhesion, indicating the important role of cytoskeletal attachment for junction morphology and function. RNAi screens have successfully mapped important regulators of cell attachment to substratum [6], [7] and potential diagnostic tools for diseases and cancer [8]. Yet, in spite of many previous studies, a comprehensive knowledge of the molecules and processes required for stable cadherin adhesion and junction morphology is not available.

Corresponding author: Luis Pizarro, email: luis.pizarro@imperial.ac.uk.

Towards a better understanding of the adhesion properties of epithelial cells, we performed a RNAi actin screen to identify specific actin-binding proteins required for assembly of cell-cell contacts and actin reorganisation. The screen produced thousands of images for about 400 proteins considered in our study. In the following, we will consider $(R, G, B) = (I_{HECD1}, I_{ACTIN}, I_{DAPI})$ image composites like the one shown in Figure 1. To investigate the way epithelial cells attach to one another, it is essential to rely on an accurate image segmentation of the constituent parts of each cell (nucleus, cytoplasm and membrane) so as to identify those proteins whose depletion by RNAi perturbed levels of E-cadherin or F-actin at junctions or in the cytoplasm. As it can be noted in Figure 1, this is not a straightforward task since RNAi depletion causes various cell distortions that need to be quantified, making any automatic segmentation attempt extremely challenging. The main challenges for automated analysis that we observed in the acquired images are their low signal-to-noise ratio (SNR), broken cell-cell contacts and multiple cell overlaps.

In this paper we propose an algorithmic pipeline for the automatic segmentation of epithelial cells that is well-suited for high-content, high-throughput RNAi screens. Our method is able to segment most cells under these challenging conditions. It is based on a simple but effective combination of well-established filters, morphological operations and watershed segmentation. We are specially keen on segmenting three types of cell structures: nuclei, cell walls and cell-cell contacts. That way, we can aim at distinguishing different actin-binding proteins by their image readouts.

The rest of the paper is organised as follows: Some related work on cell segmentation is reported in Section II. Our proposed approach is presented in detail in Section III. Section IV demonstrates the suitability of our approach to correctly segment junctional and cytoplasmic structures, which allows for the characterisation of different proteins. Finally, we discuss our contributions and sketch some future work in Section V.



(a) (b) (c) (d)

Fig. 1. (a) image composite $(R, G, B) = (I_{HECD1}, I_{ACTIN}, I_{DAPI})$; (b), (c) and (d) display zoomed-up images of each channel, respectively.

II. Related Work on Cell Segmentation

From the abundant literature on cell segmentation, here we briefly report on two recent and competing methods:

A common approach to cell segmentation is to begin by thresholding the nuclei and then propagating the obtained seeds to define the cytoplasm. Jones *et al*. [9] proposed a Voronoi-based segmentation approach that approximates the Voronoi regions of each seed on a manifold with a metric that moves pixels closer depending on the similarity of their surrounds. This method has become very popular and it has been implemented in CellProfiler [10] and in Bioconductor [11].

The method proposed by Yan *et al*. [12] is based first on a segmentation of the nuclei with the following pipeline: Otsu thresholding, distance transform and watershed on the distance transform. Then, for each nucleus an active contour is fitted to the cell boundaries by minimising an energy functional that describes both the repulsion between neighbouring cells and the competition for segmenting pixels. The authors compared their method to the Voronoi-based approach mentioned above, obtaining about 10% better F-score. Nevertheless, they did not provide the same seeds to both methods. They run a separate nuclei segmentation, thus blaming the competing method for over-segmenting the cells. Later in our experiment we also compare our approach against Jones *et al*.'s method, but we attempt doing it as fairly as possible.

There exist other methods particularly focused on nuclei segmentation [13], [14] and cytoplasm segmentation [15].

III. Proposed Segmentation Approach

Our approach to cell segmentation utilises a couple of components that are standard in the biological literature, plus a few new ingredients that allow us to cope with the challenging conditions mentioned in the introduction. Furthermore, we will be able to define three cellular objects: nuclei, cell-cell contacts and cell walls, which will be very important for masking the I_{HECD1} and I_{ACTIN} channels to distinguish between junctional and cytoplasmic cell-features. The proposed pipeline for cell segmentation is outlined in Figure 2. It consists of four main blocks: *Pre-Processing, Nuclei Processing, Edge-Map Processing,* and *Adaptive Watershed*, which are detailed below.

