Toward a Better Integration of Spatial Relations in Learning with Graphical Models

Emanuel Aldea and Isabelle Bloch

Abstract This paper deals with structural representations of images for machine learning and image categorization. The representation consists of a graph where vertices represent image regions and edges spatial relations between them. Both vertices and edges are attributed. The method is based on graph kernels, in order to derive a metrics for comparing images. We show in particular the importance of edge information (i.e. spatial relations) in the specific context of the influence of the satisfaction or non-satisfaction of a relation between two regions. The main contribution of the paper is situated in highlighting the challenges that follow in terms of image representation, if fuzzy models are considered for estimating relation satisfiability.

1 Introduction

Generic machine learning algorithms do not cope with complex data such as images directly, a preprocessing step being usually required in order for them to perform various tasks. Among the solutions used to adapt image data to algorithm inputs, we discuss in this article a representation method as a structure which encodes explicitly image parts and spatial interactions in a graphical model.

Discriminative learning algorithms that are well suited for this kind of graphical models have been created [1] and optimized [2] in view of specific applications in computational chemistry and biology. An adaptation for coping with graphical models extracted from images is required nevertheless, since the properties of the

TELECOM ParisTech, TSI Department
CNRS UMR 5141 LTCI
46 rue Barrault, Paris 75013, France
emanuel.aldea@telecom-paristech.fr
isabelle.bloch@telecom-paristech.fr

information encoded in the graphical structure is fundamentally different than in the context of biological or chemical structured data analysis.

In a larger context, image interpretation methods use primarily the visual features of low-level or high-level interest elements. However, spatial information concerning the relative positioning of these elements is equally beneficial, as it has been shown previously in segmentation and structure recognition. Therefore, an interest for the integration of spatial information in the learning framework has emerged recently. The fact that spatial information is often perceived and expressed in a manner which is close to natural language, along with the fact that the absence of a spatial interaction is also relevant, hint at the usefulness of fuzzy spatial information for image representation. Fuzzy representations actually permit to assess at the same time the imprecision degree of a relation (e.g., “close to” or “to the left of”) and the gradual transition between the satisfiability and the non-satisfiability of a relation.

The objective of this article is to explore the limits of spatial information representation and its integration in the learning process within the context of image classifiers that make use of graph kernels. In the first part of our work, we present the advantages that labeled graphs provide for representing images, along with the general learning strategy employed by the corresponding SVM classifier. We continue with a short reminder on the use of spatial information in some related graph representations, and on the particularities of spatial information for image representation. The results show that spatial information complements the visual features of distinctive elements in images and that adjusting the kernel functions for the fuzzy spatial representations is beneficial in terms of performance.

2 Knowledge representation by labeled graphs

In the domain of machine learning, generic supervised statistical algorithms accept input data in the form of numerical arrays or sequences and return a numerical value or indicate a specific category. Nowadays, input data are increasingly provided in complex configurations: as trees, graphs or other relational structures. These data arise very often from health and life sciences, but also from image processing, reasoning models for forecasting and decision making, etc. We witness accordingly the apparition of complex tasks that require the extraction of relations and structural dependencies out of input data. This situation suggests the emergence of learning methods adapted for these tasks and coping with large quantities of data.

In a structured data representation by graphical models, vertices may represent for instance atoms [1], simple chemical structures with specific properties [3], proteins [4], segmentation regions in images [5, 6], while edges encode specific interactions such as interdependence and scheduling, or spatial relations (adjacency, distance, relative localization, topology). In the context and particularly for image processing tasks, key sources of imprecision must be taken into account, concerning the objects and their imprecise delimitation and the relative essence of the interaction information, often depicted using natural language. The graph structure and

labeling integrate therefore the information that we possess concerning the basic elements that form the input objects, their features, the interactions among them but equally our confidence level for these types of information.

In the case of an image, one possible approach for the extraction of a graphical model is by building an adjacency graph upon the output of image segmentation. Graph vertices are associated to image regions and are labeled according to specific region features, related to size, color, texture. Usually, these numerical values are continuous, as opposed to discrete values that we may encounter in other applications (a chemical symbol, a protein identification reference or a nucleotide). The only structural information being used is the region adjacency, implicitly encoded by the graph edges. Extensions of this basic graphical model take into account more complex spatial and topological information using a richer labeling of the edges.

The next step consists in using a Support Vector Machine (SVM) [7] to classify the structures that were extracted. Given a positive definite function K, denoted as the *kernel function* of the classifier, a set of training objects $\mathscr{X}$ and a set of labels $\mathscr{Y}$ associated to the elements of $\mathscr{X}$, such that $y_i \in \{-1,+1\}$ for any $x_i \in \mathscr{X}$, the output of the classifier for a new object x is:

$$y(x) = \operatorname{sgn}\left(\sum_{i=1}^{|\mathscr{X}|} \alpha_i y_i K(x_i,x)\right) \tag{1}$$

where α_i is the Lagrange multiplier in the optimization solution associated to the training object x_i.

An important observation is that the classifier only needs the value of the kernel function between pairs of examples, as a similarity estimation. An additional advantage of this approch is that it allows classifying elements issued from spaces which are not naturally endowed with inner products (such as graph, tree or string spaces), as long as we use a valid kernel function.

Furthermore, we describe the specific marginalized kernels that are being used in labeled graph analysis.

