

Classification d'images par fusion d'attributs flous de graphes, relations spatiales et noyaux marginalisés

Image classification using fusion of fuzzy graph attributes, spatial relations and marginalized kernels

Emanuel Aldea⁽¹⁾, Geoffroy Fouquier⁽¹⁾, Jamal Atif⁽²⁾, Isabelle Bloch⁽¹⁾

(1) GET - Télécom Paris (ENST), Dép. TSI, CNRS UMR 5141 LTCI, Paris, France, Isabelle.Bloch@enst.fr

(2) Unité ESPACE S140, IRD-Cayenne/UAG, Guyane Française

Résumé :

Dans cet article, nous proposons une méthode de classification d'images par une méthode à noyaux. Les images sont représentées par des graphes attribués. Les attributs des nœuds représentent des caractéristiques des régions de l'image et les attributs des arcs représentent des relations spatiales, formalisées de manière floue. Les principales contributions de cet article sont d'une part la construction de noyaux marginalisés pour chaque attribut flou représentant des relations spatiales, et d'autre part l'étude de la fusion de noyaux de plusieurs attributs. L'apport de la fusion pour améliorer la classification est illustré sur des exemples de classification d'images médicales.

Mots-clés :

Relations spatiales floues, graphes, noyaux marginalisés, classification d'images.

Abstract:

Various kernel functions on graphs have been defined recently. In this article, our purpose is to assess the efficiency of a marginalized kernel for image classification. Graphs are built from image segmentations, and various types of information concerning the underlying image regions as well as the spatial relations between them are incorporated in the graph labeling. The main contribution of this paper consists in studying the impact of fusing kernels for different attributes on the classification decision, while suggesting the use of fuzzy attributes for estimating spatial relations.

Keywords:

Fuzzy spatial relations, graphs, marginalized kernels, image classification.

1 Introduction

Les méthodes d'apprentissage classiques ne prennent pas naturellement en compte les données structurées. Plutôt que d'essayer de modifier les algorithmes classiques, il est possible de chercher dans la direction opposée et de modifier les données afin de modifier la complexité structurelle tout en préservant les caractéristiques qui permettent d'affecter les données à différentes classes. C'est le type d'approches développées en bio-informatique pour

la classification des molécules par exemple.

Dans le cas particulier des images, différentes approches ont été proposées récemment. L'une d'elles traite les images comme des objets indivisibles et exploite des caractéristiques globales telles que l'histogramme des couleurs [7]. D'autres stratégies considèrent les images comme des « sacs » d'objets indépendants et s'appuient sur une vectorisation du contenu de l'image [17]. Enfin, certaines méthodes s'appuient sur une représentation de l'image comme un ensemble structuré d'objets [15, 16], faisant apparaître explicitement les constituants de l'image et leurs relations. L'approche proposée dans cet article entre dans cette dernière catégorie.

Nous avons proposé une première méthode de classification d'images à partir de noyaux marginalisés pour des graphes dans [1]. Les images sont représentées par des graphes d'adjacence à partir d'une sur-segmentation automatique en régions. Les nœuds du graphe sont attribués à partir des caractéristiques des régions. 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 de noyaux généralisés et permet de construire un classifieur d'images. Cette méthode initiale est résumée dans la section 2.

Ici, nous proposons une extension substantielle de cette méthode. La structure du graphe est enrichie et des attributs structurels représentant d'autres relations spatiales que l'adjacence sont ajoutés. Ces relations sont modélisées sous forme floue [3]. Si de tels modèles ont été assez largement utilisés pour la reconnaissance de structures dans les images, ils ne l'ont pas encore été en apprentissage, ce qui constitue donc une des originalités de notre approche. Les

