

Vers une utilisation améliorée de relations spatiales pour l'apprentissage de données dans les modèles graphiques

Emanuel Aldea*, Isabelle Bloch*

*TELECOM Paris-Tech, département TSI
CNRS UMR 5141 LTCI
46 rue Barrault, Paris 75634
nom@telecom-paristech.fr

Résumé. Nous nous intéressons dans cet article aux représentations des relations spatiales pour l'extraction d'information et la modélisation des données visuelles, en particulier dans le contexte de la catégorisation d'images. Nous montrons comment la prise en compte d'une relation spatiale entre deux éléments entraîne l'apparition d'une information supplémentaire entre ces éléments et le reste de l'ensemble à modéliser, ce qui est rarement exploité explicitement. Une représentation floue des relations dans un modèle graphique est bien adaptée pour les algorithmes d'apprentissage utilisés actuellement et permet d'intégrer ce type d'information complémentaire qui concerne l'absence d'une interaction plutôt que sa présence. Nous tentons d'évaluer les bénéfices de cette approche sur un problème de traitement d'images.

1 Introduction

Les méthodes de représentation de données et d'apprentissage classiques ne prennent pas naturellement en compte les images. Parmi les différentes méthodes qui ont été proposées récemment pour la modélisation des images, nous nous intéressons dans cet article à une représentation de l'image comme un ensemble structuré d'objets, faisant apparaître explicitement les constituants de l'image et leurs relations dans un modèle graphique.

Les algorithmes capables de gérer l'apprentissage sur ces modèles ont été créés (Kashima et al., 2003) et optimisés (Mahé et al., 2004) dans le contexte des applications bioinformatiques. Une adaptation pour l'apprentissage des graphes issus des images est nécessaire, car les propriétés des informations codées dans la structure graphique représentant une image sont fondamentalement différentes de celles utilisées en bioinformatique.

La plupart des méthodes de classification s'appuient en premier lieu sur les attributs des objets d'intérêt dans les images. Cependant, les informations spatiales liées aux relations entre ces objets sont également utiles, comme cela a été montré en segmentation et reconnaissance de structures dans les images, et leur intégration dans des méthodes d'apprentissage et de classification commence à apparaître. Le fait que ces informations soient souvent exprimées sous forme linguistique, et que l'absence d'une relation puisse constituer également une information utile pour la classification, suggère l'utilité des relations spatiales floues pour la modélisation

d'images. En effet les modélisations floues permettent de représenter à la fois l'imprécision de la relation (par exemple *proche* ou *à gauche*) et le passage graduel de la satisfaction à la non satisfaction de cette relation.

L'objectif de cet article est d'explorer les limites de la représentation et de l'intégration actuelle des informations spatiales dans les classificateurs d'images qui utilisent les noyaux de graphes. Dans une première étape, nous décrivons les avantages de la représentation des connaissances par des graphes étiquetés et la stratégie générale de l'algorithme qui est capable d'utiliser ce type de structure de données. Ensuite, nous discutons de l'utilisation de l'information spatiale dans les travaux connexes et des particularités des relations spatiales dans les images. Les expériences montrent que l'information spatiale complète les informations visuelles intrinsèques des objets dans la classification, et qu'une utilisation optimale nécessite des adaptations des méthodes de discrimination utilisées par les noyaux de graphes.

2 Représentation de la connaissance générique sous forme de graphes étiquetés

Dans le cadre de l'apprentissage artificiel, les méthodes statistiques supervisées ont jusqu'à présent donné lieu à des algorithmes permettant de traiter des données d'entrée sous forme de vecteurs ou de séquences et de prédire des sorties sous la forme de valeurs réelles ou de catégories. Aujourd'hui, de nombreuses données ont une représentation complexe sous la forme d'arbres, de graphes ou de données relationnelles, particulièrement dans les sciences du vivant mais aussi dans le traitement des images, les modèles de raisonnement, etc. De même, on voit apparaître de nouvelles tâches d'apprentissage qui nécessitent d'extraire, à partir de données, des relations ou des dépendances structurelles. Ces besoins permettent d'envisager l'apparition de méthodes d'apprentissage pour des tâches plus complexes et souvent liées à l'analyse de données de grand volume, issues de l'informatique médicale, de la biologie, du langage naturel, etc.

