

Automated cell individualization and counting in cerebral microscopic images

Zhenzhen You, Michel Vandenberghe, Yael Balbastre, Nicolas Souedet, Philippe Hantraye, Caroline Jan, Anne-Sophie Herard, Thierry Delzescaux

▶ **To cite this version:**

Zhenzhen You, Michel Vandenberghe, Yael Balbastre, Nicolas Souedet, Philippe Hantraye, et al.. Automated cell individualization and counting in cerebral microscopic images. 2016 IEEE International Conference on Image Processing (ICIP), Sep 2016, Phoenix, United States. 10.1109/ICIP.2016.7532988 . hal-01539966

HAL Id: hal-01539966
https://hal.science/hal-01539966

Submitted on 6 Jul 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

AUTOMATED CELL INDIVIDUALIZATION AND COUNTING IN CEREBRAL MICROSCOPIC IMAGES

Zhenzhen YOU[1], Michel E. VANDENBERGHE[1], Yael BALBASTRE[1], Nicolas SOUEDET[1], Philippe HANTRAYE[1], Caroline JAN[1], Anne-Sophie HERARD[1], Thierry DELZESCAUX[1]

[1]CEA-CNRS-UMR 9199, MIRCen, Fontenay-aux-Roses, France

Abstract

In biomedical research, cell counting is important to assess physiological and pathophysiological information. However, the automated analysis of microscopic images of tissues remains extremely challenging. We propose an automated processing protocol for proper segmentation of individual cells in microscopic images. A Gaussian filter is applied to improve signal to noise ratio (SNR) then an original min-max method is proposed to produce an image in which information describing both cell centers (minima) and boundaries are enhanced. Finally, a contour-based model initialized from minima in the min-max cartography is carried out to achieve cell individualization. This method is evaluated on a NeuN-stained macaque brain section in sub-regions presenting various levels of fraction of neuron surface occupation. Comparison with several methods of reference demonstrates that the performances of our method are superior. A first application to the segmentation of neurons in the hippocampus illustrates the ability of our approach to deal with massive and complex data.

Index Terms— Cell individualization, neuron counting, segmentation, microscopic images, min-max cartography

1. Introduction

Cell segmentation is a major biological research tool which can provide useful information about cell number and organization in the tissue. In neurodegenerative disease, for example, it allows to investigate neuronal loss in pathological brains. Microscopic images are commonly used in biology and provide high resolution images in which cytoarchitecture can be investigated. Traditionally, stereology methods [1] are used by the biologists to perform robust cell counting in an anatomical region based on manual segmentation performed in several samples. But this manual operation is time consuming and cumbersome. In addition, its accuracy depends on the complexity of images and the experience of the biologists. In order to overcome these limitations, an automated, accurate and robust method is requested. So far, several image processing methods have been proposed, but most of them are highly specific. Mathematical morphology [2] and concavity detection [3], [4] can only segment cells in simple images where only few cells are touching. However with histological tissue, a tremendous quantity of data is to be analyzed. Some anatomical regions like the dentate gyrus consist of massively packed neurons which are extremely difficult to individualize. Watershed algorithms [5] can segment touching cells but can easily cause over- and under-segmentation. Snakes [6]–[8] can avoid over-segmentation but the initial contour and a pertinent external force are essential for a good segmentation. More recently, the iCut algorithm [9] has been proposed to segment touching cells based on a size prior and concavity detection.

This paper presents a new image processing protocol which aims at automatically segmenting individual cells in two steps. First, segmentation of cells from tissue and background is performed using Random Forest (RF) classifier. The second step concerns the individualization of touching cells. We propose to use an original min-max filter that exacerbates local extrema to recover valuable information about cell centers and boundaries. The generated cartography can then guide a competitive region growing scheme. An optimization of the parameters of the proposed method is realized on a dataset manually segmented by experts and this method is compared to two methods of

reference (watershed, iCut). Our method is then applied to segment cells on the hippocampus region, one of the most challenging anatomical regions to study due to its size and complex organization.