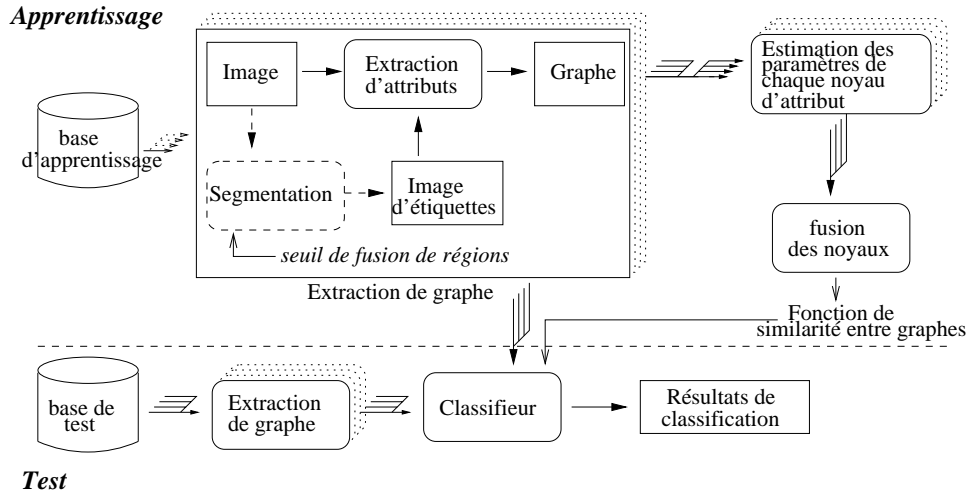


Figure 1 – Schéma général de la méthode. Phase d'apprentissage : si elles ne le sont pas déjà, les images de la base d'apprentissage sont segmentées. Un graphe est extrait de chaque image en utilisant l'image d'étiquettes correspondante. Ensuite, pour chaque attribut du graphe, les paramètres de la fonction noyau correspondante sont estimés. Enfin, les noyaux sont fusionnés. Phase de test : un graphe est extrait de chaque image de la base de test. Chaque graphe est alors comparé aux graphes de la base d'apprentissage et classifié selon la fonction de similarité apprise.

noyaux sont calculés par fusion de plusieurs attributs, en s'inspirant de [12]. Le schéma général de la méthode proposée est illustré sur la figure 1. La structure du graphe ainsi que les attributs sont décrits dans la section 3. La fusion des noyaux correspondant aux différents attributs, autre contribution originale de ce travail, est développée dans la section 4. Des résultats préliminaires, montrant l'amélioration de la classification par l'utilisation de plusieurs attributs, sont donnés dans la section 5.

2 Classification par une méthode à noyaux pour des graphes

Nous présentons ici brièvement le principe général de la méthode de classification proposée dans [1].

Une méthode de sur-segmentation utilisant un processus hiérarchique non supervisé [6, 9] est appliquée à chaque image (si elle n'est pas déjà segmentée). Les régions obtenues sont regroupées si elles sont adjacentes et si elles ont des niveaux de gris proches. Le critère d'arrêt du regroupement est fonction d'un seuil dynamique sur les différences de niveaux de gris entre régions adjacentes, mis à jour à chaque itération. Un graphe d'adjacence est alors construit : les nœuds sont les régions et les arcs représentent l'adjacence entre les régions. Dans [1] seule cette adjacence est considérée comme information structurelle et constitue implicitement un attribut d'arc. Des attributs de nœuds sont calculés tels

que la surface en pixels de la région, le rapport entre cette surface et celle de l'image (surface dite relative), le niveau de gris moyen et le niveau de gris relatif (par rapport à la dynamique de l'image), le périmètre de la région, sa compacité et le nombre de voisins dans le graphe. Ces attributs prennent des valeurs sur une échelle continue.

La 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 [8, 10], pondérée par la probabilité d'apparition de ces chemins.

