

Thèse

Présentée devant

L'INSTITUT NATIONAL DES SCIENCES APPLIQUÉES DE LYON

Pour obtenir

LE GRADE DE DOCTEUR

*École doctorale : Électronique, Électrotechnique, Automatique
Formation doctorale : Images et Systèmes*

Par

SIMON MARACHE-FRANCISCO
(Ingénieur INSA Rouen)

Évaluation de la correction du mouvement respiratoire sur la détection des lésions en oncologie TEP

Soutenue le XX XX XXXX devant la Commission d'examen

Jury :

Rapporteur	Michel Desvignes	Professeur, GIPSA-Lab, INP Grenoble
Rapporteur	Su Ruan	Professeur, LITIS, Université de Rouen
Examineur	Frédéric Lamare	Docteur, Ingénieur de Recherche, INCIA
Examineur	Christophe Odet	Professeur, CREATIS, INSA-lyon
Examineur	Jean-Michel Rouet	Docteur, Ingénieur Philips Healthcare Medisys
Co-encadrante	Carole Lartizien	Chargée de Recherche CNRS, CREATIS, Lyon
Directeur de thèse	Rémy Prost	Professeur, CREATIS, INSA-Lyon

Résumé

La tomographie par émission de positons (TEP) est une méthode d'imagerie clinique en forte expansion dans le domaine de l'oncologie. De nombreuses études cliniques montrent que la TEP permet, d'une part de diagnostiquer et caractériser les lésions cancéreuses à des stades plus précoces que l'imagerie anatomique conventionnelle, et d'autre part d'évaluer plus rapidement la réponse au traitement. Le raccourcissement du cycle comprenant le diagnostic, la thérapie, le suivi et la réorientation thérapeutique contribue à augmenter le pronostic vital du patient et maîtriser les coûts de santé.

La durée d'un examen TEP, de l'ordre de 5 à 10 minutes pour un champ de vue axial de l'ordre de 15 cm, ne permet pas de réaliser une acquisition sous apnée. La qualité des images TEP est par conséquent affectée par les mouvements respiratoires du patient qui induisent un flou dans les images. Les effets du mouvement respiratoire sont particulièrement marqués au niveau du thorax et de l'abdomen.

Plusieurs types de méthodes ont été proposés pour corriger les données de ce phénomène, mais elles demeurent lourdes à mettre en place en routine clinique. La problématique de la correction du mouvement respiratoire et le choix de la méthode appropriée sont des sujets d'actualité au sein de la communauté de médecine nucléaire. Des travaux récemment publiés proposent une évaluation de ces méthodes basée sur des critères de qualité tels que le rapport signal sur bruit ou le biais. Aucune étude à ce jour n'a évalué l'impact de ces corrections sur la qualité du diagnostic clinique. Ce problème pose des questions d'orientation stratégique et financière importantes.

Nous nous sommes focalisés sur la problématique de la détection des lésions du thorax et de l'abdomen de petit diamètre et faible contraste, qui sont les plus susceptibles de bénéficier de la correction du mouvement respiratoire en routine clinique.

Nos travaux ont consisté dans un premier temps à construire une base d'images TEP qui modélisent un mouvement respiratoire non-uniforme, une variabilité inter-individuelle et contiennent un échantillonnage de lésions de taille et de contraste variable. Ce cahier des charges nous a orientés vers les méthodes de simulation Monte Carlo qui permettent de contrôler l'ensemble des paramètres influençant la formation et la qualité de l'image. Une base de 15 modèles de patient a été créée en adaptant le modèle anthropomorphique XCAT sur des images tomodynamométriques (TDM) de patients. Cette base contient 280 lésions sphériques de 8 et 12 mm de diamètre réparties dans les poumons et le foie et dont le contraste a été calibré pour échantillonner la gamme de détection. Le signal respiratoire simulé est constitué d'un motif de 4 cycles mesurés sur des patients. Nous avons ensuite développé un protocole de simulation à l'aide du logiciel PET-SORTEO afin de générer les examens virtuels de chacun des modèles en considérant deux types d'acquisition : une acquisition dynamique en mode-liste synchronisée sur le signal respiratoire et une acquisition statique

de même durée que l'acquisition dynamique mais correspondant au cas idéal d'un patient qui pourrait retenir sa respiration.

Nous avons en parallèle développé une stratégie originale d'évaluation des performances de détection. Cette méthode comprend un système de détection des lésions automatisé basé sur l'utilisation de machines à vecteurs de support (SVM) utilisant des caractéristiques fréquentielles. Les performances sont mesurées par l'analyse psychophysique des courbes free-receiver operating characteristics (FROC) que nous avons adaptée aux spécificités de l'imagerie TEP. L'évaluation des performances est réalisée sur deux méthodes prometteuses de correction du mouvement respiratoire, en les comparant avec les performances obtenues sur les images non corrigées ainsi que sur les images idéales sans mouvement respiratoire. Les résultats obtenus sont prometteurs et montrent une réelle amélioration de la détection des lésions après correction, qui approche les performances obtenues sur les images statiques.

MOTS-CLEFS : TEP, oncologie, mouvement respiratoire, reconnaissance de formes, Computer aided detection (CAD), Séparateurs à vaste marge (SVM), Simulation Monte-Carlo

Table des matières

Résumé	3
Table des matières	5
Liste des figures	11
Liste des tables	14
Introduction	16
I Imagerie TEP	19
1 Principe Physique	23
1.1 Généralités	23
1.2 Traceur ^{18}FDG	23
1.3 Émission des photons	25
1.4 Détection des photons	25
1.5 Types de coïncidences	26
1.6 Perturbation du trajet du photon	28
1.7 Quantification en imagerie TEP	28
2 Déroulement d'une acquisition	31
2.1 TEP au ^{18}FDG	31
2.2 Acquisitions corps entier	31
2.3 Acquisition en mode 3D	32

TABLE DES MATIÈRES

2.4	Correction de l'atténuation	34
2.5	Correction des autres sources de bruit	35
2.6	Format des données	36
2.6.1	Séquentielles	36
2.6.2	Sinogramme	37
3	Algorithmes de reconstruction	39
3.1	Algorithmes itératifs	39
3.1.1	Algorithme ML-EM (Maximum Likelihood Expectation Maximisation)	39
3.1.2	Algorithme OS-EM (Ordered Subset Expectation Maximisation)	40
3.1.3	Cas des acquisitions en mode séquentiel	41
3.2	Reconstructions analytiques	42
II	Mouvement respiratoire	45
4	Respiration et influence sur les acquisitions TEP/TDM	49
4.1	Introduction	49
4.2	Mouvement respiratoire	49
4.3	Localisation et volume	50
4.4	Mesure de l'activité des tumeurs	51
4.4.1	Décalage de la carte d'atténuation	51
4.4.2	Artefacts TEP dus aux artefacts TDM	52
4.4.3	Déplacement de la tumeur au cours du cycle	53
4.5	Impact du mouvement respiratoire sur la détection	53
4.5.1	Article de Thielemans et al. [Thielemans et al., 2006]	54
4.5.2	Article de Chang et al. [Chang et al., 2010]	54
5	Processus d'estimation du mouvement	57
5.1	Introduction	57
5.2	Estimation du signal respiratoire	57
5.2.1	Spiromètre	57
5.2.2	Ceinture	58
5.2.3	Techniques basés sur des caméras vidéo	58

TABLE DES MATIÈRES

5.2.4	Techniques basées sur les images TEP	59
5.3	Estimation du champ de mouvement	60
5.3.1	Image TEP 4D	60
5.3.2	Image TDM 4D	61
5.3.3	Modèle	63
6	Correction du mouvement respiratoire	65
6.1	Introduction	65
6.2	Synchronisation respiratoire	65
6.3	Synchronisation respiratoire avec recalage	68
6.4	Correction pré-reconstruction	69
6.5	Correction pendant la reconstruction	71
6.6	Déconvolution de l'image	72
III	Evaluation des performances de détection	77
7	Performance des outils de détection	81
7.1	Généralités	81
7.1.1	Définition d'un Vrai positif / Faux positif	82
7.2	Méthodologie ROC - Receiver-Operating Curve	83
7.2.1	Courbes "Localization ROC" (L-ROC)	85
7.3	Courbes Free ROC	86
7.3.1	Courbes Free-ROC	86
7.3.2	Courbes Alternative Free-ROC	87
7.3.3	Comparaison des courbes	87
8	Systèmes de détection	89
8.1	Principe des CAD	89
8.1.1	Introduction	89
8.1.2	Les différentes étapes d'un système CAD supervisé	89
8.2	Les CAD en TEP	90
8.3	Extraction initiale des lésions	91
8.3.1	Pré-traitements de l'image	91
8.3.2	Extraction des caractéristiques	91

TABLE DES MATIÈRES

8.4	Méthodes de classification	92
8.4.1	Méthodes non supervisées	92
8.4.2	Méthodes supervisées	93
8.5	Machines à vecteur de support (SVM)	93
8.5.1	Autres classifieurs supervisés	94
9	Choix méthodologiques	101
9.1	Construction de la vérité terrain	101
9.2	Mesure des performances de détection	102
IV	Construction d'une base d'images TEP simulées respirante	105
10	Simulations	107
10.1	Principe des simulations	107
10.1.1	Simulations analytiques	107
10.1.2	Simulation Monte-Carlo	108
10.1.3	Simulation Monte-Carlo accélérées	108
10.2	État de l'art des simulateurs TEP	108
10.3	Processus de simulation avec SORTEO	109
10.4	Contribution à PET-SORTEO	110
10.4.1	Adaptation pour l'exécution sur centre de calcul	110
10.4.2	Sortie en mode Séquence	111
11	Base de données : OncoPET Respi	113
11.1	Présentation	113
11.1.1	Avantages des bases de données simulées	113
11.2	Modèles	114
11.2.1	Modèle anatomique	114
11.2.2	Activités des organes	115
11.2.3	Modèle de Respiration	115
11.2.4	Modèle de tumeur	117
11.3	Simulation	119
11.3.1	Exemples de données reconstruites	120
11.4	Données clef	120

TABLE DES MATIÈRES

11.4.1	Modèles	120
11.4.2	Lésions	121
V Évaluation de la correction du mouvement respiratoire sur la détection des lésions pulmonaires et hépatiques		123
12	Méthodes	125
12.1	Reconstruction des images	126
12.1.1	Principe de la reconstruction	126
12.1.2	Paramètres de reconstruction	126
12.2	Estimation du mouvement	129
12.2.1	Principe de l'estimation de mouvement	129
12.2.2	Paramètres de l'estimation de mouvement	129
12.3	Description des différents types d'images à évaluer	133
12.3.1	Images Statiques	133
12.3.2	Images Non corrigées	133
12.3.3	Images TE-MS	135
12.3.4	Images TE-IM	135
12.4	Système CAD	135
12.4.1	Vecteur de caractéristiques	136
12.4.2	Génération de la base d'apprentissage	138
12.4.3	Apprentissage de la Machine à Vecteur de Support (SVM)	138
12.4.4	Génération des sites présumés	139
12.4.5	Règles d'évaluation du résultat	139
12.4.6	Résumé	140
12.5	Optimisation des paramètres du système CAD	140
12.5.1	Paramètres à optimiser	142
12.5.2	Méthodes d'optimisation	143
12.5.3	Mesure des performances de détection	145
12.6	Comparaison des performances de détection des différentes méthodes de correction du mouvement respiratoire	145
13	Analyse des résultats	147
13.1	Optimisation des paramètres	147

TABLE DES MATIÈRES

13.1.1	Sélection des meilleurs paramètres du classifieur	148
13.1.2	Courbe Free-ROC	152
13.1.3	Comparaison des performances JAFROC	153
13.1.4	Conclusion	153
13.2	Comparaison des performances de la détection des tumeurs pulmonaires	154
13.2.1	Sélection des meilleurs paramètres du classifieur	155
13.2.2	Courbes Free-ROC	157
13.2.3	Comparaison des performances JAFROC	157
13.2.4	Conclusion	158
13.3	Comparaison des performances des différentes méthodes pour la détection des tumeurs hépatiques	159
13.3.1	Sélection des meilleurs paramètres du classifieur	159
13.3.2	Courbes Free-ROC	159
13.3.3	Comparaison des performances JAFROC	162
13.3.4	Conclusion	163
VI	Discussion	165
14	Discussion	167
14.1	Synthèse des résultats	167
14.2	Limites et perspectives	168
14.2.1	Système de détection	168
14.2.2	Base de données	169
14.3	Imagerie TEP/TDM	170
	Conclusion	172
	Bibliographie	172
	Bibliographie personnelle	183

Table des figures

1.1	Présentation simplifiée de la TEP	24
1.2	Émission des photons	25
1.3	Coupe d'un module de Philips Allegro	26
1.4	Photo d'un imageur Philips Gemini ouvert	27
1.5	Les différents types de coïncidences en TEP	27
1.6	Effet de l'atténuation sur les images TEP	29
2.1	Exemples d'images TEP	32
2.2	Photographie d'un scanner TEP Gemini	33
2.3	Acquisitions corps entier en TEP	33
2.4	Acquisitions 3D en TEP	34
2.5	Principe du sinogramme	37
3.1	Schéma de principe de l'algorithme MLEM	41
4.1	Modèle de mouvement intégré dans le fantôme NCAT	50
4.2	Effet du déplacement d'une tumeur sur les données acquises	51
4.3	Influence de l'influence de la carte d'atténuation sur l'activité des lésions	52
4.4	Artefacts présents sur des images TDM utilisées pour la correction d'atténuation.	53
4.5	Exemple d'image avec et sans correction de mouvement respiratoire. . .	55
5.1	Illustration du spiromètre Syn'r	58
5.2	Photographie du système RGS de Varian medical Systems en action . . .	59
5.3	Exemple de champ de mouvement	62
6.1	Illustration de la synchronisation respiratoire en phase	66