A. Pre-Processing

Screen images were acquired using a line scanning confocal system. Four projections in the z-plane were acquired at each experimental well, from which we computed the maximum intensity projection for each channel. To avoid some optical aberrations we discarded the pixels with the 2% darkest and the 1% brightest intensities, performing then a normalisation to the range $[0, 255]$. This constitutes our raw data.

The intensity variability and the low SNR observed at the cell-cell contacts and cytoplasm hinders their proper segmentation. We therefore opt for denoising the I_{HECD1} and I_{DAPI} channels, which are the ones we base our segmentation approach on. We utilise the *block matching 3D* (BM3D) filter of Dabov *et al*. [16] that is the state-of-the-art for filtering Gaussian noise. Another issue affecting the I_{HECD1} channel is the presence of broken cell-cell contacts. To partially solve this problem we use the *coherence-enhancing diffusion* (CED) filter of Weickert [17] that connects and enhances interrupted line-like structures. It is important to mention that all these filtering steps are only employed to help the segmentation process. Once we have built segmentation masks for the nuclei, the cytoplasm and the cell-cell contacts, all subsequent phenotypic measurements are taken from the raw data.

B. Nuclei Processing

With the auto-threshold method of Li and Lee [18] we obtain a first segmentation of the nuclei, which we then refine with a watershed process on the grey-scale I_{DAPI} channel to separate the joined nuclei. As a last step, touching nuclei are eroded in one pixel to accentuate their separation.

C. Edge-Map Processing

We obtain an approximate localisation of the cell-cell contacts and their thickness by computing an edge map on the I_{HECD1} channel with a Sobel filter, followed by the application of a mean filter. The latter is used to fill the gaps in the edge map. Note that a morphological closing would have served the same purpose.