2.1 Marginalized Kernels

Given a generic class of objects $\mathscr{X}$, we assume that the constituents $x \in \mathscr{X}$ are generated according to a latent variable model which consists of the visible variable x and of a hidden variable θ, being considered jointly in a pair $z = [\theta, x]$. As we need a kernel $K(x,x')$ for the visible variables, we define first a *joint kernel* $K_z(z,z')$ for the mixed pair, which is used in a *marginalized kernel* [8] defined as the expectation of the joint kernel over all the values of the hidden variable:

$$K(x,x') = \int_{\theta,\theta' \in \Theta} p(\theta|x)p(\theta'|x')K_z(z,z')d\theta d\theta' \tag{2}$$

where Θ refers to the domain of the hidden variable. In a discrete setting, the value of the marginalized kernel is estimated with:

$$K(x,x') = \sum_{\theta,\theta' \in \Theta} p(\theta|x)p(\theta'|x')K_z(z,z') \tag{3}$$

The difficulties to be considered when estimating marginalized kernels are the computational burden which is related to the dimension of Θ, and the estimation from the data of the probabilistic model $p(\theta|x)$. Therefore, the choice of the model $p(\theta|x)$ should maximize the relevance for the specific data $x \in \mathscr{X}$ under the tractability constraint of $K(x,x')$. With respect to the properties of the function $K(x,x')$, as long as the joint kernel $K_z(z,z')$ is positive semidefinite, the kernel $K(x,x')$ is also positive semidefinite, since the class of positive semidefinite kernels is closed under addition and multiplication [9]; the kernel may also be interpreted as the inner product of the two vectors $p(\theta|x)$ and $p(\theta'|x')$.

2.2 Building SVM classifiers for graphs

The graph similarity is assessed using a marginalized kernel function and is employed in a SVM classifier. This similarity, related to a specific feature a, between two graphs G and G' extracted from images is evaluated with a kernel that sums the similarities between all possible pairs of random walks in the two graphs [1], weighted by their probability of apparition.

In reference to other applications that used this type of marginalized graph kernel [2], the labeling space for vertex features is continuous and multidimensional. The similarity function for feature values has to be less discriminative than the Dirac delta function usually employed in the discrete case. Therefore, we use Gaussian kernels with variance σ^2 in order to evaluate the similarity $k_a^{rbf}(a_1,a_2)$ between two values a_1 and a_2 of the numerical feature a :

$$k_a^{rbf}(a_1,a_2) = \exp\left(-\frac{\|a_1-a_2\|^2}{2\sigma^2}\right) \tag{4}$$

Specific kernels have been shown to be adapted for other types of features. These kernels usually employ a well known distance between features that they kernelize; examples include a χ^2-kernel between histograms :

$$k_{\chi^2}(h_1,h_2) = e^{-\alpha\chi^2(h_1,h_2)} \tag{5}$$

or a L1/L2 distance based kernel between multichannel mean color levels. In case of texture features, we use a distance metric defined in [10] on the means and standard deviations of Gabor filter energy responses [11].

Given the two graphs G and G' to compare, Equation (4) is used to evaluate the similarity $k_{v,e}(h,h')$ between two random walks $h = \{x_1,\ldots,x_n\}$ in G and $h' =$

$\{x'_1,\ldots,x'_n\}$ in G', by combining the similarity functions k_v for a vertex feature v and k_e for an edge feature e along h and h':

$$k_{v,e}(h,h') = k_v(v_{x_1}, v_{x'_1}) \prod_{i=2}^{n} k_e(e_{x_{i-1}x_i}, e_{x'_{i-1}x'_i}) k_v(v_{x_i}, v_{x'_i}) \tag{6}$$

This general equation may be simplified if we take into account only a region feature (as it happens with adjacency graphs) or if we take into account only an edge feature. In order to simplify the computation, we fix the value of the missing function to 1 (however, the element that is not considered for similarity computation must exist, otherwise a random walk containing the element could not exist).

At this point, we can underline the link between our specific kernel and the formal marginalized model depicted in Section 2.1. The input graph G represents the visible variable, and the random walk h represents the hidden variable. Therefore, the graph kernel between G and G' is computed by adding the similarities between all possible random walk kernels h and h', weighted by their probability of apparition:

$$K_{v,e}(G,G') = \sum_{h} \sum_{h'} k_{v,e}(h,h') p(h|G) p(h'|G') \tag{7}$$

This function is subsequently used with a support vector machine (SVM) in order to build an image classifier. The matrix $K_{v,e}$ defines the similarity between the graphs to compare.

3 Spatial relations in the context of graph learning

The spatial context has been taken into account in computational biology and chemistry when representing structured data. With few exceptions [3], spatial relations being used are binary and model rigorously the presence of an interaction between two elements of the structure. Even under this binary relation model, it has been shown that information brought by the absence of interactions may increase prediction performance, in relevant applications. For example, in protein-protein interaction (PPI) networks the absence of protein interactions is relevant for disease prediction. Therefore, a complement graph $\tilde{G}$ of the initial interaction graph G, which encodes the absence of interactions, has been proposed [4]. The resulting composite kernel:

$$K^*(G,G') = K(G,G') + K(\tilde{G},\tilde{G}') \tag{8}$$

leads to noteworthy improvements in classification accuracies on disease outcome prediction for cancer patients.

As to the extraction of spatial information for image representations, the situation is more complex. First of all, spatial interactions present an inherent semantic variability which goes well beyond the binary case mentioned above. Secondly, the integration of fuzzy spatial information and region feature information turns out to

be more complicated than the direct method depicted in Equation 8. However, this type of spatial information has been shown to enrich the description of images and to be useful for segmentation and structure recognition purposes. Below, we examine how we can use spatial information in learning and, more specifically, categorization.