TABLE DES FIGURES

6.2	Illustration de la synchronisation respiratoire en amplitude : Le cycle respiratoire acquis est découpé selon son amplitude.	66
6.3	Illustration de l'étendue du mouvement respiratoire sur des images reconstruites	67
6.4	Images TEP/TDM superposées du poumon	68
6.5	Illustration du recalage des lignes de réponse dans l'espace du détecteur	70
6.6	Résultats de l'algorithme de recalage des LDR sur des images de patients	70
6.7	comparaison des performances des différentes techniques de correction du mouvement sur un profil d'image TEP	73
6.8	Index de CHO pour différentes techniques de correction du mouvement respiratoire en fonction du nombre d'itérations de la reconstruction	73
6.9	Exemple de filtres TLP pour la correction par déconvolution	74
6.10	Comparaison de l'erreur de sous-estimation de l'activité des lésions	75
7.1	Exemple de courbe ROC	84
7.2	Modèle de la distribution de probabilité de la variable de décision	85
7.3	Exemple de courbe Free ROC	86
8.1	Les différentes étapes d'un système CAD.	96
8.2	Fonctionnement d'un classifieur non supervisé	97
8.3	Fonctionnement d'un classifieur supervisé	98
8.4	Machine à Vecteur de Support	98
8.5	Illustration du principe d'optimisation de la marge sur le SVM	99
8.6	Changement de base pour les SVM	99
11.1	Adaptation du modèle XCAT sur une image TDM de patient	114
11.2	Exemple d'adaptation d'images du modèle XCAT sur des images TDM .	115
11.3	Exemple de courbes respiratoires régulières et irrégulières	116
11.4	Signal respiratoire utilisé pour les modèles de la base de données	116
11.5	Détectabilité des tumeurs du poumon en fonction du contraste et de la taille des tumeurs	118
11.6	Détectabilité des tumeurs du foie en fonction du contraste et de la taille des tumeurs	119
11.7	Images reconstruites tirées de la base de donnée	121
11.8	Images respirantes reconstruites tirées de la base de donnée	122

TABLE DES FIGURES

12.1	Illustration de l'évolution des lésions en fonction du nombre d'itérations	127
12.2	Évaluation du rapport contraste sur bruit des lésions du foie en fonction du nombre d'itérations	128
12.3	Évaluation du rapport "contraste sur bruit" des lésions du poumon en fonction du nombre d'itérations	128
12.4	Exemple d'interpolation par B-spline	130
12.5	Illustration du recalage obtenu	131
12.6	Performances de l'estimation de mouvement en fonction de la taille de la grille de recherche	132
12.7	Performances de l'estimation de mouvement en fonction de la régularisation	133
12.8	Exemples d'images reconstruites corrigées et non corrigées du mouvement	134
12.9	Décomposition en ondelettes par banc de filtres	136
12.10	Exemples de décomposition d'images en ondelettes	137
12.11	Fonction d'ondelette et d'échelle biorthogonale 4/4	138
12.12	Schéma du système CAD	141
12.13	Réalisation d'une validation croisée à n éléments	144
13.1	(1/2) Recherche des meilleurs paramètres du classifieurs : Fronts de Pareto	149
13.2	(2/2) Recherche des meilleurs paramètres du classifieurs : Fronts de Pareto	150
13.3	Courbes Free-ROC comparant les performances du CAD selon les différents paramètres de base d'apprentissage	153
13.4	Figures de Mérites JAFROC obtenues pour les différents paramètres de génération de la base d'apprentissage	154
13.5	Fronts de Pareto des résultats de la recherche des meilleurs paramètres du classifieur pour les différents jeux d'images pulmonaires	156
13.6	Courbes Free-ROC comparant les performances du CAD sur les lésions pulmonaires selon la technique de correction du mouvement respiratoire.	158
13.7	Les FDM (Figure de Mérite) obtenues pour les différentes techniques de correction du mouvement respiratoire sur la détectabilité des lésions du poumon.	158
13.8	(1/2) Fronts de Pareto des résultats de la recherche des meilleurs paramètres du classifieur pour les lésions du foie	160
13.9	(2/2) Fronts de Pareto des résultats de la recherche des meilleurs paramètres du classifieur pour les lésions du foie	161

TABLE DES FIGURES

13.10	Courbes Free-ROC comparant les performances du CAD sur les lésions hépatiques selon la technique de correction du mouvement respiratoire.	162
13.11	Les FDM (Figure de Mérite) JAFROC obtenues pour les différents types d'images pour la détectabilité des lésions du foie.	162

Liste des tableaux

7.1	Matrice de confusion	81
11.1	Activités des organes des patients de la base de données	116
11.2	Contraste de la base originale OncoPET_DB des lésions du poumon et du foie pour l'étude de détectabilité	118
11.3	Détectabilité estimée des lésions en fonction du contraste et de leur diamètre	119
11.4	Contraste final des lésions du foie et du poumon	120
11.5	Tableau récapitulatif des lésions	121
13.1	Paramètres (C, γ, j) sélectionnés pour l'optimisation des performances sur les différentes bases	152
13.2	Paramètres sélectionnés pour l'optimisation des performances de détection des tumeurs pulmonaires	155
13.3	Paramètres sélectionnés pour l'optimisation des performances du CAD sur les tumeurs hépatiques	160

Introduction

La tomographie par émission de positons (TEP) couplée à un tomodensitomètre (TDM) est une méthode d'imagerie clinique en forte expansion dans le domaine de l'oncologie. De nombreuses études cliniques montrent que la TEP permet, d'une part de diagnostiquer et caractériser les lésions cancéreuses à des stades plus précoces que l'imagerie anatomique conventionnelle, et d'autre part d'évaluer plus rapidement la réponse au traitement. Le raccourcissement du cycle comprenant le diagnostic, la thérapie, le suivi et la réorientation thérapeutique contribue à augmenter le pronostic vital du patient et maîtriser les coûts de santé.

L'examen TEP/TDM au FDG apparaît désormais indispensable pour évaluer la réponse thérapeutique des patients atteints de lymphome. Les résultats préliminaires sont également très encourageants dans les tumeurs solides, en particulier les cancers pulmonaires, digestifs et de la sphère ORL [Cachin et al., 2006]. La TEP pourrait prédire très précocement la réponse et ainsi permettre une réadaptation du schéma thérapeutique. Enfin de nouveaux traceurs TEP ouvrent la voie de l'imagerie moléculaire qui semble extrêmement prometteuse en thérapie génique et dans le suivi du traitement ciblé du cancer.

La durée d'un examen TEP, de l'ordre de 5 à 10 minutes pour un champ de vue axial de l'ordre de 15 cm, ne permet pas de réaliser une acquisition sous apnée. La qualité des images TEP est par conséquent affectée par les mouvements respiratoires du patient qui induisent un flou dans les images. Les effets du mouvement respiratoire sont particulièrement marqués au niveau du thorax et de l'abdomen. Deux types de méthodes ont été proposés pour corriger les données de ce phénomène. Le premier type est basé sur l'utilisation d'acquisitions synchronisées sur la respiration produisant plusieurs images non affectées par la respiration et correspondant à différents instants du cycle respiratoire [Nehmeh et al., 2002][Boucher et al., 2004]. Cette méthodologie se base sur l'hypothèse d'une grande régularité du signal respiratoire de synchronisation [Boucher et al., 2004], et conduit à des images d'une qualité statistique limitée car chacune ne contient qu'une fraction réduite du nombre de photons détectés pendant toute l'acquisition [Visvikis et al., 2004]. Le deuxième type de méthode se base sur l'utilisation d'un scanner 4D TEP/TDM du patient permettant de modéliser les mouvements respiratoires. Ce modèle est alors intégré lors du processus de reconstruction tomographique [Lamare et al., 2007b]. Ce type de méthode permet de pouvoir conserver la statistique de l'image mais nécessite une information anatomique dynamique assez lourde à acquérir en routine clinique.

La problématique de la correction du mouvement respiratoire et le choix de la méthode appropriée sont des sujets d'actualité au sein de la communauté de médecine nucléaire. Des travaux récemment publiés proposent une évaluation de ces méthodes basée sur des critères de qualité tels que le rapport signal sur bruit ou le biais [Visvikis et al., 2004]. Aucune étude à ce jour n'a évalué l'impact de ces corrections sur la qua-

lité du diagnostic clinique. Ce problème pose des questions d'orientation stratégique et financière importantes, puisque le second type de méthode requiert l'acquisition de données TEP/TDM dynamiques, très peu accessibles à l'heure actuelle.

C'est dans l'optique de résoudre ce problème que j'ai commencé cette thèse sur **l'estimation des apports de la correction du mouvement respiratoire en oncologie TEP**.

La première partie de ce manuscrit présente l'imagerie TEP. Nous détaillons les principes physiques de la TEP, le déroulement des acquisitions en nous focalisant sur l'imagerie TEP au ^{18}F FDG pour l'oncologie, et enfin les algorithmes de reconstruction utilisés pour former les images.

Ensuite, nous présentons un état de l'art sur le mouvement respiratoire, ainsi que son influence sur la visualisation des lésions en imagerie TEP. Nous parlons ensuite des techniques permettant de l'estimer, puis de le corriger dans les images reconstruites.

La troisième partie dresse un état de l'art des méthodes d'évaluation des performances de détection. Nous présentons tout d'abord les bases de l'analyse psychophysique basée sur les courbes ROC (Receiver-Operating Characteristics), puis nous présentons le principe des systèmes de détection automatique (CAD en anglais, pour Computer-Aided Detection) que nous avons choisi d'utiliser dans cette étude, comme observateurs numériques en remplacement de l'observateur humain. Le dernier chapitre de cette partie présente et justifie les choix méthodologiques que nous avons faits pour répondre à la problématique clinique.

Enfin, nous présentons nos contributions : en premier lieu, la base de données, que nous avons créée pour réaliser les mesures de performances, puis nous décrivons la méthode utilisée pour répondre à la problématique, et enfin les résultats que nous avons obtenus.

La dernière partie correspond aux discussions et perspectives sur notre travail.

Première partie

Imagerie TEP

La Tomographie par Émission de Positons (TEP) au fluoro-deoxy-glucose marqué au Fluor 18 (^{18}F FDG) est une modalité d'imagerie fonctionnelle utilisant la désintégration d'un traceur radioactif pour mettre en valeur les zones de forte activité métabolique. Elle est principalement utilisée en imagerie cérébrale, oncologique et cardiologique.

Dans ces trois chapitres, nous allons tout d'abord détailler rapidement les principes physiques amenant à la création des images TEP, de la désintégration du traceur radioactif à la détection des photons émis lorsqu'ils atteignent un capteur. Ensuite, nous parlerons des solutions techniques mises en place pour réaliser les images TEP, avec les différents types d'acquisitions. Enfin, nous détaillerons les algorithmes utilisés pour reconstruire les images TEP.

Chapitre 1

Principe Physique

1.1 Généralités

L'imagerie TEP est basée sur l'injection d'une molécule traceuse marquée par un atome radioactif émetteur de positon. Ce paragraphe présente rapidement les événements qui se produisent lors d'un examen TEP, présentés dans la figure 1.1.

Ce traceur est conçu de manière à se fixer sur les zones du corps que l'on souhaite imager. Pendant toute la durée de l'examen, les particules radioactives vont se désintégrer selon la loi de décroissance radioactive de l'équation 1.1.

$$dN = -\lambda N dt \quad (1.1)$$

N représente le nombre de particules radioactives présentes dans le corps du patient. dN représente la variation de ce nombre de particules (le nombre de désintégrations par dt) et λ est une constante dépendant de l'élément radioactif.

Chaque désintégration d'un élément radioactif va déclencher l'émission d'une particule chargée β^+ , aussi appelée positon. En oncologie, on utilise par exemple le Fluor ^{18}F qui se désintègre en Oxygène ^{18}O en émettant un positon. Cette particule va parcourir quelques mm avant de s'annihiler avec un électron en émettant 2 photons dans deux directions opposées avec une énergie de 511 KeV.

Ce sont ces photons qui sont détectés par l'imageur TEP. Ils représentent une coïncidence, car l'instant d'arrivée et leur énergie sont approximativement semblables. L'ensemble des "lignes de réponse" (LDR) correspondant aux coïncidences détectées est utilisé pour reconstruire les images.