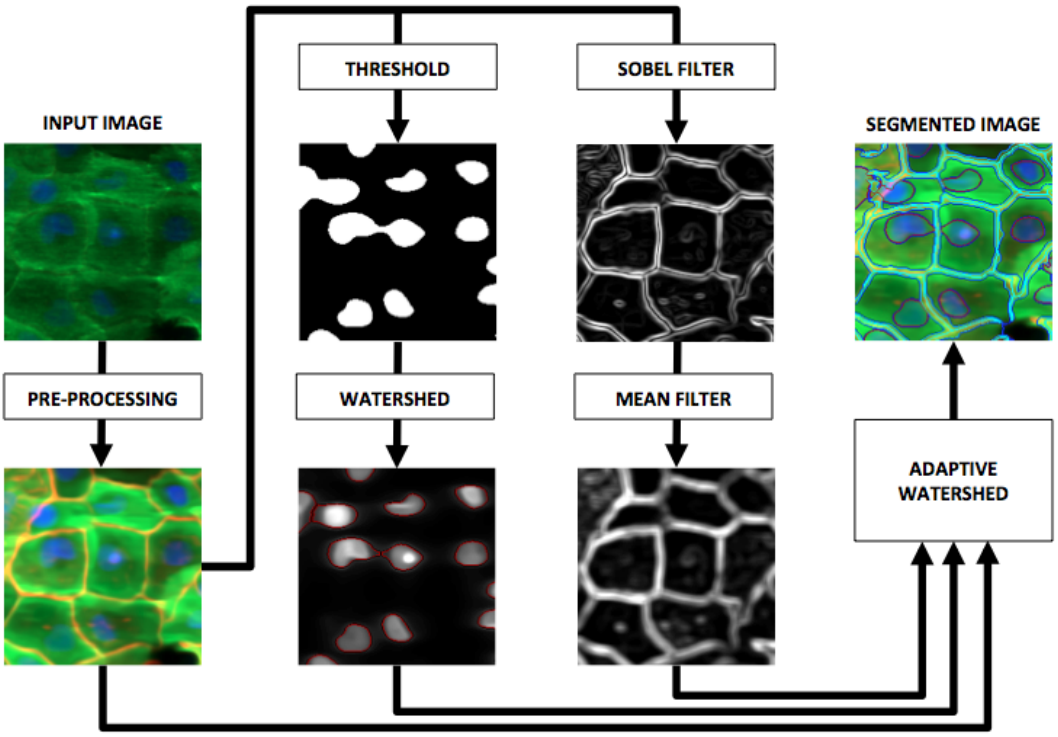


Fig. 2. The proposed pipeline for cell segmentation consists of 4 blocks: *Pre-Processing, Nuclei Processing, Edge-Map Processing,* and *Adaptive Watershed*.

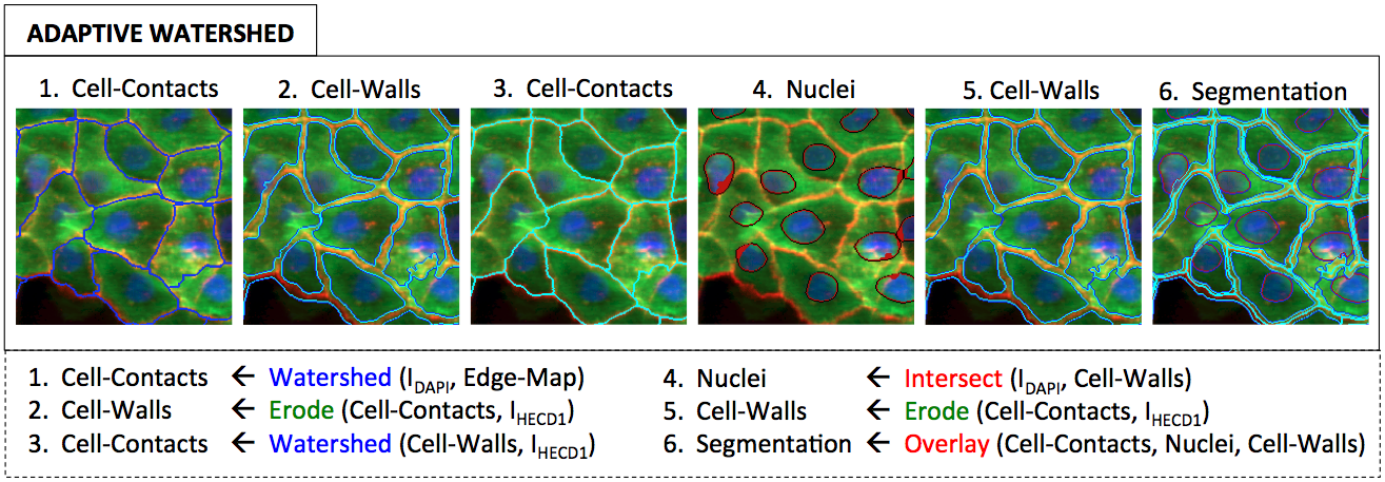


Fig. 3. *Adaptive Watershed* block that produces the final segmentation as a composition of three classes of structures: Cell-Contacts, Nuclei, Cell-Walls.

D. Adaptive Watershed

As shown stepwise in Figure 3, we aim at extracting three different types of cell structures from the images: the cell-cell contacts where two cells touch each other, the cell walls delimiting the extent of the cytoplasm, and the nuclei respecting such cell boundaries.

Using the input nuclei as seeds we perform a watershed on the edge map to obtain a first segmentation of the cell-cell contacts, which are then eroded towards the regions of high HECD1 to obtain a first segmentation of the cell walls.

Despite the fact that the smoothed edge map provides a good initialisation for estimating the cell boundaries, it does not carry precise information about the regions of high cadherin. Therefore, we optimise the cell-cell contacts by adapting another watershed to the I_{HECD1} channel using the cell walls as seed regions.

