

Segmentation of Histological Images using a Metaheuristic-based Level Set Approach

Pablo Mesejo
Stefano Cagnoni
Dept. of Information Engineering
University of Parma, Italy
{pmesejo,cagnoni}@ce.unipr.it

Alessandro Costalunga
Davide Valeriani
Dept. of Information Engineering
University of Parma, Italy
{alessandro.costalunga,
davide.valeriani}@studenti.unipr.it

ABSTRACT

This paper presents a two-phase method to segment the hippocampus in histological images. The first phase represents a training stage where, from a training set of manually labelled images, the hippocampus representative shape and texture are derived. The second one, the proper segmentation, uses a metaheuristic to evolve the contour of a geometric deformable model using region and texture information.

Three different metaheuristics (real-coded GA, Particle Swarm Optimization and Differential Evolution) and two classical segmentation algorithms (Chan & Vese model and Geodesic Active Contours) were compared over a test set of 10 histological images. The best results were attained by the real-coded GA, achieving an average and median Dice Coefficient of 0.72 and 0.77, respectively.

Categories and Subject Descriptors

I.5.4 [Pattern Recognition]: Applications—*Computer Vision*; I.2.10 [Artificial Intelligence]: Vision and Scene Understanding

General Terms

Algorithms, Experimentation

Keywords

Hippocampus, Differential Evolution, Real-Coded Genetic Algorithms, Particle Swarm Optimization, Histological Images, Level Set method, Image Segmentation

1. INTRODUCTION

Image segmentation is defined as the partitioning of an image into non-overlapping regions that are homogeneous with respect to some visual feature such as intensity, color or texture [25]. In particular, image segmentation plays a crucial

role in many biomedical imaging applications by automating and facilitating the delineation of anatomical structures and different regions of interest (ROIs). As one can easily check taking a look at the literature [10, 11, 24], most of the segmentation algorithms developed so far have been applied to MRI and CT images, not paying too much attention to other imaging modalities like ultrasound or microscopy. Even when dealing with histological images, it is surprising to check that most literature about their processing and analysis is focused on registration and 3D reconstruction of the whole brain with very few papers about segmentation. This paper presents a method aimed at segmenting anatomical structures in histological images using deformable models and metaheuristics.

Among the different anatomical structures which make up the mammalian brain, the hippocampus is particularly interesting due to its crucial role in learning and memory processes [22]. Moreover, it has recently been demonstrated that the volume of the hippocampus is an early biomarker for Alzheimer's disease. Therefore, there is great interest in understanding the cellular and molecular events that take place in this structure, under both normal and pathological conditions.

In particular, we consider the problem of segmenting the hippocampus in histological images extracted from the Allen Brain Atlas (ABA): a huge, publicly available image database, which has recently provided scientists with a gene-expression map for future study and investigation. The ABA contains a genome-scale collection of histological images (cellular resolution gene-expression profiles) obtained by In Situ Hybridization (ISH) of serial sections [1] of mouse brains.

As in most medical imaging applications, the main problems to be tackled are related with the characteristics of the images involved in the segmentation:

- natural variability of brain structure shapes in different subjects;
- fuzziness of the hippocampus boundary;
- grained patterns with many irregularities, which hamper the classification of individual pixels in the anatomical structures under consideration;
- rotation or displacement of the imaged structures on the slice with respect to a "standard" alignment;
- variable brightness of the images within the same set;

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

GECCO'13 Companion, July 6–10, 2013, Amsterdam, The Netherlands.
Copyright 2013 ACM 978-1-4503-1964-5/13/07 ...\$15.00.

- presence of artefacts: tears, scraps, bubbles, streaks in tissues, partial cut-off of regions; and
- large image size (the typical resolution of ABA images is about $15,000 \times 7,000$ pixels, and the ROIs about $2,500 \times 2,000$ pixels).

A procedure based on Active Shape Models and Random Forests (ASM+RF) [20, 21] was recently developed to accurately segment these images. The results obtained were very promising in comparison with many other segmentation techniques [19] but, despite good performance, they also highlighted some drawbacks:

- impossibility to deal with topological changes in a natural manner;
- ad hoc nature, since the procedure needs a training set of shapes to manually create the parametric template and its possible deformations, as well as a training set of textural patterns for the ensemble classifier used to refine the results of the segmentation; and
- long and complex pipeline of four independent stages (initialization of the deformable model, localization of the anatomical structure of interest, segmentation, and refinement of the segmentation using classifiers). Therefore, a more integrated and compact approach would be desirable.

The main idea of this paper is to overcome some of these problems using another kind of deformable model (the level set method) that can: a) easily handle topological changes of the contours, b) be adapted to solve large dimensional problems without great effort, and c) easily determine the areas inside and outside an active contour.