Par rapport à d'autres contextes dans lesquels ce type de méthode a été utilisé [13], le voisinage des régions dans l'image a moins d'importance que le voisinage entre constituants d'une molécule par exemple. Ici 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. Nous avons donc proposé d'utiliser des noyaux gaussiens ou triangulaires pour évaluer la similarité entre deux valeurs a_1 et a_2 d'un attribut numérique a :

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

$$K_a^\Delta(a_1, a_2) = \begin{cases} \frac{\Gamma - \|a_1 - a_2\|}{\Gamma} & \text{si } \|a_1 - a_2\| \leq \Gamma \\ 0 & \text{sinon} \end{cases}$$

Etant donnés deux graphes G et G' à comparer, les fonctions de l'équation 1 permettent d'évaluer la similarité $k_a(h, h')$ entre deux marches aléatoires h dans G et h' dans G' , en combinant les similarités des valeurs de l'attribut a de tous les nœuds (ou arcs) le long de h et h' . Dans [1], une extension du noyau de base $k_a(h, h')$ est proposée pour mieux tenir compte des spécificités des attributs extraits des images. Les valeurs de similarité continues entre les constituants du graphe sont interprétées comme des pénalités sur les probabilités de transition, qui modifient les marches aléatoires sans les arrêter prématurément.

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

$$K_a(G, G') = \sum_h \sum_{h'} k_a(h, h') p(h|G) p(h'|G') \quad (2)$$

Cette fonction est ensuite utilisée dans un séparateur à vaste marge (SVM) [12] pour construire un classifieur d'images. K définit une matrice de similarité entre les paires de graphes à comparer.

3 Représentation des images par des graphes attribués

Alors que la méthode résumée ci-dessus ne considère que des attributs de nœuds caractérisant les régions de l'image, nous proposons ici d'améliorer la structure du graphe (en y ajoutant des arcs) et d'ajouter de l'information concernant la structure de l'image sous la forme de relations spatiales.

3.1 Structure du graphe

La méthode initiale [1] utilise un graphe d'adjacence. Nous pourrions nous contenter d'ajouter de l'information structurelle sur les arcs existants,

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. La figure 2 illustre différents résultats de segmentation probables du point de rencontre de quatre régions.

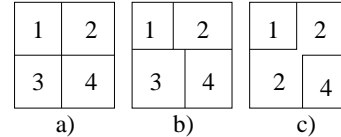


Figure 2 – Défaut de robustesse de l'adjacence : différents résultats de segmentation probables pour quatre régions. (a) Toutes les régions sont adjacentes deux à deux. (b) Les régions 1 et 4 ne sont plus adjacentes. (c) Les régions 2 et 3 ont fusionné, les régions 1 et 4 ne sont plus adjacentes.

L'ajout d'arcs comportant plus qu'une information implicite d'adjacence nous permet donc non seulement de mieux décrire les informations structurales, mais aussi d'améliorer la robustesse de la représentation. Le graphe généré ne sera donc plus un graphe d'adjacence. Il peut même devenir complet si cela ne nuit pas aux temps de calculs.

3.2 Ajout d'information structurelle

Nous proposons d'utiliser de nouveaux attributs décrivant l'information structurelle. Plus précisément, cette information sera décrite en utilisant des relations spatiales. Classiquement, les relations spatiales sont divisées en relations topologiques et relations métriques [11]. Parmi ces dernières, nous avons choisi deux exemples classiques : une relation de distance et une relation d'orientation (position relative directionnelle). Comme relation topologique, nous calculons un degré d'adjacence étendu. Nous allons maintenant présenter ces caractéristiques.

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 [4] et nous avons choisi ici de représenter cette information en utilisant des histogrammes

d'angles [14]. 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 :

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

où $\angle(\vec{i}, \vec{p_i q_j})$ représente l'angle entre un vecteur de référence \vec{i} et le vecteur $\vec{p_i q_j}$. L'extension à trois dimensions est une fonction de deux angles. Un exemple est illustré figure 3.

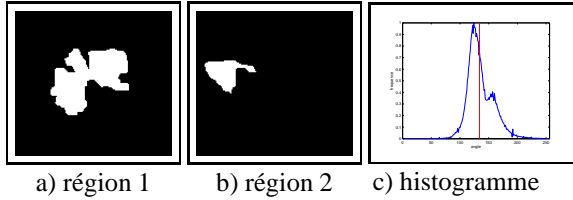


Figure 3 – Histogramme d'angles entre les régions 1 et 2. Le centre de gravité est représenté par une ligne verticale.

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 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.