Subsequently, the nuclei (I_{DAPI}) are refined such that a single nucleus is not allowed to be on both sides of a cell-cell contact. That is, they need to be intercepted with the cell walls. Finally, a last refinement takes place at the cell walls. They are re-estimated with the better localised cell-cell contacts. Note that the update of the cell walls is done after the nuclei segmentation. That way we diminish the risk of having nuclei cutting through the cell boundaries (although it might happen), so we can later make junctional measurements without bias towards the presence/absence of nuclei.

IV. Results

We would first like to emphasise the importance of the pre-processing block in our pipeline. As it was shown in Figure 1, the I_{HECD1} channel is affected by a low SNR and by incomplete cell boundaries. The joint utilisation of the BM3D and the CED filters does help obtain smoother and more accurate segmentations as it can be observed in Figure 4(a)-(b).

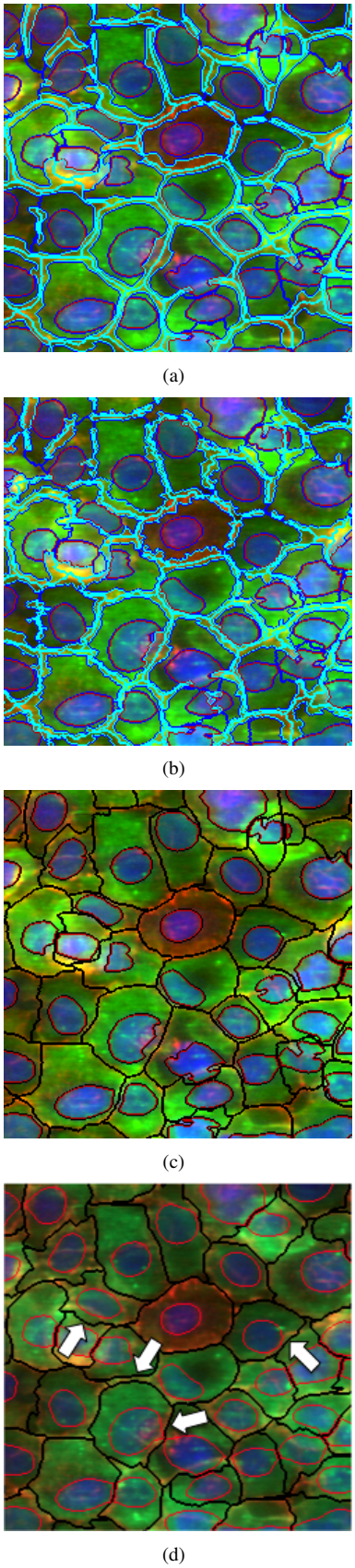
Another important characteristic of our segmentation approach is its ability to accurately localise cell-cell contacts, which is encoded in the *adaptive watershed* algorithm described above. We compare the segmentation of cell-cell contacts using our approach and the *propagate* method based on Voronoi diagrams implemented in Bioconductor. For a fair comparison, both methods were fed with the same pre-processed data and the same seed regions as in Figure 2. The results in Figure 4(c)-(d) indicate that the Voronoi-based method cannot produce as accurate segmentations of the cell-cell contacts as our approach. The white arrows point out some of the misalignments with respect to the I_{HECD1} channel obtained via *propagate*. At this point it is worth noting that it should not be difficult to implement a two-step *propagate* method to mimic what we have done with the two-step watershed algorithm.

To test the biological significance of our segmentation pipeline, we compare the results of three images from the cell-based high-throughput screen of actin-binding proteins. After segmenting all cell components, Figure 5 shows boxplots of the intensity distribution of three experimental parameters: E-cadherin (HECD1) and actin filaments (J-Actin) at the cell-cell contacts, as well as of the non-junctional (cytoplasmic) Actin (NJ-Actin). In all graphs, the first bar is the quantification of images from a positive control (Ca2+), while the second and third bars refer to images obtained after depletion of two proteins in cells: Tropomyosin-2 beta (TPM2) and VAV-2 oncogene (VAV2), respectively. TPM2 binds actin filaments and stabilise them and is thereby predicted to play a role in cytoplasmic actin and potentially at Junctional actin. VAV2 is an upstream regulator of the small GTPase Rac1, a known regulator of actin remodelling following different stimuli such as migration, growth factor stimulation and cell-cell adhesion.