1.2 Traceur ^{18}F FDG

Le glucose est considéré comme un carburant essentiel à l'organisme qui peut l'assimiler directement. Les cellules cancéreuses, qui ont un métabolisme accéléré par rap-

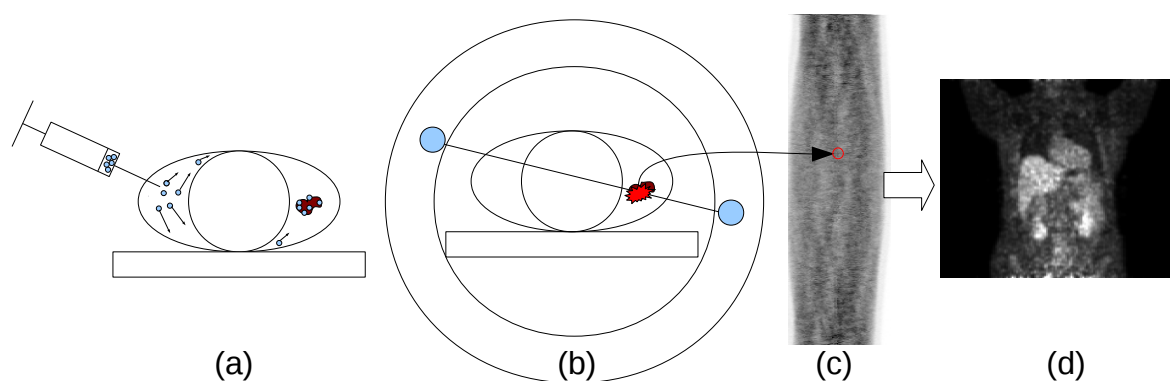


FIG. 1.1: Processus d'un examen TEP : a) Injection du traceur radioactif, qui va se fixer préférentiellement sur les zones que l'on cherche à observer. b) Désintégration d'un atome radioactif du traceur, ici dans la zone à imager. Cela entraîne l'émission de deux photons dans deux directions opposées, qui sont détectés dans la couronne de capteurs. c) Tous les événements détectés sont stockés dans la mémoire de la console, ici sous forme de sinogramme. d) L'image est reconstruite à la suite de l'acquisition pour former un volume 3D estimant la répartition du traceur dans l'organisme.

port aux autres cellules du milieu, sont particulièrement friandes de glucose. Celui-ci est transporté à l'intérieur des cellules par des transporteurs spécifiques de la membrane afin d'être dégradé par un ensemble de réactions appelé glycolyse. Le glucose est d'abord phosphorylé en glucose-6-phosphate, catalysé par l'enzyme hexokinase. Le glucose-6-phosphate est ensuite transformé en fructose-6-phosphate qui subit à son tour plusieurs transformations successives pour être finalement excrété par la cellule sous forme de lactate.

Le fluorodeoxyglucose ou FDG est un analogue du glucose avec un aspect structural très proche. Il est donc aussi capté par les cellules, mais ne peut pas être utilisé dans leur métabolisme. Le FDG suit le même processus métabolique que le glucose jusqu'à l'étape de phosphorylation en FDG-6-phosphate. Il n'est cependant pas métabolisé par l'enzyme glucose-6-phosphate et s'accumule dans la cellule sous la forme de FDG-6-phosphate. Cette accumulation permet de distinguer la présence de tumeurs, généralement beaucoup plus gourmandes en énergie que les tissus sains et donc marquées par une présence élevée de FDG-6-phosphate. Par son métabolisme et son importance pour les cellules cancéreuses, le FDG est actuellement préconisé en clinique pour les examens concernant des patients en cancérologie. Il est généralement marqué au fluor 18 donnant le ^{18}FDG .

Cependant, ce radiotracer n'est pas spécifique aux tumeurs mais se fixe sur toute cellule consommant du glucose. C'est pourquoi le coeur ainsi que le cerveau fixent toujours une grande quantité de ^{18}FDG , comme l'indique la figure 1.1.d).

1.3. ÉMISSION DES PHOTONS

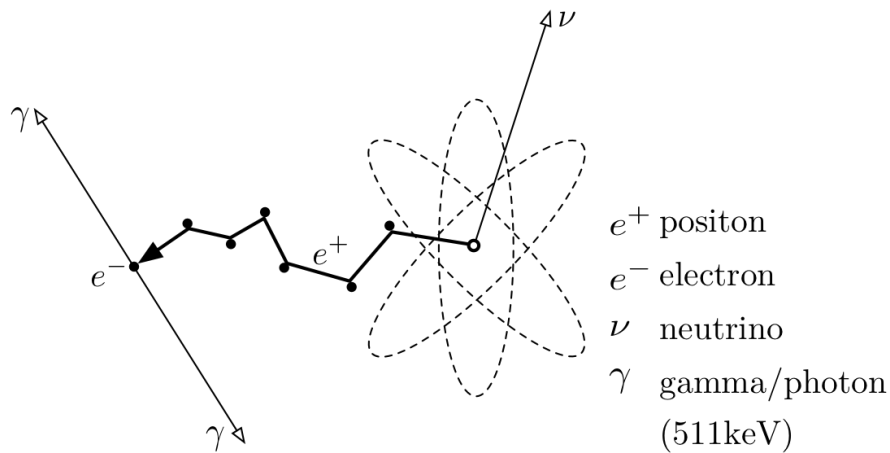


FIG. 1.2: Émission des photons [Langner, 2008] : Le radio traceur se désintègre en émettant un neutrino et un positon. Après un parcours de quelques mm dans les tissus, ce dernier s'annihile avec un électron. Cette réaction provoque la création de deux photons γ d'énergie 511 keV partant dans deux directions opposées.

1.3 Émission des photons

Nous allons maintenant détailler le processus qui déclenche l'émission des photons détectés par l'imager, présenté dans la figure 1.2.

Les émetteurs de positons utilisés en TEP sont des isotopes radioactifs présentant un excès de charge positive, ou proton, dans leurs noyaux. Un processus de désintégration β^+ , correspondant à la transformation d'un proton p en un neutron n , leur permet de migrer vers un état stable. Cette désintégration résulte en l'émission d'un neutrino ν et d'un positon e^+ selon l'équation 1.2. Le positon est une particule de même masse que l'électron mais de charge opposée.



Le positon va alors s'annihiler avec un électron après un parcours de quelques millimètres en émettant simultanément deux photons γ de même énergie (511 keV), comme indiqué dans l'équation 1.3. L'angle d'émission des photons γ est de 180° avec une incertitude de $\pm 0.25^\circ$ [Bailey, 2005].



1.4 Détection des photons

Les détecteurs les plus couramment utilisés en TEP sont constitués d'un couplage entre un cristal scintillant et un tube photomultiplicateur.

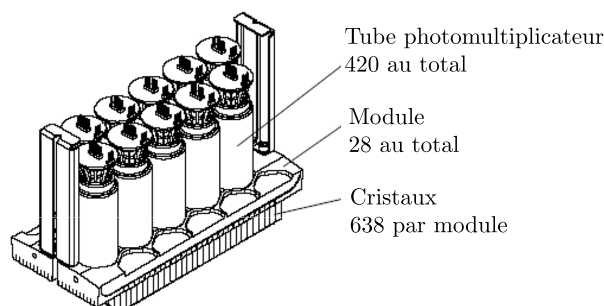


FIG. 1.3: Coupe d'un module de Philips Allegro. Les 28 modules sont placés en anneaux autour du patient.

Le cristal scintillant permet de convertir les photons γ de 511 keV en photons lumineux par les différents phénomènes d'interaction particules-matières.

Chaque photon absorbé par le matériau scintillant va entraîner une réaction en chaîne qui va déclencher une émission lumineuse. Le tube multiplicateur situé derrière va convertir cette émission lumineuse en charge électrique. Un système électronique va ensuite apparier les photons détectés pour retrouver la ligne de réponse (LDR) sur laquelle la désintégration a eu lieu. Les paires de photons assemblées correspondent à des coïncidences.

Les détecteurs sont regroupés par modules, placés en anneaux autour du patient. L'imageur TEP Gemini que nous utiliserons dans cette étude dispose par exemple de 28 modules, contenant chacun une matrice de 29 par 22 cristaux de GSO (voir figure 1.3). Une photo de l'imageur ouvert est visible sur la figure 1.4. 420 photomultiplicateurs placés autour des blocs sont utilisés pour estimer la position des détecteurs.

1.5 Types de coïncidences

La figure 1.5 représente les différents types de coïncidences rencontrés en TEP.

Les coïncidences vraies (1.5.a) correspondent aux cas où les deux détections appariées correspondent bien à une seule désintégration et où aucun des photons n'a été dévié. Les coïncidences fortuites (1.5.b) correspondent à la détection en coïncidence de deux photons issus de deux annihilations différentes, par exemple à cause de l'absorption d'un des photons d'une désintégration par les tissus. Enfin, les coïncidences diffusées (1.5.c) sont le résultat de la déviation d'un des deux photons produit engendré par des interactions rayonnement-matière dans les tissus (diffusion Compton).

Les coïncidences diffusées ainsi que les coïncidences aléatoires génèrent un bruit parasite dans les sinogrammes. L'absorption des photons sans détection va engendrer un effet d'atténuation du signal qui sera de plus en plus important en fonction de la profondeur des tissus et de leur absorption.

1.5. TYPES DE COÏNCIDENCES



FIG. 1.4: Cette photo représente la partie PET d'un imageur Philips Gemini ouverte. On distingue les cartes électroniques connectées aux tubes photomultiplicateurs.

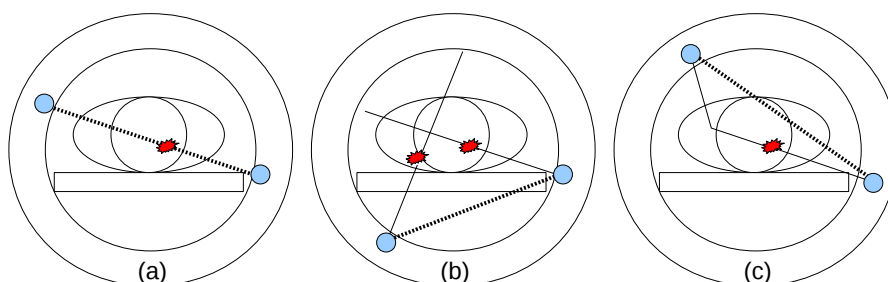


FIG. 1.5: Les différents types de coïncidences en TEP. Le trajet réel du photon est indiqué en trait fin simple, tandis que la ligne de réponse est indiquée en trait pointillé épais. On appelle les coïncidences vraies (a) lorsque l'annihilation est bien sur la ligne de réponse, fortuites (b) lorsque deux désintégrations réalisées en même temps sont considérées comme une seule et diffusée (c) lorsqu'un des photons est dévié.

1.6 Perturbation du trajet du photon

Les deux interactions les plus importantes que peuvent avoir les photons avec les tissus sont la diffusion Compton et l'absorption photo-électrique [Cherry and Dahlbom, 2006]. L'effet Rayleigh, qui correspond à la diffusion du photon sur un atome, est peu présent aux énergies concernées par la TEP.

La diffusion Compton correspond à l'interaction entre le photon émis et un électron du milieu. L'énergie cinétique de l'électron est augmentée, tandis que le photon est dévié. Il est intéressant de noter qu'une déviation de 25° engendre une perte d'énergie du photon de seulement 10% [Evans, 1955].

L'absorption photo-électrique correspond quant à elle à l'interaction entre un noyau atomique et un des photons. L'énergie du photon est absorbée par le noyau et transmis à un de ses électrons.

L'atténuation du signal est liée à une combinaison de ces effets. L'effet d'atténuation est particulièrement important en TEP, car il génère des artefacts très visibles, qui doivent être corrigés à partir d'une carte d'atténuation, déduite d'un examen de tomodensitométrie (tomographie par rayons X) ou d'une carte de transmission, comme indiqué sur la figure 1.6.

On peut représenter cette atténuation de manière analytique à l'aide de la loi de Beer-Lambert [Cherry and Dahlbom, 2006], qui montre que l'atténuation augmente exponentiellement avec la longueur du trajet dans les tissus :

$$I = I_0 e^{-\mu l} \quad (1.4)$$

Avec I_0 la quantité originale de photons, I le nombre de photons qui traversent le milieu, μ le coefficient d'atténuation du milieu et l la longueur du trajet.

Dans le cas où le milieu est complexe, et si l'on connaît l'atténuation $\mu(x)$ en chaque point x du trajet, on peut connaître précisément l'atténuation subie selon chaque ligne de réponse :

$$I = I_0 e^{-\int_0^L \mu(l) dl} \quad (1.5)$$

1.7 Quantification en imagerie TEP

L'image visualisée par le clinicien est directement liée à la concentration de l'isotope radioactif. La concentration observée est la somme de la concentration du traceur piégé dans les cellules et du traceur libre dans le compartiment vasculaire. Cette concentration mesurée est très dépendante du volume du patient, ainsi que de la dose injectée.

1.7. QUANTIFICATION EN IMAGERIE TEP

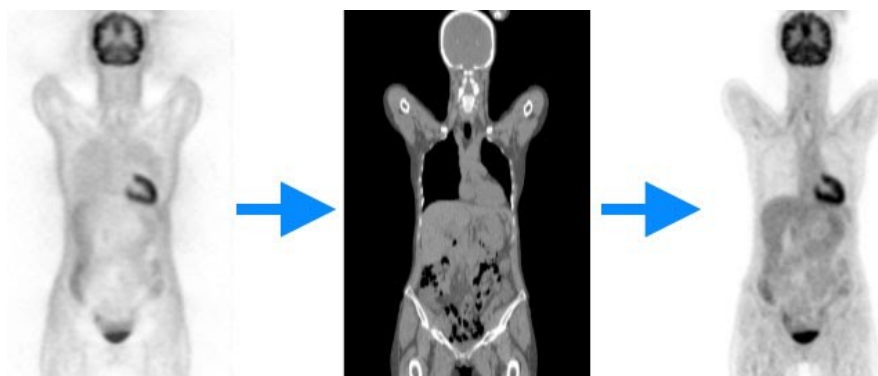


FIG. 1.6: Effets de l'atténuation sur des images TEP : L'image de gauche correspond à la distribution de radioactivité reçue par les détecteurs. On peut observer une perte progressive de l'activité vers l'intérieur des tissus, car les photons ont plus de chances d'interagir avec la matière. L'image TDM au centre est utilisée pour estimer une carte de l'atténuation des tissus et prendre en compte cette atténuation lors de la reconstruction de l'image TEP.