The work developed here is based on the one by Ghosh et al in [6, 7] which, in turn, was inspired by [28]. It is important to underline that this research represents one of the very few cases in which metaheuristics have been used to optimize a geometric deformable model, let alone their application to histological images, and that the current work also presents significant differences with respect to the papers mentioned above:

- an intensity-based term from Chan & Vese approach [3] has been included in the fitness function, trying to combine a region-based approach with prior knowledge about texture and shape;
- in order to adapt this approach to the particular nature of histological images, the textural part has also been critically modified: we do not use Laws' textural measures or Gabor wavelet transform-based features, like in the original paper, but textural features extracted from the Gray Level Co-occurrence Matrix, as explained in [20];
- in relation to the genetic operators used, the single-point crossover used in [6] has been replaced by a real-coded one like the BLX- α , due to the nature of the GA chromosomes; and
- the comparison has been extended to Particle Swarm Optimization (PSO) and Differential Evolution (DE), converting this paper into one of the few examples of

the application of these two stochastic techniques to geometric deformable models, and tests on histological images (instead of usual MRI and CT) have been performed.

The paper is structured 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 segmentation of the hippocampus. Finally, section 4 presents results on 10 real images with appropriate statistical tests, followed, in section 5, by some final remarks and a discussion about possible future developments.

2. BACKGROUND

This section briefly introduces some of the fundamental computer vision and soft computing techniques applied in this paper, as well as the main contributions to the field found in literature.

The classical gradient search techniques are effective when the problem to optimize satisfies tight constraints. However, when the search space is discontinuous, noisy, high-dimensional, or multimodal, then metaheuristics offer a number of attractive features: no requirement for a differentiable or continuous objective function, global search capability, or virtually no need of specific information about the problem to solve.

2.1 Particle Swarm Optimization

PSO [17] is a bio-inspired optimization algorithm based on the simulation of the social behavior of bird flocks. In the last fifteen years PSO has been applied to a very large variety of problems and numerous variants of the algorithm have been presented.

During the execution of PSO, a set of particles moves within a (fitness) function domain searching for its optimum (best fitness value). The motion of each particle is driven by the best positions visited so far by the particle itself and by the best positions found by the entire swarm (*gbest* PSO) or by particles in some pre-defined neighborhood (*lbest* PSO). Consequently, each particle relies both on "individual" and on "swarm" intelligence, and its motion can be described by the following two simple equations which update the particles' position and velocity:

$$\begin{aligned}
 P_n(t) &= P_n(t-1) + v_n(t) \\
 v_n(t) &= w \cdot v_n(t-1) \\
 &+ c_1 \cdot rand() \cdot (BP_n - P_n(t-1)) \\
 &+ c_2 \cdot rand() \cdot (BLP_n - P_n(t-1))
 \end{aligned}$$

where $P_n(t)$ and $v_n(t)$ are the position and velocity of the n^{th} particle in iteration t ; c_1 , c_2 and w (inertia factor) are positive constants; $rand()$ returns random values uniformly distributed in $[0, 1]$; BP_n is the best-fitness location visited so far by the particle and BLP_n is the best-fitness location visited so far by any particle of a neighborhood which may be as large as the current swarm: in this case, this position corresponds to the global best.

In particular, we have used a version of PSO already presented in [29], where the inertia factor w adapts its value to the fitness function of each particle and a particle is re-initialized in case of stagnation.

2.2 Differential Evolution

DE, first introduced by Storn and Price [26], has recently been shown to be one of the most successful Evolutionary Algorithms (EAs) applied to global continuous optimization. Unlike traditional EAs, DE perturbs the current population members with the scaled differences of randomly selected distinct individuals. In the first iterations, the elements are widely scattered in the search space and have great exploration ability while, as optimization proceeds, the individuals tend to concentrate in the regions of the search space with better values, such that the search automatically focuses on the most promising areas. In DE, every element acts as a parent vector and, for each of them, a donor vector is created. In the original version of DE, the donor vector for the i^{th} parent (X_i) is generated by combining three random distinct elements X_{r1} , X_{r2} and X_{r3} . The donor vector V_i is calculated as follows:

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

where F (scale factor) is a parameter that strongly influences DE's performances and typically lies in the interval $[0.4, 1]$. After mutation, every parent-donor pair generates an offspring (the so-called trial vector) by means of a crossover operation. Two kinds of crossover are typically used: binomial (also called uniform) and exponential. Besides F , the crossover rate Cr is another parameter which regulates the searching behavior of DE.

2.3 Real-Coded Genetic Algorithms

Genetic Algorithms (GAs) are stochastic, parallel search algorithms based on the mechanics of natural selection [8]. GAs were designed to efficiently search large, non-linear, poorly-understood search spaces where expert knowledge is scarce or difficult to encode and where traditional optimization techniques fail. They are flexible, robust, and exhibit the adaptiveness of biological systems.

GAs process a population of solutions by three operations: selection, crossover and mutation. In the initial formulation, the solutions were binary encoded; however, other encoding types have been taken into account for the representation. In particular, real encoding seems particularly natural when optimizing parameters in continuous domains. In that case, a chromosome is a vector of real numbers, each of which is one of the parameters to be optimized. GAs based on real-number representations are called real-coded GAs (RCGAs).

