



UNIVERSITÉ DE FRANCHE-COMTÉ

L2 CMI Informatique

PROJET DE RECHERCHE DOCUMENTAIRE

---

# État de l'art des ontologies médicales de l'abdomen

---

ATASSI Alain, BORDE Corentin

*Tuteurs :* BETBEDER Marie-Laure, HENRIET Julien

Année universitaire 2017-2018

## Remerciements

Nous tenons à remercier nos tuteurs Marie-Laure BETBEDER, Julien HENRIET et le docteur Yann CHAUSSY pour les précieux conseils et le suivi qu'ils nous ont apportés.

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>État de l'art</b>	<b>5</b>
2.1	Segmentation . . . . .	5
2.1.1	Segmentation par seuillage . . . . .	5
2.1.2	Segmentation fondée sur les régions . . . . .	6
2.1.3	Segmentation fondée sur les contours . . . . .	6
2.1.4	Segmentation fondée sur la division . . . . .	6
2.2	Raisonnement à partir de cas . . . . .	7
2.2.1	Constitution des cas . . . . .	7
2.2.2	Mise en forme de base de cas . . . . .	7
2.2.3	Mesure de similarité . . . . .	7
2.2.4	Adaptation et révision de la solution . . . . .	8
2.2.5	Mesure de satisfaction et capitalisation . . . . .	8
2.2.6	Les différents types de connaissances dans un RàPC . . . . .	8
2.3	Ontologie . . . . .	10
2.3.1	Définition et intérêt . . . . .	10
2.3.2	Outils . . . . .	10
<b>3</b>	<b>Contribution</b>	<b>11</b>
3.1	Recherche d'ontologies . . . . .	11
3.1.1	Nacional Cancer Institute Thesaurus Ontology (NCIT Ontology) . . . . .	11
3.1.2	Breast Cancer Grading Ontology (BCGO) . . . . .	12
3.1.3	Children's Health Exposure Analysis Ressource (CHEAR) . . . . .	13
3.1.4	Cell Ontology (CL) . . . . .	13
3.1.5	Cell Behavior Ontology (CBO) . . . . .	13
3.1.6	Human Disease Ontology (HDO) . . . . .	13
3.1.7	Tableau récapitulatif des contraintes respectées par chaque ontologie . . . . .	14
3.2	Création d'une ontologie . . . . .	14
3.2.1	Description des structures anatomiques les unes par rapport aux autres . . . . .	14
3.2.2	Instances . . . . .	17
3.2.3	Perspectives d'évolution . . . . .	17



# 1 Introduction

Le projet SAIAD (Segmentation Automatique de reins tumoraux chez l'enfant par Intelligence Artificielle Distribuée) a pour objectif de développer un outil de segmentation automatique des scanners des patients, en utilisant des techniques d'intelligence artificielle.

C'est un projet franco-suisse qui fait collaborer l'institut FEMTO-ST, le CHRU de Besançon, IDO-In (start-up bisontine), l'École Polytechnique Fédérale de Lausanne et CFI (PME Suisse).

Le néphroblastome (ou tumeur de Wilms) est une tumeur cancéreuse du rein chez l'enfant de moins de 6 ans. Le diagnostic est basé sur l'imagerie et l'étude des tissus biologiques prélevés pendant une intervention chirurgicale.

Il est important de bien analyser les données pré-opératoires pour pouvoir localiser précisément la tumeur et ainsi établir une bonne stratégie thérapeutique. Il est donc primordial que la segmentation d'image scanner soit réalisée dans de bonnes conditions et soit la plus précise possible.

Actuellement, l'étude des images médicales prend énormément de temps. En effet, la visualisation entière de l'abdomen nécessite l'analyse d'environ 250 coupes. Les méthodes actuelles de segmentation d'images médicales sont manuelles ou semi-automatiques et nécessitent plusieurs heures de travail. L'objectif du projet SAIAD est d'automatiser la segmentation efficacement et de modéliser les objets (organes, cavités et tumeurs) en images en 3 dimensions.

Réaliser la segmentation automatiquement pour ce projet soulève des problématiques :

- Les patients se trouvent dans une tranche d'âge (de 1 à 6 ans) où la croissance des organes est forte, cela engendre des différences morphologiques entre les patients ;
- À cet âge, les patients ont très peu de tissus graisseux entre la tumeur et les muscles, or ces deux entités ont le même niveau de gris ce qui les rend difficiles à distinguer ;
- La forme d'une tumeur est très aléatoire, il n'existe donc pas de modèle prédéfini ;
- Le rein se trouve très affecté par le développement de la tumeur : il peut être compressé ou déplacé ce qui rend la segmentation plus difficile ;
- Le contraste des différents scanners est faible, cela implique qu'il sera difficile de distinguer les différents objets.

Pour réaliser une segmentation automatisée, il est possible d'utiliser la méthode du raisonnement à partir de cas. Pour que cette méthode soit optimale, il faut inclure des connaissances théoriques. Les ontologies fournissent les connaissances théoriques. Elles permettent de ranger les connaissances de façon structurée.