Dans une représentation des données complexes par des graphes, les nœuds représentent par exemple des atomes (Kashima et al., 2003) ou des structures atomiques simples (Mahé et al., 2006), des protéines (Borgwardt et Kriegel, 2007), des objets que l'on cherche à segmenter dans des images ou à classifier (Aldea et al., 2007; Harchaoui et Bach, 2007), et les arêtes modélisent des interactions spécifiques, des relations d'ordonnement et de dépendance, des relations spatiales (adjacence, distance, localisation relative, topologie, ...). Dans ce contexte, et en particulier pour les applications en interprétation d'images et en raisonnement spatial, de nombreuses sources d'imprécision doivent être prises en compte : sur les objets, à cause de leurs limites imprécises et des éventuelles difficultés à les segmenter, sur les relations et sur les connaissances manipulées, souvent exprimées sous forme linguistique. L'étiquetage des graphes nous permet de synthétiser les connaissances sur la nature et les propriétés des composantes du graphe, mais aussi le degré de confiance en toutes ces informations.

Dans le cas des images, une possibilité de représentation est constituée par le graphe d'adjacence à partir d'une segmentation en régions. Les nœuds du graphe sont attribués à partir des caractéristiques des régions, telles que la surface en pixels de la région, le rapport entre cette surface et celle de l'image (surface relative), le niveau de gris moyen et le niveau de gris relatif

(par rapport à la dynamique de l'image), la compacité de la région. Généralement, ces attributs prennent des valeurs sur une échelle continue.

La seule information structurelle qui est utilisée est l'adjacence, implicitement codée par la structure du graphe. La similarité entre graphes est définie par une méthode à noyaux et permet de construire un classifieur d'images. Cette similarité, selon un attribut a , entre deux graphes G et G' , issus de deux images, est évaluée par un noyau qui mesure la similarité de tous les chemins aléatoires possibles dans les deux graphes (Kashima et al., 2003), pondérée par la probabilité d'apparition de ces chemins.

Par rapport à d'autres contextes dans lesquels ce type de méthode a été utilisé (Mahé et al., 2004), l'espace des variables devient continu et multi-dimensionnel, et constitue une part prédominante de l'information. Par conséquent, la fonction de similarité pour des attributs à valeurs continues doit être moins discriminante que le noyau de Dirac classiquement utilisé en bio-informatique. En conséquence nous utilisons des noyaux gaussiens de variance σ^2 pour évaluer la similarité $k_a^{rbf}(a_1, a_2)$ entre deux valeurs a_1 et a_2 d'un attribut numérique a :

$$k_a^{rbf}(a_1, a_2) = \exp\left(-\frac{\|a_1 - a_2\|^2}{2\sigma^2}\right) \quad (1)$$

Etant donnés deux graphes G et G' à comparer, l'équation (1) permet d'évaluer la similarité $k_{v,e}(h, h')$ entre deux marches aléatoires $h = \{x_1, \dots, x_n\}$ dans G et $h' = \{x'_1, \dots, x'_n\}$ dans G' , en combinant les similarités k_v et k_e des valeurs d'un attribut de nœuds v et respectivement les valeurs d'un attribut d'arêtes e le long de h et h' :

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

Cette formule générale peut être simplifiée dans le cas où l'on considère seulement un attribut de régions (comme c'est le cas dans les graphes d'adjacence stricte), ou dans le cas où l'on considère seulement un attribut d'arêtes. Pour la simplifier, on met la valeur de la fonction noyau manquante à 1 (l'élément dont on ne prend pas en compte l'étiquetage doit exister, car autrement la marche aléatoire dont il fait partie ne serait pas réalisable).

