

Utilisation de relations ontologiques pour la comparaison d'images décrites par des annotations sémantiques

Camille Kurtz*, Daniel L. Rubin**

* LIPADE (EA 2517), Université Paris Descartes - camille.kurtz@parisdescartes.fr

** QIL (School of Medicine), Stanford University - dlrubin@stanford.edu

Résumé. Face à la complexité des nouvelles générations d'images médicales, les processus de recherche d'images basés sur leurs contenus visuels peuvent s'avérer insuffisants. Cet article propose une nouvelle approche basée sur l'annotation des images via des termes sémantiques pouvant pallier ce problème. Elle repose sur la combinaison d'une distance hiérarchique permettant de comparer les images en considérant les corrélations entre les termes utilisés pour les décrire et d'une mesure de similarité permettant d'évaluer la proximité sémantique entre des termes ontologiques. Cette approche est validée dans le cadre de la recherche d'images tomodensitométriques.

1 Introduction

Les méthodes automatiques de recherche d'images fournissent un moyen d'aide à la décision dans de nombreux domaines d'application. Dans le domaine médical, elles permettent d'assister les radiologues lors de leur travail d'interprétation d'images en identifiant des images similaires au sein de bases de données. Un cas typique d'utilisation est la recherche par l'exemple où l'on souhaite retrouver des images similaires à un exemple d'image donné en requête correspondant à un examen médical. Pour ce faire les images sont généralement décrites par leurs caractéristiques bas-niveaux (*e.g.*, couleur, texture) induites de leurs pixels et une mesure de distance est utilisée pour rechercher des images similaires dans l'espace des caractéristiques. Cependant, face à la complexité des nouvelles générations d'images médicales, les processus de recherche d'images basés sur le contenu peuvent s'avérer insuffisants. Une des limites principales est liée au problème du saut sémantique : les caractéristiques bas-niveaux ne sont plus suffisamment discriminantes pour caractériser le contenu visuel haut-niveau des images.

Pour proposer une solution à ces problèmes, des travaux récents ont montré l'intérêt de caractériser le contenu des images par des termes sémantiques (Kwitt et al., 2012). Ces termes peuvent être utilisés pour décrire un nombre important d'informations relatives au contenu visuel des images (Fig. 1(a-b)). Ils peuvent être dérivés des observations des radiologues ou automatiquement prédits à partir de caractéristiques bas-niveaux extraits des pixels. Par conséquent, l'intégration de termes sémantiques dans les processus de recherche d'images est une solution prometteuse pour faire face aux problèmes liés au saut sémantique. Cependant, les systèmes actuels de recherche d'images basés sur les annotations sémantiques ne considèrent pas les relations intrinsèques (*e.g.*, sémantiques, anatomiques) qui existent entre ces termes lors de la comparaison des images. Dans la plupart des cas, chaque image est modélisée par un

Une approche sémantique pour la comparaison d'images

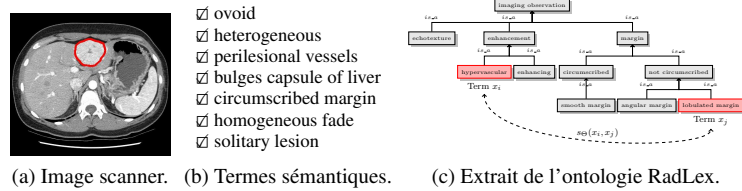


FIG. 1: Une image scanner du foie annotée avec des termes extraits d'une ontologie.

vecteur d'attributs (binaires) représentant la présence ou l'absence des termes et une distance classique (*e.g.*, Manhattan, Euclidienne) est calculée entre ces vecteurs. Cette stratégie ne permet donc pas de considérer la similarité entre les éléments sémantiques des vecteurs décrivant les images : la distance entre une image A représentant une lésion annotée avec le terme *ovoid* et une image B annotée avec le terme *round* est égale à la distance entre une image A annotée avec le terme *ovoid* et une image C annotée avec le terme *rectangular*. Cependant, l'image A est sémantiquement (et intuitivement) plus proche de l'image B qu'elle ne l'est de l'image C .