## 2 État de l'art

### 2.1 Segmentation

La segmentation [16] d'image est une opération qui permet de regrouper des pixels entre-eux suivant des critères prédéfinis. Les pixels regroupés formeront alors des régions, constituant un pavage, ou une partition de l'image. Chaque région ou groupe de pixels est supposé correspondre à un « objet » de l'image. La segmentation est une étape très importante en traitement d'image. Il existe de nombreuses méthodes de segmentation que l'on peut regrouper en plusieurs grandes catégories :

- La segmentation fondée sur la classification ou le seuillage des pixels en fonction de leur intensité : c'est le « prétraitement » ;
- La segmentation fondée sur les régions comme par exemples la croissance de région ou la décomposition/fusion. Elle est utilisée après avoir fait un seuillage ;
- La segmentation fondée sur les contours ;
- La segmentation fondée sur la division.

#### 2.1.1 Segmentation par seuillage

La segmentation par seuillage permet de partitionner une image à partir de son histogramme. L'histogramme d'une image est la représentation du nombre de pixels en fonction de leur niveau de gris. Le niveau de gris est une valeur comprise entre 0 et 255.

Il existe différentes manières de faire une segmentation par seuillage, les plus courantes sont :

- La détection de vallée ;
- La méthode du pourcentage ;
- La méthode de minimisation de variance.

La méthode par minimisation de variance consiste à regrouper un ensemble de données (par exemple des pixels) pour faire en sorte que les données d'un même groupe (aussi appelés *clusters*) soient plus proches les unes des autres que celles des autres groupes. La méthode des k-moyennes est utilisée. L'algorithme des *k-means* est composé des étapes suivantes :

- **Initialisation** : des pixels sont choisis au hasard en fonction de leur niveau de gris qui deviendront les centres des clusters ;
- **Affectation** : chaque donnée est regroupée avec le centre d'une classe dont elle est la plus proche en distance ;
- **Étape de recalage des centres** : le centre de chaque clusters est recalculé en étant la moyenne arithmétique de toutes les données du clusters.

L'algorithme, une fois terminé, permet d'obtenir une estimation des classes et des centres. Cela permet de connaître l'appartenance de chaque donnée à une classe.

### 2.1.2 Segmentation fondée sur les régions

La réalisation de ce type de segmentation nécessite le partitionnement de l'image en régions homogènes. Les principales techniques sont la croissance de région, la division-fusion ou encore la fusion de régions.

La croissance de région consiste à placer des pixels germes au centre des différentes régions de l'image et à appliquer l'algorithme de remplissage par diffusion pour chaque pixel germe. Cet algorithme admet trois paramètres : la position du pixel de départ, la couleur ciblée et la couleur de remplacement. La couleur du pixel germe est comparée avec la couleur ciblée. Si elles sont égales alors on change la couleur de ce pixel avec la couleur de remplacement et on applique récursivement l'algorithme aux pixels voisins (on peut l'appliquer seulement à 4 voisins ou à 8 voisins).

### 2.1.3 Segmentation fondée sur les contours

La segmentation fondée sur les contours [11] utilise les contours de chaque objet de l'image. Des filtres dérivateurs vont permettre d'obtenir les contours des objets. L'utilisation de filtres de Sobel, Roberts par exemple reposent sur des critères d'optimalité. Cette technique n'est pas très utilisée car la méthode ne donne pas de très bon résultat : les contours ne sont pas fermés et ne sont pas détectés.

### 2.1.4 Segmentation fondée sur la division

L'algorithme par division-fusion (*split and merge*) détecte les régions précises d'une image. L'algorithme est découpé en deux étapes. La première étape correspond à la division, on va regarder chaque région : si elle ne respecte pas le critère d'homogénéité, alors elle est divisée en plusieurs sous-régions. Cette étape est appliquée récursivement et demande l'utilisation d'un graphe. Il y a deux types de graphes : le quad-tree, où chaque nœud correspond à une région et le graphe d'adjacence des régions qui permet une gestion plus simple des régions voisines.

La deuxième étape correspond à la fusion, le graphe est parcouru : chaque région est comparée à ses voisines, et si elles respectent le même critère d'homogénéité alors elles sont fusionnées pour créer une région plus grande.

Toutes ces techniques ne sont pas forcément très précises, il faut donc faire appel à un expert qui se chargera de vérifier la segmentation réalisée par la machine.

## 2.2 Raisonnement à partir de cas

Le raisonnement à partir de cas (RàPC) est une méthode, créée par Janet Kolodner en 1992 [20], permettant de résoudre un problème grâce à l'expérience, l'analogie et la réutilisation de problèmes déjà résolus (un cas est constitué d'un problème déjà résolu et de sa solution). Pour ce faire, des mesures de similarité sont utilisées pour trouver dans une base de données des cas similaires au problème posé.

Le RàPC peut se définir par 4 R : *Retrieve*, *Reuse*, *Revise*, *Retain*, en français Remémoration, Adaptation, Révision et Capitalisation.