Enfin, le noyau entre G et G' est calculé en sommant les similarités de tous les chemins aléatoires h et h' possibles, pondérées par leurs probabilités d'apparition :

$$K_{v,e}(G, G') = \sum_h \sum_{h'} k_{v,e}(h, h') p(h|G) p(h'|G') \quad (3)$$

Cette fonction est ensuite utilisée dans un séparateur à vaste marge (SVM) pour construire un classifieur d'images. $K_{v,e}$ définit une matrice de similarité entre les paires de graphes à comparer.

3 Relations spatiales dans le contexte d'apprentissage de graphes

Le contexte spatial a été intégré avec succès en bioinformatique, pour la représentation par graphes des données complexes. Avec peu d'exceptions (Mahé et al., 2006), les relations spatiales utilisées sont de type binaire et modélisent de façon très exacte la présence ou l'absence

d'une interaction entre les constituants de la structure globale. Dans ce cadre particulier, la construction du graphe complémentaire \tilde{G} de la représentation (Borgwardt et Kriegel, 2007), qui code les absences d'interactions, et l'utilisation de cette information supplémentaire pour améliorer la fonction noyau du classifieur sont non-ambiguës et facilement réalisables :

$$K^*(G, G') = K(G, G') + K(\tilde{G}, \tilde{G}') \quad (4)$$

En ce qui concerne l'extraction de l'information spatiale à partir des images, la situation devient plus complexe. Les interactions spatiales ont une variabilité liée à l'interprétation sémantique humaine qui se situe bien au-delà du cas binaire mentionné ci-dessus. De plus, l'intégration de l'information spatiale floue et des attributs de régions s'avère aussi plus complexe que dans les situations traitées par l'équation (4). Toutefois, ces relations spatiales apportent une grande richesse dans la description des images. Cela a été largement montré dans des problèmes de segmentation et reconnaissance de structures dans les images. Ici nous montrons comment utiliser ces informations en apprentissage et classification.

Nous pourrions nous contenter d'ajouter de l'information structurelle sur les arêtes existantes, mais l'utilisation d'une adjacence stricte peut poser un problème de robustesse. En effet, lorsque l'adjacence ne dépend que de quelques pixels, le graphe obtenu pourra être différent en fonction de la méthode de segmentation utilisée et du modèle générique de la classe à analyser. L'ajout d'arêtes comportant plus qu'une information implicite d'adjacence nous permet donc non seulement de mieux décrire les informations structurelles, mais aussi d'améliorer la robustesse de la représentation.

Pour notre application, nous utiliserons une relation spatiale topologique représentée par un degré d'adjacence étendu.

Distance entre régions. La distance entre deux régions R_1 et R_2 est calculée comme le minimum des distances euclidiennes entre deux points $p_i \in R_1$ et $q_j \in R_2$:

$$\min_{p_i \in R_1, q_j \in R_2} (d_{euclidienne}(p_i, q_j))$$

Position relative directionnelle. Plusieurs méthodes ont été proposées pour définir la direction entre deux objets, qui est une notion intrinsèquement vague. Les méthodes floues sont particulièrement appropriées (Bloch et Ralescu, 2003) et nous avons choisi ici de représenter cette information en utilisant des histogrammes d'angles (Miyajima et Ralescu, 1994). Cela nous permet de représenter toutes les directions possibles entre deux régions. Si R_1 et R_2 sont deux ensembles de points $R_1 = \{p_1, \dots, p_n\}$ et $R_2 = \{q_1, \dots, q_n\}$, la position relative directionnelle entre les régions R_1 et R_2 est estimée à partir de la position de chaque point q_j de R_2 par rapport à chaque point p_i de R_1 . L'histogramme d'angles $H_{R_1 R_2}$ est défini en 2D par :

$$H_{R_1, R_2}(\theta) = \left| \{(p_i, q_j) \in R_1 \times R_2 / \angle(\vec{i}, \overrightarrow{p_i q_j}) = \theta\} \right|$$

où $\angle(\vec{i}, \overrightarrow{p_i q_j})$ représente l'angle entre un vecteur de référence \vec{i} et le vecteur $\overrightarrow{p_i q_j}$. L'extension à trois dimensions est une fonction de deux angles. Un exemple est illustré figure 1.