Des travaux récents (Tousch et al., 2012) ont montré que l'utilisation de vocabulaires contrôlés (comme les ontologies) pour l'annotation d'images pouvait fournir des solutions efficaces pour faire face à ces limites. Les ontologies peuvent être employées pour modéliser les relations sémantiques entre les termes utilisés pour la description des images. De plus, des travaux en traitement automatique des langues naturelles (Pivovarov et Elhadad, 2012) ont déjà proposés des approches pour mesurer la proximité sémantique entre des termes issus d'une ontologie (Fig. 1(c)). Par ailleurs, dans le domaine de la comparaison de vecteurs, des distances ont été proposées pour considérer la proximité relative entre les éléments composant les vecteurs, permettant de prendre en compte leurs corrélations (Turney et Pantel, 2010). En couplant ces trois stratégies, il existe ainsi une opportunité de considérer les relations sémantiques entre les termes lors de la comparaison d'images et d'améliorer les processus de recherche d'images.

Cet article propose une nouvelle approche basée sur la proximité sémantique du contenu des images, dédiée à la recherche d'images similaires au sein de bases de données (Sec. 2). Elle combine une distance hiérarchique permettant de prendre en compte les relations entre les éléments d'un vecteur (Kurtz, 2012) et une mesure de similarité ontologique permettant d'évaluer automatiquement la proximité sémantique entre des termes extraits d'une ontologie. Cette stratégie offre un moyen de capturer les corrélations sémantiques entre les annotations décrivant les images lors de leur comparaison. Cette approche est générique et peut ainsi être utilisée pour un large champ d'applications allant de la recherche d'images par l'exemple à la classification. Les résultats obtenus dans le domaine de la recherche d'images tomodensitométriques du foie se sont montrés encourageants et justifient l'intérêt de cette approche (Sec. 3).

2 Approche sémantique proposée

L'approche proposée est composée de deux étapes (Fig. 2) : (1) annotation d'une image requête avec des termes sémantiques extraits d'une ontologie et construction d'un vecteur de termes modélisant l'image, et (2) comparaison de l'image à d'autres images précédemment annotées via le calcul d'une distance entre vecteurs de termes couplée à une mesure ontologique.

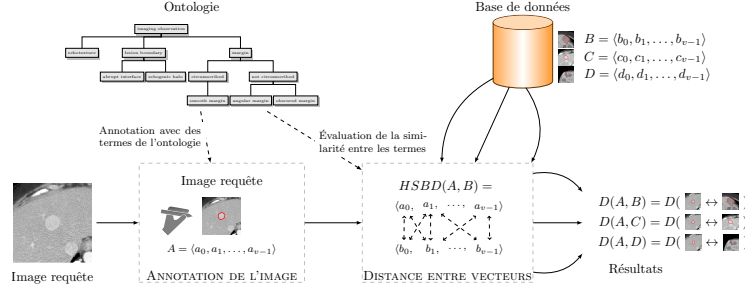


FIG. 2: Schéma de l'approche proposée.

2.1 Annotation des images

Soit \mathcal{I}_A une nouvelle image requête. La première étape consiste à décrire le contenu visuel de l'image (ou d'une région d'intérêt spécifique de l'image) par le biais d'un ensemble de termes sémantiques extraits d'une ontologie Θ . Cette étape peut être réalisée de façon automatique (par un algorithme d'apprentissage prédisant les termes à partir de caractéristiques extraites de la région d'intérêt) ou de façon manuelle. Dans le cadre de ces travaux, chaque image a été annotée par un radiologue via des termes de Θ .

Pour rendre les descriptions d'images comparables, l'approche proposée nécessite la création d'un sous-vocabulaire de termes extraits de Θ . Ce sous-vocabulaire, spécifique à l'application considérée, est défini en accord avec les radiologues. Il est noté $\mathcal{X} = \{x_0, x_1, \dots, x_{k-1}\}$ où les k termes $x_i \in \Theta$ sont des termes potentiellement utilisables pour décrire le contenu des images. Une fois l'image \mathcal{I}_A annotée, il devient possible de la caractériser via un vecteur de termes sémantiques $A = \langle a_0, a_1, \dots, a_{k-1} \rangle$ où chaque élément $a_i \in A$ est une valeur binaire représentant la présence ou l'absence du terme $x_i \in \mathcal{X}$.