### 2.2.1 Constitution des cas

Le but est de représenter simplement et efficacement le problème et la solution. Ils sont représentés par des concepts (ou des critères) qui ont une valeur. Il faut distinguer les concepts qui représentent le problème de ceux qui représentent la solution. Ces concepts peuvent avoir un poids qui va déterminer leur niveau d'importance.

### 2.2.2 Mise en forme de base de cas

La base de cas est une structure de données qui permet de stocker les cas résolus. Des cas initiaux ont été ajoutés au tout début dans la structure, puis des nouveaux cas vont être ajoutés en utilisant le système de RàPC. Il y a différentes manières de structurer la base, le choix de la structure s'effectue en fonction du nombre de cas, des besoins ou encore de la vitesse de recherche. La structure doit cependant respecter des contraintes :

- Stockage des cas ;
- Accès aux cas ;
- Mise à jour ;
- Pérennité ;
- Lecture rapide.

### 2.2.3 Mesure de similarité

Cette étape est la plus importante pour le processus de RàPC. Il faut parcourir la base pour choisir le cas le plus similaire à notre problème. Le cas trouvé est utilisé pour paramétrer la solution à notre problème. Des formules sont utilisées pour calculer la similarité entre deux cas. Pour pouvoir les utiliser, il faut



que les données soient numériques ou qu'elles puissent être converties. Il existe plusieurs types de formules :

- Formules de similarité classique : ces formules utilisent une fonction distance (généralement distance de Manhattan, de Hamming ou de Minkowski). Il faut ensuite utiliser une formule de conversion vers la similarité ;
- Formule de similarité de Perner : Perner utilise des fonctions de similarité différentes des formules classiques. Perner est une chercheuse qui a couplé le RàPC à une technique de segmentation d'image : la segmentation par ligne de partage des eaux. Sa formule est constituée d'une partie concernant les niveaux de gris et d'une partie concernant les métadonnées.
- Formule de similarité SSIM : l'indice SSIM (*Structural SIMilarity*) est utilisé principalement dans le domaine de la compression d'image (pour vérifier si l'image post-traitement et l'image originale sont les mêmes) mais peut aussi être utilisée en RàPC car ses caractéristiques mathématiques se rapprochent des données statistiques de la RàPC.

#### 2.2.4 Adaptation et révision de la solution

Cette étape est très difficile à automatiser, il existe peu de travaux ayant une phase d'adaptation automatisée. Le plus souvent, les critères sont modifiés manuellement en fonction du problème étudié.

#### 2.2.5 Mesure de satisfaction et capitalisation

Cette étape est réalisée à la fin du processus de RàPC, elle est très compliquée à automatiser. Elle permet d'évaluer la qualité de la segmentation qui vient d'être réalisée. Pour l'évaluer, des indices comme DICE ou IU (ou indices de Jaccard) sont utilisés afin de comparer les résultats obtenus à ceux obtenus à la main à l'aide d'un logiciel de segmentation. Les indices de Jaccard sont utilisés en statistique pour comparer la similarité et la diversité entre échantillons. Ces mesures de satisfaction sont réalisées par des experts.

#### 2.2.6 Les différents types de connaissances dans un RàPC

Nous listons ici les différents types de connaissances utilisés dans un RàPC et les moments où ceux-ci interviennent :

- les cas : lors de la recherche, dans la base de cas, de cas similaires au problème posé. Ce type de connaissances intervient donc au moment de la Remémoration (*Retrieve*)
- les règles d'adaptation : elles permettent d'adapter la solution d'un cas, elles peuvent être vécues ou apprises. Ce type de connaissances intervient au moment de l'Adaptation (*Reuse*)
- l'expert qui révise la solution et la met en œuvre. Ce type de connaissances intervient au moment de la Révision (*Revise*)

- l'ontologie de domaine : elle intervient notamment au moment de la définition des concepts, c'est-à-dire au moment de la définition d'un problème et au moment de la définition d'une solution.

## 2.3 Ontologie

### 2.3.1 Définition et intérêt

Il existe de nombreuses définitions d'une ontologie. Une des plus courtes et des plus célèbres est donnée par Thomas Gruber [19] :

« An ontology is a specification of a conceptualization. »

Thomas Gruber est un informaticien, inventeur et entrepreneur américain connu pour sa définition d'ontologie dans le contexte d'intelligence artificielle. Son article "*A Translation Approach to Portable Ontology Specifications*", daté d'avril 1993 est souvent cité pour la définition de l'ontologie en tant que spécification formelle d'une conceptualisation.