Parmi les méthodes possibles, 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 [2] : 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 4 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.

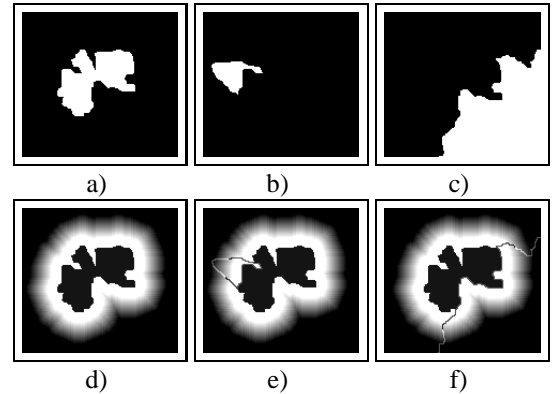


Figure 4 – (a) Région 1. (b) Région 2. (c) Région 3. (d) Sous-ensemble flou correspondant à la relation « proche de la région 1 ». (e) Même chose avec la frontière de la région 2 en sur-impression. (f) De même avec la région 3.

Nous allons à présent présenter deux critères permettant l'évaluation de la corrélation entre cette représentation floue et l'autre région. Parmi les différentes mesures de comparaison floues possibles, nous avons choisi une M-mesure de satisfaisabilité [5] définie par :

$$Sat(proche(R_1), R_2) = \frac{\sum_{x \in \mathcal{S}} \min(\mu_{proche(R_1)}(x), \mu_{R_2}(x))}{\sum_{x \in \mathcal{S}} \mu_{proche(R_1)}(x)} \quad (3)$$

Cette mesure est maximale si l'objet complet est inclus dans le noyau de $\mu_{proche(R_1)}$. Notons que la taille de la région où la relation est satisfaite n'est pas limitée et pourrait couvrir tout le domaine spatial. Si l'objet R_2 n'est pas flou, cette mesure se simplifie en $\frac{\sum_{x \in R_2} \mu_{proche(R_1)}(x)}{\sum_{x \in S} \mu_{proche(R_1)}(x)}$, c'est-à-dire la portion de $\mu_{proche(R_1)}$ qui est couverte par l'objet.

La figure 4 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 » est présenté en (d) et le contour des régions adjacentes a été ajouté respectivement dans (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.

Nous avons également choisi une mesure symétrique, contrairement à la satisfiabilité, une M-mesure de ressemblance [5] définie par :

$$Res(proche(R_1), R_2) = \frac{\sum_{x \in S} \min(\mu_{proche(R_1)}(x), \mu_{R_2}(x))}{\sum_{x \in S} \max(\mu_{proche(R_1)}(x), \mu_{R_2}(x))}. \quad (4)$$

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.

4 Fusion des attributs

Nous présentons maintenant comment différents attributs sont combinés dans le noyau de graphes. La méthode de fusion décrite est générale, et sera appliquée dans la suite aux trois exemples d'attributs structurels introduits dans la section 3 ainsi qu'à des attributs de régions.

L'intérêt de la fusion est de fournir un seul noyau pour la classification à partir de caractéristiques hétérogènes, telles que les attributs retenus. À partir d'un ensemble d'apprentissage et des graphes associés, la première étape est de construire les matrices des noyaux de base $\{K_{a_1}, \dots, K_{a_n}\}$ correspondant à chaque attribut a_i . Chacune d'entre elles, calculée selon l'équation 2, ne représente qu'une vue étroite des données. La classification d'images

fondée sur une seule vue n'est en général pas performante car une bonne discrimination ne peut pas être obtenue à partir d'un seul attribut. La fusion des informations apportées par chaque noyau est alors nécessaire. La manière la plus simple de combiner les noyaux est d'en faire une combinaison linéaire :

$$K = \sum_{i=1}^n \lambda_i K_{a_i} \quad (5)$$

Cette forme de fusion réalise un compromis qui permet une compensation entre différents points de vue sur les données, et améliore ainsi la souplesse de la classification.

