Automatic Hippocampus Localization in Histological Images using Differential Evolution-Based Deformable

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

Automatic Hippocampus Localization in Histological Images using Differential Evolution-Based Deformable Models

Pablo Mesejo[a,*], Roberto Ugolotti[a], Ferdinando Di Cunto[c], Mario Giacobini[b,c], Stefano Cagnoni[a]

[a] *Department of Information Engineering, University of Parma, Parma, Italy*
[b] *Department of Veterinary Sciences, University of Torino, Italy*
[c] *Molecular Biotechnology Center, University of Torino, Italy*

Abstract

In this paper, the localization of structures in biomedical images is considered as a multimodal global continuous optimization problem and solved by means of soft computing techniques. We have developed an automatic method aimed at localizing the hippocampus in histological images, after discoveries indicating the relevance of structural changes of this region as early biomarkers for Alzheimer's disease and epilepsy. The localization is achieved by searching the parameters of an empirically-derived deformable model of the hippocampus which maximize its overlap with the corresponding anatomical structure in histological brain images. The comparison between six real-parameter optimization techniques (Levenberg-Marquardt, Differential Evolution, Simulated Annealing, Genetic Algorithms, Particle Swarm Optimization and Scatter Search) shows that Differential Evolution significantly outperforms the other techniques in this task, providing successful localizations in 90.9% and 93.0% of two test sets of real and synthetic im-

Preprint submitted to Pattern Recognition Letters *October 8, 2012*

*Corresponding Author. Tel. +39 0521 905785

Email addresses: pmesejo@ce.unipr.it (Pablo Mesejo), rob_ugo@ce.unipr.it (Roberto Ugolotti), ferdinando.dicunto@unito.it (Ferdinando Di Cunto), mario.giacobini@unito.it (Mario Giacobini), cagnoni@ce.unipr.it (Stefano Cagnoni)

ages, respectively.

Keywords: Hippocampus, Deformable Models, Automatic Localization, Histological Images, Global Continuous Optimization, Differential Evolution

1. Introduction

Among the different anatomical structures which make up the mammalian brain, the hippocampal formation (HPF) (see Figure 1) is particularly interesting. From an anatomical viewpoint, the HPF, composed by the Hippocampus and the Subiculum (SUB), is located within the medial temporal lobe. In turn, the Hippocampus is composed by the Dentate Gyrus (DG) and Ammon's Horn (CA), which is further composed by three different regions (CA1, CA2, and CA3).

The hippocampus has long been known for its crucial role in learning and memory processes [1]. Moreover, it has recently been demonstrated that the volume of the hippocampus is an early biomarker for Alzheimer's disease. Therefore, there is a great interest in understanding the cellular and molecular events that take place in this structure, under both normal and abnormal conditions. From this point of view, a precise gene expression map at the cellular and subcellular level within this region can provide crucial information for understanding such biological mechanisms. A very promising data source to derive this map has recently been provided by the Allen Brain Atlas (ABA), a huge, publicly available database that contains high-resolution images mapping the expression patterns of most genes contained in

the genomes of the analyzed organisms. The first release of the ABA was focused on mouse, and contained the expression patterns of about 20,000 genes obtained by In Situ Hybridization (ISH) of mouse brain sections [2]. In a more recent release, a similar resource based on spatially mapped microarray data has also been provided for the human brain [3]. There is wide availability of brain images containing morphological and functional information on the hippocampus in different organisms. Thus, it has become extremely important to design image analysis methods that accurately, robustly, and reproducibly identify the hippocampus region, to automatize any relevant analytic procedure. In the ABA, for each gene under consideration, several images are provided, corresponding to different sections of the brain. Each image is labelled according to the corresponding images of a reference atlas comprising 132 coronal and 21 sagittal sections spaced at 100 μm and 200 μm intervals, respectively.

In this paper, we describe a general model-based method for the automatic localization of anatomical districts in histological images, which could be also applied to other structures, as well as to different imaging modalities. In particular, we describe an application in which such a method is used to locate the hippocampus in ISH images from the ABA. This method allows us to automatically extract image parameters from corresponding regions of a huge number of images, to cluster genes with similar expression patterns and subcellular mRNA distribution. It can be hypothesized that genes whose expression is mapped in images with similar visual features are likely to be

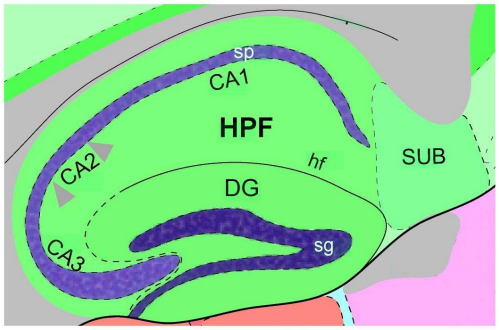


Figure 1: Regions in Hippocampal Formation (HPF)

also functionally similar.

In our method, the hippocampus is located by detecting, as landmarks, two regions which are usually well distinguishable within the structure (see Figure 1): the pyramidal (*sp*) and granule (*sg*) cell layers, which belong to the CA and DG regions, respectively.

As in most medical imaging applications, the problems we have to solve are mainly related with the characteristics of the images involved in the process. The main problematic features of these images in general, and of the hippocampus region in particular, related with biological features or with the image acquisition process, can be summarized as follows:

- natural variability of brain structure shapes in different subjects;
- fuzziness of the hippocampus boundary;
- limited relevance of color for detecting anatomical structures: regions with similar colors may represent different structures and vice versa, depending on the dye used as well as on local image acquisition settings;

- contrast variability between structures: different genes are not expressed equally in the same anatomical region, making it difficult to construct a consistent model for each landmark in all images. Moreover, grained patterns with many irregularities hamper the classification of individual pixels as belonging to the anatomical structures under consideration;

- orientation issues: the imaged structures may be rotated or displaced on the slice with respect to a "standard" alignment;

- lighting issues: within the same set of images, some are much brighter than others.

