Automatic segmentation of hippocampus in histological images of mouse brains using deformable models and random forest

Pablo Mesejo, Roberto Ugolotti, Ferdinando Di Cunto, Stefano Cagnoni, Mario Giacobini

Automatic Segmentation of Hippocampus in Histological Images of Mouse Brains using Deformable Models and Random Forest

Pablo Mesejo, Roberto Ugolotti, Stefano Cagnoni
Department of Information Engineering
University of Parma

Ferdinando Di Cunto, Mario Giacobini
Molecular Biotechnology Center
University of Torino

Abstract

We perform a two-step segmentation of the hippocampus in histological images. First, we maximize the overlap of an empirically-derived parametric Deformable Model with two crucial landmark sub-structures in the brain image using Differential Evolution. Then, the points located in the previous step determine the region where a thresholding technique based on Otsu's method is to be applied. Finally, the segmentation is expanded employing Random Forest in the regions not covered by the model. Our approach showed an average segmentation accuracy of the 92.25% and 92.11% on test sets comprising 15 real and 15 synthetic images, respectively.

1. Introduction

Among the different anatomical structures which make up the mammalian brain, the hippocampus is particularly interesting, due to its role in learning and memory, and as early biomarker for Alzheimer's disease and epilepsy. It is composed by the Dentate Gyrus (DG) and the Ammon's Horn (CA), which is further composed by CA1, CA2, and CA3 (see Figure 1).

A precise gene expression map, at the cellular and subcellular level, can provide crucial information for understanding the biological mechanisms that underlie the cellular and molecular events that take place in this structure. A promising data source to derive this map has recently been provided by the Allen Brain Atlas (ABA) [1], a huge, public database that contains high-resolution brain images mapping the expression patterns of most genes contained in the genomes of the analyzed organisms. Considering the wide availability of brain images containing morphological and functional information on the hippocampus in different organisms, it has become extremely important to design image analysis methods that may accurately, robustly, and reproducibly extract the hippocampus from them, in order to automatize any relevant analytic procedure.

Figure 1: From left: regions in the hippocampus, a real and a synthetic image

2. Theoretical background and related work

Classically, image segmentation is defined as the partitioning of an image into non-overlapping regions that are homogeneous with respect to some visual feature. In this work, we deal with the segmentation of the hippocampus in sagittal images from the ABA, but this method could be easily adapted to other anatomical planes or subcortical structures. In this section, we describe the main techniques used and their essential working.

Deformable Models (DM), first introduced by Terzopoulos [10], are curves or surfaces defined within an image domain, that move under the influence of "internal" and "external" forces. Internal forces keep the model smooth during deformation, while the external ones attract the model toward an object of interest within the image. The DM used in this work are mainly inspired by Active Shape Models (ASM) [3], that can be seen as a way to add more prior knowledge to DM, since these shape models represent objects by sets of labelled points selected by an expert in a set of training images.

In machine learning, Ensemble Classifiers (ECL) use multiple classification methods to obtain better predictive performance than could be obtained from any of the constituent models. In particular, a Random Forest (RF) [2] is an ECL that consists of a combination of Decision Trees (DT), and has been shown to perform very well compared to many other classifiers [4].

Regarding segmentation of structures in histological images, our work is related to the investigations of Senyukova

et al. [8]. There, the segmentation is treated as a classification problem using RF and Markov Random Fields, which refine the results at the pixel level, but the method requires previous knowledge about the reference slice associated to that image. In [7], the segmentation using three-dimensional Gaussian mixtures and Level Sets is carried out slice-by-slice, where the successful segmentation of one section provides a prior for the subsequent one, assuming that the segmentation of few sparsely sampled slices is done manually (so it is not a completely automatic method).

3. Global Overview of the Method

This section describes the general pipeline of the system. Details about the basics can be found in [11], which describes the implementation of a fully automatic 2D localization method based on atlas-based registration and DM. The main contributions and differences of this work with respect to the above one are the use of Differential Evolution [9] as optimization method, which leads to much better performance than the techniques previously considered, and the development of a segmentation algorithm seeded by the points located in the first step, as well as the expansion of the segmented areas using RF.