2. Materials and Methodology

2.1. Biological material

This work is based on a macaque brain which was cut after euthanasia into 40-µm-thick sections. About 150 brain sections were produced and stained using a neuronal marker NeuN. They were digitized using an AxioScan.Z1 (Zeiss) with an in plane resolution of 0.22 µm (x20-magnification). In this preliminary work, sub-parts of one brain section corresponding to 120 GB were extracted and processed. A *training set* of ten representative images (512x512 pixels) were extracted and manually segmented into 3 classes (neurons, non-stained tissue and background) by an expert. To validate our individualization method, a *centroid set* of eighty images (256x256 pixels) presenting different levels of fraction of cell surface occupation were extracted. In these images, each cell was manually segmented by 2 experts by marking a point in its center providing information about cell number and cell center location. Finally, the proposed method was applied to an image of the left hippocampus corresponding to about 5 GB (44000x28000 pixels) in order to count neurons and compute neuron density in complex manually segmented sub-regions (CA1, CA2, CA3, CA4, dentate gyrus, subiculum).

2.2. Tissue segmentation

A RF model (100 trees) that we've previously described [10] is generated using the *training set*. The features considered are H, S, V color channels and local intensity (gray level intensity computed in disk of 5 pixels radius). Then the *centroid set* are segmented into 3 classes with the RF model. Additionally, color images are converted to gray level images [11] noted I_{gray} in the following sections.

2.3. Cell individualization

To individualize touching neurons, we propose a new method by computing an image in which neuron centers and boundaries can be visualized at the same time. To achieve this purpose, I_{gray} is first filtered by a Gaussian kernel with a standard deviation σ to improve SNR. This parameter σ is optimized to minimize the count error (see section 2.4). The filtered image I_g is then masked with the RF segmentation so that all background and non-stained pixels are set to 0.

The min-max cartography I_e is then computed. For each pixel p of I_g, let $D(d, p)$ the disk (not including p) of radius d centered on p, let $N(D)$ the number of pixels in $D(d, p)$, let $S(p)$ the number of values in $D(d, p)$ smaller than $I_g(p)$ and $G(p)$ the number of values greater or equal to $I_g(p)$. The output value is then:

$$I_e(p) = \frac{G(p) - S(p)}{N(D)} \tag{1}$$

Consequently, pixels whose filtered value is -1 are absolutely minima in the local disk while those whose value is 1 are absolutely maxima. d is fixed at 10 pixels as it is the expected minimum neuron radius (2.5 µm [12]). This two-step process (Gaussian smoothing, min-max filtering) can be repeated n times to sharpen the extrema. n is optimized to minimize the count error (see section 2.4). Pixels of value -1 are selected as neuron centroids and each centroid is assigned to a unique label as the identification (*id*) of neurons.

The final segmentation is obtained with a discrete contour-based model. Contours are initialized as 5-pixel-radius circles centered on each centroid and all pixels inside them are labeled with the same *id* of their centroid. As neuron boundaries are close to maxima in I_e, we propose to give each contour point an expanding speed that depends on the contour curvature and I_e's intensity. Let p a contour point, $\kappa(p)$ the curvature-dependent term, $\mu(p)$ the intensity-dependent term, o the position of the neuron centroid. The next position of p noted p' is obtained through equation (2).

$$\overrightarrow{op'} = \overrightarrow{op} + \frac{\overrightarrow{op}}{\left\|\overrightarrow{op}\right\|} \times \kappa(p) \times \mu(p) \tag{2}$$

Let $k(p)$ the contour curvature at point p and c an optimal curvature. $\kappa(p)$ is obtained through equation (3), making the contour expanding for curvatures smaller than c and shrinking for curvatures greater than c.