- variable resolution even within the same image: high-resolution regions coexist with low-resolution ones;

- presence of artefacts: tears, scraps, bubbles, streaks in tissues, partial cut-off of regions;

- large image size (the typical resolution of ABA images is about 15,000 $\times$ 7,000 pixels).

These problems significantly hamper tasks like localization and segmentation of structures in such images.

The fully automatic 2D localization method we propose is based on atlas-based registration and on the optimization of the parameters of a parametric deformable model. Our method can be divided in two stages: (i) selection

of the corresponding slice in the reference atlas based on a two-step affine registration, and (ii) proper localization of the hippocampus.

This technique was initially presented in [4], where the stochastic search of the hippocampal region was performed using Particle Swarm Optimization (PSO) [5]. In this paper, we give more details about the second stage of the method, we delve into the rationale of its operation and we widen its experimental evaluation performing more extensive tests. Moreover, we compare six real-parameter stochastic optimization techniques on the task under consideration; in particular, we show that Differential Evolution (DE) significantly outperforms the other methods taken into consideration: Levenberg-Marquardt (LM) [6], Simulated Annealing (SA) [7], Scatter Search (SS) [8], Genetic Algorithms (GA) [9] and PSO. In our tests, we evaluate the localization of the hippocampus in real and synthetic sagittal images, but this method could as well be applied to other subcortical structures, image modalities or anatomical planes.

The paper is organized as follows: in section 2 we provide the theoretical foundations of our work, as well as an overview of previous related work. In section 3, a general overview of the method is presented, providing details about the implementation of the localization of the hippocampus. Finally, section 4 presents results on both synthetic and real images with appropriate statistical tests, followed, in section 5, by some final remarks and a discussion about possible future developments.

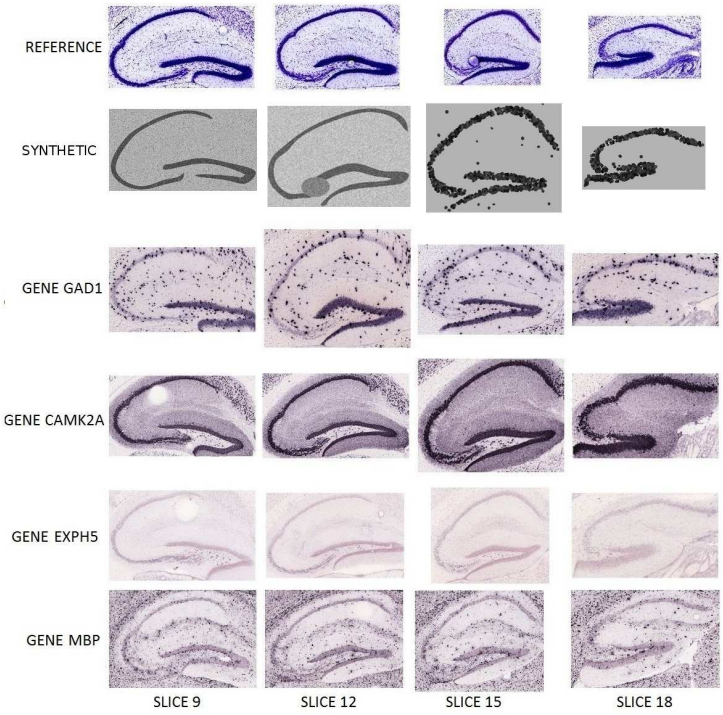


Figure 2: Hippocampus variability. Horizontal-wise: hippocampus sections taken at different levels look different, and different genes can as well produce very different visual features. Vertical-wise: corresponding sections from different brains maintain some shape similarity.

2. Theoretical background and related work

2.1. Deformable Models

Deformable models are curves or surfaces, defined within the image domain, that are deformed under the influence of "internal" forces, related with the curve features, and "external" forces, related with the surrounding image. Internal forces enforce regularity constraints and keep the model smooth during deformation, while external forces are defined such that the model is attracted toward an object or other features of interest within the image. The term "deformable models" was first used in the late eighties [10, 11], and one of the first examples, called "snakes" or Active Contour Models, was presented shortly after in [12].

Active Shape Models (ASMs) [13] add more prior knowledge to deformable models. These shape models derive a "point distribution model" from sets of labelled points (landmarks) selected by an expert in a training set of images. Thus, the final model of the object of interest can be derived by examining the distribution of the positions of the labelled points. The model considers the points' average positions and the main modes of variation observed in the training set. It is important to notice that an instance of the model can only take into account deformations which appear in the training set: because of this, the model may have problems with unexpected shapes, but it is robust with respect to noise and image artefacts, like missing or damaged parts.

Although originally developed for computer vision applications to natural scenes and computer graphics problems, the applicability of deformable

models in medical image analysis has already been proven [14, 15].

2.2. Differential Evolution

In the last two decades, research on global optimization has been very active [16, 17, 18], and many different deterministic and stochastic algorithms for continuous optimization have been developed. Among the stochastic approaches, Evolutionary Algorithms (EAs) [19, 20] offer a number of advantages that make them attractive: implicit parallelism, robust and reliable performance, global search capability, no need of specific information about the problem to solve, easy implementation, good insensitivity to noise, and no requirements for differentiable or continuous objective functions. DE, first introduced by Storn and Price [21], has recently been one of the most successful evolutionary algorithms. Unlike traditional EAs, DE perturbs the current population members with the scaled differences of randomly selected and distinct individuals [22]. This way, in the first iterations, the elements are widely scattered in the search space and have a great exploration ability. During optimization, the individuals tend to concentrate in the regions of the search space with the best fitness values [23].