3.1 Localization of the Hippocampus

It is well-known that one of the main problems of DM is their initialization. In this work, the initialization, in terms of the choice of a model and its initial position, is solved by an atlas-based affine registration using the ABA reference atlas. The main idea, explained in detail in [11], is to find the sagittal reference slice of the atlas which best matches the target image. This phase produces two results: firstly, it makes it possible to determine the position within the brain of the section represented in the target image and, consequently, to choose the right model. Secondly, based on the information contained in the corresponding reference atlas image, it allows us to extract the Region of Interest (ROI) where the hippocampus is expected to be located.

The parametric representation of the model is moved and deformed, guided by Differential Evolution, according to an intensity-based similarity function between the model and the object which is being sought. In our method, the hippocampus is located by detecting, as landmarks, two regions which are usually well distinguishable within the structure: the pyramidal (sp) and granule (sg) cell layers, which belong to the CA and DG regions, respectively (see Figure 1). Two models are used in the localization process, to locate each of the two separated structures (one for sp and another for sg). The process maximizes the intensity differences between the hippocampus and the surrounding structures while keeping the shape of the model plausible.

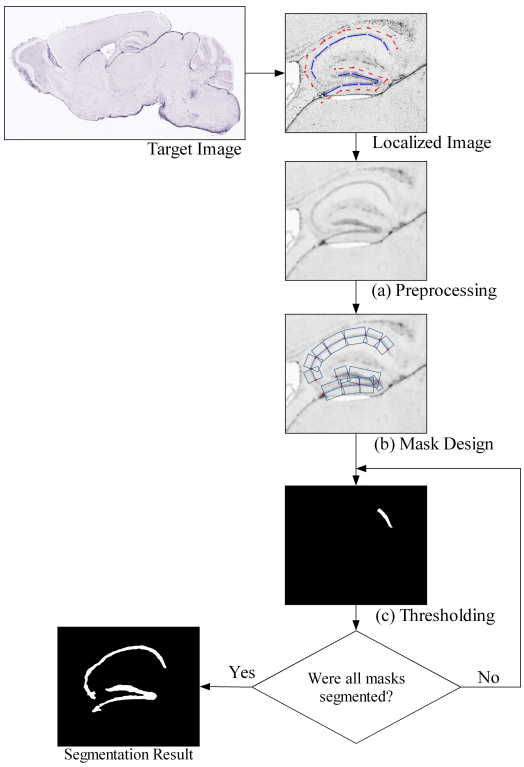

Figure 2: Segmentation pipeline.

3.2 Segmentation of the Hippocampus

The results of the localization are the ROI (whose size is usually around 2.500×2.000 pixels) and the points of the models which overlap the hippocampus. The segmentation, whose pipeline is shown in Figure 2, is developed using a combination of classical computer vision techniques.

(a) Preprocessing: a Median Filter (with a 25×25 pixels mask) is applied to the ROI in order to reduce the noise that can affect the segmentation. The median filter in constant time [6] was chosen since it obtains the same results as the classical one but faster, exhibiting $O(1)$ runtime complexity.

(b) Mask Design: we create window masks around the points located with the previous step and use them as seeds. Doing so, we can apply the thresholding technique at a local level, without applying it to the whole image.

(c) Thresholding: We apply Otsu's method [5] iteratively on every window mask created from the seeds and, in every iteration, we keep only the largest segmented component. Therefore, the localization process can be seen as an intelligent technique for localizing areas where the segmentation method is to be applied.

3.3 Expansion of the Segmentation

Two main problems may affect segmentation results: situations in which the located points do not overlap with the hippocampus, and situations in which the hippocampus is not completely covered by the final model. Even if experiments on a test set of 320 images yielded a perfect (all points over the hippocampus and covering it completely) or good localization (the points do not cover it entirely or, at most, two points are slightly outside it) in 90.9% of cases, yielding better results than previously obtained by Particle Swarm Optimization [11], we can further refine our results by introducing a method to detect/classify points belonging to the hippocampus, in order to extend the segmentation to regions not yet included.

In order to select the most adequate classifier we tried several of them, starting from the simplest ones (like Naive Bayes, 1-Nearest Neighbour or DT) up to more advanced techniques (like RF, Support Vector Machines, Multi Layer Perceptrons and Adaboost).