$$\kappa(p) = c - k(p) \tag{3}$$

The intensity-dependent term is inspired by [13]. Let t a coefficient empirically set to $0.8 \times \max(I_e)$. $\mu(p)$ is obtained through equation (4).

$$\mu(p) = \exp\left(-\left(\frac{I_e(p)+1}{2t}\right)^2\right) \tag{4}$$

After each iteration, when the distance between two consecutive points p and q is superior to a predefined maximal distance d_{max}, new points are interpolated automatically according to equation (5).

$$\begin{cases} N = \left\lceil \dfrac{\left\|\overrightarrow{pq}\right\|}{d_{\max}} \right\rceil - 1 & \\ \overrightarrow{op_i} = \overrightarrow{op} + \dfrac{\overrightarrow{pq}}{N+1} \times i & 1 \le i \le N \end{cases} \tag{5}$$

Then pixels surround each contour point p within d_{max} are examined. For those nearer to their centroid than p and not yet labeled are assigned to the same *id* of their centroid.

All cell contours are simultaneously expanded, and contour crossing is forbidden. In our experiment, we fixed the number of expanding iterations at 30. This number can be arbitrarily high, and should be chosen so that the computation time is low enough but contours have enough time to reach the cell boundaries. Once the process is finished, the remaining unlabeled pixels are assigned according to their neighbors by majority voting.

2.4. Parameters optimization and validation

To optimize parameters (σ and iteration number n), our protocol is applied to the *centroid set* with σ values ranging from 1 to 8 pixels and iteration number from 1 to 3. The relative count error ε is defined as the absolute difference between the automatic count (N_a) and the expert count (N_e) divided by the expert count:

$$\varepsilon = \frac{|N_a - N_e|}{N_e} \tag{6}$$

In addition, a classification score that takes into account the position of the segmented neurons is defined. For each automatically segmented neuron, the number of expert centroids contained in the neuron is computed. If exactly one expert centroid is contained, this neuron is considered as a true positive. Else, it is either over segmented (zeros expert centroids) or under segmented (more than 1 expert centroids). Recall (R), precision (P) and F-score (F) are calculated as follows:

$$R = \frac{N_t}{N_e}; \quad P = \frac{N_t}{N_a}; \quad F = 2\frac{R \times P}{R + P} \tag{7}$$

where N_t is the number of correctly automatic segmented neurons.

The bigger the values of those criteria are, the better the performance of method is.

3. Results and Discussion

3.1. Parameters optimization

Fig. 1 shows an example of different iterations applied to touching neurons (σ=4). A number of two iterations of min-max cartography computation is optimal because one iteration lead to under-detection of neurons and three iterations to over-detection. A consistent estimation between the expert and automated method is obtained for two iterations (10 centroids detected). This setting enables to force the detection of small, blurred or brighter neurons which are very difficult to detect.

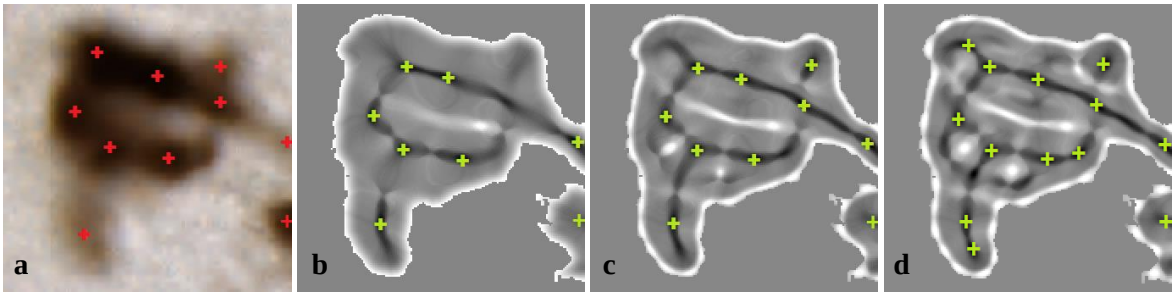


Figure 1: Creation of min-max cartography (σ=4). a) Original color image. b) One iteration I_e. c) Two iterations I_e. d) Three iterations I_e. Red crosses are neuron centroids marked by expert while green crosses are the detected neuron centroids.