En clinique, on utilise habituellement le SUV (Standard Uptake Value), qui est un index semi-quantitatif permettant de s'affranchir de la dose injectée et des caractéristiques morphologiques du patient. Cela permet de réaliser des comparaisons entre des fixations radioactives de plusieurs examens, voire de plusieurs patients.

Le SUV est défini comme suit :

$$SUV = \frac{C}{\frac{\text{dose injectée}}{\text{poids corporel}}} \quad (1.6)$$

Où C est la concentration du traceur radioactif.

Des méthodes, plus complexes, basées sur des acquisitions dynamiques et la connaissance *a priori* d'un modèle de distribution du traceur dans l'organisme, permettent de remonter à des paramètres biologiques, tels que la consommation de glucose par gramme de tissus et par unité de temps pour le FDG. Ces méthodes sont en pratique peu utilisées en routine clinique en raison de la difficulté et du coût de réalisation des acquisitions dynamiques.

CHAPITRE 1. PRINCIPE PHYSIQUE

Chapitre 2

Déroulement d'une acquisition

Nous allons maintenant décrire rapidement quelques points techniques permettant de mieux comprendre comment sont réalisées les acquisitions TEP en routine clinique. Nous parlerons de la méthode utilisée pour réaliser des acquisitions corps entier alors que l'imageur a un champ de vue axial limité et des contraintes associées. Enfin, nous aborderons la problématique du format des données et ses implications.

2.1 TEP au ^{18}F FDG

En oncologie, les patients qui ont une suspicion de cancer se voient proposer un examen TEP pour réaliser un diagnostic initial, qui va viser à déterminer le degré de malignité de la pathologie détectée. Il a été montré [Gould et al., 2001] que les performances de la TEP sur la détection des lésions pulmonaires étaient supérieures à l'imagerie anatomique. Les performances de la TEP sont aussi reconnues pour les cancers hépatiques et la détection précoce des lymphomes. Plusieurs images sont visibles dans la figure 2.1.

La TEP est ensuite utilisée pour suivre l'évolution des lésions, afin d'adapter le traitement à l'évolution de la maladie.

2.2 Acquisitions corps entier

Les acquisitions TEP réalisées pour l'oncologie sont habituellement des acquisitions corps entier, de manière à pouvoir avoir une vue de l'étendue des lésions.

Or le champ de vue axial de la plupart des caméras TEP est de l'ordre de 15 à 20 cm environ (18 cm pour le Philips allegro [Lamare et al., 2006]). En routine clinique, le patient est placé sur un lit qui se déplace par incréments successifs lors de l'acquisition. Une photographie de l'imageur avec le lit est présentée dans la figure 2.2.

Le nombre de positions du lit est déterminé par la taille du patient ainsi que par le champ de vue du tomographe. Chaque lit est reconstruit séparément puis assemblé

CHAPITRE 2. DÉROULEMENT D'UNE ACQUISITION

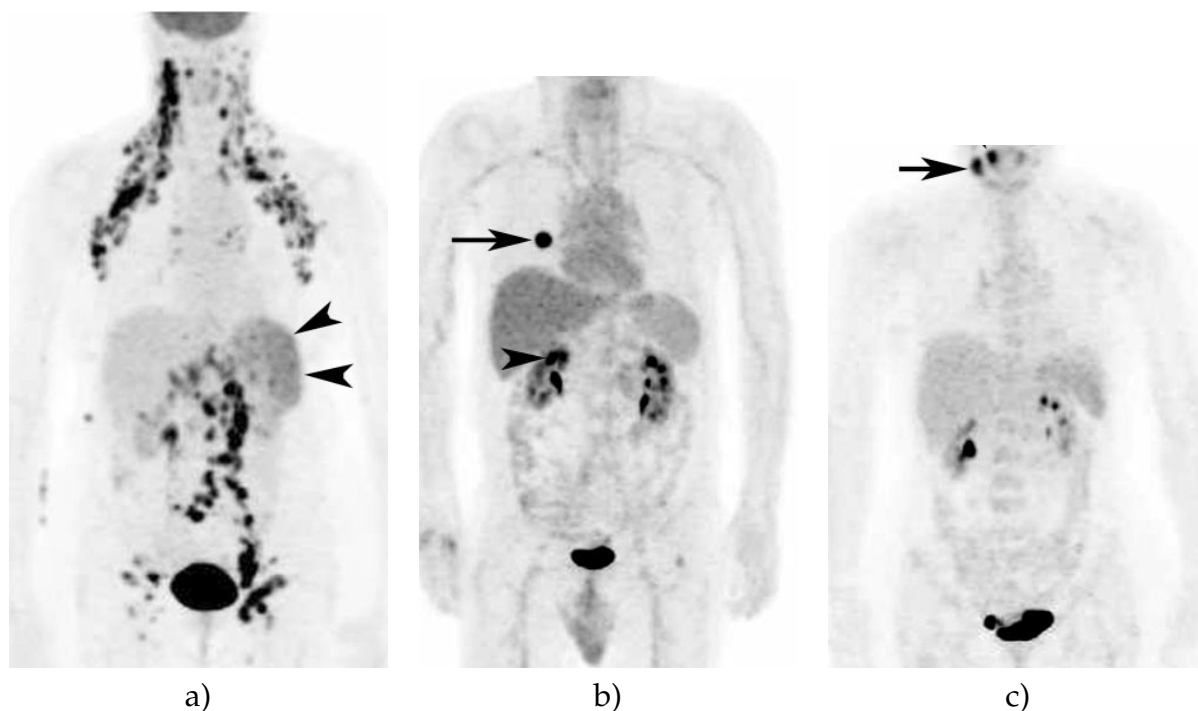


FIG. 2.1: Exemples d'images TEP au ^{18}F FDG provenant d'examens de patients atteints d'un cancer a) des ganglions (lymphome), b) du poumon et c) du cou. Certaines lésions sont représentées par les flèches noires. Les images sont extraites de l'article [Rohren et al., 2004].

avec les lits suivants et précédents, comme indiqué dans la figure 2.3. Étant donné que les caméras TEP ont une sensibilité plus faible à leurs extrémités, on introduit un chevauchement plus ou moins important dans les lits.

Les acquisitions corps entier utilisent typiquement plus de 10 lits sur la caméra Philips Gemini, ce qui demande un temps considérable. Il faut donc faire des compromis entre le temps d'acquisition et la qualité des images. Sachant que les examens TEP sont coûteux et que l'immobilité est inconfortable pour les patients, le temps d'acquisition est calibré pour durer moins d'une heure au total. Les nouvelles générations d'imageurs permettent cependant de réduire les temps d'acquisition à 2 minutes par lit à qualité équivalente.

2.3 Acquisition en mode 3D

Historiquement, les données étaient acquises par coupe : des septas (plaque destinées à absorber les photons) étaient placés entre les détecteurs dans la direction axiale de manière à éliminer les photons qui n'étaient pas émis dans le plan transverse (perpendiculaire à l'axe du détecteur). Aujourd'hui, la totalité des scanners fonctionnent en mode 3D uniquement. Le retrait des septas (voir la figure 2.4) permet d'augmenter le rapport signal sur bruit pour les gammes d'activité injectées aux patients, même si la détection des coïncidences vraies s'accompagne d'une augmentation de bruit parasite

2.3. ACQUISITION EN MODE 3D

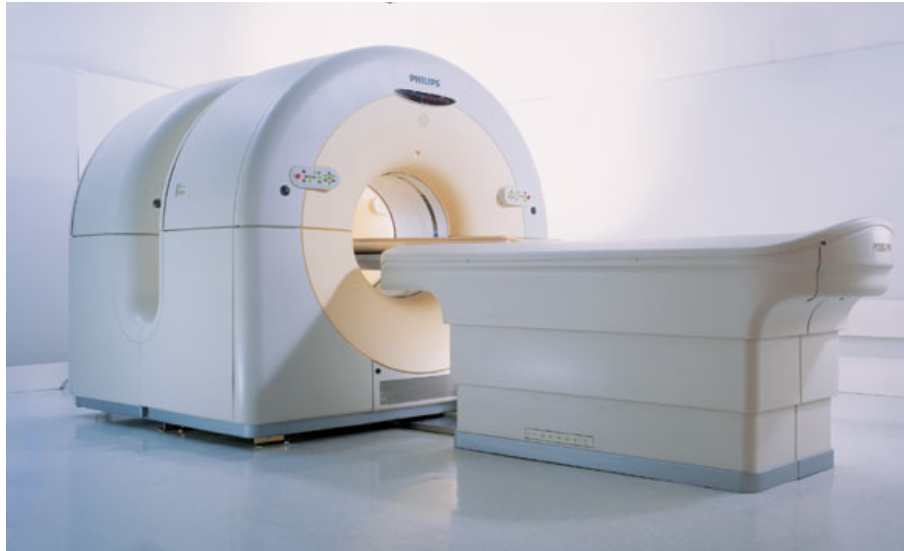


FIG. 2.2: Photographie d'un scanner TEP/TDM Gemini. Le lit motorisé est présenté au premier plan et permet de faciliter le déplacement du patient au cours de l'acquisition.

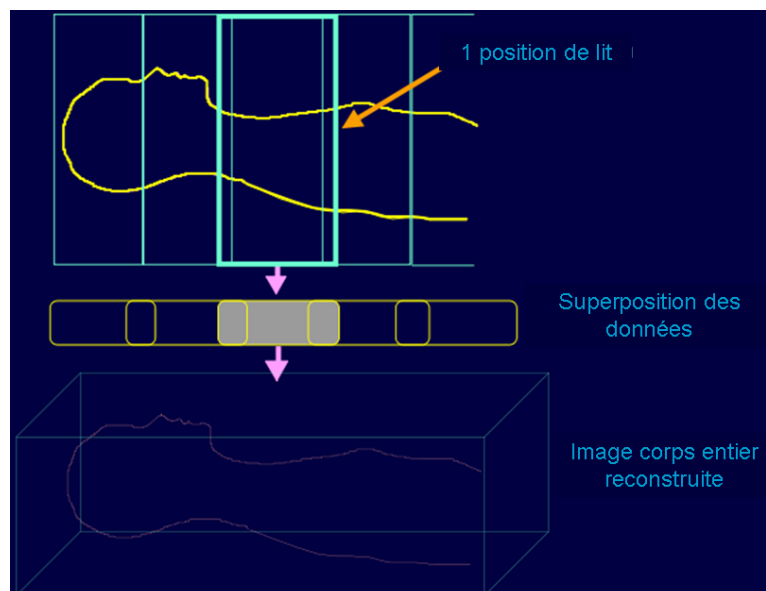


FIG. 2.3: Réalisation d'une acquisition corps entier en TEP : Chaque lit est acquis et reconstruit séparément avec un chevauchement. Les lits sont ensuite assemblés pour produire l'image finale.

CHAPITRE 2. DÉROULEMENT D'UNE ACQUISITION

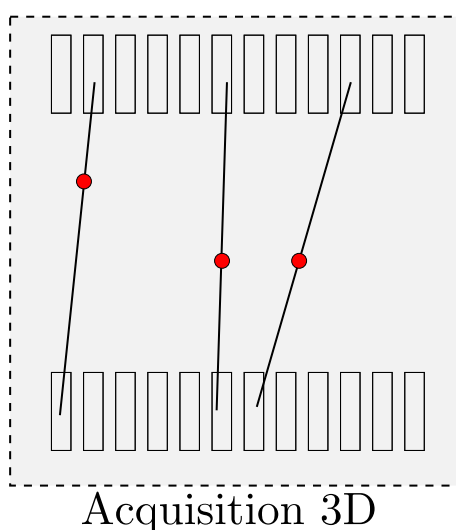


FIG. 2.4: Coupe schématique axiale d'un imageur TEP. Les annihilations représentées en rouge émettent des photons qui sont détectés par des cristaux correspondant à des plans de coupes différents.

liée à l'augmentation de la détection des coïncidences fortuites et diffusées.

2.4 Correction de l'atténuation

Le patient passe tout d'abord un examen TDM sur les scanners couplés TEP/TDM. Les valeurs des coefficients Hounsfield correspondant à l'atténuation des rayons X sont transformées pour correspondre à l'atténuation des photons de 511keV de la TEP. Cette image est ensuite utilisée pour réaliser la correction d'atténuation.

Une technique plus ancienne d'estimation de la carte d'atténuation est basée sur une image acquise en transmission, où une source radioactive émettrice de positons est placée dans l'imageur et tourne autour du patient. Une acquisition préalable est réalisée "à blanc", sans le patient. Le rapport entre la quantité de photons reçue par les détecteurs avec et sans le patient permet de générer la carte d'atténuation du corps du patient, de la même manière que pour l'imagerie TDM. Cependant, la généralisation des scanners couplés TEP/TDM limite l'utilisation de cette technique.

$$I(x) = I_0(x)e^{-\mu x} \quad (2.1)$$

En connaissant la radioactivité reçue sans le patient ("à blanc") I_0 , la radioactivité reçue avec le patient I , on peut en déduire la valeur de μ .