The training set was formed by 189 images from our database, in each of which 20 points over the hippocampus and 20 points outside the hippocampus were selected. In total, 7560 patterns were used during the training process (employing a 5x2 cross-validation). After training, the best models were tested using 1200 patterns. Such a test set was extracted from 30 completely different images, downloaded randomly from the ABA, selecting, for each image, 20 points within the hippocampus and 20 points outside the hippocampus. All these patterns were encoded as a vector of 11 textural features (first order measures: mean, standard deviation, skewness, kurtosis, entropy, coefficient of variation and energy; second order: contrast, correlation, energy and homogeneity from the Gray Level Co-ocurrence Matrix using (1,1) as spatial relationship) employing 5 window sizes (30×30, 90×90, 150×150, 210×210 and 300×300). RF (with 500 trees and 7 variables randomly sampled as candidates at each split) obtained better results than the other methods, achieving an accuracy and a false positive rate of 97.75% and 1.75%, respectively.

The classifier trained as described is then used to extend the segmentation towards the parts that have not been considered yet (top right frame in Figure 3). This expansion acts as a region growing method, that first detects the intersection between the boundaries of the window masks and the segmented region (bottom left of Figure 3). Then we create a new window mask centered in each intersection. After that, this new window is segmented, and several random points are selected and classified (see bottom right of Figure 3) in the new segmented part. If the majority of the points are classified as hippocampus, this new segmentation is accepted. The expansion of the window mask continues until no further step is possible, i.e. no more intersections between any window mask and the segmented areas are found, or all new segmentations are discarded by the classifier.

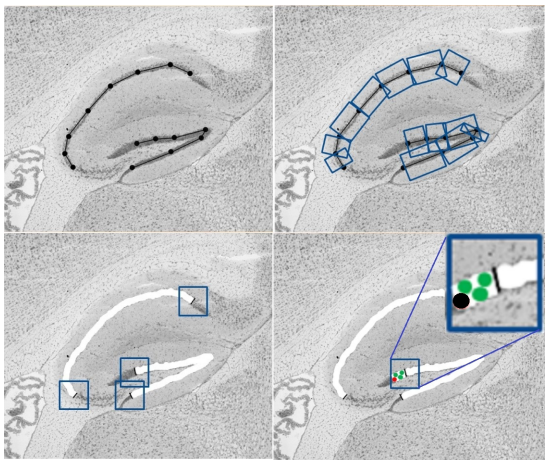

Figure 3: Expansion of the segmentation.

4. Experimental Results

In order to study the behavior of this method, tests over real and synthetic images were run on an Intel® Core™ i5-2410M CPU @ 2.30GHz with 4.00 GB of RAM. Table 1 reports the results obtained using synthetic and real images employing 25 runs per image (375 independent executions).

The first two rows in Table 1 show statistical information about the results obtained in terms of true positives (TP) and false positives (FP), respectively. Such statistics are the average on all runs, the mean of the medians computed on each image, and the standard deviation on all runs. The third row reports the same statistics referred to the execution time, while the last one shows the p-value obtained using the Wilcoxon signed-rank test, with a level of confidence of 0.01, for the Null-Hypothesis "There are no differences between the median TP/FP percentages obtained with and without expansion".

We firstly tested our approach on a synthetic version of the problem using 15 images, to evaluate the performance of our system when dealing with noisy images with numerous artefacts. In these images, the hippocampus is formed by small circles of random radius and color. Small and big ellipses (between 1000 and 3000) were included trying to simulate cells, and gaussian (mean $\in$ [0.0, 0.2] and variance $\in$ [0.01, 0.15]) and salt and pepper noise (density $\in$ [0.05, 0.25]) were added to introduce fuzziness into the images.

The p-values obtained were less than the level of confidence, giving a statistical proof that the results with the expansion were better than the ones obtained using only seg-

	Synthetic Images						Real Images					
	Segmentation			Expansion			Segmentation			Expansion		
	Avg	Median	Std	Avg	Median	Std	Avg	Median	Std	Avg	Median	Std
TP (%)	88.38	89.17	6.31	92.11	92.38	5.13	88.57	89.33	7.30	92.25	93.46	6.98
FP (%)	11.01	10.64	5.88	11.08	10.82	6.09	22.12	22.02	9.51	22.29	21.85	9.70
Time (sec)	36.39	36.49	5.36	7.92	7.89	2.98	51.52	51.56	4.73	15.79	14.65	9.20
Wilcoxon test	p-value (TP): 3.19E-033, p-value (FP): 7.42E-005						p-value (TP): 1.84E-022, p-value (FP): 2.37E-001					

Table 1: Results of segmentation with synthetic and real images.

mentation, which supports the idea of introducing such a step in the segmentation pipeline.