Relative count errors resulting from the optimization of σ and iteration number n are calculated in Table 1 for expert 1 (similar results are obtained for expert 2).

Table 1: Relative count error ε based on n and σ.

Error Expert 1		Iteration number n		
		1	2	3
σ	1	0.20±0.19	0.55±0.44	1.48±0.86
	2	0.22±0.19	0.32±0.30	0.95±0.62
	3	0.23±0.19	0.16±0.19	0.48±0.41
	4	0.24±0.19	**0.12**±0.16	0.25±0.26
	5	0.26±0.19	0.14±0.17	0.16±0.19
	6	0.29±0.19	0.18±0.17	0.14±0.17
	7	0.31±0.20	0.22±0.18	0.16±0.18
	8	0.35±0.20	0.26±0.19	0.21±0.18

Based on the minimal relative count error calculated, parameters σ=4 and 2 iterations are selected as optimal.

3.2. Neuron individualization

We have applied the proposed method to the *centroid set*. Fig. 2 presents the results obtained on 3 different images presenting different levels of fraction of neuron surface occupation. Images 1 represent simple images with a few individual neurons, images 2 represent moderately dense images in which several neurons touch and images 3 represent extremely dense images in which many neurons touch. A simple classification (neuron and unstained tissue) illustrated in Fig 2. b1-b3 is insufficient to segment neurons when touching neurons are present. Fig 2. c1-c3 present two iterations min-max cartographies in which neuron centers appear in black and neuron boundaries in bright (visually very close compared to the final neuron contours automatically detected). The proposed individualization method provides good results visually assessed in Fig 2. d1-d3.