2.5. CORRECTION DES AUTRES SOURCES DE BRUIT

2.5 Correction des autres sources de bruit

Correction de temps mort

Le temps mort correspond à l'intervalle de temps durant lequel le dispositif est insensible après le passage d'une impulsion électrique. Lorsque deux impulsions surviennent avec un décalage temporel inférieur au temps d'intégration, seule la première est traitée et la deuxième est perdue. La correction de ce temps mort consiste à modéliser les pertes de comptage en événements pour chaque module de détection. En effet, le taux de photons simples s et le facteur de temps mort, noté d_{tk} , sont liés par la relation 2.2 où les constantes A et B apparaissant sont déterminées expérimentalement.

$$d_{tk} = 1 + As + Bs^2 \quad (2.2)$$

Correction de coïncidences diffusées

Les coïncidences diffusées, bien que déjà triées par les fenêtres énergétiques, peuvent excéder 50% des coïncidences détectées. Les images non corrigées présentent une diminution du contraste, du rapport signal sur bruit et une perte en résolution spatiale. Trois types de méthodes existent afin de corriger les données de ces coïncidences diffusées :

- Combinaison des données acquises dans au moins deux fenêtres en énergie.
- Exploitation de l'information spatiale de localisation erronée des coïncidences diffusées.
- Calcul direct de la distribution des coïncidences diffusées de manière analytique pour un patient donné, ou en utilisant des méthodes de simulation de type Monte Carlo.

Correction des coïncidences aléatoires

Les coïncidences aléatoires provenant de la détection de deux photons issus de deux annihilations différentes mais détectées dans la même fenêtre temporelle bruitent les données mesurées. La correction de ce phénomène consiste à soustraire au niveau du sinogramme les coïncidences aléatoires estimées des coïncidences vraies. Ces coïncidences aléatoires peuvent être estimées de trois façons [Brasse et al., 2005] :

- Utilisation d'une fenêtre temporelle décalée. L'activité y est considérée comme constante par sa taille réduite, mais les événements détectés proviennent forcément d'annihilations différentes par sa taille suffisamment élevée.
- Estimation du taux de coïncidences aléatoires en connaissant le nombre total de photons détectés par chaque module : la distribution des coïncidences aléatoires est plus ou moins uniforme dans le champ de vue. Le taux d'événements aléatoires mesuré sur une LDR est fonction de la largeur de la fenêtre de coïncidences temporelles (2τ) et des taux de photons simples (aussi appelé singles) détectés par les deux modules considérés.

CHAPITRE 2. DÉROULEMENT D'UNE ACQUISITION

- Estimation à partir de la distribution des coïncidences dans les projections en dehors du patient.

La normalisation

Des différences de sensibilité de détection d'une source uniforme existent d'une LDR à l'autre, d'une part à cause de la forme des systèmes de détections, en général à anneaux circulaires ; d'autre part à cause des différences d'incidence des photons par rapport à la face d'entrée du cristal et de l'efficacité individuelle du système de détection. Les méthodes de correction peuvent être basées sur l'inversion directe des données acquises à partir d'une source irradiant uniformément toutes les lignes de réponse. Elles peuvent aussi être basées sur des mesures traitées par un modèle mathématique reliant le facteur de normalisation pour une LDR entre deux cristaux à l'efficacité intrinsèque des détecteurs et une fonction de la sensibilité dépendant de l'arrangement géométrique des détecteurs.

2.6 Format des données

Les données acquises par une caméra TEP peuvent être stockées sous deux formes principales : séquentielles et sinogramme.

2.6.1 Séquentielles

Ce format correspond à un enregistrement "brut" des données issues de l'électronique de la caméra. Il est aussi appelé mode liste.

Ce format de fichier est en fait un enregistrement séquentiel des événements, dans leur ordre de détection. On peut enregistrer chaque photon détecté indépendamment, ou encore uniquement les coïncidences. Chaque événement est daté, ce qui permet de conserver l'information temporelle. On peut l'utiliser notamment pour synchroniser les données acquises avec le temps, pour la correction de mouvement par exemple, ou pour observer comment se répartit le radio-traceur au cours du temps.

Il existe plusieurs formats publics de fichiers pour le stockage de ces données, notamment le format LMF (List-Mode Format) développé pour le projet ClearPET et le format ROOT développé par le CERN. Cependant, chaque constructeur utilise son propre format de fichier propriétaire.

Ce mode de stockage consomme une très grande quantité de ressources, étant donné qu'un examen PET est constitué de plusieurs millions d'événements. Cela engendre de fortes contraintes en espace disque et complique de beaucoup la manipulation des données.

2.6. FORMAT DES DONNÉES

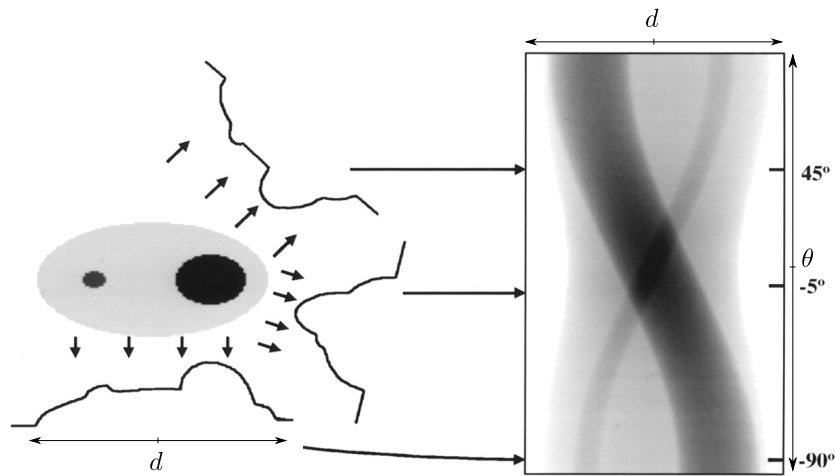


FIG. 2.5: Exemple de sinogramme : chaque ligne correspond à une projection de l'image selon un angle θ correspondant à la position de la ligne dans le sinogramme.

2.6.2 Sinogramme

Le sinogramme est une matrice indiquant pour chaque ligne de réponse le nombre de détections réalisées, comme présenté dans la figure 2.5. En ordonnées sont représentés les angles de la ligne de réponse, et en abscisse leur distance au centre du détecteur [Fahey, 2002]. Dans le mode 3D, on détecte des paires de photon qui ne sont pas directement dans le plan du détecteur. Une dimension supplémentaire correspondant à l'angle d'inclinaison de plans par rapport à l'axe du scanner est donc ajoutée pour prendre en compte cet effet.

Le principal avantage du sinogramme est qu'il permet de stocker les données acquises lors d'un examen TEP de manière beaucoup plus compacte que le format séquentiel. Cependant, il ne conserve aucune information temporelle ni énergétique.

CHAPITRE 2. DÉROULEMENT D'UNE ACQUISITION

Chapitre 3

Algorithmes de reconstruction

La reconstruction des images TEP correspond à un problème inverse : à partir de l'ensemble des lignes de réponse, il faut déduire la répartition du radio traceur dans l'organisme du patient. Deux classes d'algorithmes existent pour résoudre ce type de problèmes, mais actuellement seuls les algorithmes itératifs sont utilisés en oncologie clinique.

3.1 Algorithmes itératifs

La relation liant l'image TEP f et les projections sur les détecteurs p est la suivante :

$$p = Rf + e \quad (3.1)$$

Avec R la matrice de projection et e correspondant aux phénomènes parasites. Ce problème est mal posé, car la perte d'information due aux projections ne garantit pas la résolution de ce problème. De plus, le bruit est un élément perturbateur supplémentaire. Il nécessite donc l'utilisation d'algorithmes spécialisés.

3.1.1 Algorithme ML-EM (Maximum Likelihood Expectation Maximisation)

L'algorithme ML-EM [Shepp and Vardi, 1982] est initialisé avec une image de départ f_0 , et considère que les comptages des lignes de réponses p_i sont indépendants et bruités selon une loi de Poisson. Il se déroule en deux phases principales : une phase de comparaison de l'image au niveau n avec les lignes de réponses observées, puis une phase de mise à jour de l'image pour prendre en compte les différences calculées à l'étape précédente, comme présenté dans la figure 3.1.

On considère les $(p_i)_{i=1..L}$, avec L le nombre de projections, comme des variables

CHAPITRE 3. ALGORITHMES DE RECONSTRUCTION

aléatoires suivant une loi de Poisson dont la vraisemblance s'écrit :

$$proba(p|f) = \prod_{i=1}^L \frac{\langle p_i \rangle^{p_i}}{\langle p_i \rangle!} e^{-\langle p_i \rangle} \quad (3.2)$$

Où $\langle p_i \rangle$ est le paramètre de la loi de Poisson (moyenne) de la ligne de réponse p_i . L'objectif des algorithmes de reconstruction itératifs est d'estimer la distribution f des $(f_j)_{j=1..P}$, avec P le nombre de pixels, conduisant à l'ensemble $(\langle p_i \rangle)_{i=1..L}$. L'équation 3.1, en négligeant les erreurs, s'écrit donc :

$$\langle p_i \rangle = \sum_{j=1}^P R_{ij} f_j \quad (3.3)$$

On maximise la log-vraisemblance de l'équation 3.2 pour évaluer la cohérence entre les pixels de l'image $(f_j)_{j=1..P}$ et les projections observées $(p_i)_{i=1..L}$. Cette fonction de coût est notée $Q(f_1 \dots f_P, q_1 \dots q_L)$. On cherche l'estimation \hat{f} de f qui maximise cette fonction de coût :

$$\hat{f}_j^{(n+1)} = \frac{\hat{f}_j^{(n)}}{\sum_{i'} R_{i'j}} \sum_i R_{ij} \frac{p_i}{\sum_k R_{ik} \hat{f}_k^{(n)}} \quad (3.4)$$

En pratique, la matrice système R peut inclure des informations supplémentaires pour corriger les différences de sensibilité entre les différents capteurs, et réaliser la correction des diffusés [Shepp and Vardi, 1982, Chornoboy et al., 1990] ou la correction de l'atténuation des tissus. Cette matrice "système" R_{ij} indique la probabilité de détecter une désintégration issue du voxel j de l'image dans la ligne de réponse i .

Le nombre d'itérations à réaliser avant d'atteindre la convergence dépend de l'image, mais l'ordre de grandeur est d'environ 20 à 50. Il existe plusieurs publications proposant un critère d'arrêt [Bissantz et al., 2006], mais en pratique la reconstruction est souvent réalisée pour un nombre d'itérations élevées pour assurer la convergence [Bailey, 2005] et suivi d'un filtrage passe-bas [Daube-Witherspoon et al., 2001] pour atténuer l'amplification du bruit. Cet algorithme de reconstruction prend un temps important car il faut réaliser l'opération de projection et de rétroprojection sur l'ensemble des données à chaque itération.

3.1.2 Algorithme OS-EM (Ordered Subset Expectation Maximisation)

L'algorithme OS-EM est une évolution de l'algorithme ML-EM qui permet une accélération substantielle de la convergence en réalisant les itérations de l'équation 3.4 sur des sous-ensemble des données acquises [Hudson, 1994].

3.1. ALGORITHMES ITÉRATIFS

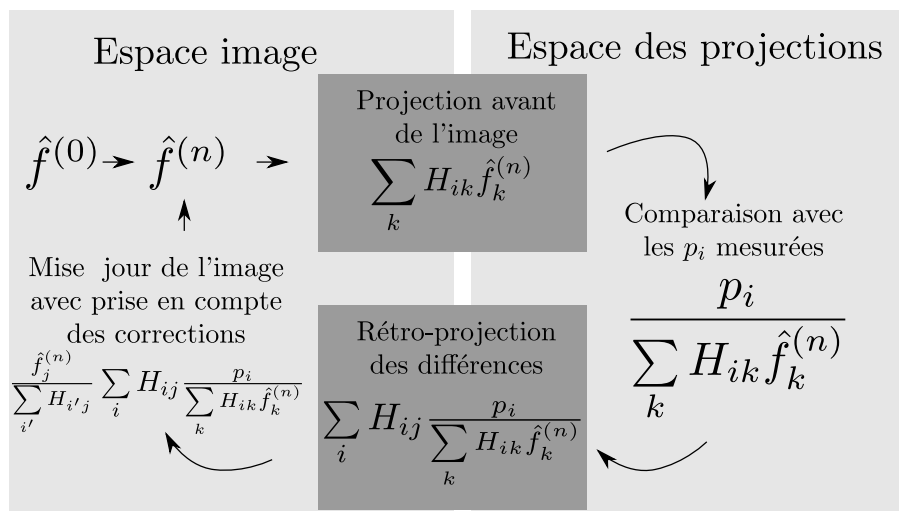


FIG. 3.1: Schéma décrivant l'algorithme itératif ML-EM. $\hat{f}_j^{(n)}$ est l'estimation du voxel j de l'image à l'itération n . R_{ij} est la matrice système, et représente la probabilité de détecter une désintégration dans le voxel j à la ligne de réponse i .

Ces sous-ensembles de données sont réalisés en échantillonnant de manière régulière les lignes de réponses. Leur nombre est un des paramètres de l'algorithme, mais la convergence n'est plus garantie contrairement à ML-EM. En pratique, l'algorithme converge toujours approximativement, et le nombre d'itérations est déterminé de manière empirique [Bailey, 2005]. Des évolutions de OS-EM ont été réalisées pour garantir une convergence, notamment l'algorithme Row-Action Maximum Likelihood (RAMLA) [Browne and De Pierro, 1996, Chiang et al., 2004].