In DE, new individuals that will be part of the next generation are created by combining members of the current population. Every individual acts as a parent and is associated to a donor vector. In the basic version of DE, the donor vector V_i for the i^{th} parent (X_i) is generated, by combining three random and distinct population members (X_{r1}, X_{r2}) and X_{r3}, as follows:

$$V_i = X_{r1} + F \cdot (X_{r2} - X_{r3})$$

where F (scale factor) is a real-valued parameter that strongly influences DE's performances and typically lies in the interval $[0.4,\ 1]$. Other mutation strategies have been applied to DE, experimenting with different base vectors and different numbers of vectors for perturbation.

After mutation, every parent-donor pair generates an offspring O_i by means of a crossover operation. The newly-generated offspring is evaluated and its fitness and its parent's are compared. The better survives and will be part of the next generation.

2.3. Related Work

We shortly review some of the most relevant and recent approaches which combine deformable models and evolutionary computation techniques. In [24], "Genetic snakes" are active contour models with an optimization procedure based on genetic algorithms. In [25] a GA evolves a population of medial-based shapes, using prior shape knowledge to produce feasible deformations while also controlling the scale and localization of these deformations. In [26] a GA is used to perform level set curve evolution using texture and shape information to automatically segment the prostate in CT and MRI pelvic images. Finally, in [27], the authors describe a GA-based method that minimizes the Topological Active Nets energy, evaluating the segmentation results of a greedy, a genetic, and a memetic algorithm over seven synthetic

images and three CT scans of human bones.

Recently, PSO has been successfully used in conjunction with various types of deformable models. Asl and Seyedin [28] apply the technique proposed in [29] using PSO instead of a GA, obtaining similar results in terms of precision but in shorter time. In [30] the initial segmentation based on the level set method is refined using swarms of intelligent agents. Finally, in [31], a customized PSO algorithm overcomes the drawbacks of snakes, including initialization, concave boundaries, sensitivity to noise and local minima.

To the best of our knowledge, one of the few works that uses DE and deformable models for solving medical imaging problems is presented in [32], where DE, in combination with a greedy search algorithm, is used to evolve Topological Active Nets, a discrete implementation of an elastic mesh, for CT image segmentation.

3. DE-based hippocampus localization

Our accurate and automatic structure localization method consists of two phases:

1. *Best Reference Slice Selection*, implemented as a two-step affine registration method which: (i) determines the position of the section displayed in the target image according to a reference atlas, and (ii) extracts the region of interest (ROI) where the hippocampus is more likely to be located. Hence, it represents the initialization of the deformable

model: it determines which template should be used and where it should be applied.

2. *Structure Localization*, in which the point distribution model selected in the previous step is adapted by a stochastic optimization procedure based on DE, to fit the image region where the structure of interest (if any) is located.

3.1. Best Reference Slice Selection

The initialization of the Deformable Model, in terms of choosing a model and its starting position, is obtained using an atlas-based affine registration with the reference images of the ABA. The main idea is to find the sagittal reference slice of the atlas which best matches the target image. This phase produces two results: firstly, based on the information contained in the corresponding reference atlas image, it allows one to extract the Region of Interest (ROI) where the hippocampus is expected to be located. Secondly, it makes it possible to determine the position, within the brain, of the section represented in the target image. Consequently, one can select the corresponding point distribution model of the hippocampus (derived empirically as described for the ASMs), which is to be applied in the following step within the selected ROI. This can be achieved by the two-step matching method presented in [4].

The first step performs a "global" affine registration between the target image and each image in the reference atlas, considering the shapes of the

whole sections and ignoring the internal structures. Then, using the atlas, the ROI where the hippocampus is more likely to be located is extracted, and a "local" affine registration is performed in order to match the hippocampus in the reference and the target image. After these two registration steps, two similarity measures between the target image and each reference image are computed. The reference image that achieves a better trade off between these two measures is selected as the best reference.

In order to evaluate the correctness of this two-step approach, we computed the distance (in slices) between the slice selected by our system and the one suggested by the ABA on a test set of 320 images. In 45% of cases the slice selected was exactly the same, while in 43%, 10% and 2% of cases there was a difference of one, two, and three or more slices, respectively. In those cases in which our system selected a reference slice different from the one proposed by the ABA, we also compared the fitness values obtained by the hippocampus localization algorithm. The results of these experiments showed that there are no statistically significant differences between the two choices. Therefore, these results demonstrate that our initialization method is effective and can be used also in problems where no ground truth is available.

3.2. Hippocampus Localization

While in the standard ASMs external forces are driven by the contours in the image, in our approach we parametrically deform the model shape to

match as closely as possible the shape of the hippocampus in the region we want to locate. The model is moved and deformed, by altering its parametric representation using an optimization heuristic like DE or PSO, which maximizes a function which measures the similarity between the model and the object itself. Since, in our problem, color and shapes are greatly variable, the only common information on which we can rely is that all hippocampi have two substructures (*sg* and *sp*) whose shapes have more consistent features, besides being usually characterized by lower intensity values and higher color saturation with respect to the surrounding structures.

To enhance the contrast between *sg*, *sp* and the surrounding parts, we pre-process the ROI containing the hippocampus by stretching its histogram. We invert pixel intensities, to consistently use the convention by which, as for binarized images, the most relevant regions in the image have higher intensities. These operations are not applied to the original images of the ABA we show, to avoid altering their actual appearance.

3.2.1. Templates

Obviously, our main concern is to adapt a model that permits an efficient localization, but we want, as well, to obtain this goal with the simplest possible model, i.e. having the least possible number of points in each template. In this work, for every slice of the reference atlas, two templates (one composed of 7 points for *sg*, one with 8 points for *sp*) have been created by manually selecting significant corresponding landmarks in all images in a

training set.