Afin d'obtenir une valeur réelle, nous calculons le centre de gravité de l'histogramme. Le calcul de l'histogramme d'angles étant coûteux en termes de temps de calculs, nous pouvons



FIG. 1 – Histogramme d’angles entre les régions 1 et 2. Le centre de gravité est représenté par une ligne verticale.

être amenés à travailler avec des images réduites, ce qui affecte peu le calcul du centre de gravité.

Mesure d’adjacence fondée sur une comparaison floue. La distance et l’orientation ne sont pas toujours significatives. Par exemple, la distance ne distingue pas deux régions adjacentes par un unique pixel de deux régions imbriquées. Dans ce dernier cas, le centre de gravité de l’histogramme n’a pas beaucoup de sens. Nous proposons donc une autre caractéristique, topologique, qui estime un degré d’adjacence entre deux régions.

Nous estimons le degré d’adjacence entre deux régions en mesurant la corrélation entre la portion de l’espace ‘proche’ de la première région dite de référence et la deuxième. Cette mesure est maximale lorsque la région de référence est imbriquée dans la région cible. Elle est nulle si les deux régions sont trop éloignées. Une valeur moyenne implique que deux régions sont adjacentes pour au moins la moitié du contour de la région de référence.

Les représentations floues sont appropriées pour modéliser l’imprécision intrinsèque de plusieurs relations telles que ‘proche’. Ici, nous considérons la formulation proposée dans Bloch (2000) : étant donné un objet de référence, la portion de l’espace dans laquelle la relation à cet objet est satisfaite est définie comme un ensemble flou spatial, dont la fonction d’appartenance en chaque point donne le degré de satisfaction de la relation en ce point. Cette représentation dans l’espace de l’image est propice à sa combinaison avec d’autres informations de type image.

La relation spatiale ‘proche’ peut être déduite d’un intervalle flou de forme trapézoïdale f sur \mathbb{R}^+ . Un sous-ensemble flou $\mu_{proche(R_1)}$ de l’espace de l’image \mathcal{S} est alors construit en combinant f avec une carte de distance d_{R_1} à l’objet de référence R_1 : $\forall x \in \mathcal{S}, \mu_{proche(R_1)}(x) = f(d_{R_1}(x))$, où $d_{R_1}(x) = \inf_{y \in R_1} d(x, y)$. La figure 2 présente une région (a) et le sous-ensemble flou correspondant à la relation ‘proche de la région 1’ (d). Dans nos expériences, la fonction d’appartenance f a été définie avec les valeurs fixes suivantes : 0, 0, 10, 30.

Nous avons choisi une fonction symétrique, une M-mesure de ressemblance (Bouchon-Meunier et al., 1996) définie par :

$$Res(proche(R_1), R_2) = \frac{\sum_{x \in \mathcal{S}} \min(\mu_{proche(R_1)}(x), \mu_{R_2}(x))}{\sum_{x \in \mathcal{S}} \max(\mu_{proche(R_1)}(x), \mu_{R_2}(x))}. \quad (5)$$

Cette mesure est maximale si l’objet et la relation sont identiques : cette mesure de ressemblance prend en compte le positionnement de l’objet, mais également la précision du sous-ensemble flou.