D'après Bruno Bachimont [17], enseignant-chercheur à l'Université de Compiègne et ancien directeur scientifique à l'Institut national de l'audiovisuel :

« On peut caractériser une ontologie comme une structuration des concepts d'un domaine. Ces concepts sont rassemblés pour fournir les briques élémentaires et exprimer les connaissances dont on dispose dans ce domaine. »

Une ontologie donne aux chercheurs un vocabulaire commun, et permet ainsi un partage de l'information. "*Ontology Development 101 : A Guide to Creating Your First Ontology*", une publication de Natalya F. Noy et Deborah L. McGuinness [15], traduite en français par Anila Angjeli (BnF, Bureau de normalisation documentaire), donnent les principales raisons pour lesquelles on peut développer une ontologie :

« Partager la compréhension commune de la structure de l'information entre les personnes ou les fabricants de logiciels, permettre la réutilisation du savoir sur un domaine, expliciter ce qui est considéré comme implicite sur un domaine, distinguer le savoir sur un domaine du savoir opérationnel, analyser le savoir sur un domaine »

Une ontologie est composée de classes, de propriétés et d'instances. Les classes permettent de décrire les concepts dans le domaine étudié. Ces classes sont rangées sous forme d'arborescence. Les propriétés permettent de décrire les classes et les instances (et permettent aussi de les lier entre elles). Une instance est un objet constituant un exemplaire d'une classe. Elle possède des valeurs précises. En pratique, développer une ontologie comprend de définir les classes, de définir les propriétés et de créer les différentes instances.

### 2.3.2 Outils

Il existe plusieurs outils qui permettent de développer et gérer des ontologies. Un des plus populaires est Protégé [10] : gratuit et open-source, il a été développé à l'université de Stanford.

## 3 Contribution

Comme nous l'avons vu dans la section 2.2.6, le RàPC permet d'inclure des connaissances théoriques, et notamment des ontologies. C'est pourquoi il est intéressant, dans le cadre du projet SAIAD, de réaliser une ontologie médicale de l'abdomen. Cette ontologie doit inclure des données fournies par les experts. Les médecins peuvent notamment fournir des informations floues suite à une échographie : ces informations aideront à situer les organes les uns par rapport aux autres, ce qui peut permettre de distinguer clairement deux cas très similaires au niveau des informations statistiques de l'image (moyenne, variance, etc.).

### 3.1 Recherche d'ontologies

Nous avons tout d'abord recherché des ontologies médicales qui devaient respecter au moins une de ces contraintes :

- Ontologie sur l'abdomen (pas forcément exclusivement)
- Ontologie sur l'enfant (pas forcément exclusivement)
- Description de tumeur
- Descriptions anatomiques
- Description, représentation de cellules

Nous avons effectué la plupart de nos recherches sur le site *Biportal*. Ce site coopère avec le NCBO (National Center for Biomedical Ontology) [14]. Ce centre permet de soutenir les chercheurs biomédicaux en leur permettant d'accéder à de nombreuses ontologies.

Nous avons trouvé 7 ontologies qui satisfont au moins un critère.

#### 3.1.1 Nacional Cancer Institute Thesaurus Ontology (NCIT Ontology)

Le National Cancer Institute (NCI) [8] est le principal organisme gouvernemental américain pour la recherche sur le cancer.

Le National Cancer Institute Thesaurus (NCIT) [7] fournit une terminologie de référence sur le cancer. Le projet NCIT OBO Edition a pour objectif d'accroître l'intégration du NCIT avec les ontologies OBO. Il existe donc une ontologie basée sur le NCIT.

Le NCIT fournit un large vocabulaire des soins cliniques, de la recherche, de l'information publique et des activités administratives. Il recense plus de 100 000 définitions textuelles et plus de 400 000 liens entre les concepts. Le contenu est régulièrement mis à jour par des experts. Le NCIT est un standard largement reconnu, il est notamment utilisé par le Clinical Data Interchange Standards Consortium (CDISC).

L'ontologie basée sur le NCIT [6] se distingue des autres ontologies que nous avons trouvées par sa taille : elle contient 138 291 classes. Sa bonne organisation

permet néanmoins un accès rapide aux informations recherchées. Elle contient notamment de nombreuses informations sur différentes tumeurs et sur les cellules cancéreuses, y compris chez l'enfant, mais ne décrit pas la situation spatiale des structures anatomiques.

### 3.1.2 Breast Cancer Grading Ontology (BCGO)

Cette ontologie [1] est utilisée pour la détection de cancer du sein. Elle permet d'attribuer une note à la tumeur.

Avec 133 classes, cette ontologie propose une représentation spatiale des éléments anatomiques de la poitrine. Elle possède de nombreux descripteurs spatiaux dont nous nous sommes inspirés pour la création de notre ontologie. Elle est intégrée dans une plateforme de microscope virtuel cognitif (qui permet d'exploiter des images issues de microscopie) et permet d'annoter sémantiquement les images histopathologiques. Elle conduit la phase d'exploration d'image et aide au pronostic en utilisant des requêtes sémantiques.

La méthodologie suivie pour construire cette ontologie est illustrée par la figure 1. On peut distinguer 3 phases : acquisition des connaissances, traduction de celles-ci en une représentation formelle et révision des connaissances. Le Nottingham Grading System et le National Cancer Institute Thesaurus (NCIT) fournissent les connaissances médicales nécessaires pour la phase d'acquisition des connaissances.