There is a compelling evidence indicating that classical discrete crossover operators (DCOs), i.e., all the crossover operators used for binary encoding which are directly applicable to real coding (like the simple, two-point and uniform crossover operators), are ineffective for RCGAs [12, 13]. Since the crossover operators that exploit the numerical nature of the real coding (aggregation-based operators, like arithmetic or geometric, and neighborhood-based crossover operators, like BLX- α or SBX) consistently outperform the classical DCOs, these kinds of operators have been considered in this work.

2.4 Deformable Models

Deformable Models (DMs) [27] 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 image regions surrounding the curve. 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.

Basically DMs can be parametric or geometric. The former represent curves and surfaces explicitly in their parametric forms during deformation, leading to a fast interaction with the model and a compact representation of the object to segment. On the other hand, geometric DMs can handle topological changes naturally, since they are based on the theory of curve evolution and represent curves implicitly as the level set of a higher-dimensional scalar function [23]. Among the parametric DMs we may mention: Active Shape Models [4], Deformable Templates [31], Topological Active Nets [15] and Active Contour Models ("snakes") [16]. Examples of geometric DMs are Geodesic Active Contours (GAC) [2] and the Chan&Vese model (CV) [3].

Although originally developed for computer vision applications to natural scenes and computer graphics problems, the potential of deformable models in medical image analysis has already been proven [10, 11].

2.5 Related Work

We shortly review some of the most relevant and recent approaches which combine geometric deformable models and evolutionary computation techniques. In [7], 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. In [19], an Active Shape Model optimized by DE is used as the initial contour for two well-known geometric DMs (CV and GAC). In [5], the initial segmentation based on the level set method is refined using swarms of intelligent agents. Finally, in [14], a GA is used to find an optimal set of parameters for a level set method used to segment MRI and CT.

3. OVERVIEW OF THE SYSTEM

As explained in the introduction, our work essentially adapts the ideas introduced by Ghosh and Mitchell in [6] to the segmentation of the hippocampus in histological images. Such an adaptation solves some of the problems detected in the original proposal and takes into account the specificities of the case under study.

On the one hand, we introduce a real-value crossover operator conforming to the nature of our chromosomes (see Section 2.3). On the other hand, we use textural features that differ from the initially proposed ones, since it has been shown that the Grey-level Co-occurrence Matrix (GLCM) obtained very good results with the same image modality [20]. Due to the huge size of these images, the textural segmentation is only applied for a limited number of iterations, to guide the search to promising regions. Finally, the structure orientation (pose) is not taken into account during optimization because of two main reasons: it increases the temporal and computational cost of the optimization, adding a burden that is not justified by the results obtained.

The method consists of two phases: training and proper segmentation (see Figure 1). In the first one, from a set of manually segmented training images we compute the average shape, the main modes of variation, and the median texture, trying to characterize an "ideal" hippocampus. An implicit representation of the segmenting curve is created by applying principal component analysis to a collection of

signed distance representations of the training data. During segmentation, the metaheuristic searches for the best weights to linearly combine the mean shape and shape variabilities, and the texture enclosed by the evolving contour is compared with the “representative” texture.

3.1 Training Phase

3.1.1 Shape

First of all, the shape priors are derived from the training set. To do so, each contour from the training data is represented as the zero level set of the signed distance function $\psi_i(x,y)$, $i = 1 \dots n$, where (x,y) are the pixel coordinates, and n is the number of training contours used to assess shape variability. The signed distance function is the shape representation: the boundaries of each shape are embedded as the zero level set of a signed distance function with negative distances assigned to the inside and positive distances assigned to the outside of the object.

The mean level set function is defined as:

$$\bar{\Phi}(x,y) = \frac{1}{n} \sum_{i=1}^n \psi_i(x,y)$$

Mean offset functions are then derived by subtracting the mean from the signed distance representations of the training contours:

$$\tilde{\psi}_i = \psi_i - \bar{\Phi}$$

Assume that the images are of size $N = N_1 \times N_2$. Given the huge size of the images under consideration, all images have been resized to $N_1 \times N_2 = 500 \times 500$ pixels. The columns of the mean offset functions ($|N| = N_1 \times N_2$) are then serially stacked to form one column vector β_i of size $1 \times N$. The shape variability matrix S (of size $N \times n$) is obtained from these n column vectors:

$$S = [\beta_1, \beta_2, \beta_3, \dots, \beta_n]$$

The variance in shape is then computed by an eigenvalue decomposition on this shape variability matrix:

$$\frac{1}{n} S S^T = U \Sigma U^T$$

where U is an $N \times n$ matrix whose columns represent n orthogonal modes of shape variation, Σ is an $n \times n$ diagonal matrix of eigenvalues, and the columns of $U = [u_i]$ are the corresponding eigenvectors. In this paper, instead of computing the eigenvectors of such a large matrix, we have considered the smaller one, $\frac{1}{n} S^T S$, because it is more computationally efficient [18] to obtain the n different eigen-shapes $\{\Phi_1, \Phi_2, \Phi_3, \dots, \Phi_n\}$.