A template aims at taking into account all possible positions and deformations with respect to the prototypical shape of the structure it represents. The template does not refer to the absolute positions of the points, but describes the relative positions (or shifts) between consecutive points in polar coordinates. This is uncommon for a DM, since the majority of the implementations use cartesian coordinates. Every template is composed of two parts:

- an "inner set" of points that lies on the anatomical part we want to locate;
- an "outer set" of points that lies just outside it, obtained by rigidly shifting the previous set.

For a proper localization, the "outer set" is fundamental; otherwise, using only the "inner set", a completely dark and large area would always achieve the highest values of the target function.

A template is fully described by four $2 \times n$ matrices, where n is the number of points in the template:

$$\mathbf{S} = \begin{bmatrix} \rho_{m1} & \vartheta_{m1} \\ \rho_{m2} & \vartheta_{m2} \\ \vdots & \vdots \\ \rho_{mn} & \vartheta_{mn} \end{bmatrix} \quad \mathbf{\Delta} = \begin{bmatrix} \Delta_{\rho 1} & \Delta_{\vartheta 1} \\ \Delta_{\rho 2} & \Delta_{\vartheta 2} \\ \vdots & \vdots \\ \Delta_{\rho n} & \Delta_{\vartheta n} \end{bmatrix}$$

$$\mathbf{L} = \begin{bmatrix} \rho_{l1} & \vartheta_{l1} \\ \rho_{l2} & \vartheta_{l2} \\ \vdots & \vdots \\ \rho_{ln} & \vartheta_{ln} \end{bmatrix} \quad \mathbf{U} = \begin{bmatrix} \rho_{u1} & \vartheta_{u1} \\ \rho_{u2} & \vartheta_{u2} \\ \vdots & \vdots \\ \rho_{un} & \vartheta_{un} \end{bmatrix}$$

S is the "standard template" and represents the standard coordinates of the inner set, $\mathbf{\Delta}$ is the displacement of the outer set with respect to the inner set, **L** and **U** are the minimum and maximum values allowed for every parameter that describes the inner set. It should be noticed that ρ_1 and ϑ_1 represent the positions of the first point with respect to the upper left corner of the image. After that point, proceeding row-wise, every (ρ, ϑ) pair represents the shifts of the subsequent point in terms of distance and angle with respect to the previous point, respectively. As a consequence, the number of parameters to optimize for each model is twice the number of the points in the template (14 parameters for *sg* and 16 for *sp*).

The matrices **S**, **L** and **U** have been computed by manually selecting the reference points in a training set. The first one is the median of the selected shifts and the other two are the minimum and the maximum values

observed in the training set, respectively. To improve the templates, a manual refinement of the parameters has been performed. The matrix $\mathbf{\Delta}$ has been manually built based on the observation of several hippocampi.

From a computational point of view, a model is a vector of $2 \times n$ elements that ranges within the "boundary templates" $\mathbf{L}$ and $\mathbf{U}$; its elements (coordinates) are evolved by a metaheuristic which optimizes the target function described in section 3.3. The aims with which the target function was defined are to match the model with the target while keeping it as close as possible to the shape of the standard template $\mathbf{S}$.

3.3. Target Function

As in classical deformable models, we consider the model to be subjected to external forces (driven by the image features) and internal forces (driven by the model itself).

The target function F, which is to be maximized, has three components: external energy E, internal energy I, and contraction factor C:

$$F = E - (I + C)$$

The external forces move (and deform) the model to maximize the intensity of pixels in the inner set, while minimizing the intensity of pixels in the outer set. For both sets, we evaluate the intensity of the image within a 3×3 neighborhood N_3 of all points in the model (Punctual Energy, PE) and in p intermediate points along the segment between two consecutive points

(Continuous Energy, CE), i.e.

$$PE = \sum_{i=1}^{n} [T(N_3(I_i)) - T(N_3(O_i))]$$

where n is the number of points in the model, $I_i = \{x_i, y_i\}$ is the i^{th} point of the inner set (in cartesian coordinates), $O_i = \{x_i + \Delta x_i, y_i + \Delta y_i\}$ is the i^{th} point of the outer set, $T(P)$ is the intensity of the image in P, if P is a point, or the average intensity, if P is a neighborhood; $\Delta x_i, \Delta y_i$ are the elements of $\mathbf{\Delta}$ in cartesian coordinates. As well,

$$\begin{aligned} CE &= \sum_{i=2}^{n} \sum_{j=1}^{p} T(I_{i-1} + \frac{j}{p+1}(I_i - I_{i-1})) \\ &- \sum_{i=2}^{n} \sum_{j=1}^{p} T(O_{i-1} + \frac{j}{p+1}(O_i - O_{i-1})) \end{aligned}$$

where p is the number of points to evaluate in every segment. In our case we set $p = 20$. The final external energy is computed as

$$E = \gamma_P \cdot PE + \gamma_C \cdot CE$$

where the weights γ_P and γ_C have been empirically set to 5 and 1 for all tests, respectively (see Figure 3).