2.2 Comparaison d'images

Les vecteurs de termes caractérisant les images peuvent ensuite être comparés via le calcul d'une distance hiérarchique nommée HSB (Hierarchical Semantic-Based Distance) (Kurtz, 2012). Lors de son calcul, HSB permet de considérer des valeurs de proximité entre les éléments des vecteurs comparés. Ces valeurs de proximité sémantique sont préalablement calculées par l'intermédiaire d'une mesure de similarité ontologique.

2.2.1 Similarité entre termes sémantiques

On trouve dans la littérature différents types de mesures permettant d'évaluer la similarité entre des termes appartenant à une ontologie. Parmi ces mesures, les mesures structurelles permettent de quantifier la similarité entre termes en se basant sur la structure de l'ontologie. Notre contexte applicatif nécessitant la proposition d'outils automatiques, notre étude s'est ainsi penchée sur l'utilisation de telles mesures.

Considérons deux termes sémantiques $x_i, x_j \in \mathcal{X}$. Un ensemble de liens connectant x_i et x_j dans l'ontologie est défini par $path(x_i, x_j) = \{l_0, \dots, l_{n-1}\}$. Pour quantifier une va-

leur de similarité sémantique entre deux termes x_i et x_j , une méthode intuitive a été proposée dans (Al-Mubaid et Nguyen, 2006). Elle est basée sur une stratégie par branches qui évalue la longueur du chemin minimum entre les termes et leur profondeur taxonomique au sein des branches considérées : plus le chemin entre les termes est long, plus les termes sont différents sémantiquement. L'idée est d'évaluer la spécificité commune (SC) des deux termes en soustrayant la profondeur de leur premier ancêtre commun (PAC) à la profondeur D_c de la branche principale auquel ces termes appartiennent. La spécificité commune est utilisée pour prendre en compte le fait que les couples de termes se trouvant dans les niveaux les plus bas de la hiérarchie sont plus similaires que les couples situés à un niveau plus élevé. Nous avons modifié la définition originale de cette mesure pour la normaliser et pour attribuer un poids égal à la longueur du chemin entre les termes et à leur spécificité commune :

$$s_{\Theta}(x_i, x_j) = \frac{\log \left(\sqrt{\min_{\forall p} |path_p(x_i, x_j)| \cdot SC(x_i, x_j)} + \gamma \right)}{\omega_{norm}^{\Theta}} \quad (1)$$

où ω_{norm}^{Θ} est un facteur de normalisation évaluant la valeur de similarité maximale entre termes et $SC(x_i, x_j) = D_c - depth(PAC(x_i, x_j))$ représente la spécificité commune des termes.

Pour modéliser l'ensemble des valeurs de similarité sémantique entre les k termes x_i appartenant au vocabulaire \mathcal{X} , nous définissons une matrice de dissimilarité notée \mathcal{M}^{sem} (de taille $k \times k$) où la valeur de chaque case $\mathcal{M}_{i,j}^{sem}$ est obtenue via la mesure $s_{\Theta}(x_i, x_j), \forall x_i, x_j \in \mathcal{X}$.