La figure 5 illustre l'étape d'apprentissage par fusion d'attributs.

Une méthode d'obtention du vecteur de poids optimal λ a été proposée dans [12]. Elle considère le cône convexe P des matrices symétriques et définies positives :

$$P = \{X \in \mathbb{R}^{p \times p} \mid X = X^T, X \succeq 0\}.$$

Sur ce cône, le problème dual du SVM est optimisé globalement :

$$\min_{\lambda \in \mathbb{R}^n, K \in P} \max_{\alpha \in \mathbb{R}^m} 2\alpha^T e - \alpha^T D(y) K D(y) \alpha \quad (6)$$

sous la contrainte :

$$C \geq \alpha \geq 0, \text{ trace}(K) = c, K = \sum_{i=1}^n \lambda_i K_{a_i}, \alpha^T y = 0,$$

où m désigne la taille de la base d'apprentissage, $e \in \mathbb{R}^m$ le vecteur dont les éléments sont tous égaux à 1, et $D(y) \in \mathbb{R}^m \times \mathbb{R}^m$ la matrice dont les éléments sont nuls sauf sur la diagonale où ils prennent les étiquettes (+1 ou -1) des exemples d'apprentissage ($D(y)_{ii} = y_i$).

Dans ce problème d'optimisation, C représente le paramètre définissant la marge du SVM et $c \geq 0$ définit la trace de la matrice résultante. L'intérêt de cette méthode est qu'elle minimise la fonction de coût du classifieur en termes à la fois de frontières entre classes et des paramètres λ_i . Le résultat est un ensemble de poids et une fonction discriminante séparant les classes, combinant l'information de plusieurs noyaux.

Ce problème peut être transposé en un programme quadratique contraint [12], dont la solution primale donne les poids optimaux λ_i :

$$\min_{\alpha, t} 2\alpha^T e - ct \quad (7)$$

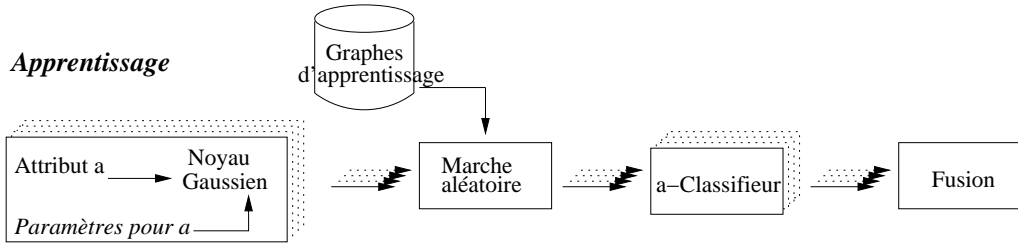


Figure 5 – Fusion des noyaux pour plusieurs attributs lors de l’apprentissage. Pour chaque attribut, un noyau gaussien est calculé. La similarité entre les marches aléatoires calculée pour chaque attribut définit un classifieur. Les classifieurs sont ensuite combinés linéairement.

sous la contrainte

$$t \geq \frac{1}{\text{trace}(K_{a_i})} \alpha^T D(y) K_{a_i} D(y) \alpha \quad i = 1, \dots, n$$

$$C \geq \alpha \geq 0, \alpha^T y = 0.$$

Pour chaque attribut, nous définissons donc un noyau (gaussien ou triangulaire), dont les paramètres sont choisis en fonction de la variabilité de l’attribut dans les données. Plus précisément, le seuil de la fonction de discrimination doit correspondre approximativement à la plus petite distance entre des valeurs d’attributs qui ne permettraient pas à un observateur de prendre une décision. Ce seuil dépend du type d’attributs et des données.

Pour chaque noyau d’attribut, nous construisons un noyau de graphe qui permet de mesurer la similarité entre graphes à partir de ce seul attribut comme décrit dans la section 2. Certaines caractéristiques sont plus discriminantes que d’autres pour certaines données. La fusion permet alors de construire une fonction de décision exploitant l’hétérogénéité des attributs et qui pondère l’importance de chacun en fonction de sa pertinence pour la classification, via les paramètres estimés λ_i . Cette fusion permet d’améliorer les performances de la classification dans le cas où aucun attribut ne permet seul une bonne classification (ce qui est généralement le cas). Cette propriété est illustrée expérimentalement dans la section suivante.