Le principal avantage de OS-EM est qu'il permet d'augmenter la vitesse de la reconstruction d'un facteur correspondant au nombre de sous-ensembles utilisés.

3.1.3 Cas des acquisitions en mode séquentiel

Les acquisitions en mode séquentiel génèrent une suite d'enregistrements correspondant aux événements détectés par l'imageur. La lecture de tous ces événements un par un demande des temps de calculs extrêmement importants, c'est pourquoi Reader et al. [Reader et al., 2002] ont proposé une adaptation de l'OS-EM au mode séquentiel appelée One-Pass List-Mode EM (OPL-EM). L'algorithme proposé est conçu pour ne parcourir qu'une seule fois la liste des événements. Comme pour OS-EM, les données sont groupées en sous-ensembles, qui sont utilisés les uns après les autres pour chaque itération, comme indiqué dans l'équation 3.5.

Ici, s correspond à un numéro de sous-ensemble, j à un voxel et i à une ligne de réponse. Lorsque tous les sous-ensembles ont été parcourus, une nouvelle itération est déclenchée. Cet algorithme est conçu pour permettre de parcourir les données une seule fois, ce qui revient à réaliser une seule itération, cependant, il reste possible d'en

réaliser plusieurs.

$$f_j^{(s+1)} = \frac{\hat{f}_j^{(s)}}{\sum_{i'} R_{i'j}} \sum_{i \in T^s} R_{ij} \frac{1}{\sum_k R_{ik} \hat{f}_k^{(s)}} \quad (3.5)$$

La différence entre OS-EM et cet algorithme est la manière dont sont sommées les différences. Dans OPL-EM, la sommation est réalisée sur l'ensemble des événements T^s correspondant au sous-ensemble d'événements s . Cette somme est réalisée non pas pour chaque ligne de réponse p comme précédemment, mais pour chaque événement détecté, ce qui explique le remplacement de p_i par la valeur 1.

3.2 Reconstructions analytiques

Ces algorithmes sont supplantés en oncologie par les algorithmes itératifs qui sont jugés plus performants sur l'évaluation du SUV et moins bruités [Schöder et al., 2004].

Les méthodes analytiques utilisent la rétroprojection des projections du sinogramme pour reconstruire l'image. Si P est l'opération de projection d'une distribution $f(x, y)$ d'un objet vers $p(r, \theta)$, selon l'angle azimutal θ , et r la distance entre la projetée de la LDR dans le plan (x, y) et l'axe x , alors la projection peut s'écrire comme la transformée de Radon de cette distribution :

$$p(u, \theta) = P(f(x, y)) = \int_{-\infty}^{+\infty} f(u \cos \theta - v \sin \theta, u \sin \theta + v \cos \theta) dv \quad (3.6)$$

avec la relation suivante entre (x, y) et (u, v) :

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \quad (3.7)$$

L'opération de rétroprojection notée RP approxime l'opération inverse de la projection et permet d'obtenir l'estimation $\hat{f}(x, y)$ en fonction des projections acquises :

$$\hat{f}(x, y) = RP(p(u, \theta)) \quad (3.8)$$

La rétroprojection simple consiste à ajouter les contributions de chaque ligne de réponse à l'image estimée :

$$\hat{f}(x, y) = \int_0^\pi p(u, \theta) d\theta \quad (3.9)$$

Cependant cette rétroprojection génère des artefacts en "étoile" autour de l'objet. La rétroprojection filtrée [Kinahan et al., 1988] permet de limiter ces artefacts. Elle est

3.2. RECONSTRUCTIONS ANALYTIQUES

basée sur le théorème de la coupe centrale pour réaliser la rétroprojection des données. Ce théorème démontre que la transformée de Fourier monodimensionnelle de la projection $p(u, \theta)$ selon la direction θ est égale à la transformée bidimensionnelle de la distribution $f(x, y)$ selon un plan perpendiculaire à la direction de projection et passant par l'origine :

$$[TF_{2D}(f)](v_x, v_y) = [TF_{2D}(b)](v_x, v_y)|v| \quad (3.10)$$

avec $b(x, y)$ correspondant au résultat de la rétroprojection non filtrée de l'équation 3.9.

Les projections mesurées sont d'abord filtrées par le filtre rampe v dans le domaine de Fourier. On rétroprojète ensuite la transformée de Fourier inverse de ces données filtrées. On obtient alors l'estimation de la distribution originale \hat{f} .

CHAPITRE 3. ALGORITHMES DE RECONSTRUCTION

Deuxième partie

Mouvement respiratoire

Cette partie présente tout d'abord la problématique du mouvement respiratoire en TEP. Puis il aborde successivement les différentes techniques permettant d'estimer soit le signal respiratoire, soit le champ de mouvement 3D du patient. Enfin, le dernier chapitre montre comment cette information sur le signal respiratoire ou le champ de mouvement peut être prise en compte pour corriger les phénomènes de dégradation engendrés par la respiration.

Chapitre 4

Respiration et influence sur les acquisitions TEP/TDM

4.1 Introduction

Le mouvement respiratoire en imagerie TEP engendre plusieurs effets sur les images reconstruites, qui seront détaillés ci-après. Il occasionne notamment une diminution de la qualité des images, ce qui peut perturber le travail des praticiens.

Nous ferons tout d'abord une introduction générale sur le mouvement respiratoire, puis nous enchaînerons ensuite sur les différents effets que peut avoir ce mouvement sur les images TEP, avec notamment des incertitudes sur la forme et la localisation des lésions et une diminution de l'activité estimée. Nous parlerons ensuite de l'impact de ce mouvement sur la détectabilité des lésions.

4.2 Mouvement respiratoire

Ce mouvement est la succession d'une phase inspiratoire, suivi d'une phase expiratoire. Chacune de ces phases combine plusieurs mouvements élémentaires [Servant-Laval, 2007] :

1. Thoracique, avec un déplacement des côtes.
2. Abdominal, avec un déplacement du diaphragme.
3. En cas d'inspiration forcée, action des pectoraux.

La figure 4.1 représente 3 instants du cycle respiratoire utilisé dans le NCAT [Segars, 2001], un modèle de corps humain respirant. On voit clairement tous les effets présentés précédemment, notamment le relèvement des côtes et du sternum par rapport aux guides de référence présents sur les images.

La variabilité inter et intra-patient de ce mouvement est très importante : le volume d'air inspiré peut varier de 500 ml à 1200 ml selon que la personne a une respiration

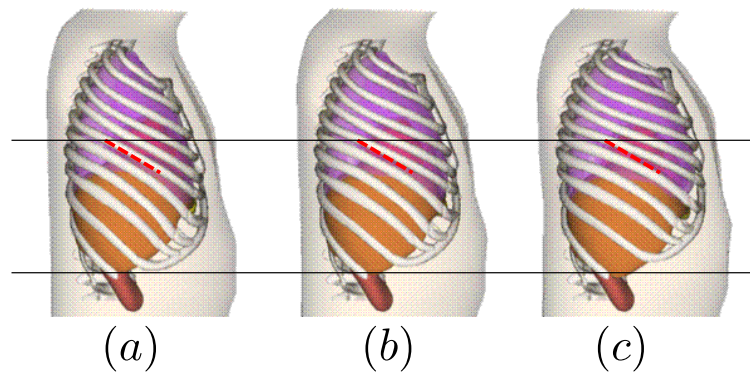


FIG. 4.1: Modèle de mouvement intégré dans le fantôme NCAT. La ligne pointillée rouge sert de référence pour la position d'une des côtes du modèle en expiration. L'image (a) correspond à l'expiration complète, et la (c) à la fin de l'inspiration. (b) correspond à un instant intermédiaire du cycle.

normale ou profonde. Pour ces deux extrêmes, la fréquence respiratoire varie de 5 cycles/min. à 20 cycles/min. [Sherwood, 2006].

Or, une acquisition TEP a une durée de plusieurs minutes par lit, ce qui provoque des artefacts dû aux mouvements réalisés pendant l'acquisition, notamment au niveau de la localisation de la tumeur, de son activité mesurée et, par extension, de sa détection. De la même manière, des artefacts (voir figure 4.4) apparaissent au niveau des zones de fort mouvement dans les images TDM, lorsque les organes bougent pendant une rotation du détecteur.

Nous allons tout d'abord introduire les effets visibles sur les images, puis nous nous attarderons sur les mesures quantitatives utilisées pour étudier l'apport de la correction du mouvement sur la détection des tumeurs. Enfin, nous dresserons un état de l'art plus précis des publications s'intéressant à l'impact du mouvement respiratoire sur la détection.

4.3 Localisation et volume

La localisation et le volume apparent des tumeurs sur les images TEP peuvent être modifiés par le mouvement engendré par la respiration (voir figure 4.2). D'après l'étude de [Lamare et al., 2007a], réalisée sur des simulations Monte-Carlo à l'aide du logiciel de simulation GATE [Jan et al., 2004] en utilisant le modèle NCAT [Segars, 2001], la largeur à mi-hauteur des lésions peut être modifiée de 48% (équation 4.1 : Différence relative) dans le cas d'une lésion de 7mm de diamètre dans la partie basse du poumon. L'imprécision axiale sur le positionnement de la tumeur peut atteindre

4.4. MESURE DE L'ACTIVITÉ DES TUMEURS

9% dans les mêmes conditions.

$$\%DR = \left| \frac{\text{Critère sur image non corrigée} - \text{Critère sur image de référence}}{\text{Critèresurimage de référence}} \right| \quad (4.1)$$

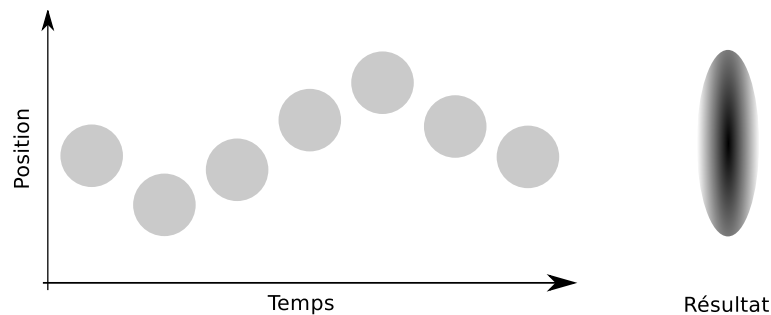


FIG. 4.2: Effet du déplacement d'une tumeur sur les données acquises. La position de la tumeur change en fonction du temps, ce qui provoque l'acquisition d'une tumeur équivalente présentée à droite.

4.4 Mesure de l'activité des tumeurs

Le contraste des tumeurs par rapport au fond est un critère important pour déterminer la malignité des tumeurs [Dimitrakopoulou-Strauss et al., 2002, Krak et al., 2005]. De plus ce contraste joue aussi un rôle important sur la détectabilité de la tumeur en facilitant sa distinction du fond, comme nous allons le voir ci-après.

L'activité peut être influencée par la respiration de deux manières :

- Par un mauvais ajustement de la carte d'atténuation.
- Par le moyennage de la position de la tumeur.

4.4.1 Décalage de la carte d'atténuation

La carte d'atténuation utilisée pour corriger les images d'émission est basée sur une image TDM prise à un instant donné du cycle. Or l'atténuation de la zone correspondant à une tumeur peut être différente de celle des tissus environnants, ce qui peut occasionner des sous-estimations ou des sur-estimations de l'activité de la tumeur si la carte d'atténuation est mal positionnée.

D'après l'étude [Erdi et al., 2004], on peut observer des variations très importantes de SUV_{max} (voir équation 4.2) sur les images TEP reconstruites. L'étude a été réalisée sur 5 patients totalisant 8 lésions. L'acquisition TDM 4D a été synchronisée sur le mouvement à l'aide d'un dispositif placé sur l'abdomen du patient. Ce dispositif était constitué de deux marqueurs réfléchissant les ondes infrarouges, placés dans le champ

CHAPITRE 4. RESPIRATION ET INFLUENCE SUR LES ACQUISITIONS TEP/TDM

de vue d'une caméra. Les données synchronisées acquises en TDM 4D ont été utilisées pour générer un cycle de 10 images TDM. Chacune de ces 10 images a ensuite servi de carte d'atténuation lors de la reconstruction des images d'émission, générant ainsi 10 images TEP.

$$\% \text{ variation } SUV_{max} = 100 \times \frac{|SUV_1 - SUV_2|}{(SUV_1 + SUV_2)/2} \quad (4.2)$$

Les données TEP reconstruites à partir des 10 images ont montré de grandes différences en terme de SUV selon l'instant du cycle à partir duquel les lésions ont été extraites. Ces variations allant jusqu'à 24% pour une lésion de 19.8 mm^3 placée dans l'espace médiastinal, entre une reconstruction TEP réalisée à partir d'une image TDM de fin d'expiration et une image TDM de fin d'expiration. En recherchant la variation la plus grande sur l'ensemble du cycle (en non pas en se limitant aux extrêmes), la variation peut atteindre 30%.

La figure 4.3 présente la variation de SUV_{max} en fonction de la phase utilisée pour la reconstruction.