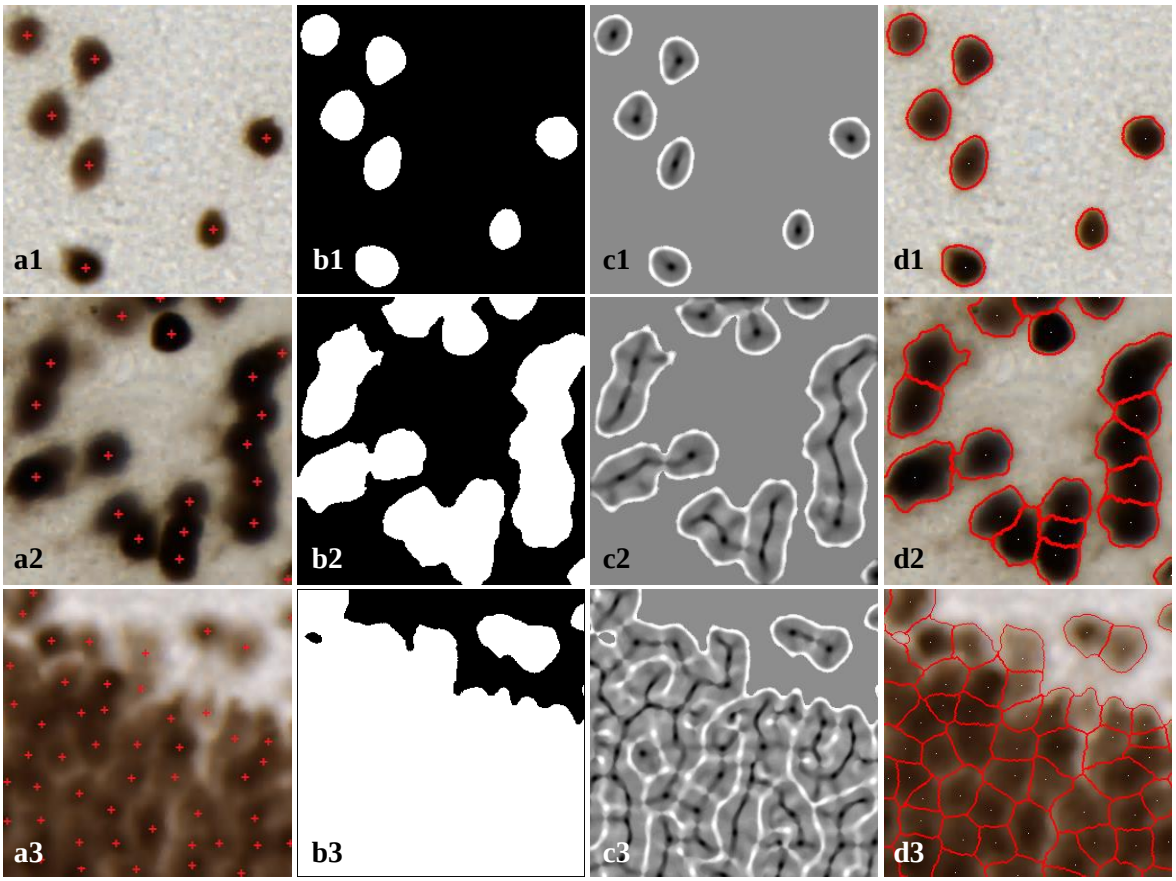


Figure 2: Neuron segmentation results obtained in images presenting different levels of fraction of neuron surface occupation. a1-a3) Original images. Red crosses are neuron centroids marked by expert. a1) a few individual neurons, a2) several touching neurons, a3) lots of touching neurons. b1-b3) Classification results by random forest. c1-c3) Two iterations min-max cartographies Ie. d1-d3) Individualization results obtained by our method.

Fig. 3 presents the results obtained on the same images using two methods of reference: iCut and watershed. All these methods perform well in the simple cases. However, in the more complicated situations, the results of iCut method and watershed are not accurate. iCut separates touching neurons based on the features taking into account neuron size, spatial location, intervening contours and concave contours. In dense neuronal regions, information of concavity contours is not helpful anymore. iCut estimates the number of neurons by their mean area and gives a homogeneous segmentation result which doesn't correspond visually to real neurons (Fig 2. a1-a3). The watershed results present

over- and under-segmentation making the final contours unnatural (straight borders, large neurons detected). Visually, the proposed method provides a better neuron individualization result.

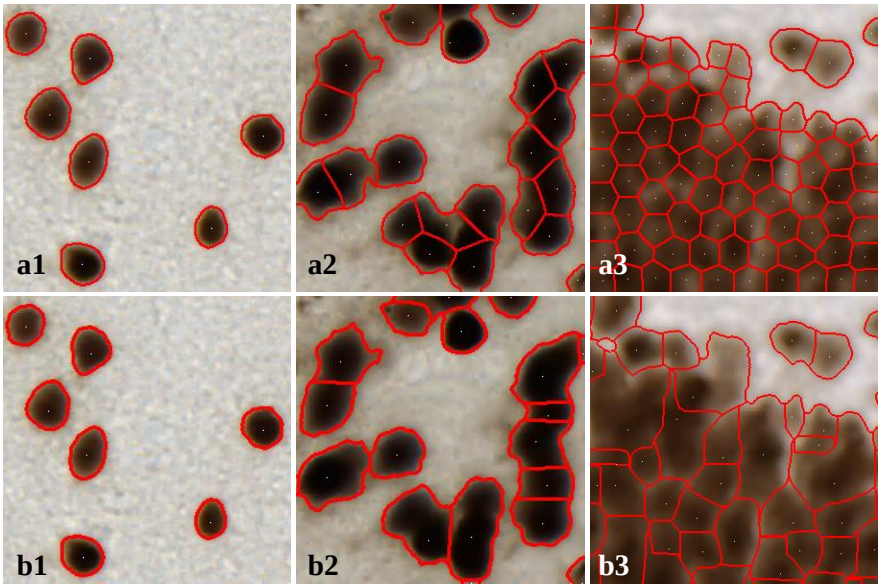


Figure 3: Segmentation results of 2 reference methods. a) represents iCut method, b) Watershed method. Images 1, 2, 3 represent different levels of fraction of neuron surface occupation.