Regarding real images, a ground truth image was created by manually segmenting the hippocampus in 15 significant images. In order to avoid erroneous or incomplete manual segmentations, these were supervised by an expert in molecular biology. Every image was manually segmented 5 times and, for each group of 5 manual segmentations, the intersection and union images were calculated. The difference between the maximum (union) and minimum (intersection) area outlined in equally biologically valid segmentations was computed and showed an average difference of 25.26%. Hence, two segmentations of the same image can differ by around 25% of the segmented area while being both "correct". In all these experiments, the image considered as reference (ground truth) was the intersection of the five manual segmentations.

As in the previous case, the best results were obtained including the expansion phase. Although the percentage of FP may seem very high (around 20%), it is necessary to consider the intrinsic uncertainty mentioned in the previous paragraph and that the ground truth used was the intersection image, that considers a smaller area as gold standard.

5. Conclusions and Future Work

In this paper, we have presented a method for segmenting anatomical structures in histological images. This method was applied to real and synthetic images, obtaining an average accuracy of 92.25% in the first case, and 92.11% in the second one. To achieve this goal, the method only needs an anatomical atlas and a parametric model, associated to the atlas, representing the structure of interest. It automatically selects the most suitable reference slice, localizes the hippocampus using a parametric deformable model, and segments this anatomical structure using an iterative version of Otsu's thresholding method. For the expansion of the segmentation, when the localization was not perfect, an ensemble classifier (Random Forest) has been used. The use of the expansion increases the TP rate, keeping almost constant the FP percentage and the standard deviation. Therefore, the segmentation with expansion has been shown to be a better method for tackling these images.

6. Acknowledgements

This work is funded by Compagnia di San Paolo, Torino, Italy (Neuroscience Programme). Pablo Mesejo is funded by the European Commission (MIBISOC Marie Curie Initial Training Network, FP7 PEOPLE-ITN-2008, GA n. 238819). All the images were downloaded from the Allen Mouse Brain Atlas [http://mouse.brain-map.org]. Seattle (WA): Allen Institute for Brain Science. ©2009.

References

[1] Allen Institute for Brain Science. Allen Reference Atlases. `http://mouse.brain-map.org`, 2004-2006.

[2] L. Breiman. Random forests. *Maching Learning*, 45:5–32, 2001.

[3] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models-their training and application. *Computer Vision and Image Understanding*, 61:38–59, 1995.

[4] L. I. Kuncheva and J. J. Rodríguez. Classifier ensembles for fMRI data analysis: an experiment. *Magnetic Resonance Imaging*, 28(4):583 – 593, 2010.

[5] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans. on Systems, Man and Cybernetics*, 9(1):62 –66, 1979.

[6] S. Perreault and P. Hébert. Median filtering in constant time. *IEEE Trans. on Image Processing*, 16(9):2389–2394, 2007.

[7] T. Riklin-Raviv, N. Sochen, N. Kiryati, N. Ben-Zadok, S. Gefen, L. Bertand, and J. Nissanov. Propagating distributions for segmentation of brain atlas. In *Proc. IEEE International Symposium on Biomedical Imaging, ISBI '07*, pages 1304–1307, 2007.

[8] O. V. Senyukova, A. S. Lukin, and D. P. Vetrov. Automated atlas-based segmentation of Nissl-stained mouse brain sections using supervised learning. *Programming and Computing Software*, 37:245–251, 2011.

[9] R. Storn and K. Price. Differential Evolution- a simple and efficient adaptive scheme for global optimization over continuous spaces. Technical report, International Computer Science Institute, 1995.

[10] D. Terzopoulos and K. Fleischer. Deformable models. *The Visual Computer*, 4:306–331, 1988.

[11] R. Ugolotti, P. Mesejo, S. Cagnoni, M. Giacobini, and F. Di Cunto. Automatic Hippocampus Localization in Histological Images using PSO-Based Deformable Models. In *Proc. Genetic and Evolutionary Computation Conference, GECCO '11 (Companion)*, pages 487–494, 2011.