The internal energy I, related with the forces that oppose the deformation of the model, is computed as:

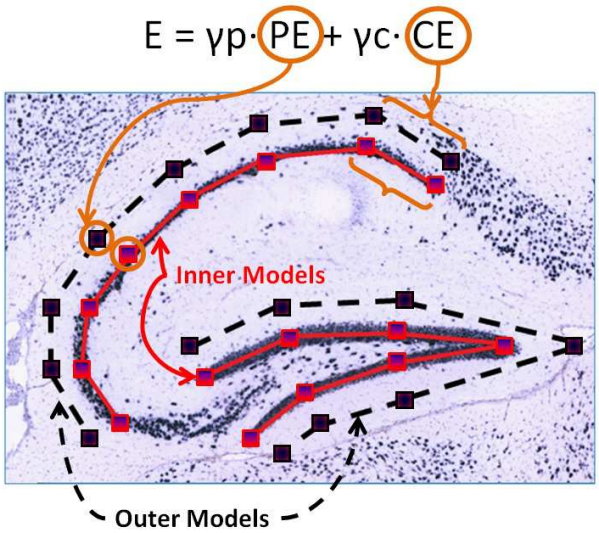


Figure 3: Inner and outer models superimposed to a hippocampus image, showing the components of the external energy.

$$I = \xi_\rho \cdot \sqrt{\sum_{i=2}^{n} (\rho_i - \rho_{mi})^2} + \xi_\vartheta \cdot \sqrt{\sum_{i=2}^{n} (\vartheta_i - \vartheta_{mi})^2}$$

where ξ_ρ and ξ_ϑ are two positive weights that regulate the deformability of the model. The higher their values, the less the model deforms and the more it keeps the shape of the standard template $\mathbf{S}$. Please notice that the index i starts from 2. This way, the first shift (which actually represents the starting position of the model) is not taken into account and the model is independent of its position in the image.

Finally, the contraction factor C also regulates the model's deformability to avoid unfeasible situations (e.g. having the extremes of sg too close to each other), and is defined as follows:

$$C = \xi_c \cdot \|I_n - I_1\|$$

If $\xi_c < 0$ the two extremes of the model repel each other, if $\xi_c > 0$ they attract each other. In our case we set $\xi_c > 0$ for the *sp* models and $\xi_c < 0$ for the *sg* models.

4. Experimental Results

To test the system we used both real and synthetic images. As for the former, we randomly selected 320 images (corresponding to 320 different genes) from the ABA within subsets of representative samples of all possible hippocampi, featuring high and low levels of gene expression, good-quality and low-quality images, and so on. We also created 20 synthetic images that represent simplified versions of the real ones. In these images, representing all reference slices, the hippocampus is made up of small circles having random radius and color; small and big ellipses were also added trying to simulate cells and, finally, gaussian ($\mu \in [0.0, 0.2]$ and $\sigma \in [0.01, 0.15]$) and salt and pepper noise (density $\in [0.05, 0.25]$) were introduced to add fuzziness to the images.

Tests were run on a computer equipped with a 64-bit Intel® Core™i7 CPU running at 2.67 GHz with 4 Gb of RAM.

We compared the results obtained by DE with those obtained by LM, a classical combination of the gradient descent and the Gauss-Newton methods, and the following stochastic optimization techniques: a variation of PSO,

used in [4], where it was shown to give better results than the standard PSO on this task; SS with local search based on the line search algorithm and the blx-α crossover, as described in [33]; GA and SA. In LM and SA we used, as initial value, the best solution in a randomly generated set of the same size as the other methods' populations.

The parameters (as well as the population topologies and the mutation and crossover operators) set in the tests have been initially chosen based on the most commonly used, and then refined during the development of the system. The values chosen for the most relevant parameters are shown in Table 1. For all techniques, a limit of 20000 function evaluations has been set as the only termination criterion, and the limits for the deformation were used as constraints in creating new solutions by crossover and mutation.

Table 1: Parameters used in testing different optimization techniques.

DE	PSO	SS	GA	SA	LM
$Cr = 0.9$ $F = 0.7$ Uniform Crossover Mutation: DE/target-to-best/1 Population Size = 80 Iterations = 250	$c_1 = 2.05$ $c_2 = 1.75$ $w_{max} = 1.0$ $w_{min} = 0.2$ Population Size = 80 Iterations = 250	$b_1 = 7$ $b_2 = 8$ $\alpha = 0.5$	$P_{crossover} = 0.6$ $P_{mutation} = 0.09$ Population Size = 80 Iterations = 250 Tournament of size 4 Uniform Crossover	$T_{start} = 1.5$ $T_{end} = 1E-9$ exponential cooling schedule: $T_{k+1} = 0.8 \cdot T_k$	$\lambda_0 = 0.01$

For every image, we first performed the extraction of the ROI and the selection of the best model, as described in Section 3.1. After this, for each optimization technique, we ran 25 tests on every image, for a total of 8000 experiments on real images and 500 on synthetic images.

Table 2: Comparative results of *sp* and *sg* localization for synthetic and real images (20000 function evaluations per experiment; the higher the values, the better the localization).

Synthetic Images						
sp localization						
Method	Average	StdDev	Avg. Worst	Avg. Best	Avg. Median	Wilcoxon test
DE	120.41	5.49	106.30	126.37	121.68	-
PSO	105.86	8.67	84.86	118.61	107.71	<1.00E-16
SS	105.70	6.16	91.66	115.69	106.37	<1.00E-16
GA	105.93	15.64	54.93	120.40	110.11	<1.00E-16
SA	101.80	8.19	80.87	110.64	103.76	<1.00E-16
LM	35.43	30.86	19.96	94.24	34.11	<1.00E-16
sg localization						
Method	Average	StdDev	Avg. Worst	Avg. Best	Avg. Median	Wilcoxon test
DE	128.91	2.92	121.18	132.99	129.49	-
PSO	118.29	5.85	105.55	127.63	119.05	<1.00E-16
SS	115.56	6.00	101.37	124.81	116.08	<1.00E-16
GA	115.03	9.67	87.86	127.20	116.49	<1.00E-16
SA	111.23	3.87	102.49	117.91	111.40	<1.00E-16
LM	64.76	19.58	33.79	103.34	61.42	<1.00E-16
Real Images						
sp localization						
Method	Average	StdDev	Avg. Worst	Avg. Best	Avg. Median	Wilcoxon test
DE	141.47	5.59	124.36	146.06	145.21	-
PSO	133.56	8.24	110.94	143.89	138.64	<1.00E-16
SS	131.06	6.04	118.54	141.66	135.81	<1.00E-16
GA	132.00	14.49	86.07	143.70	138.58	<1.00E-16
SA	131.00	3.05	123.83	137.71	134.91	<1.00E-16
LM	79.32	28.26	9.60	126.11	85.02	<1.00E-16
sg localization						
Method	Average	StdDev	Avg. Worst	Avg. Best	Avg. Median	Wilcoxon test
DE	148.24	1.54	143.58	149.73	149.30	-
PSO	144.65	3.69	135.37	148.70	145.52	<1.00E-16
SS	140.07	4.27	129.10	146.23	140.69	<1.00E-16
GA	140.89	6.83	122.56	148.50	143.06	<1.00E-16
SA	138.31	1.70	134.60	141.51	139.28	<1.00E-16
LM	105.58	20.01	65.54	141.61	110.52	<1.00E-16