In addition, a quantitative evaluation applied to the *centroid set* among the 3 methods is illustrated in Table 2. It shows that the proposed method provides the best performances. Moreover, the standard deviation of our method is smaller, demonstrating that the proposed method is more robust for both simple and complicated cases.

Table 2: Comparison of performances among 3 neuron individualization methods based on 2 experts segmentation. Highest scores are shown in bold.

Method		Performances	
		Expert 1	Expert 2
iCut	Recall	0.88±0.13	0.83±0.14
	Precision	0.87±0.14	0.90±0.11
	F-score	0.86±0.11	0.85±0.10
Watershed	Recall	0.88±0.18	0.82±0.14
	Precision	0.82±0.13	0.86±0.09
	F-score	0.84±0.09	0.83±0.09
Proposed method	Recall	0.92±0.09	0.85±0.11
	Precision	0.94±0.07	0.97±0.05
	F-score	0.92±0.06	0.90±0.07

Another important fact to mention for iCut and watershed methods is that despite relatively high scores observed all superior to 0.8, the segmentation of the neurons is questionable (visual evaluation). It is thus important to define a more sensitive quality index.

3.3. Application to the study of hippocampal sub-regions

Fig. 4a presents the left hippocampus and the sub-regions manually segmented (CA1 to CA4, dentate gyrus (DG) and subiculum (S) shown in Fig. 4b) to which our method is applied. Table 3 presents the fraction of neuron surface occupation (*F*), the number of neurons detected (*N*) and neuron density (*D*, the number of neurons per mm²) in each sub-region using the optimal parameters. Our method performs well on large images of several GB in reasonable computational time (about 5 hours in this case on a recent workstation).

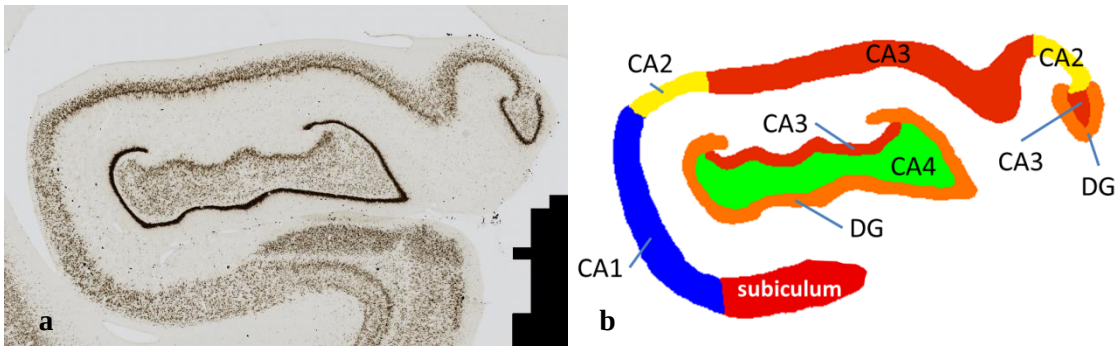


Figure 4: a) Hippocampus of one macaque brain section. b) 6 sub-regions segmented by expert.

Table 3: F, N and D obtained in each sub-region.