After this procedure, introduced in [28], the mean shape and shape variability derived as described are used to define a level set function that implicitly represents the segmenting curve:

$$\Phi[w] = \bar{\Phi} + \sum_{i=1}^k w_i \Phi_i$$

Thus, the task of the metaheuristic will be to find the values of w that minimize a fitness function to be defined in the test phase.

3.1.2 Texture

A texture is a set of visual elements occurring in some regular or repeated pattern. GLCM was introduced by Haralick [9] and is a feature-based method that characterizes a texture as a homogeneous distribution of feature values. A co-occurrence matrix describes how often a gray level appears in a specified spatial relationship to another gray level. The entry at (i, j) of the GLCM indicates the number of occurrences of the pair of gray levels i and j which are a

distance d apart along a given direction θ . The values of d and θ are parameters for constructing the GLCM.

The textural priors used here are the same as in [20], which yielded very good performance. The training patterns, as well as the test values calculated during the segmentation, 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 measures: contrast, correlation, energy and homogeneity from the GLCM, using (1,1) as spatial relationship (i.e. $\theta = 315^\circ$ and $d=1$ pixel), employing a window of size 30×30 pixels.

The main idea is to capture the textural essence of the hippocampus and compare, during the segmentation phase, the texture enclosed by the evolving surface of our DM with the “ideal” texture of the training set. To create this median texture, p points per image in the training set were randomly selected inside the hippocampus, and t textural patterns were extracted using those points as pixels-of-interest. This procedure creates a $(t \cdot p) \times n$ matrix, n being the number of training images, and $t \cdot p$ the number of textural features multiplied by the number of selected points. In this case, $t = 11$ and $p = 100$. From this matrix, the median is calculated to obtain a “general representation” of the texture in the hippocampus: all points which are closer to these values are assumed to belong to the hippocampus.

3.2 Test Phase

The test phase corresponds with the proper segmentation of the object of interest. The implicit representation of the contour is deformed guided by a metaheuristic which tries to fit the boundaries of the hippocampus. To do so, such a metaheuristic stochastically generates weights to combine the mean and the variability to create new shapes.

The fitness function combines region- and texture-based terms. The former represents the Chan&Vese model [3], while the latter is the euclidean distance between the texture enclosed by our contour and the median texture found in our training set.

The energy functional to minimize is the following:

$$\begin{aligned} F(i_1, i_2, C) = & \alpha \cdot (\mu \cdot \text{Length}(C) + \nu \cdot \text{Area}(\text{inside}(C))) \\ & + \lambda_1 \int_{\text{inside}(C)} |u_0(x,y) - i_1|^2 dx dy \\ & + \lambda_2 \int_{\text{outside}(C)} |u_0(x,y) - i_2|^2 dx dy \\ & + \beta \cdot (d(t(C), \bar{T})) \end{aligned}$$

where u_0 is the image to segment, formed by two regions (internal and external), C is the evolving curve, i_1 and i_2 are constants depending on C (pixel intensity averages of u_0 inside and outside C , respectively), and $d(t(C), \bar{T})$ the euclidean distance between the median texture extracted from the training set (\bar{T}) and the actual texture enclosed by the evolving contour ($t(C)$). Using this functional, our model simultaneously takes into account intensity and texture criteria. In this case, both terms have been weighted equally ($\alpha = \beta = 1$).

Finally, a fast refinement step is applied. This refinement sequentially applies 50 iterations of a local implementation of Chan&Vese [30], that takes into account only a neighbourhood of the boundary, and the removal of connected components whose area is smaller than a threshold (in this case 500 pixels).