2.2.2 Distance entre vecteurs de termes

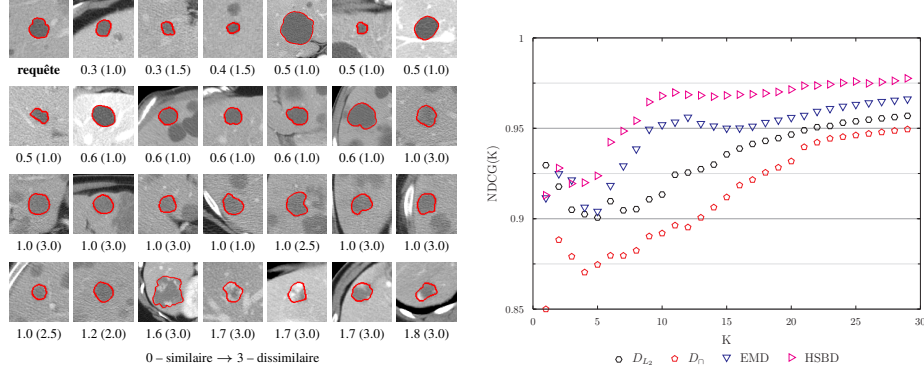
Le calcul de HSBD, entre deux vecteurs A et B , requière une matrice de dissimilarité modélisant les corrélations entre les éléments des vecteurs composant A et B . Pour ce faire, HSBD est initialisée avec la matrice de similarité sémantique \mathcal{M}^{sem} définie précédemment.

Avant de pouvoir calculer HSBD, la stratégie adoptée (basée sur un modèle fin-à-grossier) nécessite de définir un moyen de fusionner hiérarchiquement les différents termes représentés par les vecteurs en « clusters » de termes (*i.e.*, des termes de niveaux sémantiques plus élevés). Cette étape de pré-traitement repose sur la construction d'un dendrogramme \mathcal{D} induit par \mathcal{M}^{sem} et modélisant la hiérarchie de fusion des termes. La construction de \mathcal{D} est réalisée par le biais d'un algorithme de clustering hiérarchique ascendant. Il est à noter que cette étape de pré-traitement ne doit être effectuée qu'une seule fois pour une matrice \mathcal{M}^{sem} donnée.

Une fois que le dendrogramme \mathcal{D} a été construit, la distance HSBD peut être calculée. Son calcul se décompose en deux étapes principales :

- **Étape 1. Calcul des sous-distances hiérarchiques** Durant un processus de fusion itératif scannant chaque étage du dendrogramme (de ses feuilles jusqu'à sa racine), les vecteurs liés à A et B , et induits par la fusion des termes composant chaque cluster de l'étage courant, sont construits. Après chaque itération, la distance de Manhattan D_{L_1} est calculée entre chaque couple de vecteurs créé. Ces « sous-distances » permettent d'évaluer la similarité entre A et B à différents niveaux de sémantiques.
- **Étape 2. Fusion des sous-distances** Les sous-distances hiérarchiques calculées pour tous les étages du dendrogramme, et l'énergie sémantique nécessaire pour aller d'un étage à l'autre, sont ensuite fusionnées en une fonction qui est finalement intégrée pour fournir la valeur de la distance HSBD.

Pour plus de détails sur le calcul de HSBD, les lecteurs peuvent se référer à (Kurtz et al., 2013).



(a) Images similaires à une requête dans la base de données. Pour chaque image, la valeur calculée par HSBD et la valeur de similarité de référence (entre parenthèses) sont présentées.

(b) Scores NDCG des 4 distances évaluées.

FIG. 3: Évaluation des résultats de recherche d'images tomodynamométriques similaires.

3 Validation expérimentale

L'approche proposée a été validée dans le contexte de la recherche d'images similaires au sein d'une base de données de $N = 77$ images tomodynamométriques. Chaque image représente une coupe de foie affectée d'une lésion (e.g., tumeur, kyste). Chaque lésion a été annotée (12 termes en moyenne) par un radiologue par l'intermédiaire de termes appartenant à un sous-vocabulaire de 72 termes extraits de l'ontologie médicale RadLex (Langlotz, 2006).

Les expériences ont consisté à considérer successivement chaque image de la base comme une requête afin de rechercher des images similaires via la distance HSBD combinée à la mesure ontologique. Pour ce faire, l'image requête a été comparée à l'ensemble des $N - 1$ images qui ont ensuite été classées par ordre de similarité (Fig. 3(a)). Les valeurs de similarité ainsi calculées ont ensuite été comparées via l'indice NDCG (*Normalized Discounted Cumulative Gain*) à des valeurs de similarité de référence définies par un groupe de radiologues pour 30×30 couples d'images. L'indice NDCG permet d'estimer la pertinence d'un résultat d'expérience de classement par rapport à un classement de référence. Pour chaque image requête, la moyenne de cet indice a été calculée pour chaque $K = 1, \dots, 30$, évaluant ainsi la qualité du classement pour différents nombres d'images recherchées. Les résultats obtenus par HSBD ont également été comparés à ceux obtenus par d'autres distances de l'état de l'art : la distance Euclidienne D_{L_2} , l'intersection de vecteurs D_{\cap} et la distance EMD (Rubner et al., 2000).