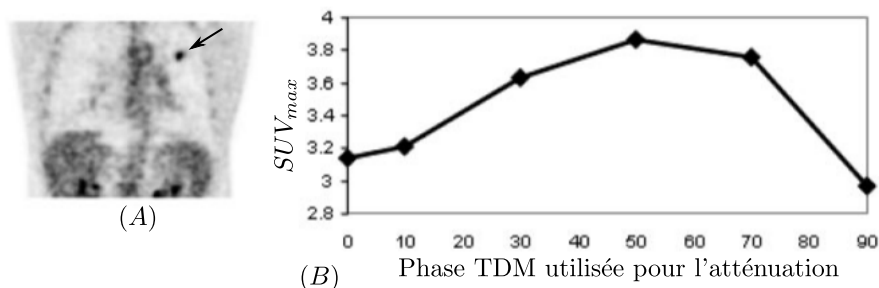


FIG. 4.3: A) Vue coronale d'un patient avec une lésion clairement visible en haut à droite (flèche noire). B) SUV_{max} observé sur les images reconstruites en fonction de l'image TDM utilisée pour la reconstruction. Le nombre en abscisse représente la position dans le cycle respiratoire en pourcentage.

4.4.2 Artefacts TEP dus aux artefacts TDM

On peut voir sur les images de la figure 4.4 des artefacts présents sur les images TDM utilisées pour la correction d'atténuation. Ces artefacts proviennent de la manière dont les images sont acquises : le couple "source de rayons X-détecteur" tourne autour du sujet dans un mouvement hélicoïdal, et l'algorithme de reconstruction va ensuite utiliser les acquisitions pour reconstruire une image complète. Or, en cas de respiration rapide, des incohérences peuvent survenir quand le mouvement du diaphragme est plus important que celui de la caméra. Ce type d'artefacts peut créer des incohérences dans les images TEP reconstruites, car le poumon sera corrigé de l'atténuation avec des coefficients non uniformes.

4.5. IMPACT DU MOUVEMENT RESPIRATOIRE SUR LA DÉTECTION



FIG. 4.4: Artefacts présents sur des images TDM utilisées pour la correction d'atténuation.

4.4.3 Déplacement de la tumeur au cours du cycle

Le mouvement respiratoire va avoir pour effet de déplacer la tumeur pendant l'acquisition, ce qui va moyenniser la quantité de radioactivité sur l'ensemble du cycle, comme indiqué sur la figure 4.2. Si le déplacement de la tumeur est suffisamment grand par rapport à son diamètre, la réduction de radioactivité va être importante.

L'étude [Boucher et al., 2004] réalisée sur un fantôme elliptique (modèle "Elliptical Jaszczak Phantom") montre qu'un déplacement d'une source radioactive de 6 mm sur un cycle respiratoire moyen entraîne une sous-estimation de l'activité maximale de la source de 41% pour un objet de 1.2 ml et de 21% pour une sphère de 19.4 ml.

4.5 Impact du mouvement respiratoire sur la détection

Peu de travaux ont été réalisés sur l'impact du mouvement respiratoire sur la détection des tumeurs. Globalement, les publications traitent principalement de la mesure de la quantification des lésions (SUV_{max} , profils de lésions, ...). Très peu d'articles utilisent des critères orientés sur les performances de détection.

Voici une liste des critères utilisés dans différentes publications pour évaluer les performances d'algorithmes de correction du mouvement respiratoire :

1. Basés sur la récupération de l'activité des lésions, SUV_{max} , contraste : [Chang

CHAPITRE 4. RESPIRATION ET INFLUENCE SUR LES ACQUISITIONS TEP/TDM

et al., 2010, Lamare et al., 2007b, Nehmeh et al., 2002, Detorie and Dahlbom, 2008].

2. Sur la visualisation des profils de lésion : [Chang et al., 2010, Thielemans et al., 2006, Lamare et al., 2007b].
3. Sur la récupération du volume et de la position de la lésion : [Chang et al., 2010, Lamare et al., 2007b, Nehmeh et al., 2002].
4. Sur le rapport signal sur bruit (SNR) : [Chang et al., 2010].
5. Sur l'observateur de Hotelling (CHO) : [Thielemans et al., 2006].

Comme on peut le voir, les deux seuls critères qui pourraient s'approcher d'une étude sur la détection sont le SNR et le CHO, mais ils sont largement sous-représentés.

Je vais me concentrer sur les deux publications de Thielemans et Chang qui utilisent un observateur.

Un autre document par Rahmim Arman [Rahmim et al., 2011] propose d'utiliser le CHO pour évaluer l'amélioration de la détection des défauts dans de l'imagerie cardiaque corrigée du mouvement respiratoire. Ce document n'a pas encore donné lieu à publication.

4.5.1 Article de Thielemans et al. [Thielemans et al., 2006]

Dans sa publication, Thielemans et al. utilisent le "Channelized Hotelling Observer" (CHO) [Barrett et al., 1993], qui est un observateur dérivé du classifieur linéaire LDA présenté en 8.5.1, utilisé en conjonction avec des informations fréquentielles.

L'étude se base sur des simulations analytiques d'un fantôme thoracique dans lequel sont insérées des lésions sphériques de gros diamètre (13 mm) et de contraste élevé (4.25 :1). Cette étude qui compare plusieurs méthodes de correction du mouvement respiratoire est détaillée dans le chapitre suivant en 6.5.

4.5.2 Article de Chang et al. [Chang et al., 2010]

Cette publication est une évaluation clinique d'un système de découpage automatique du signal respiratoire selon l'amplitude. Elle est basée sur 13 patients totalisant 21 tumeurs du poumon de taille diverses (estimée de 1 à 27 cm³). Le signal respiratoire est acquis à l'aide d'un dispositif propriétaire (Anzai AZ-733V) basé sur une ceinture abdominale avec capteur de pression. Seules les données TEP acquises autour d'une amplitude présélectionnée (+/-10% de l'amplitude correspondant à l'acquisition TDM) sont utilisées pour la reconstruction des images corrigées.

Le jeu d'images "témoin", non corrigées, est réalisé en extrayant des données acquises le même nombre d'évènements N que le jeu d'images synchronisées en considérant les N premiers du fichier en mode séquentiel.

L'auteur compare le SUV_{max} , SUV_{moyen} , Rapport Signal/Bruit (SNR) et volume apparent de la lésion pour les deux jeux d'images. Le volume de la lésion et le SUV_{moyen}

4.5. IMPACT DU MOUVEMENT RESPIRATOIRE SUR LA DÉTECTION

sont calculés sur une région d'intérêt correspondant à 40% du SUV_{max} de la lésion. Le SNR est calculé en divisant le SUV moyen de la tumeur par l'écart-type d'une zone d'intérêt dans le poumon, dont ni la taille ni la localisation ne sont précisées.

Les résultats présentés montrent une amélioration importante des valeurs du SNR pour les images corrigées du mouvement par rapport aux images témoin, comme présenté dans la figure 4.5. En moyenne, cette augmentation est de 26.3%, mais deux tumeurs sur les 21 de l'étude voient leur SNR augmenter de plus de 66%.



FIG. 4.5: Exemple d'image avec et sans correction de mouvement respiratoire.

Il est étonnant de constater que les mesures moyennes sur les 21 tumeurs de SUV_{max} et SUV_{moyen} voient leurs valeurs augmenter respectivement de 26.8% et 26%, ce qui est tout à fait semblable avec l'augmentation de SNR. Cependant, les valeurs individuelles d'amélioration pour chaque tumeur ne montrent pas de corrélation avec le SNR. Par exemple, une lésion va montrer une augmentation de 24% de son SUV_{max} , de 20% de son SUV_{moyen} et de 37% de son SNR.

Il est intéressant de constater qu'il y a un point où le SNR diminue de 3.4%, mais que le SUV_{max} et le SUV_{moyen} augmentent tout de même de presque 18%. Cela semble indiquer une erreur car le SUV de la zone d'intérêt n'est pas censé changer de manière importante.

CHAPITRE 4. RESPIRATION ET INFLUENCE SUR LES ACQUISITIONS TEP/TDM

Chapitre 5

Processus d'estimation du mouvement

5.1 Introduction

Cette partie présente tout d'abord la problématique du mouvement respiratoire en TEP. Puis il aborde successivement les différentes techniques permettant d'estimer soit le signal respiratoire, soit le champ de mouvement 3D des organes internes du patient. Enfin, le dernier chapitre montre comment cette information sur le signal respiratoire ou le champ de mouvement peut être prise en compte pour corriger les phénomènes de dégradation engendrés par la respiration.

5.2 Estimation du signal respiratoire

Le signal respiratoire est une grandeur caractérisant la position du patient sur le cycle respiratoire, entre la fin d'inspiration et la fin d'expiration. Il est habituellement fourni par des capteurs externes qui génèrent un signal corrélé avec la respiration. Cela permet de faire correspondre les données acquises par un imageur avec une phase particulière du mouvement respiratoire.

5.2.1 Spiromètre

Le spiromètre est un capteur externe placé sur la bouche du patient et qui permet de mesurer les déplacements d'air dans le système respiratoire [Guivarc'h et al., 2004a]. Les spiromètres mesurent un débit ou un volume d'air inspiré/expiré (voir illustration figure 5.1). À partir de l'une des grandeurs (volume ou débit), il est possible d'estimer l'autre facilement. L'avantage du spiromètre est qu'il permet d'accéder à une mesure caractérisant directement la respiration du patient, et n'est pas sujet à des perturbations externes (mouvements involontaires par exemple). Par contre cela demande un appareillage qui peut être assez invasif pour le patient.

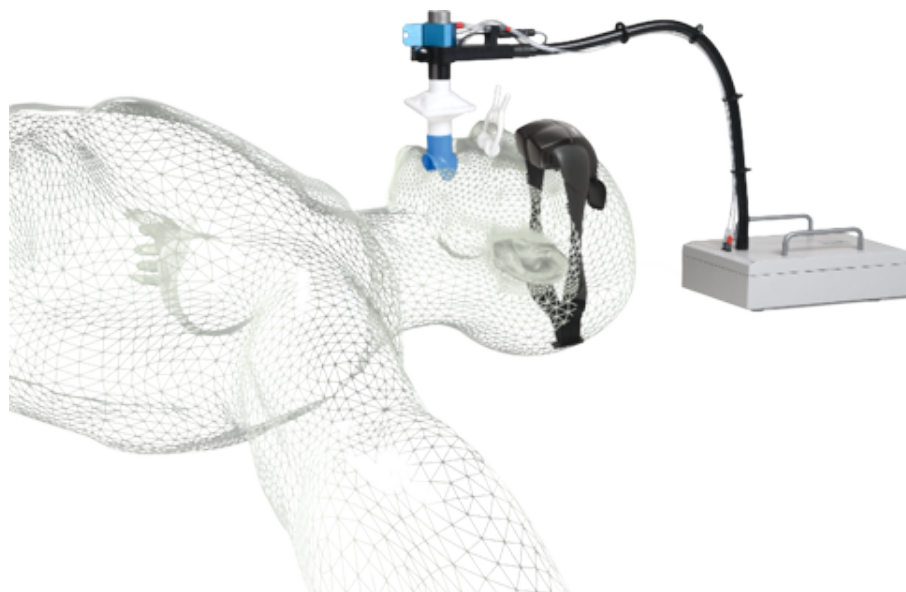


FIG. 5.1: Illustration du spiromètre Syn'r : on peut voir le système de mesure de la respiration ainsi qu'un système de moniteurs implantés dans les lunettes pour aider le patient à contrôler sa respiration.

5.2.2 Ceinture

Pour mesurer le signal respiratoire, il est possible d'utiliser un capteur qui va mesurer le périmètre du thorax. L'extension de cette ceinture va correspondre aux mouvements de la cage thoracique et de l'abdomen pendant la respiration du patient. C'est une mesure indirecte de l'amplitude du mouvement respiratoire utilisée couramment en routine clinique.

Différentes technologies existent pour mesurer cette information (RespiTrace R250 de Studley. Data Systems, Respiratory Belt Transducer de ADInstruments, ...). Elles sont basées sur plusieurs effets (résistif, inductif...) et ont l'avantage d'avoir un faible coût et de ne pas perturber le patient.

Bien qu'elles puissent être influencées par les mouvements involontaires du patient, il a été montré dans [Guivarc'h et al., 2004b] que les données acquises selon les méthodes par spiromètre et par ceinture sont équivalentes.

5.2.3 Techniques basés sur des caméras vidéo

Des caméras peuvent être utilisées pour estimer le mouvement respiratoire. Une des techniques consiste à utiliser des informations surfaciques en reconstruisant en 3D certaines parties du corps à l'aide de plusieurs caméras [Beach et al., 2004], ou en utilisant des caméras 3D temps de vol [Fayad et al., 2009]. Cela permet d'avoir plus d'informations sur la respiration.

5.2. ESTIMATION DU SIGNAL RESPIRATOIRE

Une autre technique consiste à installer un marqueur sur le corps du patient et à relever les déplacements de ce marqueur selon plusieurs axes à l'aide d'une ou plusieurs caméras. Un tel système est décrit dans [Nehmeh et al., 2002]. L'un de ces systèmes, le "Respiratory Gating System" de Varian Medical Systems est illustré sur la figure 5.2.