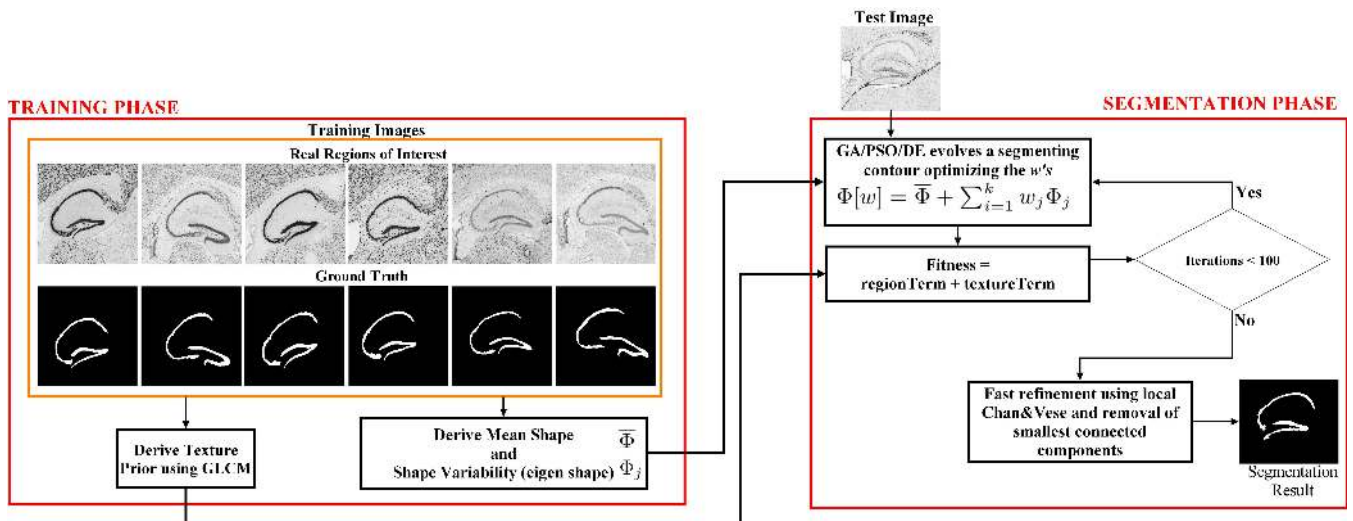


Figure 1: General Overview of the Hippocampus Segmentation System

4. EXPERIMENTAL RESULTS

To check the performance of the new method, 25 and 10 manually segmented images were used as training and test set, respectively. The 10 images selected as test are representative of the problem and include different scenarios (see Figure 2). Every segmentation method was run 15 times per image (excluding C&V and GAC because they are deterministic methods and always produce the same result). The performance of GA (LS-GA), PSO (LS-PSO) and DE (LS-DE) was compared with the one obtained by two classical geometric DMs: a region-based method (CV) and an edge-based one (GAC). With respect to the population and the number of iterations used in the metaheuristic, the same configuration as the original paper was used (50 individuals, 100 iterations), and the images were resized to 500×500 pixels.

Standard segmentation metrics were used to evaluate performance and the results were analyzed using proper statistical tests. The Jaccard similarity index (JI) and the Dice coefficient (DSC) measure set agreement: a value of 0 indicates no overlap with the ground truth, and a value of 1 indicates perfect agreement. In turn, the Hausdorff distance (HD) represents a measure of the spatial distance between two sets of points: it is the largest of all the distances from any point in the resulting segmentation to the closest point in the ground truth.

In order to perform a fair comparison with the deterministic/classical methods, the latter have been intensively tuned using an exhaustive search among the most commonly used parameter values over the complete dataset (training and test sets). Also, for these two methods, the input images were pre-processed using a median filter (5×5 in CV and 10×10 in GAC) and, as post-processing, the removal of connected components smaller than 500 pixels was also applied.

Since the normality and homoscedasticity assumptions were not accomplished, as checked through the application of Kolmogorov-Smirnov and Bartlett's tests, non-parametric tests were used. To check the statistical significance of the results obtained, a Friedman test was performed with a level of confidence of 0.01 for the null hypothesis that all samples

are drawn from distributions with the same median. Since the p-value was near zero, the statistical test suggested that, at least, one sample median is significantly different from the others. Pairwise statistical differences were studied using Friedman test with the Tukey-Kramer correction, and real-coded GA was found to be the best method among the ones in the comparison.

Some conclusions can be derived from the numerical data in Table 2 and the visual information contained in Figure 3. The first conclusion could be the difficulty of tackling satisfactorily these images, since established and well-known segmentation methods, like C&V and GAC, did not obtain as good results as one could expect. In particular, GAC presents problems when the difference, in terms of gradient values, between the hippocampus and the background is small. In fact, between the two classical methods, the region-based one (CV) obtained better results than the one relying on edge information. This can be justified by the fuzzy boundaries of the hippocampus that may have hampered the performance of GAC. Also, the dissatisfactory results obtained by CV can be explained by the complex background of these images, since it is well-known that the performance of CV is better when dealing with images where foreground and background determine two regions with clearly different average intensity levels.

Considering the stochastic approaches, the best results were obtained by LS-GA, which appeared to be the best in all metrics. Real-coded GA obtained the best results in all images but in one, while the worst results were obtained by DE. Obviously, this is only a preliminary study, so the convergence problems of DE and PSO, which in some cases are not able to obtain good results, should be investigated more deeply. For example, the PSO topology used here has been the global best; tests with other topologies should be performed to check if there is some improvement in the results. Also, in the case of DE, tests with other crossover operators could be run to study the influence of this factor in the general performance of the segmentation algorithm. Finally, experiments with a larger population size or a larger number of iterations should be tested, considering the huge search space these techniques have to explore.