La figure 2 présente trois régions : la région de référence (a) et deux régions qui lui sont adjacentes (b) et (c). Le sous-ensemble flou correspondant à la relation ‘proche de la région 1’

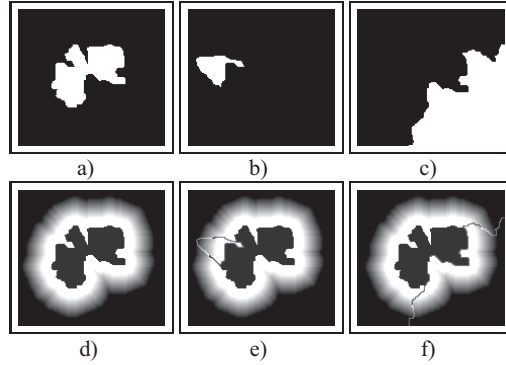


FIG. 2 – (a) Région 1. (b) Région 2. (c) Région 3. (d) Sous-ensemble flou correspondant à la relation 'proche de la région 1'. (e) Même chose avec la frontière de la région 2 en sur-impression. (f) De même avec la région 3.

est présenté en (d) et le contour des régions adjacentes a été ajouté respectivement en (e) et (f), permettant de visualiser l'intersection entre les régions adjacentes et la relation 'proche de la région 1'. La satisfiabilité mesurée entre le sous-ensemble flou de la relation et la région 2 est de 0,06 et il est de 0,29 pour la région 3. La figure 3(a) présente la partie du modèle graphique créée avec ces informations : les nœuds correspondent aux régions et les arêtes sont étiquetées par les valeurs de satisfiabilité calculée ci-dessus.

4 Expériences

La base d'images IBSR¹ contient des données cliniques réelles acquises par imagerie par résonance magnétique (IRM) et est couramment utilisée en tant que base d'images de cerveaux sains en trois dimensions. Elle contient les segmentations de 18 cerveaux, réalisées manuellement par un expert, et chacune est disponible selon trois plans de coupe (ou vues) différents : axial, coronal et sagittal. Chaque élément de la base de données est un ensemble de coupes qui couvre le cerveau dans son ensemble.

Cette base de données est principalement utilisée comme un moyen d'évaluer les performances d'algorithmes de segmentation, mais la qualité des segmentations proposées et le fait qu'elle soit disponible gratuitement la rend utile pour nos expériences. La plupart des attributs proposés permettent de classifier les images de deux vues différentes avec 100% de réussite. En conséquence, nous devons construire une base plus difficile à classifier. Nous avons donc tenté de classifier des images provenant de la même vue. Pour chaque sujet, nous sélectionnons pour chaque vue trois coupes successives pour la première classe, puis trois autres coupes successives pour la seconde. Cela donne pour chaque vue et chaque classe 54 images (3x18) dont 15 sont sélectionnées aléatoirement pour l'apprentissage. Le reste des images sera utilisé pour tester le classifieur. Le tableau 1 recense les indices des coupes utilisées dans chacune

¹Internet Brain Segmentation Repository, disponible sur <http://www.cma.mgh.harvard.edu/ibsr/>

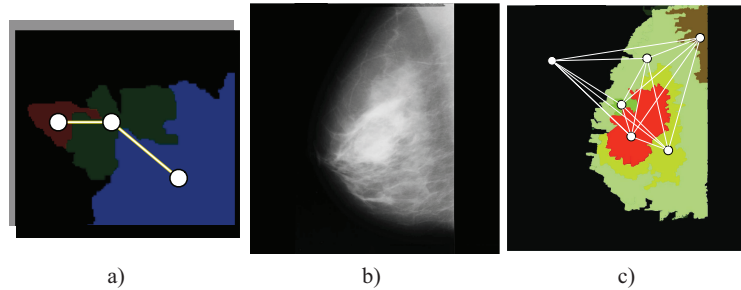


FIG. 3 – Un sous-ensemble du modèle graphique constitué des nœuds qui représentent les trois régions de la figure 2, et des arêtes qui encodent les relations spatiales entre la région centrale 2(a) et les autres. En itérant pour l'ensemble des régions, un graphe complet étiqueté sera créé. Pour la mammographie présentée en 3(b), on voit en 3(c) une segmentation automatique par régions, avec le modèle graphique correspondant.

des trois vues pour définir les classes. Les coupes sont sélectionnées parmi celles qui disposent d'un nombre suffisant de régions après segmentation.