We could just add fuzzy information on the existing edges of the image representation graph, but using strict adjacency for the underlying structure may pose robustness issues. Indeed, in cases when adjacency relies on a small number of pixels, the resulting graph may differ according to the segmentation method. Adding edges that represent more than the implicit strict adjacency relation does not only help with encoding structural information, but at the same time improves the robustness of the representation.

For our application, we use a topological spatial relation represented by an extended degree of adjacency, described below. Note that other relations could be added as well, using the same framework.

3.1 Distance between regions.

The distance between two regions R_1 and R_2 is computed as the minimal Euclidean distance between two points $p_i \in R_1$ and $q_j \in R_2$:

$$d(R_1, R_2) = \min_{p_i \in R_1, q_j \in R_2} (d_{Euclidean}(p_i, q_j)) \tag{9}$$

Distance, as well as orientation, may not always be relevant, for instance the distance between two regions is the same if those two regions are adjacent by only one pixel, or if a region is surrounded by another region. Therefore we propose to consider a topological feature that measures the adjacency length between two regions.

3.2 Adjacency measure based on fuzzy satisfiability.

One way to estimate this measure is to compute the matching between the area "near" a reference region and another region. This measure is maximal in the case where the reference region is embedded into the second, and is minimal if the two regions are far away from each other.

Fuzzy representations are appropriate to model the intrinsic imprecision of several relations (such as "near") and the necessary flexibility for spatial reasoning [12]. We define the region of space in which a relation to a given object is satisfied as a fuzzy set. The membership degree of each point to this fuzzy set corresponds to the satisfiability degree of the relation at that point [12]. Note that this representation

is in the image space and thus may be more easily merged with a representation of another region.

The spatial relation "near" is defined as a distance relation. A distance relation can be defined as a fuzzy interval f of trapezoidal shape on $\mathbb{R}^+$. A fuzzy subset μ_d of the image space $\mathscr{S}$ can then be derived by combining f with a distance map d_R to the reference object R: $\forall x \in \mathscr{S},\ \mu_d(x) = f(d_R(x))$, where $d_R(x) = \inf_{y \in R} d(x,y)$. In our experiments, the fuzzy interval f is defined with the following fixed values: 0, 0, 10, 30 (Figure 1a). We exemplify using a butterfly image (Figure 1b) and the result of a segmentation exhibiting four distinct regions (Figure 1c). We illustrate the distance map to the region represented by the left wing (Figure 1d) and the fuzzy subset corresponding to the relation "near the left wing" (Figure 1e) which uses the distance map and the fuzzy interval defined above. Similarly, we compute fuzzy subsets for the right wing (Figure 1f) as well as for any other regions designated by the segmentation.

So far we have defined the area of the image in which the relation "near to" a reference object is defined. The next step consists in estimating the matching between this fuzzy representation and the other region. Among all possible fuzzy measures, we choose as a criterion a *M-measure of satisfiability* [13] defined as:

$$Sat(near(R_1), R_2) = \frac{\sum_{x \in \mathscr{S}} \min(\mu_{near(R_1)}(x), \mu_{R_2}(x))}{\sum_{x \in \mathscr{S}} \mu_{near(R_1)}(x)} \tag{10}$$

where $\mathscr{S}$ denotes the spatial domain. It measures the precision of the position of the object in the region where the relation is satisfied. It is maximal if the whole object is included in the kernel of $\mu_{near(R_1)}$. Note that the size of the region where the relation is satisfied is not restricted and could be the whole image space. If the object R_2 is crisp, this measure reduces to $\frac{\sum_{x \in R_2} \mu_{near(R_1)}(x)}{\sum_{x \in \mathscr{S}} \mu_{near(R_1)}(x)}$, i.e. the portion of $\mu_{near(R_1)}$ that is covered by the object R_2.

3.3 Adjacency measure based on fuzzy resemblance.

Beside satisfiability, we also choose a symmetric measure, the *M-measure of resemblance* [13] defined as :

$$Res(near(R_1), R_2) = \frac{\sum_{x \in \mathscr{S}} \min(\mu_{near(R_1)}(x), \mu_{R_2}(x))}{\sum_{x \in \mathscr{S}} \max(\mu_{near(R_1)}(x), \mu_{R_2}(x))}$$

This measure is maximal if the object and the relation are identical: this resemblance measure accounts for the positioning of the object and for the precision of the fuzzy set as well.

In Figure 1(e) and Figure 1(f) we have illustrated the fuzzy subsets corresponding to the two wings. With the fuzzy satisfiability measure defined above, we get a response of 0.100 for "right wing near the left wing" and 0.109 for "left wing near