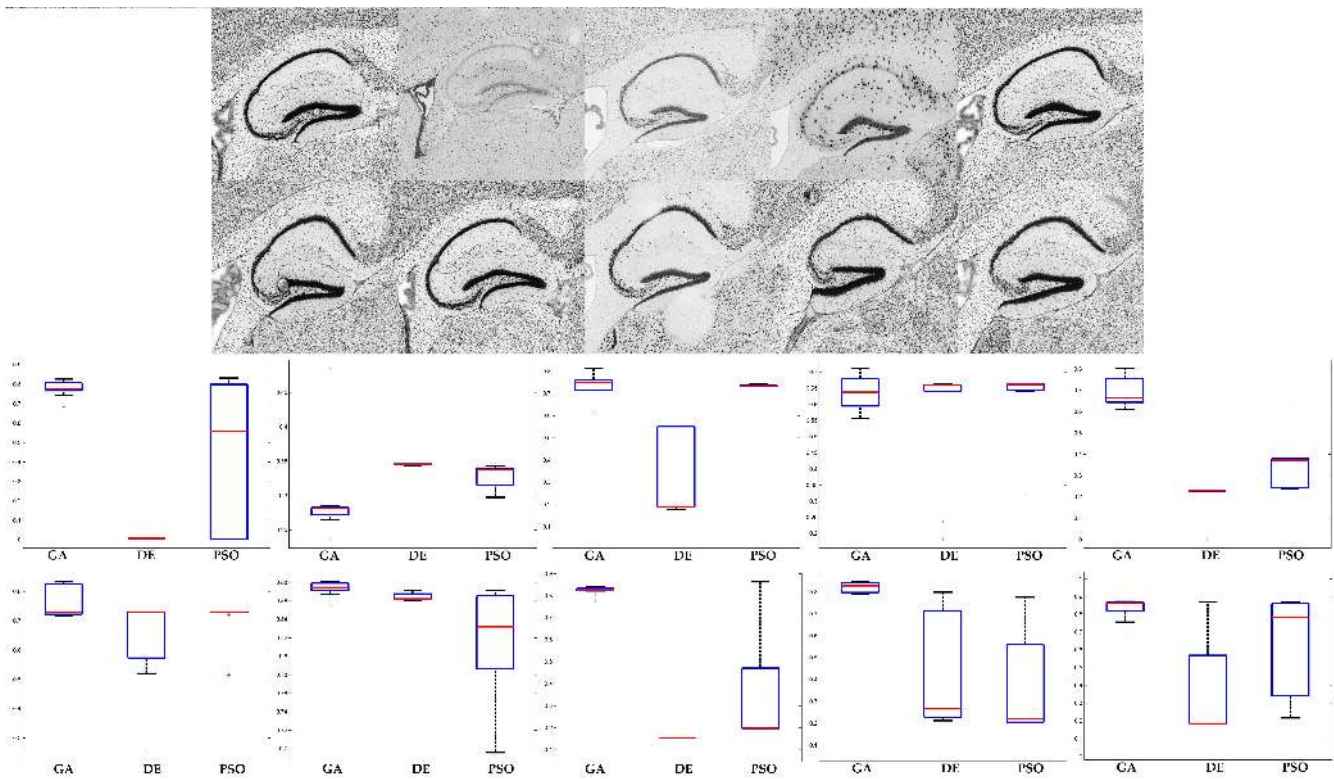


Figure 2: Test set used in experiments and boxplots of the DSC results obtained per image by the stochastic methods under study. Each plot refers to the image having the corresponding position in the upper row.

From the values of HD, interesting conclusions can also be drawn. It is interesting to check how our shape-based approach has a smaller (best) average and median HD with respect to C&V and GAC. This could be justified by the absence in the latter of any kind of shape restriction, that impose that the pixels are segmented only taking into account intensity and boundary information, respectively.

5. CONCLUSIONS AND FUTURE DEVELOPMENTS

This paper investigates hybridizations between metaheuristics and the level set method, and represents a first approach to the segmentation of histological images using eigenshapes, whose linear combination is optimized using PSO, DE and real-coded GA. On the one hand, metaheuristics can optimize the energy function of the deformable model or find the most suitable parameters for such a model. On the other hand, the level set method represents an elegant solution for the main drawbacks of parametric deformable models, like the possibility of managing topological changes in a natural manner. With these concepts in mind, this paper studies ways of using a training set of shapes and textures to solve a difficult problem, like segmenting the hippocampus in histological images, and compare different global search optimization techniques. Finally, this work represents one of the very few cases in which PSO and DE have been used to optimize the level set method, and also one of the few examples of application of geometric DMs to the segmentation of histological images.

It is important to remark that the work presented here

is a first approach that should be refined. First, it is slow compared to classic approaches based on only one feature (intensity, edge)¹. Obviously, if someone is looking for a fast method this would not be the best option: what this method can offer is accuracy. For instance, segmentations obtained by LS-GA have a median Dice coefficient which is three times better than GAC. Also, a more sophisticated use of the textures could be taken into account (for example, differentiating the two structures of the hippocampus), or including more prior information and restrictions (for instance, the relative position of the hippocampus with respect to other organs). In any case, this initial approximation has shown good performance using three different standard metrics and the results obtained have been better than the ones obtained by very well-established segmentation algorithms.