5 Expériences et résultats

La base d’images IBSR¹ contient des données cliniques réelles et est couramment utilisée en tant que base d’images par résonance magnétique (IRM) de

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

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 réussissent à classifier les images de deux vues différentes avec 100% de réussite. En conséquence, nous devons construire une base plus difficile à classifier². Nous avons donc tenté de classifier des images provenant de la même vue. Chaque élément de la base d’une vue va fournir trois coupes successives pour la première classe, puis trois autres coupes successives pour la seconde. Chaque ensemble de 54 images constitue une classe dont 15 éléments, sélectionnés de manière aléatoire, seront utilisés pour l’apprentissage du système. Le reste des images sera utilisé pour tester le classifieur. Le tableau 1 recense les indices des coupes utilisées dans chacune des trois vues pour définir les classes. Les coupes sont sélectionnées parmi celles qui dispose d’un nombre suffisant de régions après segmentation.

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

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

²certes de manière un peu artificielle

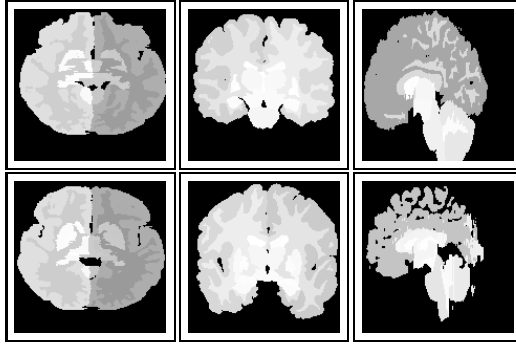


Figure 6 – Exemples d’images provenant de la base IBSR. Les niveaux de gris représentent des étiquettes. À gauche, en haut et en bas : deux coupes provenant d’une vue axiale du même volume IRM 3D et appartenant à chacune des classes. Au centre : vue coronale. À droite : vue sagittale.

Pour estimer la similarité pour un attribut, nous utilisons des noyaux gaussiens avec une valeur de seuil relativement petite qui les rend sensibles aux différences d’étiquetage. Chaque noyau d’attribut est injecté dans un noyau marginalisé pour graphe utilisé par l’algorithme du SVM. 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_2 : $\log_2 C \in \{-5, \dots, 15\}$. Pour chaque tâche de classification, nous utilisons $N = 30$ graphes pour l’apprentissage et $T = 78$ graphes pour le test, les deux valeurs étant également divisées entre les deux classes.

D’autre part, la fusion est réalisée pour des paires d’attributs (un attribut « région » et un attribut « relation spatiale »), à partir de leurs noyaux marginalisés respectifs. Afin de fusionner k noyaux, nous fixons le paramètre de contrainte de la trace de l’algorithme de fusion $c = kN$ et nous calculons les poids $\lambda_1, \dots, \lambda_k$ des noyaux en entrée de la fonction de fusion (équation 5), en résolvant le système d’équations 7 avec cvx ³. Enfin, les performances du noyau résultant sont testées avec un classifieur SVM. Quelques tests avec trois attributs ont également été réalisés selon le même principe.

Dans la plupart des cas, les résultats préliminaires regroupés dans le tableau 3 montrent une amélioration des performances par rapport aux taux de classification initiaux (tableau 2), ce qui

Att.	Axiale		Coronale		Sagittale	
	Par.	%	Par.	%	Par.	%
sa	0,01	0,79	0,05	0,74	0,05	0,85
re			0,01	0,82	0,01	0,95
su	0,01	0,69			0,01	0,96
co	0,01	0,87	0,01	0,81	0,01	0,91
ng	0,10	0,65	0,10	0,86	0,10	0,81

Tableau 2 – Performance de la classification pour différents attributs (sa : satisfaisabilité ; re : ressemblance ; su : surface relative ; co : compacité ; ng : niveau de gris relatif). La vue indique l’ensemble d’images utilisé. Les colonnes 2, 4 et 6 présentent les paramètres des noyaux gaussiens, les colonnes 3, 5 et 7 présentent les performances de classification pour chacun des plans de coupe et chaque attribut séparément.