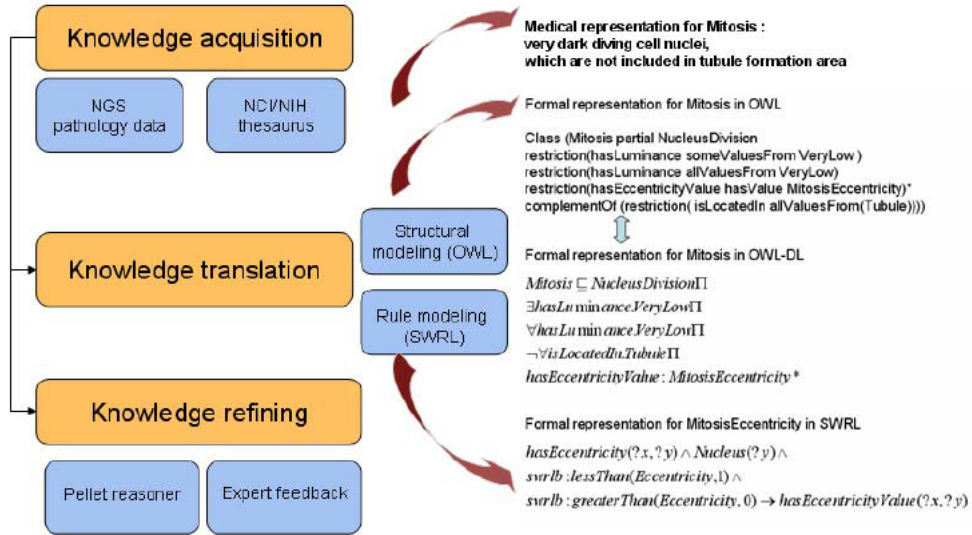


FIGURE 1 – Méthodologie pour la construction de l'ontologie BCGO

### 3.1.3 Children’s Health Exposure Analysis Ressource (CHEAR)

La santé des enfants est influencée par des facteurs environnementaux. Pour aider les chercheurs qui s’intéressent à la santé des enfants, le National Institute of Environmental Health Sciences (NIEHS) [9] a mis au point une ontologie : Children’s Health Exposure Analysis Resource (CHEAR) [4]. Celle-ci leur fournit les données nécessaires pour ajouter ou étendre leur considération des facteurs environnementaux dans leurs recherches.

### 3.1.4 Cell Ontology (CL)

Cette ontologie [3] décrit les types cellulaires d’animaux (elle n’est donc pas spécifique à un organisme, mais décrit des types cellulaires allant des procaryotes aux mammifères). Elle est utilisée dans des bases de données bio-informatiques où elle répond à un besoin de vocabulaire standard et contrôlé pour la description de cellules animales. Elle est également utilisée par de nombreuses autres ontologies (BioAssay Ontology, Cell line ontology, Plant Ontology).

Elle possède 6 636 classes et inclut notamment des descriptions de cellules cancéreuses et quelques descriptions anatomiques.

### 3.1.5 Cell Behavior Ontology (CBO)

Avec 246 classes, la Cell Behavior Ontology (CBO) [2] décrit l’organisation spatiale de groupes cellulaires et les comportements de ces cellules (croissance, mouvement, adhésion, mort, *etc.*).

### 3.1.6 Human Disease Ontology (HDO)

Avec 12 694 classes, la Human Disease Ontology (HDO) [5] propose un vocabulaire pour la représentation des maladies humaines. De nombreux cancers sont décrits dans cette ontologie, et notamment le cancer du rein chez l’enfant. Cependant, l’ontologie ne propose aucune description spatiale des structures anatomiques.

### 3.1.7 Tableau récapitulatif des contraintes respectées par chaque ontologie

	NCIT	BCGO	CHEAR	CL	CBO	HDO
Abdomen	+					+
Enfant	+		+			+
Description tumeur	+	+		+		+
Descriptions anatomiques		+	+	+		
Niveau cellulaire	+	+	+	+	+	+
Date de dernière mise à jour	27/03/2018	24/05/2011	15/02/2018	12/12/2017	10/10/2017	02/03/2018

Les contraintes que nous avons préalablement établies ont axé nos recherches. Ce tableau fait la synthèse de celles qui sont respectées, pour chaque ontologie que nous avons trouvée. Il donne également l'information de la date de dernière mise à jour de l'ontologie.

## 3.2 Création d'une ontologie

Aucune des ontologies que nous avons trouvées ne décrit les emplacements des structures anatomiques de l'abdomen les unes par rapport aux autres. Nous devons donc créer une nouvelle ontologie en nous inspirant de celles que nous avons étudiées. Nous utilisons le logiciel Protégé pour créer notre ontologie.

Nous adoptons la représentation décrite par la figure 2. Nous distinguons 2 sous-classes de la classe principale "Patient" : une sous-classe servant à définir les différents éléments anatomiques de l'abdomen et une autre qui permettra d'instancier des scans avec toutes les coupes qu'ils contiennent.