4.1. Algorithm Comparison

Table 2 summarizes the results of our tests. In the upper part of the table we report the results for the 20 synthetic images and, in the lower part, the ones related to the 320 real images. The second column reports the average fitness for all tests, while the third reports the average standard deviation of fitness in the 25 experiments performed on each image. The fourth and fifth columns report the average of the worst and best results obtained for each image. Similarly, the sixth column reports the average of the medians.

We focused our analysis on the comparison between the performance of DE and of the other methods. DE achieved the best average fitness as well as the lowest standard deviation, which indicates a more robust behavior, as it is able to produce more consistent results over different trials; the performance of the other algorithms is more dependent on their random initialization. To assess the significance of this result, we performed a statistical test with confidence level of 0.01. We used non-parametric tests because the assumption of normality was not met. The Kruskal-Wallis test [34] was statistically significant in all cases (p-value $<$ 1.00E-16) and proved the existence of differences between the sets of results, where at least one sample median was significantly different from the others. After that, the paired Wilcoxon signed-rank test was performed [35], assuming as null hypothesis that the median of DE results is less than or equal to the median of the other methods. This is a one-tailed test in which the alternate hypothesis is that the median of any other method is less than the median of DE. The corre-

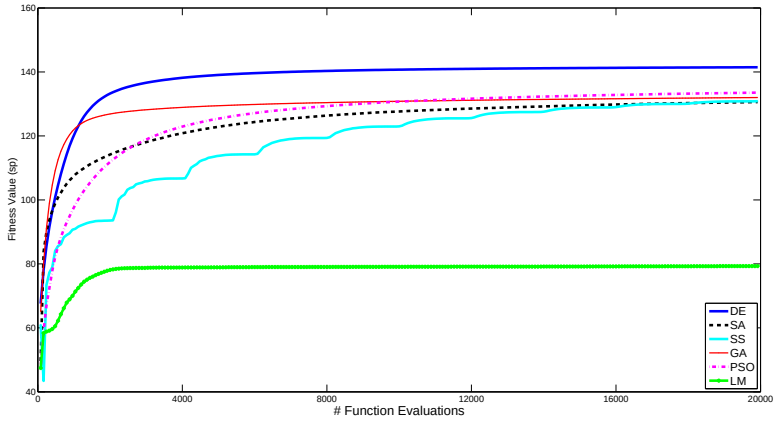
sponding p-values (shown in the last column of Table 2) reject, in all cases, the null hypothesis, showing that significant differences exist between the performance of DE and of the other methods. Moreover, the average worst values show that DE avoids local minima and is able to get good results more consistently than the other methods.

Figure 4 describes the behavior of the five metaheuristics and a gradient-based local search method by plotting the average fitness value versus the number of function evaluations. It can be seen how DE and GA start better than the other techniques but, while the former continues to improve its performance, the latter is not able to refine the solution found. At the same time, the evolution of SS is more discontinuous due to the premature convergence of the reference set, followed, after each reconstruction, by an abrupt improvement of the fitness values. The worst results were obtained by the gradient-based local search method (LM), due to the multimodality and non-differentiability of the function to optimize, providing a reason for using metaheuristics to solve this kind of problems.

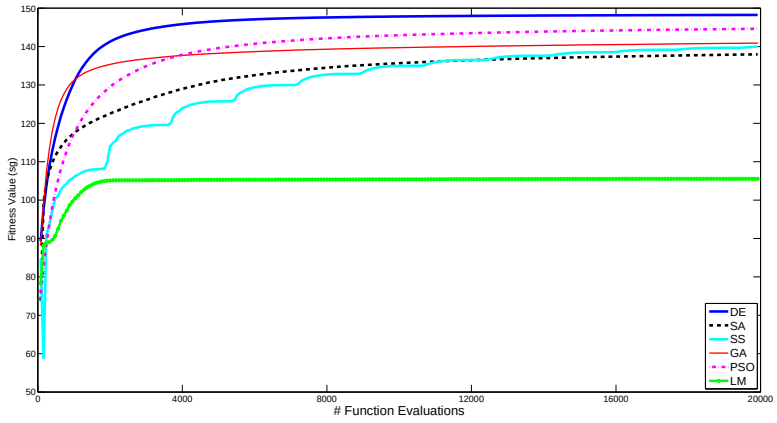
4.2. Overall System Performance

In a second test we evaluated the results of the entire localization method by dividing the outcomes into three quality classes (see Figure 5):

1. Perfect Match: all points of the two models are over the corresponding parts and cover them almost entirely;
2. Good Localization: (i) all points of the two models belong to the regions



(a) *sp*



(b) *sg*

Figure 4: Evolution of the fitness function for the localization of *sp* (above) and *sg* (below).

which must be detected, but they do not cover them entirely or (ii) at most two points are slightly outside of them;

3. Error: all other possibilities, from three or more misplaced points to models which are located in a completely different position of the brain.