Vue		Att.	Fusion		Att.	Fusion
	2	sa,co	0,90	5	sa,su,ng	0,96
	3	sa,ng	0,94			
C	6	re,ng	0,99	9	ng,co	0,87
	7	sa,co	0,83	10	sa,ng,co	0,87
	8	sa,ng	0,90			
S	11	re,su	0,96	15	sa,co	0,91
	12	re,ng	0,83	16	ng,co	0,95
	13	sa,su	0,96	17	sa,ng,co	0,95
	14	sa,ng	0,83			

Tableau 3 – Performance de la classification pour différentes combinaison d’attributs et leur fusion. La vue indique l’ensemble d’images utilisé. Les colonnes 3 et 6 présentent les attributs utilisés pour la fusion, les colonnes 4 et 7 les scores obtenus avec le noyau issu de la fusion.

démontre l’intérêt de la fusion. Les attributs des résultats 1, 3, 4, 6 et 16 semblent donner des points de vue complémentaires sur l’image et la fusion permet d’améliorer les performances. Les résultats 5 et 17 montrent que la fusion de 3 noyaux obtient le meilleur résultat des fusions deux à deux des 3 attributs concernés, ce qui correspond à un effet de saturation lorsque les taux de classification sont déjà élevés. Il y a également quelques cas (les résultats 10, 12 et 14) où la fusion diminue les performances du noyau par rapport à celles obtenues pour une combinaison d’un sous-ensemble des attributs impliqués, produisant un estimateur de type moyenne. En effet, l’optimisation globale du problème convexe de fusion garantit uniquement une meilleure limite statistique du nombre d’erreurs, et donc pas de meilleures performances sur chaque exemple de test. Notons que si la fusion des deux attributs du résultat 14 correspond à cette situation, la fusion de ces mêmes attributs avec un troisième au résultat 17 permet de dépasser les performances de chaque attribut. Il est donc possible d’éviter ce comportement en considérant plus d’attributs dans la fusion. Un autre aspect important à prendre

³Matlab Software for Disciplined Convex Programming, available at <http://www.stanford.edu/~boyd/cvx/>

en compte : la fusion accroît la dimension de l'espace des caractéristiques et il y a donc plus de risque de sur-apprentissage avec des ensembles d'entraînement de taille réduite.

Du point de vue de la complexité et du temps de calcul, le calcul des noyaux K_{a_i} entre deux graphes G et G' (équation 2) constitue la partie la plus coûteuse dans l'exécution de l'algorithme. La complexité du calcul de $K_{a_i}(G, G')$ est en $O((|G||G'|)^3)$, ce qui demande quelques millisecondes pour les graphes issus des images de la base IBSR et approximativement une minute pour des graphes plus complexes, de 60-70 nœuds.

6 Conclusion

Nous avons proposé une méthode de classification d'images fondée sur des noyaux marginalisés. En particulier, nous avons montré qu'une représentation des images sous forme de graphe, enrichi par des attributs numériques caractérisant les régions de l'image et les relations spatiales entre celles-ci, et une méthode de fusion d'attributs sont deux éléments clés pour améliorer les performances de la classification. Une représentation floue des relations spatiales est utilisée pour construire les noyaux correspondant aux attributs structurels. La fusion de noyaux est réalisée par une moyenne pondérée dont les poids sont estimés de manière automatique afin de donner plus d'importance aux attributs les plus pertinents. Des résultats préliminaires sur une base réduite d'images médicales montrent l'intérêt de l'approche proposée.

La suite de ce travail porte sur l'extension de l'étude expérimentale à d'autres bases d'images et pour des problèmes concrets plus pertinents. D'un point de vue méthodologique, d'autres types de fusion sont à l'étude, afin de mieux mettre en évidence l'intérêt de la combinaison de représentations structurelles floues et de méthodes à noyaux sur des graphes pour l'apprentissage et la classification d'image.