These plots confirm that the RNAi perturbation of both TPM2 and VAV2 has a strong effect on the response of cadherin receptors. In terms of the actin measurements, both proteins distinguish themselves from the positive control with higher and lower intensity levels, respectively. Thus the new pipeline proposed here is able to identify cell borders efficiently and provides reliable quantification. They are also consistent with the images and predicted phenotype of the proteins investigated (see discussion in Section V).

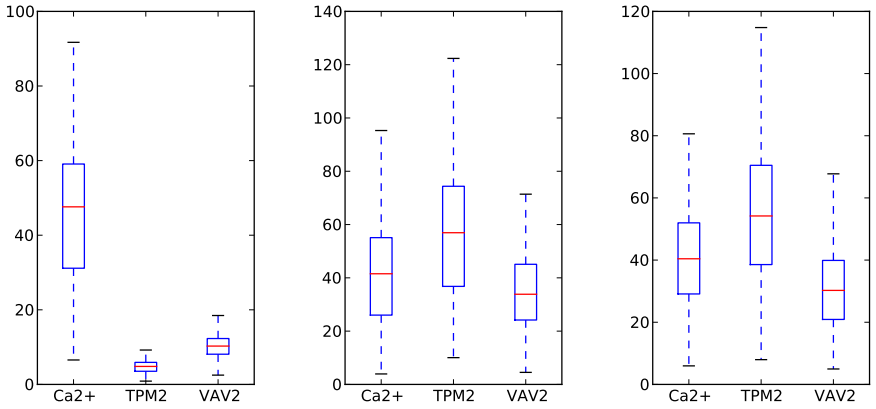
V. Discussion and Conclusions

We have presented a general approach to cell segmentation that is well-suited for image-based high-throughput RNAi screens. The most challenging issues were: i) to accurately extract well-localised cell-cell contacts –which sometimes appear fuzzy– respecting the tightly packed organisation of cell



(a) (b) (c) (d)

Fig. 4. Cell segmentation (b) with and (c) without pre-processing; localisation of cell-cell contacts (black lines) after segmentation with (d) our approach and (e) Bioconductor. Please refer to the text for details.



(a) (b) (c)

Fig. 5. Boxplots of the intensity distribution of E-cadherin (HECD1), junctional actin (J-actin) and non-junctional actin (NJ-actin) for the positive control (Ca2+) and the proteins Tropomyosin-2 beta (TPM2) and VAV-2 oncogene (VAV2).

conglomerates, and ii) to determine the thickness of the cell walls –even fuzzier. Both issues have a significant impact on the definition of cell phenotypes that can reliably distinguish between actin-binding proteins inhibiting or promoting actin reorganisation and the establishment of cadherin receptors.

The method proposed here has shown its usefulness in identifying modulation of the levels of cell-cell adhesion receptors and changes in cytoskeleton caused by removal of specific proteins from cells. The data suggest that, upon depletion of TPM2 or VAV2, cell-cell contacts are not very stable and cadherin levels at junctions severely perturbed as shown by a strong reduction in HECD1 levels. Potential explanations are, in the case of TPM2, that stabilisation of actin filaments adjacent to cell-cell contacts are important to stabilise the membrane at junctions and to cluster cadherin receptors in place. Regarding VAV2, its downstream partner Rac1 is well-known to regulate actin recruitment to cadherin receptors engaged in adhesion. Thus by removing VAV2, it is feasible that Rac1 cannot be activated and thus unable to stabilise contacts by providing actin to support adhesion. Consistent with this, VAV2 depletion leads to a partial reduction in the actin levels in their respective images.

These encouraging preliminary results are currently followed by a more extensive validation with further controls and other cell samples in order to apply the proposed methodology to a whole set of circa 400 RNAi-depleted actin-binding proteins.

REFERENCES

[1] Christophe J. Echeverri and Norbert Perrimon, "High-throughput RNAi screening in cultured cells: a user's guide," *Nat Rev Genet*, vol. 7, no. 5, pp. 373–384, 05 2006.