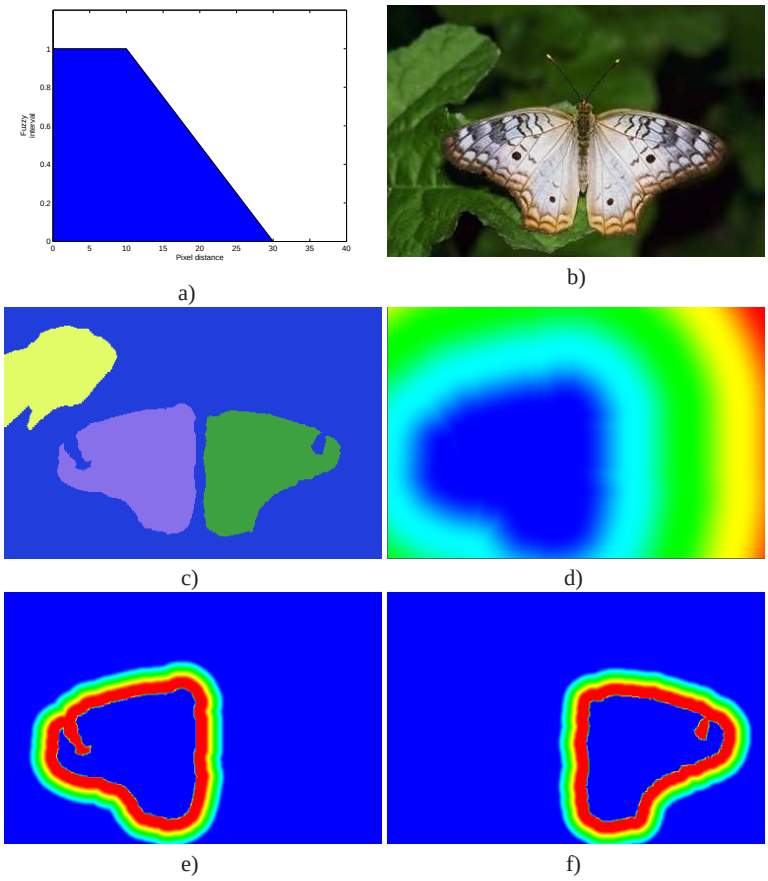


Fig. 1 (a) Fuzzy interval for the distance relation. (b) Input image. (c) Segmentation result, four distinct regions. (d) Distance map to the region represented by the left butterfly wing. (e) Fuzzy subset corresponding to the relation "near the left wing" (red corresponds to highest values). (f) Fuzzy subset corresponding to the relation "near the right wing".

the right wing". It is equally worth noting that the two regions are disconnected with respect to the strict pixel adjacency.

In the remaining sections, we will denote by a spatial relation R one of these measures of fuzzy adjacency, but we underline the fact that R could be substituted for other functions that estimate the interaction between elements of the image structure. The choice of the spatial relation of adjacency for our illustration is immediate because fuzzy adjacency information extends naturally one of the most simple and pertinent relations between image regions, the strict adjacency. However, taking into account more complex spatial relations such as "parallel to" or "along", along with their fuzzy measures of satisfiability [14, 15], is possible as long as these spatial relations are appropriate for the content of the input images.

4 Fuzzy spatial information and discriminative models

We can see that fuzzy spatial relations for images extend very conveniently the binary relations that are being used in other domains and achieve to merge at the same time information concerning the presence and the absence of an interaction. However, there are difficulties that arise when using this type of relations for discriminative learning.

In the context of image representation using space relations, related work has been done using binary relations supported by a specific ontology [16], using a count vector which estimates simple relative positioning [17], or using fuzzy spatial relations [18]. Each independent spatial relation builds in itself a novel data representation, therefore additional work may be necessary in order to make use of different spatial features simultaneously and efficiently. In this part of the article, we focus on the fact that a single fuzzy spatial relation creates *by itself* an infinite set of different representations. Rather than using multiple spatial relations for learning, we try to underline the specific challenges that a discriminative learning algorithm has when using a family of representations generated by the same fuzzy spatial relation.

Very often, there is a correlation between the value of a fuzzy spatial relation and the information gain: if the response is high, it means that the relation that the function has been designed for is much present. Consequently, low responses may be frequent (e.g. in the case of the "near" spatial relation) and may not bring the same amount of information. Discriminative learning, and discriminative learning for labeled graphs in particular, makes intensive use of similarity assessments between input objects. The similarity score between two graphs increases if these graphs exhibit many similar substructures. Complete graphs must be used if we compute a spatial relation value between all possible regions; therefore, if very low relation values are frequent (close to, or equal to 0), the graph kernel function will overincrease the graph similarity measure.

A straightforward solution to this situation is to threshold the spatial relation values, so that edges will exist only when the fuzzy adjacency estimation between two vertices is beyond a minimum value θ. However, the strict adjacency graph does not necessarily belong to this set $\mathscr{G}$ of threshold graphs. In terms of *graph edit distance* [19], let us consider for an image the adjacency graph G and an element $G_\theta \in \mathscr{G}$:

$$G_\theta = \{\mathscr{V}(G); (v_1; v_2) \in \mathscr{V}^2(G) | R(v_1, v_2) \geq \theta\} \tag{11}$$

where $R(v_1, v_2)$ is the generic spatial relation function between regions (vertices) v_1 and v_2.

Obviously, the vertex sets of G and G_θ are identical, as representations of the same image segmentation. We denote by $\mathscr{V}(G)$ the vertex set of graph G, and by $\mathscr{E}(G)$ its edge set. Under these circumstances, the graph edit distance $d_{g.e.}(G, G_\theta)$ between G and G_θ is generated by strictly adjacent regions that are not close enough in terms of θ-fuzzy adjacency and non strictly adjacent regions that are close in terms of θ-fuzzy adjacency:

$$d_{g.e.}(G,G_\theta) = \text{card}\{(v_1;v_2) \in \mathscr{V}^2(G) | (v_1,v_2) \in \mathscr{E}(G) \wedge R(v_1,v_2) < \theta\} + \text{card}\{(v_1;v_2) \in \mathscr{V}^2(G) | (v_1,v_2) \notin \mathscr{E}(G) \wedge R(v_1,v_2) \geq \theta\} \quad (12)$$

In practice this means that if we extend the spatial information beyond the intuitive strict adjacency between regions, we get for each spatial relation R a graph set $\mathscr{G}$ instead of a single graph representation of the image, and learning using spatial relations should be adapted to this situation. More precisely, we should know which of these graphs is more appropriate for learning.