Références

- [1] E. Aldea, J. Atif, and I. Bloch. Image classification using marginalized kernels for graphs. In *6th IAPR TC-15 Workshop on Graph-based Representations in Pattern Recognition*, Alicante, Spain, June 2007.
- [2] I. Bloch. Spatial Representation of Spatial Relationships Knowledge. In A. G. Cohn, F. Giunchiglia, and B. Selman, editors, *7th International Conference on Principles of Knowledge Representation and Reasoning KR 2000*, pages 247–258, Breckenridge, CO, 2000. Morgan Kaufmann, San Francisco, CA.
- [3] I. Bloch. Fuzzy Spatial Relationships for Image Processing and Interpretation : A Review. *Image and Vision Computing*, 23(2) :89–110, 2005.
- [4] I. Bloch and A. Ralescu. Directional Relative Position between Objects in Image Processing : A Comparison between Fuzzy Approaches. *Pattern Recognition*, 36 :1563–1582, 2003.
- [5] B. Bouchon-Meunier, M. Rifqi, and S. Bothorel. Towards general measures of comparison of objects. *Fuzzy sets and Systems*, 84(2) :143–153, 1996.
- [6] L. Brun, M. Mokhtari, and F. Meyer. Hierarchical watersheds within the combinatorial pyramid framework. In *12th International Conference on Discrete Geometry for Computer Imagery*, volume 3429, pages 34–44, Poitiers, France, April 2005. Springer.
- [7] O. Chapelle, P. Haffner, and V. Vapnik. Support vector machines for histogram-based image classification. *IEEE Transactions on Neural Networks*, 10(5) :1055–1064, 1999.
- [8] T. Gaertner, P. Flach, and S. Wrobel. On graph kernels : Hardness results and efficient alternatives. In *16th Annual Conference on Computational Learning Theory*, pages 129–143, Washington, DC, USA, Aug. 2003.
- [9] K. Haris, S. N. Estradiadis, N. Maglaveras, and A. K. Katsaggelos. Hybrid image segmentation using watersheds and fast region merging. *IEEE Transactions on Image Processing*, 7(12) :1684–1699, 1998.
- [10] H. Kashima, K. Tsuda, and A. Inokuchi. Marginalized kernels between labeled graphs. In *20st Int. Conf. on Machine Learning*, pages 321–328, 2003.
- [11] B. Kuipers. Modeling spatial knowledge. *Cognitive Science*, 2 :129–153, 1978.
- [12] G.R.G. Lanckriet, N. Cristianini, P. Bartlett, L. El Ghaoui, and M.I. Jordan. Learning the Kernel Matrix with Semidefinite Programming. *Journal of Machine Learning Research*, 5 :27–72, 2004.
- [13] P. Mahé, N. Ueda, T. Akutsu, J.-L. Perret, and J.-P. Vert. Extensions of marginalized graph kernels. In *ICML '04 : 21st Int. Conf. on Machine Learning*, 2004.
- [14] K. Miyajima and A. Ralescu. Spatial organization in 2D segmented images : representation and recognition of primitive spatial relations. *Fuzzy Sets and Systems*, 65(2-3) :225–236, 1994.
- [15] M. Neuhaus and H. Bunke. Edit distance based kernel functions for attributed graph matching. In *5th IAPR TC-15 Workshop on Graph-based Representations in Pattern Recognition*, pages 352–361, Poitiers, France, April 2005.
- [16] M. Neuhaus and H. Bunke. A random walk kernel derived from graph edit distance. In *11th International Workshop on Structural and Syntactic Pattern Recognition*, volume 4109, pages 191–199, Hong-Kong, China, 2006. springer.
- [17] J. Sivic, B. C. Russell, A. A. Efros, A. Zisserman, and W. T. Freeman. Discovering objects and their location in images. In *IEEE Int. Conf. on Computer Vision*, volume 1, pages 370–377, 2005.