La figure 3(b) présente les scores NDCG obtenus pour les 4 distances évaluées. À partir de ces résultats, on observe que les scores NDCG sont toujours supérieurs quand la comparaison des résultats est effectuée avec la distance HSBD que quand elle est effectuée avec les 3 autres distances. En particulier, l'approche proposée conduit à de meilleurs résultats que ceux obtenus avec la distance EMD qui permet également de prendre en compte les relations sémantiques entre les termes (via une initialisation avec la mesure ontologique présentée). Par ailleurs, les moins bons scores ont été obtenus avec les distances D_{L_2} et D_{\cap} qui ignorent durant leur calcul les relations sémantiques entre les éléments des vecteurs.

4 Conclusion et perspectives

Cet article a présenté une approche basée sur les relations ontologiques dédiée à la recherche d'images caractérisées par des annotations sémantiques. La principale originalité de cette approche est de prendre en compte les relations de proximité sémantique entre les termes employés pour décrire les images lors de la comparaison de ces dernières. Pour ce faire, cette approche est basée sur une distance hiérarchique entre vecteurs couplée à une mesure ontologique permettant d'estimer automatiquement les relations de proximité sémantique. Par la suite, nous prévoyons d'étudier de nouveaux processus pour l'annotation automatique d'images, ainsi que des mécanismes combinant l'utilisation d'ontologies multiples.

Références

- Al-Mubaid, H. et H. A. Nguyen (2006). A cluster-based approach for semantic similarity in the biomedical domain. In *Proceedings of the IEEE EMBC 2006 conference*, pp. 2713–2717.
- Kurtz, C. (2012). Une distance hiérarchique basée sur la sémantique pour la comparaison d'histogrammes nominaux. In *Actes de la Conférence EGC 2012*, pp. 77–88.
- Kurtz, C., P. Gançarski, N. Passat, et A. Puissant (2013). A hierarchical semantic-based distance for nominal histogram comparison. *Data & Knowledge Engineering* 87(1), 206–225.
- Kwitt, R., N. Vasconcelos, N. Rasiwasia, A. Uhl, B. Davis, M. Häfner, et F. Wrba (2012). Endoscopic image analysis in semantic space. *Medical Image Analysis* 16(7), 1415–1422.
- Langlotz, C. P. (2006). RadLex : A new method for indexing online educational materials. *RadioGraphics* 26(6), 1595–1597.
- Pivovarov, R. et N. Elhadad (2012). A knowledge-based and data-driven approach to identifying semantically similar concepts. *Journal of Biomedical Informatics* 45(3), 471–481.
- Rubner, Y., C. Tomasi, et L. J. Guibas (2000). The Earth Mover's Distance as a metric for image retrieval. *International Journal of Computer Vision* 40, 99–121.
- Tousch, A. M., S. Herbin, et J. Y. Audibert (2012). Semantic hierarchies for image annotation : A survey. *Pattern Recognition* 45(1), 333–345.
- Turney, P. D. et P. Pantel (2010). From frequency to meaning : Vector space models of semantics. *Journal of Artificial Intelligence Research* 37(1), 141–188.

Summary

Automated image retrieval applications could assist radiologists by identifying similar images in large archives. However, the semantic gap between low-level image features and their high level semantics may impair the system performances. This article proposes a new approach relying on image annotation with semantic terms that can be considered to deal with this issue. This approach combines a hierarchical distance used to take into account the correlations between the terms and an ontological similarity measure enabling to automatically assess term similarities. We validate this approach in the context of the retrieval of CT images.