### 3.2.1 Description des structures anatomiques les unes par rapport aux autres

Toutes les classes de la figure 2 représentant des éléments anatomiques de l'abdomen sont reliées entre elles par des propriétés permettant de situer ces éléments les uns par rapport aux autres. Ces propriétés sont des descripteurs spatiaux (à droite de, devant, au dessus de, proche de, *etc.*). Par exemple, la figure 3 montre toutes les propriétés qui permettent de situer le rein gauche par rapport aux autres structures de l'abdomen. On note également d'autres propriétés qui caractérisent le rein gauche : la forme, le volume, et le niveau de gris sur image scanner. La figure 4 montre une représentation graphique des liens entre les classes qui permettent de situer le rein gauche dans l'abdomen. La figure 5 montre tous les liens entre les classes qui permettent de situer les structures anatomiques de l'abdomen les unes par rapport aux autres. Nous nous sommes servi d'ouvrages d'anatomie afin de recueillir les informations nécessaires à la description spatiale des structures anatomiques de l'abdomen.

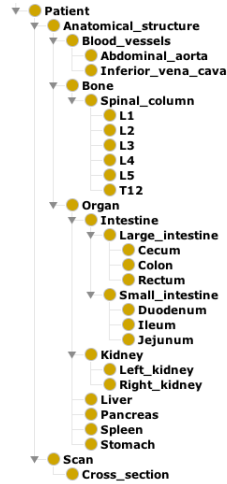


FIGURE 2 – Classes de l'ontologie pour le projet SAIAD

Description: Left_kidney
● above some Colon
● backward some Colon
● backward some Pancreas
● backward some Stomach
● closeTo some Abdominal_aorta
● closeTo some Pancreas
● closeTo some Spleen
● closeTo some Stomach
● forme some xsd:string
● forward some L1
● forward some L2
● forward some L3
● forward some T12
● Kidney
● onTheLeftOf some Abdominal_aorta
● onTheLeftOf some L1
● onTheLeftOf some L2
● onTheLeftOf some Pancreas
● onTheLeftOf some T12
● onTheRightOf some Spleen
● shadesOfGrey some owl:real
● volume some owl:real

FIGURE 3 – Propriétés de la classe "Left kidney"



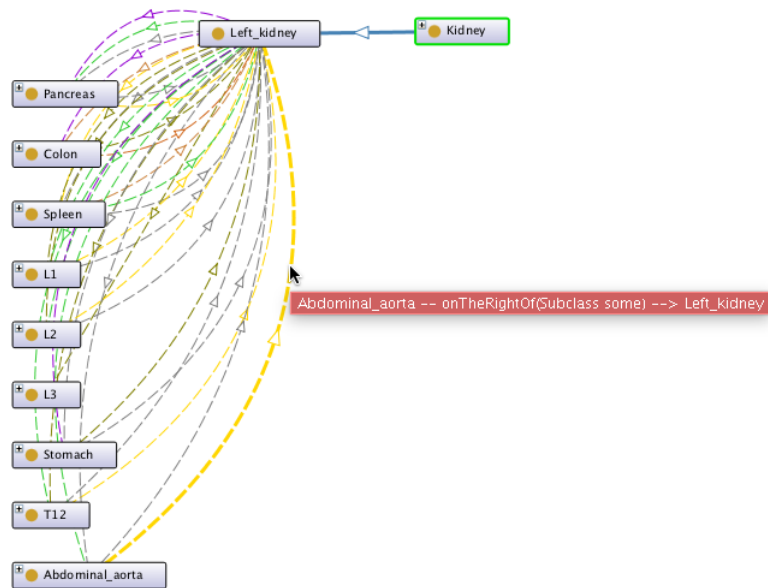


FIGURE 4 – Les liens entre les classes permettant de situer le rein gauche dans l'abdomen