6. ACKNOWLEDGMENTS

Pablo Mesejo is funded by the European Commission through the MIBISOC project (“Medical Imaging using Bio-Inspired and Soft Computing”, MIBISOC Marie Curie Initial Training Network, FP7 PEOPLE-ITN-2008, GA n. 238819).

¹In order to give some reference in this regard, the average time to perform one run with our metaheuristic-based approach on a MacBook Pro i7 dual core @ 2,7GHz with 8GB of RAM, and using MATLAB as programming environment, it is around 10 times slower than C&V and GAC, that take approximately 5 minutes.

Table 1: Parameters used in testing the different algorithms. The values for LS-GA, LS-DE and LS-PSO were based on the literature and on a brief empirical study about the suitability of different combinations of parameters.

LS-DE	LS-PSO	LS-GA	C&V	GAC
$Cr = 0.9$ $F = 0.7$ Uniform Crossover DE/target-to-best/1	$w_{min} = 0.2$ $w_{max} = 1.0$ $c_1 = 2.05$ $c_2 = 1.75$	$Cr = 0.9$ $BLX = 0.3$ $Mut = 0.09$ <i>Tournament</i> (4)	number of iterations = 500 $\mu = 0.01$ (length term) $\nu = 0$ (area term) $\lambda_1 = \lambda_2 = 1$	number of iterations = 500 $\beta = -1$ (expansion weight) $\alpha = 3$ (contour weight)

Table 2: Segmentation results using three different metrics: Dice Similarity Coefficient (DSC), Jaccard Index (JI), and Hausdorff Distance (HD).

		Methods				
		LS-GA	LS-DE	LS-PSO	C&V	GAC
DSC	Average	0.7204	0.4145	0.5383	0.5044	0.3966
	Median	0.7695	0.3446	0.6622	0.5294	0.2758
	Std	0.1672	0.2842	0.2614	0.1546	0.2151
	Best	0.8806	0.8716	0.8711	0.6648	0.7950
	Worst	0.2363	0.0059	0.0015	0.1107	0.2404
JI	Average	0.5871	0.3058	0.4120	0.3485	0.2710
	Median	0.6271	0.2081	0.4950	0.3600	0.1601
	Std	0.1745	0.2515	0.2378	0.1212	0.1962
	Best	0.7867	0.7725	0.7717	0.4979	0.6597
	Worst	0.1340	0.0029	0.0007	0.0586	0.1366
HD	Average	420.8674	617.6873	549.0358	762.4849	622.4544
	Median	176.0064	372.8288	338.5550	553.3118	507.1295
	Std	518.0279	549.3174	541.2218	354.1919	242.5857
	Best	28.3456	45.0627	45.0627	403.6087	395.1266
	Worst	1.7878e+003	2.1191e+003	1.7196e+003	1.4290e+003	1.0341e+003

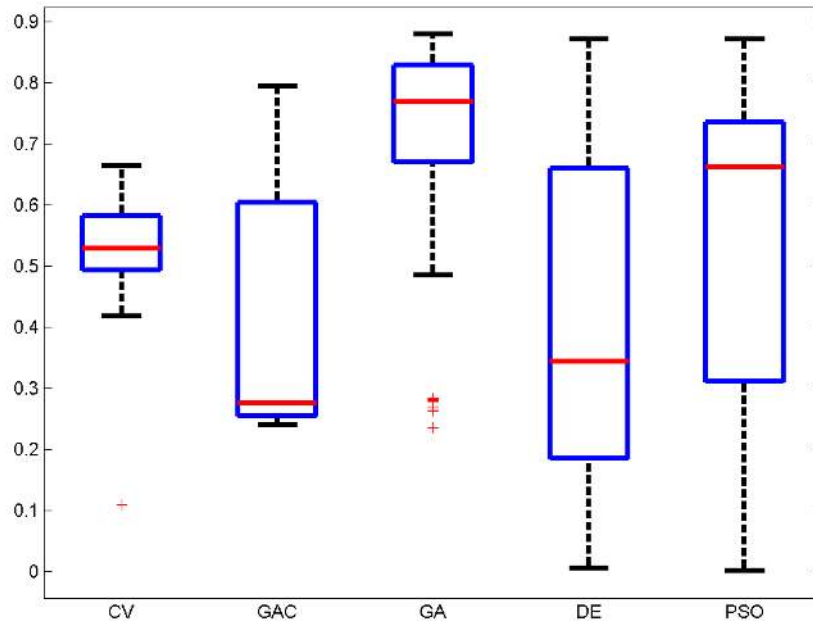


Figure 3: Box-plots representing the global DSC results obtained by the five methods compared over the whole dataset.