The graph or graphs $G_* \in \mathscr{G}$ that minimize $d_{g.e.}$ are the closest (structurally) to the adjacency graph G. These are the projections of G in the set $\mathscr{G}$, and are ideally robust generalizations of G with respect to the spatial relation R. However, G_* and G might still exhibit various differences (the edge sets $\mathscr{E}(G)$ and $\mathscr{E}(G_*)$ are not identical), therefore the structural information within might still be partially disjunct.

The element that bridges the informational gap between G and $\mathscr{G}$ is the complete graph G_f, which includes (structurally) any element $G_\theta \in \mathscr{G}$, as well as the strict adjacency graph G. Ideally, the learning algorithm should exhibit the best performance with G_f, but then it should be able to cope well with the noise generated by a lot of similar low-information edge labels.

5 Experiments and results

5.1 Data set

The Internet Brain Segmentation Repository(IBSR) data set [1] contains real clinical data and is a widely used 3D healthy brain magnetic resonance image (MRI) database. It provides eighteen manually-guided expert brain segmentations, each of them being available for three different views, along reference planes: axial, sagittal and coronal. Each element of IBSR is a set of slices that cover the whole brain.

The main purpose of the data set is to provide a tool for evaluating the performance of segmentation algorithms. However, the fact that it is freely available and that it offers high quality segmentations as input data makes it also useful for our experiments.

5.2 Experimental setup

Image categorization between images belonging to different views in the data set (sagittal, coronal, axial) is performed with a 100% success rate for many of the

[1] The MR brain data sets and their manual segmentations were provided by the Center for Morphometric Analysis at Massachusetts General Hospital and are available at http://www.cma.mgh.harvard.edu/ibsr/.

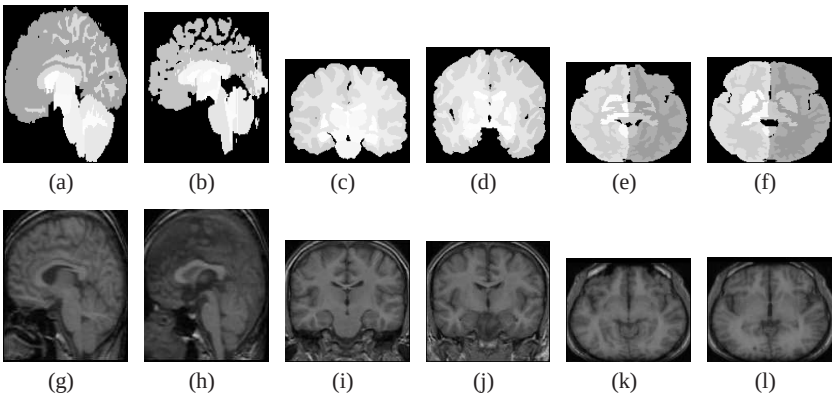


Fig. 2 Samples from IBSR data set. 2(a), 2(b) Two slices of the sagittal view of the same 3D MRI volume representing the two categories. 2(c), 2(d) Coronal view. 2(e), 2(f) Axial view. The original images are presented below their corresponding manual segmentations.

Table 1 Identification of the slices composing the database in each view of the 3D volume, for the three possible views: sagittal (S), coronal (C) and axial (A)

View	Slices	Slices cat. 1	Slices cat. 2
S	256	121, 122, 123	126, 127, 128
C	128	58, 59, 60	64, 65, 66
A	256	121, 122, 123	126, 127, 128

features that we take into account; as a result, we build a more challenging categorization problem between images belonging to the same view; a secondary benefit of this approach is that by choosing certain slices we can control the difficulty of the task. Since the brain is made up of consecutive slices in any of the three views and the brain structure varies progressively, we want to create one category using three consecutive slices which are at the same level over all the eighteen 3D brain segmentations. A second category is being built using three consecutive slices which are positioned at a certain distance from the first block; as the distance between the two blocks of slices decreases, the difficulty of the categorization task increases. We found out that choosing a distance of only two or three slices between the training blocks, along with category intra-variability, would account for a difficult categorization task. Table 1 references the total number of slices in each 3D brain view and the indices of slices being used for defining each category; Figure 2 presents typical category elements for all views. Each brain view will provide three images for each category, thus creating a category definition of 54 images.

Concerning the graph construction and labeling, nodes are represented by manually segmented regions while edges account for spatial relations between regions. For vertex labeling, we use normalized region visual features: the mean gray level (which is normalized according to the lightest and darkest regions in the image), the

relative region area (normalized according to the total image area) and the normalized region compacity, defined as the normalized ratio between its surface and its squared perimeter. For this work specifically, we will experiment with the coronal view and with the mean gray level as region feature. Spatial relations based on adjacency measures being considered between image regions build up the edge labeling, respectively.

We perform n-fold cross validation on the training set ($n = 10$), and we repeat the classification task m times ($m = 10$); the performance given below is the mean value of these m executions.