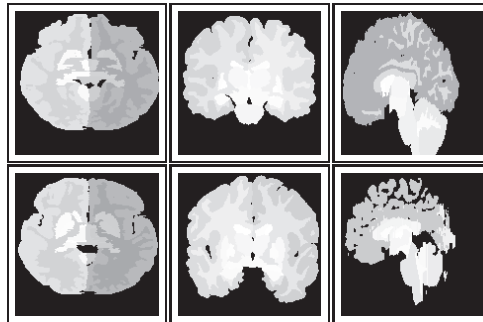


FIG. 4 – Exemples d'images provenant de la base IBSR. Les niveaux de gris représentent des étiquettes. À gauche, en haut et en bas : deux coupes provenant d'une vue axiale du même volume IRM 3D et appartenant à chacune des classes. Au centre : vue coronale. À droite : vue sagittale.

Pour estimer la similarité des images selon un attribut, nous utilisons des noyaux gaussiens avec une valeur du seuil relativement petite ($\sigma = 0.05$ pour les attributs des régions et $\sigma = 0.01$ pour les relations spatiales, dont les valeurs numériques sont plus basses) qui les rend sensibles aux différences d'étiquetage. Chaque noyau d'attribut est injecté dans un noyau marginalisé pour graphe utilisé par l'algorithme SVM décrit dans la section 2. Pour le paramètre de régularisation C du SVM, qui contrôle le compromis entre la maximisation de la marge et la minimisation de la pénalisation des points mal classés, nous avons réalisé une recherche dans une grille de résolution uniforme dans l'espace $\log_{10} : \log_{10} C \in \{-3, \dots, 6\}$. Nous avons observé que pour deux valeurs, $C = 10^3$ (SVM1) et $C = 10^5$ (SVM2), les classificateurs réalisent un compromis performant entre les objectifs décrits ci-dessus ; dans les tables 2 et 3

Vers une utilisation améliorée de relations spatiales pour l'apprentissage

Vue	Nb coupes	Classe 1	Classe 2
A	255	121, 122, 123	126, 127, 128
C	128	58, 59, 60	64, 65, 66
S	255	121, 122, 123	126, 127, 128

TAB. 1 – Recensement des coupes composant les bases de données des volumes 3D pour chaque vue : axiale (A), sagittale (S) et coronale (C).

qui suivent, les résultats sont présentés pour les deux SVMs. Pour la tâche de classification, nous utilisons la vue coronale, avec $N = 30$ graphes pour l'apprentissage et $T = 78$ graphes pour le test, les deux ensembles étant également répartis entre les deux classes.

Attribut	Représentation	SVM1	SVM2
Niveau de gris	Graphe d'adjacence	(13,4)	(11,4)
Niveau de gris	Graphe complet	(15,8)	(13,10)
Compacité	Graphe d'adjacence	(5,7)	(10,4)
Compacité	Graphe complet	(9,7)	(4,8)
Surface	Graphe d'adjacence	(13,1)	(10,5)
Surface	Graphe complet	(10,1)	(13,1)
Ressemblance	Graphe complet	(4,7)	(3,0)

TAB. 2 – Performances des attributs utilisés individuellement dans deux contextes différents de représentation (graphe d'adjacence et graphe complet). Les deux machines SVM utilisées diffèrent par leur paramètre de régularisation. Les valeurs entre parenthèses représentent le nombre d'exemples de l'ensemble de test mal classés pour chacune des deux classes.