7. REFERENCES

- [1] Allen Institute for Brain Science. Allen Reference Atlases. <http://mouse.brain-map.org>, 2004-2006.
- [2] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International Journal of Computer Vision*, 22:61–79, 1997.
- [3] T. Chan and L. Vese. Active contours without edges. *IEEE Trans. on Image Processing*, 10(2):266–277, 2001.
- [4] T. F. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham. Active shape models-their training and application. *Comput. Vis. Image Underst.*, 61:38–59, 1995.
- [5] D. Feltell and L. Bai. 3d level set image segmentation refined by intelligent agent swarm. In *Proc. of IEEE Congress on Evolutionary Computation, CEC '10*, pages 1–8, 2010.
- [6] P. Ghosh and M. Mitchell. Segmentation of medical images using a genetic algorithm. In *Proc. Genetic and Evolutionary Computation Conference, GECCO '06*, pages 1171–1178, 2006.
- [7] P. Ghosh, M. Mitchell, J. A. Tanyi, and 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*, pages 127–149. IGI Global, 2010.
- [8] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edition, 1989.
- [9] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural Features for Image Classification. *IEEE Trans. on Systems, Man and Cybernetics*, SMC-3(6):610–621, 1973.
- [10] L. He, Z. Peng, B. Everding, X. Wang, C. Y. Han, K. L. Weiss, and W. G. Wee. A comparative study of deformable contour methods on medical image segmentation. *Image and Vision Computing*, 26(2):141–163, 2008.
- [11] T. Heimann and H.-P. Meinzer. Statistical shape models for 3D medical image segmentation: a review. *Medical Image Analysis*, 13(4):543–563, 2009.
- [12] F. Herrera, M. Lozano, and A. M. Sánchez. A taxonomy for the crossover operator for real-coded Genetic Algorithms: An experimental study. *International Journal of Intelligent Systems*, 18:309–338, 2003.
- [13] F. Herrera, M. Lozano, and J. L. Verdegay. Tackling real-coded genetic algorithms: Operators and tools for behavioural analysis. *Artificial Intelligence Review*, 12:265–319, 1998.
- [14] M. Heydarian, M. Noseworthy, M. Kamath, C. Boylan, and W. Poehlman. Optimizing the Level Set Algorithm for Detecting Object Edges in MR and CT Images. *IEEE Trans. on Nuclear Science*, 56(1):156–166, 2009.
- [15] O. Ibáñez, N. Barreira, J. Santos, and M. G. Penedo. Genetic approaches for topological active nets optimization. *Pattern Recognition*, 42(5):907–917, May 2009.
- [16] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.
- [17] J. Kennedy and R. Eberhart. Particle Swarm Optimization. In *Proc. IEEE International Conference on Neural Networks*, volume 4, pages 1942–1948, 1995.
- [18] M. Leventon. *Statistical Models for Medical Image Analysis*. PhD thesis, MIT, 2000.
- [19] P. Mesejo and S. Cagnoni. An experimental study on the automatic segmentation of in situ hybridization-derived images. In *Proc. on 1st International Conference on Medical Imaging using Bio-Inspired and Soft Computing (MIBISOC'13)*, 2013. In Press.
- [20] P. Mesejo, R. Ugolotti, S. Cagnoni, F. Di Cunto, and M. Giacobini. Automatic Segmentation of Hippocampus in Histological Images of Mouse Brains using Deformable Models and Random Forest. In *Proc. on 25th International Symposium on Computer-Based Medical Systems (CBMS'12)*, pages 1–4, 2012.
- [21] P. Mesejo, R. Ugolotti, F. Di Cunto, M. Giacobini, and S. Cagnoni. Automatic hippocampus localization in histological images using Differential Evolution-based deformable models. *Pattern Recognition Letters*, 34(3):299–307, 2013.
- [22] K. A. Norman. How hippocampus and cortex contribute to recognition memory: revisiting the complementary learning systems model. *Hippocampus*, 20:1217–1227, 2010.
- [23] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations. *Journal of Computational Physics*, 79(1):12–49, 1988.
- [24] C. Petitjean and J.-N. N. Dacher. A review of segmentation methods in short axis cardiac MR images. *Medical Image Analysis*, 15(2):169–184, 2011.
- [25] D. L. Pham, C. Xu, and J. L. Prince. Current Methods in Medical Image Segmentation. *Annual Review of Biomedical Engineering*, 2:315–337, 2000.
- [26] 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.
- [27] D. Terzopoulos and K. Fleischer. Deformable models. *The Visual Computer*, 4(6):306–331, 1988.
- [28] A. Tsai, A. J. Yezzi, W. M. W. III, C. M. Tempny, D. Tucker, A. C. Fan, W. E. L. Grimson, and A. S. Willsky. A Shape-Based Approach to the Segmentation of Medical Imagery Using Level Sets. *IEEE Trans. on Medical Imaging*, 22(2):137–154, 2003.
- [29] 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*, pages 487–494, 2011.
- [30] X.-F. Wang, D.-S. Huang, and H. Xu. An efficient local Chan-Vese model for image segmentation. *Pattern Recognition*, 43(3):603–618, 2010.
- [31] Y. Zhong and A. K. Jain. Object tracking using deformable templates. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 22:544–549, 2000.