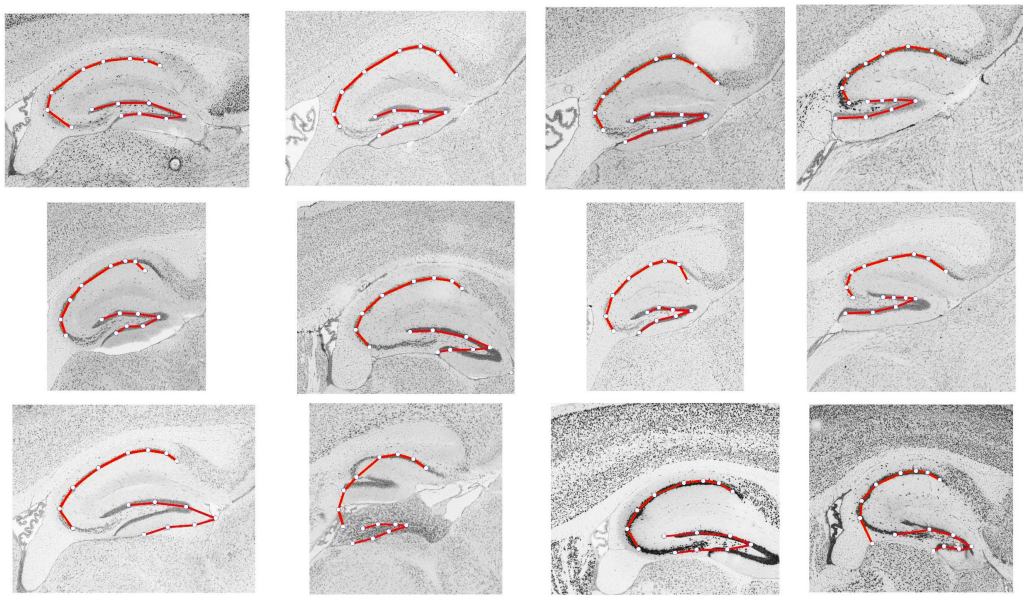


Figure 5: Upper row: perfect matches; middle row: good results; lower row: erroneous localizations.

If the results obtained by the model chosen in the Best Reference Slice Selection phase were not good enough (using a threshold for the target function values), we repeated the procedure using the model which ranked second in the previous phase, and took the best result.

Our method was able to perfectly localize the hippocampus in 58.0% of synthetic images and in 47.8% of real images, and reached a good localization in 35% and 43.1% of cases, respectively. This means that our method was

able to localize the hippocampus satisfactorily in 93.0% with synthetic images (20 images and 25 runs per image) and in 90.9% of cases with real images (320 images and 1 run per image).

The most problematic situations that appear to drive the system to bad localizations are substantially three:

- low quality of images (for instance, images that are damaged or have very low contrast): this can affect the results of ROI extraction or of the Best Reference Slice Selection phase;
- images where the hippocampus has low levels of gene expression and is surrounded by other anatomical parts with higher expression levels;
- hippocampi having shapes which differ substantially from the typical cases included in the training set.

Most of these errors could most probably be avoided by increasing the size of the training set for the point distribution models, or improving the preprocessing phase of the images.

On the other hand, the method we adopted shows very good performances when dealing with corrupted images of hippocampi.

With regard to time complexity, the localization task, implemented in C, employs on average 2.88 seconds per image, applying the Best Reference Slice Selection process on 10 reference slices.

5. Discussions and Conclusions

In this work we presented a two-step algorithm aimed at automatically localizing the hippocampus in histological images. The first phase roughly locates the anatomical structure by comparing the image under consideration with images taken from the ABA reference atlas. In the second phase, the structure is more precisely localized using a DE-based parametric deformable model that adapts its shape to match the anatomical structure of interest. The method is able to deal with imprecise and incomplete images, and, in our tests on actual images from the ABA, has been successful in 90.9% of cases.

Since classical gradient-based local search methods (like LM) are not able to solve this problem satisfactorily, due to its multimodality and high dimensionality, we have applied and studied different metaheuristics: DE, SA, GA, PSO, and SS. DE achieved the best results, for both average and standard deviation. This result can be explained by a better balance between exploration/diversification and exploitation/intensification and upholds the results obtained in other cases by DE on benchmarks containing multimodal non-separable functions [36]. Moreover, DE has shown greater robustness: it is able to reproduce good results more consistently over many trials, whereas the performance of other algorithms, like GA or SS, is more dependent on the stochastic initialization of individuals and parameters. DE is also, together with GA, the metaheuristic that converges faster to good solutions.

Possible future work includes automating the configuration step which

sets the algorithm parameters. Another interesting development could consider applying this method to other subcortical structures and anatomical planes in order to evaluate its general applicability. The automatic segmentation of the structure and the texture analysis of the segmented areas will lead, in a subsequent phase of the project within which this method was developed, to identifying sets of genes whose expression generates similar textural patterns in corresponding regions. We aim to cluster similar genes according to such visual features to detect correlations in their expression, since it can be argued that genes with similar expression patterns are also functionally correlated.

Acknowledgments

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

References

[1] K. A. Norman, How hippocampus and cortex contribute to recognition memory: revisiting the complementary learning systems model, Hippocampus 20 (2010) 1217–1227.

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

[3] M. J. Hawrylycz, et al., An anatomically comprehensive atlas of the adult human brain transcriptome, Nature 489 (7416) (2012) 391–399.

[4] R. Ugolotti, P. Mesejo, S. Cagnoni, M. Giacobini, F. Di Cunto, Automatic Hippocampus Localization in Histological Images using PSO-Based Deformable Models, in: Proc. Genetic and Evolutionary Computation Conference, GECCO '11, 2011.

[5] J. Kennedy, R. Eberhart, Particle Swarm Optimization, in: Proceedings of IEEE International Conference on Neural Networks, vol. 4, 1942–1948, 1995.