[2] Roy Wollman and Nico Stuurman, "High throughput microscopy: from raw images to discoveries," *J Cell Sci*, vol. 120, no. Pt 21, pp. 3715–22, Nov 2007.

[3] Hanchuan Peng, "Bioimage informatics: a new area of engineering biology," *Bioinformatics*, vol. 24, no. 17, pp. 1827–36, Sep 2008.

[4] W. James Nelson, "Adaptation of core mechanisms to generate cell polarity," *Nature*, vol. 422, no. 6933, pp. 766–774, 04 2003.

[5] Manuel R. Amieva, Roger Vogelmann, Antonello Covacci, Lucy S. Tompkins, W. James Nelson, and Stanley Falkow, "Disruption of the epithelial apical-junctional complex by Helicobacter pylori CagA," *Science*, vol. 300, no. 5624, pp. 1430–1434, 05 2003.

[6] Kaylene J. Simpson, Laura M. Selfors, James Bui, Angela Reynolds, Devin Leake, Anastasia Khvorova, and Joan S. Brugge, "Identification of genes that regulate epithelial cell migration using an siRNA screening approach," *Nat Cell Biol*, vol. 10, no. 9, pp. 1027–1038, 09 2008.

[7] Sabina E. Winograd-Katz, Shalev Itzkovitz, Zvi Kam, and Benjamin Geiger, "Multiparametric analysis of focal adhesion formation by RNAi-mediated gene knockdown," *The Journal of Cell Biology*, vol. 186, no. 3, pp. 423–436, 08 2009.

[8] J Mullenders and R Bernards, "Loss-of-function genetic screens as a tool to improve the diagnosis and treatment of cancer," *Oncogene*, vol. 28, no. 50, pp. 4409–4420, 09 2009.

[9] Yanxi Liu, Tianzi Jiang, Changshui Zhang, Thouis Jones, Anne Carpenter, and Polina Golland, *Voronoi-Based Segmentation of Cells on Image Manifolds*, vol. 3765, pp. 535–543–543, Springer Berlin / Heidelberg, 2005.

[10] CellProfiller, "http://www.cellprofiler.org," .

[11] Bioconductor, "http://www.bioconductor.org," .

[12] Pingkun Yan, Xiaobo Zhou, M. Shah, and S. T. C. Wong, "Automatic segmentation of high-throughput RNAi fluorescent cellular images," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 12, no. 1, pp. 109–117, Jan. 2008.

[13] Fuhui Long, Hanchuan Peng, Xiao Liu, Stuart K Kim, and Eugene Myers, "A 3D digital atlas of C. elegans and its application to single-cell analyses," *Nat Meth*, vol. 6, no. 9, pp. 667–672, 09 2009.

[14] Cris L Luengo Hendriks, Soile V E Keränen, Charless C Fowlkes, Lisa Simirenko, Gunther H Weber, Angela H DePace, Clara Henriquez, David W Kaszuba, Bernd Hamann, Michael B Eisen, Jitendra Malik, Damir Sudar, Mark D Biggin, and David W Knowles, "Three-dimensional morphology and gene expression in the Drosophila blastoderm at cellular resolution I: data acquisition pipeline," *Genome Biol*, vol. 7, no. 12, pp. R123, 2006.

[15] Carolina Wählby, Joakim Lindblad, Mikael Vondrus, Ewert Bengtsson, and Lennart Björkesten, *Algorithms for cytoplasm segmentation of fluorescence labeled cells.*, vol. 24:2-3, pp. 101–111, Analytical Cellular Pathology, 2002.

[16] K Dabov, A Foi, V Katkovnik, and K Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *Image Processing, IEEE Transactions on*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[17] Joachim Weickert, "Coherence-enhancing diffusion filtering," *International Journal of Computer Vision*, vol. 31, no. 2, pp. 111–127–127, 1999-04-01.

[18] C. H. Li and C. K. Lee, "Minimum cross entropy thresholding," *Pattern Recognition*, vol. 26, no. 4, pp. 617–625, 4 1993.