5.3 Categorization with strict adjacency structures

Given a region feature and a spatial relation within the strict adjacency graph, we use RBF kernels (Equation 4) with thresholds that are adapted to the range $[0,1]$ of these normalized features (see Figure 3), and we set up a grid search in the σ parameter space for each of the two kernel functions. For each element of the grid, we try multiple values for the regularization parameter C of the SVM, $C \in \{10^{-2}, 10^{-1}, \ldots, 10^{6}\}$. Figure 3 presents the best classification performance for each pair $(\sigma_{vertex}, \sigma_{edge})$, for the values and features specified on the axis.

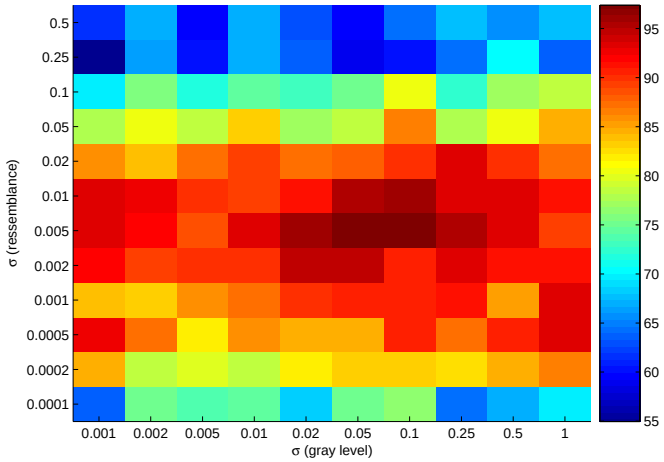
Categorization with fuzzy adjacency structures

Next, we analyze the impact of adding structural information which is not necessarily tied to the strict adjacency between image regions. For a given segmentation and for a certain spatial relation R, the complete graph encodes all the possible relations between vertices, as edge labels.

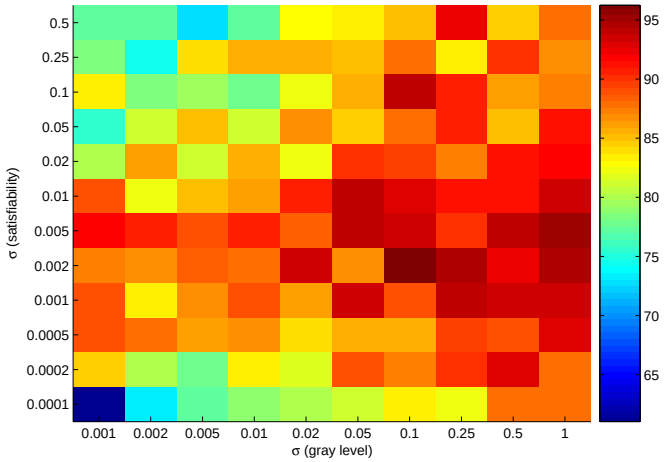
The histograms in Figure 4 present the satisfiability and ressemblance values encoded within all the complete graphs in the dataset. From these figures, we notice that the first type of measure takes the maximum value more often, while the frequency of low values is very significant. For the second measure, maximum values are quite low even for adjacent regions (the maximum value in all the dataset being 0.39), and low values are very frequent, too.

In order to estimate the impact of different spatial relation thresholds θ on the structure of elements in the threshold graph set $\mathscr{G}$, we compute the number of differences between the set of strict adjacency edges and θ-thresholded edges with respect to the relation R, using Equation 12. The difference profiles for the satisfiability and ressemblence relations are presented in Figure 5(a) and Figure 5(b). Given any of the adjacency graphs G in the dataset, the threshold θ that would minimize the structural difference between G and $G_\theta \in \mathscr{G}$ is given by the value corresponding to the minimum in the difference profile, θ_*.

For our dataset, the optimal threshold for the satisfiability measure is $\theta_*^{sat} = 0.911$. This high value proves the fact that most of the times, strictly adjacent re-

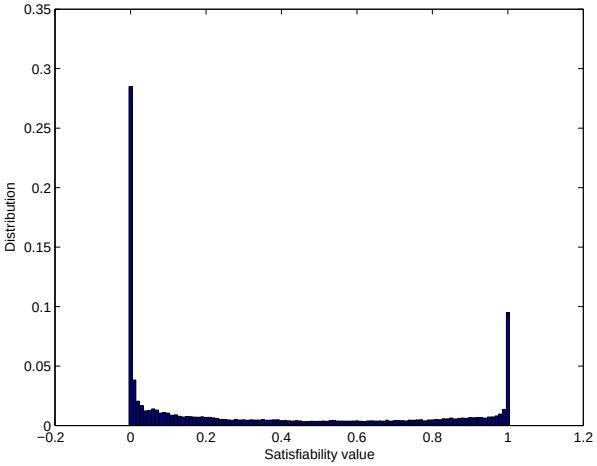


(a) Using the gray level region attribute and the ressemblance measure, the best performance, 97.72%, is attained for $\sigma_{vertex} = 0.1$ and $\sigma_{edge} = 0.005$.