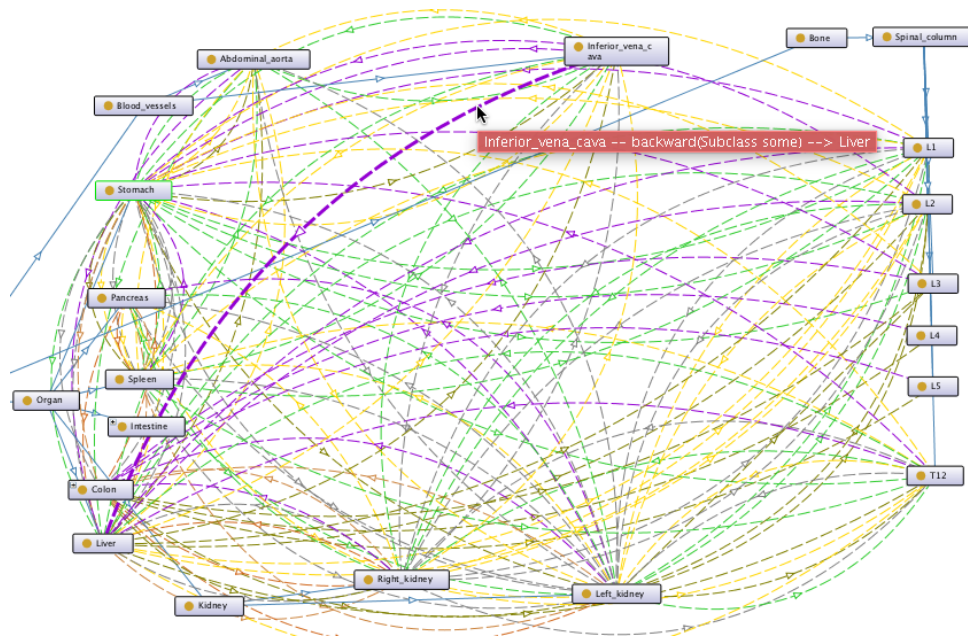


FIGURE 5 – Les liens permettant de situer les structures anatomiques de l'abdomen les unes par rapport aux autres

### 3.2.2 Instances

Des propriétés permettent de relier les instances des classes Patient, Scan et Cross-section. Ces propriétés permettent d'attribuer des scans à un patient et des coupes à un scan, comme le montre la figure 6. Les instances de coupes ont dans leurs propriétés un lien qui mène vers l'image correspondant à la coupe, comme le montre la figure 7 pour la coupe CS-1.

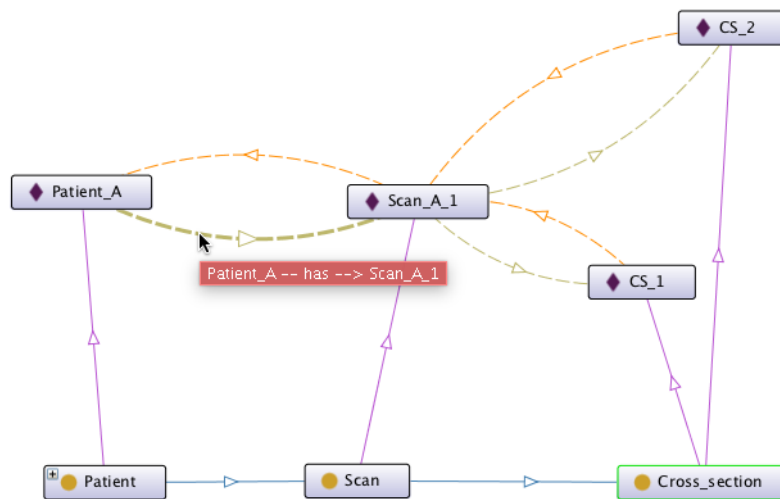


FIGURE 6 – Liens entre les instances

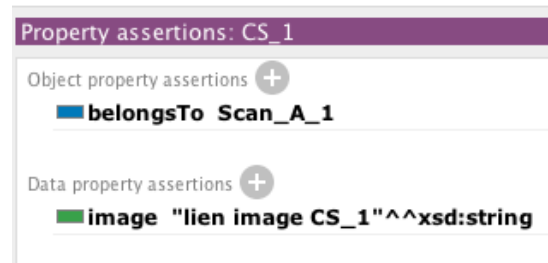


FIGURE 7 – Instance d'une coupe

### 3.2.3 Perspectives d'évolution

La situation spatiale des structures anatomiques de l'abdomen peut évoluer en cas de pathologie. Notre ontologie ne permet pas de traiter ces éventuelles évolutions et c'est donc la principale amélioration qui pourrait être apportée.

D'autre part, nous constatons qu'instancier un patient, ainsi que les scans qui lui appartiennent avec pour chaque scan toutes les coupes qui y sont rattachées

est plutôt fastidieux dans Protégé : il faut veiller, à chaque fois, à établir les liens entre les instances pour préciser que tel scan appartient à tel patient et que tel patient possède tel scan. La création d'un logiciel qui utiliserait l'ontologie que nous avons créée, mais proposerait une interface d'utilisation plus confortable, en automatisant certaines tâches, serait une amélioration significative.

## 4 Conclusion

Pour segmenter automatiquement les images médicales, et faciliter ainsi le diagnostic de néphroblastome, on peut recourir au raisonnement à partir de cas. Mais pour que cette technique d'intelligence artificielle soit optimale pour la résolution de notre problème, il faut inclure des connaissances théoriques. C'est là que les ontologies interviennent : elles constituent le format idéal pour exprimer les connaissances que l'on possède dans un domaine de façon structurée.

Notre recherche d'ontologies s'est appuyée sur plusieurs critères définis préalablement, dont la nécessité de pouvoir décrire spatialement des éléments anatomiques. Si certaines des ontologies que nous avons trouvées proposent des descriptions anatomiques, aucune d'entre elles ne propose une description des structures anatomiques de l'abdomen les uns par rapport aux autres. Nous avons donc créé notre propre ontologie dans ce but, en nous appuyant sur l'expérience acquise lors de l'étude de nombreuses ontologies.