Region	CA1	CA2	CA3	CA4	DG	S
F	0.17	0.24	0.18	0.12	0.29	0.15
N	5160	3079	12323	4409	11689	3641
D	2241	3310	2598	1515	4079	1852

4. Conclusion

This paper presents an efficient method to individualize touching cells in cerebral microscopic images. Some promising segmentation results have been obtained above on neuronal staining in macaque. The accuracy of our method is higher compared to state-of-the-art methods as assessed qualitatively and quantitatively. The method has been successfully applied to a hippocampus of macaque brain, one of the most challenging anatomical regions to study in neuroscience. Perspectives of this work will be to extend this study to the entire brain sections or brains to alleviate the issue of big data processing [14]. The automated detection of millions of cells is still challenging and an in-depth comparison of strengths and complementarities between stereology and image processing methods should be carefully addressed in the future. This generic tool should also help to address major biological issues such as understanding of pathology mechanisms or evaluation of novel therapies.

REFERENCES

[1] M. J. West, L. Slomianka, and H. J. Gundersen, “Unbiased stereological estimation of the total number of neurons in thesubdivisions of the rat hippocampus using the optical fractionator,” Anat. Rec., vol. 231, no. 4, pp. 482–497, Dec. 1991.

[2] A. Nedzved, S. Ablameyko, and I. Pitas, “Morphological segmentation of histology cell images,” 15th International Conference on Pattern Recognition (ICPR), vol. 1, pp. 500–503, 2000.

[3] X. Bai, C. Sun, and F. Zhou, “Touching Cells Splitting by Using Concave Points and Ellipse Fitting,” Digital Image Computing: Techniques and Applications (DICTA), pp. 271–278, 2008.

[4] S. Kothari, Q. Chaudry, and M. D. Wang, “Automated cell counting and cluster segmentation using concavity detection and ellipse fitting techniques,” IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ISBI), pp. 795–798, Sep. 2009.

[5] J. Cousty, G. Bertrand, L. Najman, and M. Couprie, “Watershed Cuts: Minimum Spanning Forests and the Drop of Water Principle,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 8, pp. 1362–1374, Aug. 2009.

[6] C. Xu and J. L. Prince, “Gradient vector flow: a new external force for snakes,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pp. 66–71, 1997.

[7] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active Contour Models,” International Journal of Computer Vision, pp. 321–331, 1988.

[8] J. S. Suri, K. Liu, S. Singh, S. N. Laxminarayan, X. Zeng, and L. Reden, “Shape recovery algorithms using level sets in 2-D/3-D medical imagery: a state-of-the-art review,” IEEE Trans. Inf. Technol. Biomed. Publ. IEEE Eng. Med. Biol. Soc., vol. 6, no. 1, pp. 8–28, Mar. 2002.

[9] Y. He, H. Gong, B. Xiong, X. Xu, A. Li, T. Jiang, Q. Sun, S. Wang, Q. Luo, and S. Chen, “iCut: an Integrative Cut Algorithm Enables Accurate Segmentation of Touching Cells,” Sci. Rep., vol. 5, Jul. 2015.

[10] M. E. Vandenberghe, Y. Balbastre, N. Souedet, A.-S. Herard, M. Dhenain, F. Frouin, and T. Delzescaux, “Robust supervised segmentation of neuropathology whole-slide microscopy images,” 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3851–3854, 2015.

[11] International Telecommunication Union, “Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios,” Recommendation ITU-R BT601-7, 2011.

[12] P. Andersen, R. Morris, D. Amaral, T. Bliss, and J. O’Keefe, The Hippocampus Book. Oxford University Press, 2016.

[13] P. Perona and J. Malik, “Scale space and edge detection using anisotropic diffusion,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 629–639, 1990.

[14] M. E. Vandenberghe, A.-S. Herard, N. Souedet, E. Sadouni, M. D. Santin, D. Briet, D. Carre, J. Schulz, P. Hantraye, P.-E. Chabrier, T. Rooney, T. Debeir, V. Blanchard, L. Pradier, M. Dhenain, and T. Delzescaux, “High-throughput 3D whole-brain quantitative histopathology in rodents,” Sci. Rep., vol. 6, Feb. 2016.