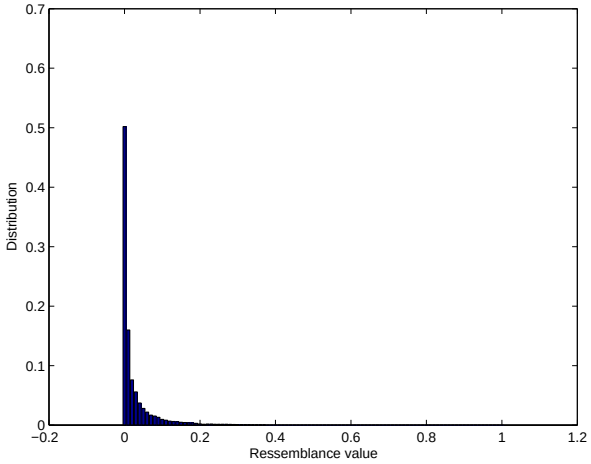


(b) Using the gray level region attribute and the satisfiability measure, the best performance, 96.51%, is attained for $\sigma_{vertex} = 0.1$ and $\sigma_{edge} = 0.002$.

Fig. 3 Categorization performance for the gray level region attribute and two different measures of fuzzy adjacency, using grid search in the space of kernel parameters. At this point, we model input data using strict adjacency graphs.

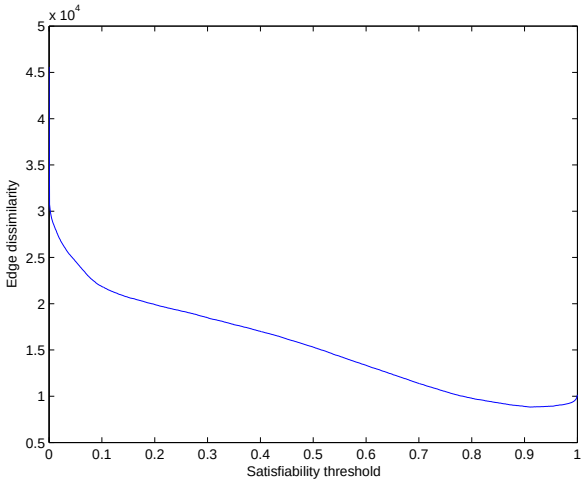


(a) Histogram of the satisfiability measure. Null values are the most frequent ones, but maximal values are frequent too.

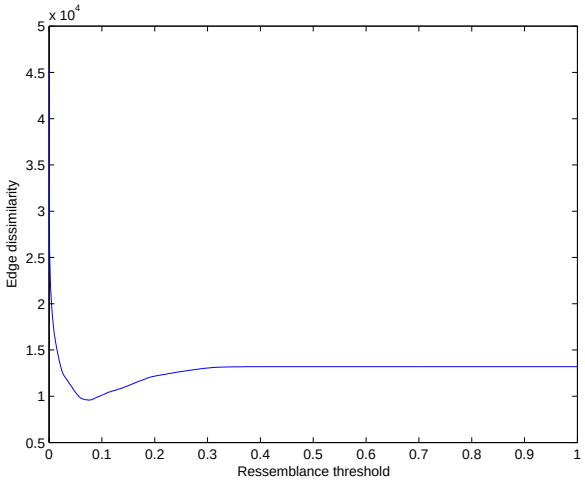


(b) Histogram of the ressemblance measure. Null values are equally the most frequent ones, but this measure penalizes very fast the absence of a strong adjacency and the maximum value in all the dataset is 0.39.

Fig. 4 Distribution of the two measures of fuzzy adjacency for all the edges in the set of complete graphs representing the dataset.



(a) Satisfiability edge dissimilarity count



(b) Ressemblance edge dissimilarity count

Fig. 5 For a given measure (satisfiability or ressemblance) threshold θ, we show the number of different edges between the set of strict adjacency graphs of the dataset and the set of θ-threshold graphs associated to the strict adjacency graphs. The minimal value accounts for the highest structural similarity between the strict adjacency graphs and the θ-threshold graphs.

Table 2 Categorization performance for (gray level - spatial relation) image information. The parameters for the kernel functions are the optimal values found using the grid search. In the third column, the strict adjacency graph is used, but no spatial relation labels are added to the graph. In the fourth column, we use fuzzy adjacency labeling on the strict adjacency graph. Afterwards, we use different θ-threshold fuzzy adjacency graphs.

Region feature	Spatial relation	Strict adj. No relation	Strict adj. Fuzzy labeling	Fuzzy graphs
Gray level	Ressemblance	58.18%	97.72%	θ=0.00 (82.90%) θ=0.01 (82.92%) θ=0.02 (86.43%) θ=0.05 (84.38%) θ_*=0.075 (74.43%) θ=0.10 (76.23%) θ=0.20 (69.90%) θ=1.00 (61.73%)
Gray level	Satisfiability	57.04%	96.51%	θ=0.00 (79.07%) θ=0.1 (70.83%) θ=0.25 (68.18%) θ=0.5 (76.75%) θ=0.75 (77.51%) θ_*=0.911 (77.29%) θ=1.00 (62.85%)

gions account for satisfiability values that go beyond the threshold, as it is perceivable from the high proportion of maximum values. The second measure has a different behavior; it penalizes very fast the absence of a strong adjacency. In this case, the R values associated to the strict adjacency relation are scattered on a larger interval, thus the optimal threshold is situated further from the maximum value: $\theta_*^{ress} = 0.075$.

In Table 2 we compare the categorization performances for different settings involving spatial relations. As a reference, we use the best classifier detected for a certain region feature-spatial relation pair, using grid search. This classifier relies on the strict adjacency graph extracted from the image, but the edges are labeled using the spatial relation value between the corresponding vertices. The interest of incorporating spatial relation information to the labeling is proven by the weak performance of the classifier on the adjacency graph which uses only the region feature information (the edge kernel k_e being fixed set as $k_e = 1$, cf. Section 2.2).