Les résultats obtenus satisfont notre objectif principal : situer les éléments de l'abdomen les uns par rapport aux autres. Nous avons planifié une séance avec le docteur Yann Chaussy pour l'aider à prendre en main l'ontologie créée sous le logiciel Protégé.

La principale amélioration qu'il faudrait apporter à notre ontologie est la prise en compte du cas pathologique où la situation spatiale des éléments peut évoluer. De plus, intégrer cette ontologie à un logiciel permettant un meilleur confort d'utilisation, en automatisant certaines tâches, serait intéressant.

## Références

- [1] Breast cancer grading ontology (bcgo), <https://bioportal.bioontology.org/ontologies/BCGO>.
- [2] Cell behavior ontology (cbo), <http://bioportal.bioontology.org/ontologies/CBO>.
- [3] Cell ontology (cl), <http://bioportal.bioontology.org/ontologies/CL>.
- [4] Children's health exposure analysis resource (chea), <https://bioportal.bioontology.org/ontologies/CHEAR>.
- [5] Human disease ontology (hdo), <https://bioportal.bioontology.org/ontologies/D0ID>.
- [6] National cancer institute thesaurus ontology, <http://bioportal.bioontology.org/ontologies/NCIT>.
- [7] National cancer institute thesaurus, <https://ncit.nci.nih.gov/ncitbrowser/>.
- [8] National cancer institute, <https://www.cancer.gov>.
- [9] National institute of environmental health sciences (niehs), <https://www.niehs.nih.gov>.
- [10] Protégé, <https://protege.stanford.edu>.
- [11] Herbulot Ariane. Mesures statistiques non-paramétriques pour la segmentation d'images et de vidéos et minimisation par contours actifs. *Thèse de fin d'étude, Université de Nice*, 2007.
- [12] A E. Tutac, Vladimir Cretu, and Daniel Racocanu. Spatial representation and reasoning in breast cancer grading ontology. 06 2010.
- [13] Femto-St. Projet INTERREG SAIAD : améliorer, par méthode informatique, l'analyse de l'imagerie médicale dans le traitement des tumeurs rénales chez l'enfant, Janvier 2016.
- [14] NCBO. Bioportal NCBO : National center for biomedical ontology, <https://www.bioontology.org/>.
- [15] Natalya F. Noy and Deborah L. McGuinness. Ontology development 101 : A guide to creating your first ontology, [https://protege.stanford.edu/publications/ontology\\_development/ontology101-noy-mcguinness.html](https://protege.stanford.edu/publications/ontology_development/ontology101-noy-mcguinness.html).
- [16] Boissard Olivier and Lavier Antoine. Segmentation d'images médicales, Janvier 2016.
- [17] Technolangu. Qu'est-ce qu'une ontologie?, [http://www.technolangu.net/imprimer.php3?id\\_article=280](http://www.technolangu.net/imprimer.php3?id_article=280), Juillet 2006.
- [18] Delavelle Thibault. Démonstrateur basé sur le raisonnement à partir de cas pour l'aide au traitement d'images médicales du cancer du rein chez l'enfant. *Rapport de stage de recherche et développement*, Août 2017.

- [19] Gruber Thomas. A translation approach to portable ontology specifications, <http://tomgruber.org/writing/ontolingua-kaj-1993.htm>, Avril 1993.
- [20] Wikipédia. Janet l. kolodner, [https://en.wikipedia.org/wiki/Janet\\_L.\\_Kolodner](https://en.wikipedia.org/wiki/Janet_L._Kolodner).

### Résumé

Le projet SAIAD (Segmentation Automatique de reins tumoraux chez l'enfant par Intelligence Artificielle Distribuée) a pour objectif d'utiliser des techniques d'intelligence artificielle (IA) pour segmenter automatiquement des scanners de patients et ainsi faciliter le diagnostic de néphroblastome (cancer du rein chez l'enfant). Le raisonnement à partir de cas (RàPC) est une des techniques d'intelligence artificielle utilisée. Pour optimiser cette technique, il faut inclure des connaissances théoriques et les ontologies constituent le format idéal pour exprimer les connaissances que l'on possède dans un domaine de façon structurée. Nous avons recherché des ontologies médicales du corps humain, puis créer notre propre ontologie de l'abdomen.

**Mots-clés** : ontologie, cancer, rein, enfants, intelligence artificielle, segmentation

### Résumé

The SAIAD project (automatic segmentation of children's kidney tumors by distributed artificial intelligence) aims to use artificial intelligence (AI) to automatically segment kidney images, and thus facilitate the diagnosis of nephroblastoma (cancer of the kidneys in children). Case-based reasoning (CBR) is one the IA concepts used. In order to optimize CBR, we have to include theoretical knowledge and ontologies are the perfect tools to express the knowledge that we possess in a field in a structured way. We searched for medical ontologies of the human body and then created our own abdominal ontology.

**Keywords** : ontology, cancer, kidney, children, artificial intelligence, segmentation