[6] K. Levenberg, A method for the solution of certain nonlinear problems in least squares, Quarterly of Applied Mathematics 2 (2) (1944) 164–168.

[7] S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, Optimization by Simulated Annealing, Science 220 (1983) 671–680.

[8] F. Glover, M. Laguna, M. Rafael, Scatter search, in: Advances in Evolutionary Computation: Theory and Applications., Springer-Verlag, 519–537, 2003.

[9] D. E. Goldberg, Genetic Algorithms in Search, Optimization and Ma-

chine Learning, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edn., 1989.

[10] D. Terzopoulos, K. Fleischer, Deformable Models, The Visual Computer 4 (1988) 306–331.

[11] D. Terzopoulos, A. Witkin, M. Kass, Constraints on Deformable Models: Recovering 3D Shape and Nonrigid Motion, Artificial Intelligence 36 (1988) 91–123.

[12] M. Kass, A. Witkin, D. Terzopoulos, Snakes: Active contour models, Int J of Computer Vision 1 (1988) 321–331.

[13] T. F. Cootes, C. J. Taylor, D. H. Cooper, J. Graham, Active shape models-their training and application, Comput. Vis. Image Underst. 61 (1995) 38–59.

[14] L. He, Z. Peng, B. Everding, X. Wang, C. Y. Han, K. L. Weiss, W. G. Wee, A comparative study of deformable contour methods on medical image segmentation, Image and Vision Computing 26 (2008) 141–163.

[15] T. Heimann, H.-P. Meinzer, Statistical shape models for 3D medical image segmentation: a review, Medical Image Analysis 13 (2009) 543–563.

[16] N. Hansen, A. Ostermeier, Completely Derandomized Self-Adaptation in Evolution Strategies, Evolutionary Computation 9 (2001) 159–195.

[17] F. Herrera, M. Lozano, J. L. Verdegay, Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis, Artificial Intelligence Review 12 (1998) 265–319.

[18] K. Deb, A. Anand, D. Joshi, A computationally efficient evolutionary algorithm for real-parameter optimization, Evolutionary Computation 10 (2002) 371–395.

[19] A. E. Eiben, J. E. Smith, Introduction to Evolutionary Computing, Springer Verlag, 2003.

[20] A. P. Engelbrecht, Computational Intelligence: An Introduction, Wiley Publishing, 2nd edn., 2007.

[21] R. Storn, K. Price, Differential Evolution- A Simple and Efficient Adaptive Scheme for Global Optimization over Continuous Spaces, Tech. Rep., International Computer Science Institute, 1995.

[22] S. Das, P. Suganthan, Differential Evolution: A Survey of the State-of-the-Art, IEEE Transactions on Evolutionary Computation 15 (2011) 4–31.

[23] F. Neri, V. Tirronen, Recent advances in differential evolution: a survey and experimental analysis, Artif. Intell. Rev. 33 (2010) 61–106.

[24] L. Ballerini, Genetic snakes: active contour models by genetic algo-

rithms, in: Genetic and Evolutionary Computation in Image Processing and Computer Vision, EURASIP Book Series on SP & C, 177–194, 2007.

[25] C. McIntosh, G. Hamarneh, Medial-based Deformable Models in Non-convex Shape-spaces for Medical Image Segmentation using Genetic Algorithms, IEEE Trans. on Medical Imaging 31 (1) (2012) 33–50.

[26] P. Ghosh, M. Mitchell, J. A. Tanyi, A. Hung, A Genetic Algorithm-Based Level Set Curve Evolution for Prostate Segmentation on Pelvic CT and MRI Images, in: Biomedical Image Analysis and Machine Learning Technologies: Applications and Techniques, IGI Global, 127–149, 2010.

[27] Ó. Ibáñez, N. Barreira, J. Santos, M. G. Penedo, Genetic approaches for topological active nets optimization, Pattern Recognition 42 (5) (2009) 907–917.

[28] M. Asl, S. Seyedin, Active Contour Optimization using Particle Swarm Optimizer, in: Information and Communication Technologies, ICTTA '06, vol. 1, 1522–1523, 2006.

[29] L. MacEachern, T. Manku, Genetic algorithms for active contour optimization, in: Proc. IEEE International Symposium on Circuits and Systems, ISCAS '98, vol. 4, 229–232, 1998.

[30] D. Feltell, L. Bai, 3D level set image segmentation refined by intelligent

584 agent swarm, in: Procs. of IEEE Congress on Evolutionary Computa-
585 tion, CEC '10, IEEE, 1–8, 2010.

586 [31] E. Shahamatnia, M. Ebadzadeh, Application of particle swarm opti-
587 mization and snake model hybrid on medical imaging, in: Procs. of
588 IEEE Internation Workshop on Computational Intelligence In Medical
589 Imaging, 1 –8, 2011.

590 [32] J. Novo, J. Santos, M. G. Penedo, Topological Active Models opti-
591 mization with Differential Evolution, Expert Systems with Applications
592 39 (15) (2012) 12165–12176.

593 [33] F. Herrera, M. Lozano, D. Molina, Continuous scatter search: An anal-
594 ysis of the integration of some combination methods and improvement
595 strategies, European Journal of Operational Research 169 (2006) 450–
596 476.

597 [34] W. H. Kruskal, W. A. Wallis, Use of Ranks in One-Criterion Variance
598 Analysis, Journal of the American Statistical Association 47 (1952) 583–
599 621.

600 [35] F. Wilcoxon, Individual comparisons by ranking methods, Biometrics
601 Bulletin (1945) 80–83.

602 [36] J. Vesterstrom, R. Thomsen, A comparative study of differential evolu-
603 tion, particle swarm optimization, and evolutionary algorithms on nu-

merical benchmark problems, in: Proc. Congress on Evolutionary Computation, CEC '04, 1980–1987, 2004.