Next, we pass to the threshold graphs G_θ in the set $\mathscr{G}$. In our setting, the spatial relations R are represented using values in $[0,1]$, therefore the threshold θ is also a number in $[0,1]$. We estimate the categorization performance along the set $\mathscr{G}$; reference elements are the complete graph $G_f = G_0$, G_{θ_*} the projection of G in $\mathscr{G}$, and G_1.

Results in Table 2 show that once we pass to a structure which is based entirely on thresholded fuzzy spatial relations, we do not improve the best performance witnessed on the strict adjacency graph structure. Within the set $\mathscr{G}$, the projection G_{θ_*}

of G performs well and the classifier performance may be improved by lowering slightly the threshold below the value of θ_*, which accounts for adding edges with spatial information. However, θ values that are far from θ_*, including the value $\theta = 0$ that corresponds to the complete graph, account for a poorer performance. This shows that in the presence of a richer information, the performance does not necessarily improve. The explanation is that the high frequency of low values for the edge labels leads to artificially high similarity estimations between graphs and masks the similarity of meaningful high label values. While the spatial information is definitely helpful in image interpretation, its generic integration into graphical models remains a difficult task and kernel functions for SVMs that cope with spatial information should be adapted specifically to different types of spatial relations.

6 Conclusion

In this article, we studied the benefits offered by image representations using labeled graphical models, as well as by employing fuzzy descriptors for spatial information. Graphical models allow for a flexible integration between intrinsic visual features of image parts and the spatial interactions taking place. We showed that fuzzy information is highly beneficial for the learning process when we use it to enrich the labeling of strict adjacency graphical structures, but that loose spatial interactions may screen more relevant spatial information and that generic kernel functions are not well adapted to take into account the entirety of spatial relations within images. Future work will try to adapt the graph similarity estimation to the specificity of spatial relations in order to benefit from information concerning the presence and the absence of interactions.

References

1. Kashima, H., Tsuda, K., Inokuchi, A.: Marginalized kernels between labeled graphs. In: 20st Int. Conf. on Machine Learning. (2003) 321–328
2. Mahé, P., Ueda, N., Akutsu, T., Perret, J.L., Vert, J.P.: Extensions of marginalized graph kernels. In: ICML '04: 21st Int. Conf. on Machine Learning. (2004)
3. Mahé, P., Ralaivola, L., Stoven, V., Vert, J.P.: The pharmacophore kernel for virtual screening with support vector machines. J. Chem. Inf. Model. **46** (2006) 2003–2014
4. Borgwardt, K.M., Kriegel, H.P.: Graph kernels for disease outcome prediction from protein-protein interaction networks. In: Pacific Symposium on Biocomputing. (2007) 4–15
5. Aldea, E., Atif, J., Bloch, I.: Image Classification using Marginalized Kernels for Graphs. In: 6th IAPR-TC15 Workshop on Graph-based Representations in Pattern Recognition, GbR'07, Alicante, Spain (2007) 103–113
6. Harchaoui, Z., Bach, F.: Image classification with segmentation graph kernels. In: Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on. (2007) 1–8
7. Vapnik, V.: Statistical Learning Theory. Wiley-Interscience (1998)
8. Tsuda, K., Kin, T., Asai, K.: Marginalized kernels for biological sequences. Bioinformatics **18** (2002) 268–275

9. Genton, M.: Classes of kernels for machine learning: A statistics perspective (2000)
10. Arivazhagan, S., Ganesan, L., Priyal, S.P.: Texture classification using gabor wavelets based rotation invariant features. Pattern Recogn. Lett. **27** (2006) 1976–1982
11. Bernardino, A., Santos Victor, J.: Fast iir isotropic 2-d complex gabor filters with boundary initialization. IP **15** (2006) 3338–3348
12. Bloch, I.: Fuzzy Spatial Relationships for Image Processing and Interpretation: A Review. Image and Vision Computing **23** (2005) 89–110
13. Bouchon-Meunier, B., Rifqi, M., Bothorel, S.: Towards general measures of comparison of objects. Fuzzy sets and Systems **84(2)** (1996) 143–153
14. Vanegas, C., Bloch, I., Maître, H., Inglada, J.: Approximate Parallelism Between Fuzzy Objects: Some Definitions. In: International Workshop on Fuzzy Logic and Applications WILF. Volume LNAI 5571., Palermo, Italy (2009) 12–19
15. Takemura, C.M., Cesar, R.M., Bloch, I.: Fuzzy modeling and evaluation of the spatial relation "along". In: CIARP. (2005) 837–848
16. Deruyver, A., Hodé, Y., Brun, L.: Image interpretation with a conceptual graph: Labeling over-segmented images and detection of unexpected objects. Artif. Intell. **173** (2009) 1245–1265
17. Lebrun, J., Philipp-Foliguet, S., Gosselin, P.H.: Image retrieval with graph kernel on regions. In: ICPR. (2008) 1–4
18. Aldea, E., Fouquier, G., Atif, J., Bloch, I.: Kernel fusion for image classification using fuzzy structural information. In: ISVC (2). (2007) 307–317
19. Riesen, K., Neuhaus, M., Bunke, H.: Bipartite graph matching for computing the edit distance of graphs. In: GbRPR07. (2007) 1–12