Le point de départ de l'interprétation des images par le contenu est constitué par les attributs des objets d'intérêt dans l'image, ici représentés par les régions. Un modèle de représentation sous forme de graphe d'adjacence est intuitif mais sensible à la variabilité de la classe et également aux étapes de détection de régions et d'extraction de la structure des régions dans l'image. A ce stade, le graphe complet comme modèle de représentation (si on ne met pas d'attribut sur les arêtes) enlève toute information spatiale et son intérêt est très limité dans un contexte de raisonnement spatial. Les relations spatiales viennent enrichir la représentation de l'image au-delà de la simple adjacence stricte. Pour la représentation de cette information spatiale, un graphe complet, dont les nœuds sont représentés par des régions et les arêtes portent des attributs décrivant une certaine relation spatiale, est bien adapté. Les performances des attributs de régions et de la relation spatiale présentée en section 3 sont données dans la table 2. Le nombre d'exemples de test mal classés est significatif, ce qui montre les limites de l'emploi d'un seul type d'information pour l'analyse d'images. On peut également remarquer une performance inférieure des attributs de régions dans le cadre d'une modélisation sans information spatiale de base (graphe complet sans attribut sur les arêtes).

Afin d'utiliser les deux types d'informations dans le but de classification, nous avons ajouté progressivement l'information spatiale sur le graphe d'adjacence initial et rajouté des arêtes en fonction de la valeur numérique de la relation spatiale associée. Dans une première étape, nous avons gardé la structure du graphe d'adjacence entre les régions et nous avons placé sur les

arêtes déjà présentes la valeur de la mesure de ressemblance entre les régions reliées. Dans les deux étapes suivantes, nous avons choisi deux seuils de plus en plus permissifs (0.01 et 0.001) et nous avons rajouté au graphe d'adjacence des arêtes dont la valeur de la relation spatiale associée est au-dessus du seuil. Finalement, nous avons rajouté au graphe initial d'adjacence toutes les arêtes étiquetées possibles, ce qui correspond à l'emploi d'un seuil 0.

Les résultats des expériences sont présentés dans la table 3. Au fur et à mesure que les arêtes associées aux valeurs faibles de la mesure de ressemblance sont ajoutées au graphe, on remarque une diminution des performances de classification, qui peut même tendre vers les niveaux observés sans utiliser l'information spatiale (le cas des graphes complets de la table 2). Cette observation vient accentuer l'importance d'une bonne adaptation de l'usage des relations spatiales dans un contexte d'apprentissage de type discriminatif, qui emploie des noyaux stationnaires (invariants par translation) comme le noyau gaussien de base. Puisque l'algorithme de classification cherche à analyser les similarités entre les graphes par leur étiquetage, le fait que la relation spatiale augmente le nombre d'étiquettes de valeur très faible peut faire diminuer l'importance des informations non nulles et perturber le processus de comparaison. L'avantage évident des relations spatiales floues est de pouvoir donner en même temps une information sur la présence et sur l'absence d'une interaction, mais les résultats des expériences montrent que l'apport de l'information sur l'absence d'une interaction n'augmente pas automatiquement les performances d'un classifieur. Une solution possible pour cette situation est de limiter l'information sur l'absence des interactions (par exemple par un seuil adapté au type de relation et au jeu de données). Une autre stratégie à envisager est d'adapter le calcul de la similarité des étiquettes spatiales pour prendre en compte non seulement la différence numérique des valeurs à comparer, mais aussi leur distribution par rapport au support de la relation spatiale. Ainsi si une relation est bien satisfaite par deux paires de régions dans les graphes à comparer, la mesure de similarité devra avoir plus d'importance que si la relation n'est pas satisfaite par les deux paires de régions.

5 Conclusion

Dans cet article, nous avons montré l'intérêt d'une part de l'utilisation de représentations graphiques des images pour l'apprentissage et la classification par des méthodes à noyau, et d'autre part de la fusion d'attributs de régions et de relations spatiales pour améliorer la classification. L'influence de la prise en compte de relations plus ou moins bien satisfaites a été illustré dans le cas de l'adjacence entre des régions issues d'IRM du cerveau.

Plutôt que de sélectionner les arêtes en fonction d'un seuil sur le degré de satisfaction de la relation, nous proposons pour nos travaux futurs de construire une fonction de similarité qui traite différemment les relations bien satisfaites et les relations non satisfaites, de manière graduelle, sans imposer de seuil.

Des expériences sur d'autres bases de données d'images et d'autres types de relations sont également en cours.

Attributs	Représentation	SVM1	SVM2
Ress. & niveau de gris	Graphe d'adjacence	(4,0)	(4,0)
Ress. & niveau de gris	Graphe (ress. > 0.01)	(10,1)	(7,1)
Ress. & niveau de gris	Graphe (ress. > 0.001)	(12,3)	(9,3)
Ress. & niveau de gris	Graphe complet	(15,4)	(12,9)
Ress. & compacité	Graphe d'adjacence	(3,0)	(3,0)
Ress. & compacité	Graphe (ress. > 0.01)	(4,0)	(3,0)
Ress. & compacité	Graphe (ress. > 0.001)	(4,0)	(4,0)
Ress. & compacité	Graphe complet	(7,2)	(4,0)
Ress. & surface	Graphe d'adjacence	(1,0)	(2,1)
Ress. & surface	Graphe (ress. > 0.01)	(2,0)	(2,0)
Ress. & surface	Graphe (ress. > 0.001)	(2,0)	(3,0)
Ress. & surface	Graphe complet	(2,0)	(2,0)

TAB. 3 – Performances d'une paire (attribut de région, relation spatiale) dans des représentations graphiques seuillées par la valeur numérique de la relation spatiale. Les deux machines SVM utilisées diffèrent par leur paramètre de régularisation. Les valeurs entre parenthèses représentent le nombre d'exemples de l'ensemble de test mal classés pour chacune des deux classes.

Références

- Aldea, E., J. Atif, et I. Bloch (2007). Image Classification using Marginalized Kernels for Graphs. In *6th IAPR-TC15 Workshop on Graph-based Representations in Pattern Recognition, GbR'07*, Alicante, Spain, pp. 103–113.
- Bloch, I. (2000). Spatial Representation of Spatial Relationships Knowledge. In A. G. Cohn, F. Giunchiglia, et B. Selman (Eds.), *7th International Conference on Principles of Knowledge Representation and Reasoning KR 2000*, Breckenridge, CO, pp. 247–258. Morgan Kaufmann, San Francisco, CA.
- Bloch, I. et A. Ralescu (2003). Directional Relative Position between Objects in Image Processing : A Comparison between Fuzzy Approaches. *Pattern Recognition 36*, 1563–1582.
- Borgwardt, K. M. et H.-P. Kriegel (2007). Graph kernels for disease outcome prediction from protein-protein interaction networks. In *Pacific Symposium on Biocomputing*, pp. 4–15.
- Bouchon-Meunier, B., M. Rifqi, et S. Bothorel (1996). Towards general measures of comparison of objects. *Fuzzy sets and Systems 84(2)*, 143–153.
- Harchaoui, Z. et F. Bach (2007). Image classification with segmentation graph kernels. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pp. 1–8.
- Kashima, H., K. Tsuda, et A. Inokuchi (2003). Marginalized kernels between labeled graphs. In *20st Int. Conf. on Machine Learning*, pp. 321–328.
- Mahé, P., L. Ralaivola, V. Stoven, et J.-P. Vert (2006). The pharmacophore kernel for virtual screening with support vector machines. *J. Chem. Inf. Model. 46(5)*, 2003–2014.

- Mahé, P., N. Ueda, T. Akutsu, J.-L. Perret, et J.-P. Vert (2004). Extensions of marginalized graph kernels. In *ICML '04 : 21st Int. Conf. on Machine Learning*.
- Miyajima, K. et A. Ralescu (1994). Spatial organization in 2D segmented images : representation and recognition of primitive spatial relations. *Fuzzy Sets and Systems* 65(2-3), 225–236.

Summary

This paper deals with structural representations of images for machine learning and image categorization. The representation consists of a graph where nodes represent image regions and edges spatial relations between them. Both nodes 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) and the influence of the satisfaction or non-satisfaction of a relation between two regions.