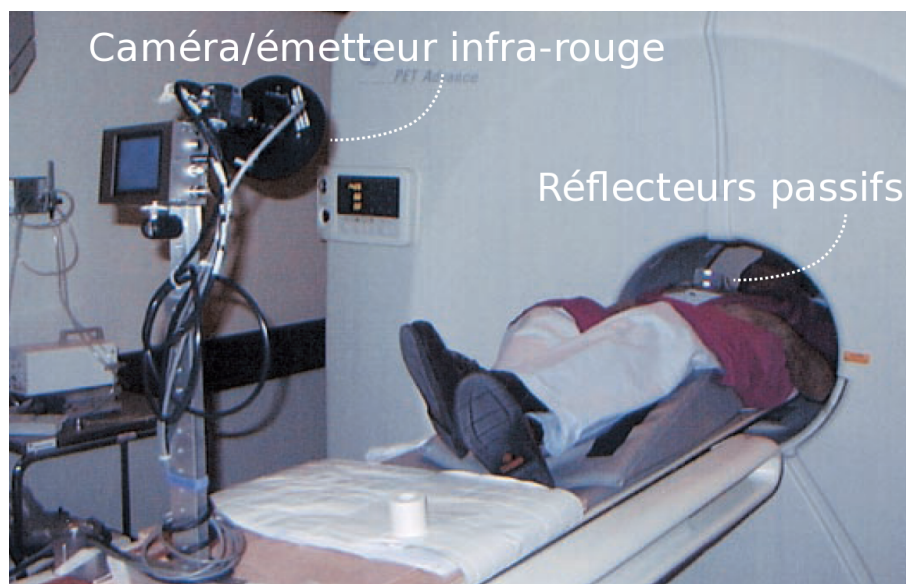


FIG. 5.2: Photographie du système RGS de Varian Medical Systems en action : une caméra va détecter le déplacement d'une zone du thorax en mesurant le déplacement de marqueurs placés sur un bloc plastique.

Ces techniques ont l'avantage d'être moins invasives et plus facilement acceptées par le patient. Cependant, elles sont beaucoup plus sensibles aux mouvements involontaires du patient. Ce risque est faible pour l'imagerie TDM car de faible durée (inférieure à la minute), mais il devient important pour les acquisitions TEP qui peuvent durer en tout plusieurs dizaines de minutes. Ces mouvements n'étant probablement pas corrélés avec le mouvement respiratoire, ils vont perturber le signal obtenu.

5.2.4 Techniques basées sur les images TEP

En TEP, la publication de Bundschuh et al. [Bundschuh et al., 2007] utilise les données dynamiques pour estimer le signal respiratoire sans avoir besoin de capteur externe. Ce processus est réalisé en 5 étapes :

1. Les données TEP sont acquises en mode séquence (aussi appelé mode liste, défini en 2.6.1) : toutes les désintégrations détectées sont enregistrées dans un fichier de manière séquentielle.
2. Une image TEP statique est reconstruite. Elle permet de localiser une lésion dans l'image qui servira d'amer pour l'estimation du mouvement respiratoire.

CHAPITRE 5. PROCESSUS D'ESTIMATION DU MOUVEMENT

3. Les données dynamiques sont regroupées en blocs de 0.5 secondes d'acquisitions. Chaque bloc est reconstruit séparément.
4. La zone d'intérêt choisie précédemment est sélectionnée dans chaque image reconstruite. La position axiale du barycentre à chaque instant temporel va donner une estimation du mouvement respiratoire pour le volume donné.

Cette technique a été évaluée sur 10 patients et les signaux ont été comparés avec ceux obtenus avec des ceintures abdominales. Pour 3 patients, les courbes respiratoires obtenues par les deux techniques étaient très fortement corrélées. Pour deux autres patients, l'estimation de mouvement obtenue par la TEP était trop bruitée, mais montrait une bonne corrélation avec le signal obtenu par les ceintures abdominales après filtrage. Pour 3 autres patients, il n'a pas été possible de trouver une corrélation entre les deux signaux. Les deux derniers patients ont bougé pendant l'acquisition TEP, ce qui a perturbé le signal.

Cette technique est intéressante mais contrainte par la qualité de l'image TEP. En pratique, la moitié des acquisitions TEP n'a pas permis d'obtenir un signal respiratoire satisfaisant.

5.3 Estimation du champ de mouvement

Le champ de mouvement est une information beaucoup plus riche que le signal respiratoire car il permet de suivre les déplacements des organes et des lésions à l'intérieur du corps du patient au cours d'un cycle respiratoire.

Le signal respiratoire acquis par les méthodes précédemment citées est utilisé pour décomposer les données acquises en TDM ou TEP en plusieurs phases, chacune correspondant à un instant du cycle. Ces informations sont utilisées pour rassembler les données acquises lors de chaque phase, reconstruites indépendamment. Ces reconstructions vont être utilisées pour estimer le champ de mouvement à l'aide de techniques de recalage.

5.3.1 Image TEP 4D

L'estimation du champ de mouvement respiratoire peut être faite à partir des données TEP, comme il a été montré par [Dawood et al., 2008, Dawood et al., 2006].

Les techniques utilisées sont analogues à celle présentée en 5.2.4. Elles consistent à réaliser une acquisition séquentielle des données en même temps qu'une acquisition du signal respiratoire, puis à réorganiser les données acquises pour reconstruire les différents instants du cycle indépendamment et sans correction d'atténuation. Le mouvement est ensuite estimé à partir de ces images.

Pour obtenir le signal respiratoire, Dawood [Dawood et al., 2008] utilise une caméra vidéo qui enregistre le mouvement d'un marqueur placé sur l'abdomen du

5.3. ESTIMATION DU CHAMP DE MOUVEMENT

patient. Ce marqueur est un point blanc placé sur un disque noir, et sa position axiale est détectée en calculant le barycentre des pixels dépassant un certain seuil.

La synchronisation avec l'acquisition PET est réalisée à l'aide d'une LED située dans le champ de vue de la caméra qui s'allume lors du début de l'acquisition. Une fois l'acquisition terminée, le signal respiratoire obtenu par la caméra est utilisé pour répartir les événements acquis par l'imageur TEP le long du cycle respiratoire. Ce cycle est décomposé en 8 parties qui seront reconstruites séparément, sans correction d'atténuation.

Les auteurs utilisent ensuite un algorithme de flux optique 3D présenté dans [Da-wood et al., 2006] et [Horn and Schunck, 1981] pour déterminer le champ de mouvement : Ils estiment un déplacement entre chaque image du cycle et l'image de référence correspondant à la quatrième phase (mi-expiration) dans le cas de cet article. Le résultat de l'algorithme est donc une séquence de champ de mouvement 3D, formant un champ de déformation 4D.

L'algorithme a été testé sur le fantôme XCAT [Segars and Tsui, 2009] ainsi que sur les données de 16 patients. La performance de l'estimation de mouvement a été évaluée selon trois critères sur les images corrigées : la correction du déplacement axial du coeur, le coefficient de corrélation des images corrigées du mouvement, ainsi que le bruit obtenu. Ces résultats sont détaillés dans la présentation de cet article de la partie 6.3.

Bai a présenté une technique d'estimation semblable [Bai and Brady, 2009]. L'étude a été réalisée sur des images simulées à l'aide du logiciel PET-SORTEO [Reilhac et al., 2005] en utilisant le modèle NCAT [Segars and Tsui, 2009], ainsi que sur une acquisition clinique. 24 lésions de forts contrastes (8 :1) ont été introduites dans les modèles, pour une durée d'examen simulé de 10 minutes. L'erreur moyenne du diamètre longitudinal des lésions est très fortement améliorée par la correction, avec une erreur passant de 36% à 16% en moyenne sur l'ensemble des lésions. Le principe est le même que pour la publication précédente, à la différence que l'estimation de mouvement a été réalisée sur des images corrigées de l'atténuation à partir de la carte d'atténuation de la phase de référence. De plus, cette estimation du mouvement a été réalisée à l'aide d'une interpolation par B-splines [Thévenaz and Unser, 2000].

Un exemple de champ de déformation obtenu sur l'acquisition clinique est présenté en 5.3

5.3.2 Image TDM 4D

Les images TDM peuvent être acquises en mode dynamique de manière à obtenir une suite d'images couvrant tout le cycle respiratoire [Lamare et al., 2007b, Qiao et al., 2006]. Les données sont reconstruites indépendamment, avec un rapport signal sur bruit plus faible que sur l'image originale.

Des algorithmes de recalage sont utilisés de manière à déduire le champ de mouvement. Le principal avantage de l'utilisation des images TDM 4D par rapport à la TEP est la précision des images. En effet, alors que les images TEP ont une résolution de

CHAPITRE 5. PROCESSUS D'ESTIMATION DU MOUVEMENT

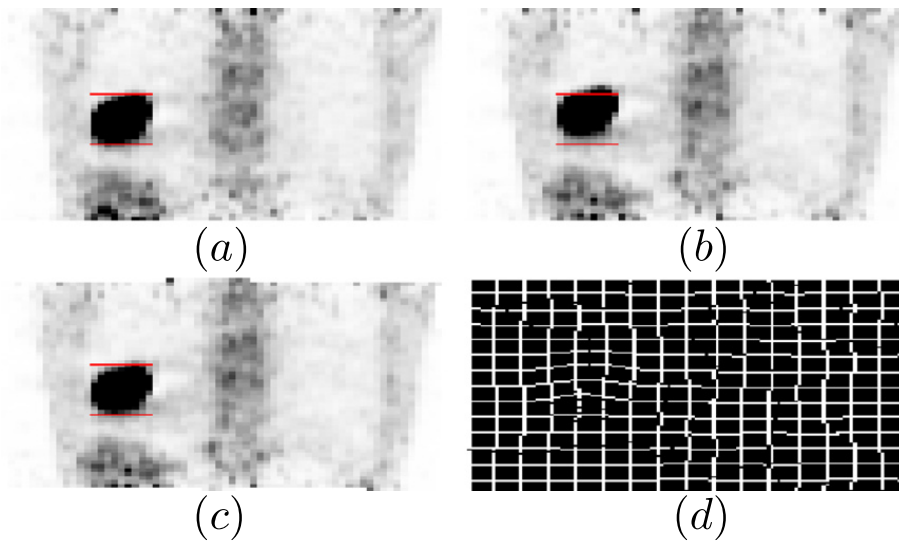


FIG. 5.3: Champ de déformation calculé sur des données cliniques à l'aide de la méthode présenté dans [Bai and Brady, 2009]. a) Image TEP obtenue à partir des données de la phase de référence (mi-expiration), b) Image reconstruite pour les données d'expiration complète, et c) Résultat du recalage de l'image de fin d'expiration sur l'image de référence. d) représente le champ de mouvement résultat. Les traits rouges représentent l'extension axiale de la tumeur dans l'image de référence.

l'ordre de 5mm, les images TDM atteignent des résolutions inférieures au mm, ce qui permet d'obtenir un champ de mouvement beaucoup plus précis. Cependant, utiliser des images TDM 4D a plusieurs inconvénients :

- Une possible incohérence entre le cycle acquis en TDM 4D et la respiration du patient en TEP. Bien que les imageurs TEP/TDM couplent les deux sur la même machine, de manière à éviter les mouvements du patient entre les acquisitions, le cycle acquis en TDM ne représente qu'un cycle, tandis que l'acquisition TEP va donner un "cycle moyen" plus proche de la réalité.
- Une dose de radiation émise plus importante. Bien que les technologies récentes permettent de réduire les doses de manière importante pour l'acquisition dynamique, elles restent tout de même plus importantes que pour un CT 3D.

Il faut noter que plusieurs publications basées sur des simulations se servent des cartes de labels utilisées par le simulateur pour réaliser les estimation de mouvement [Lamare et al., 2007b]. Cela donne une estimation dans le "meilleur des cas", où l'image TDM est parfaitement en phase avec les images TEP. On utilise alors la méthode décrite précédemment pour estimer le champ de mouvement. L'intérêt de cette méthode est que les données anatomiques sont moins bruitées que les données TEP et contiennent plus d'information pour l'estimation des déformations. Cependant, l'acquisition d'un scanner TDM 4D en routine clinique peut être considérée comme très irradiante.

5.3. ESTIMATION DU CHAMP DE MOUVEMENT

5.3.3 Modèle

Une autre voie en cours de développement est basée sur la création d'un modèle de respiration qui est adapté à chaque patient à partir d'une quantité réduite de données. Fayad [Fayad et al., 2010] propose un modèle basée sur l'analyse en composantes principales du champ de mouvement.

Ce modèle est adapté à un patient à partir de deux images TDM prises à des instants différents du cycle, et d'un maillage dynamique de la surface du corps du patient obtenu pendant un cycle respiratoire complet. Dans son implémentation, le maillage est obtenu à l'aide d'une caméra vidéo temps de vol.

L'avantage de ce modèle est qu'il est totalement continu, et permet l'extraction d'un nombre arbitraire de phases sous la forme de matrices de déformation. Ce modèle a été testé sur des images simulées (2 fantômes XCAT) et 6 patients.

L'article [Vandemeulebroucke et al., 2009] présente une autre technique adaptée à l'imagerie TDM. Dans ce cas, un modèle est généré à partir d'une acquisition TDM 4D en créant une image de mouvement moyenne, puis ce modèle est utilisé pour déduire le mouvement interne à partir d'acquisition de fluoroscopies en temps réel.

CHAPITRE 5. PROCESSUS D'ESTIMATION DU MOUVEMENT

Chapitre 6

Correction du mouvement respiratoire

6.1 Introduction

Nous allons maintenant présenter les techniques de correction du mouvement respiratoire proposées dans la littérature.

Deux approches existent pour la correction du mouvement : les techniques dites prospectives, qui consistent à réaliser la correction pendant l'acquisition en sélectionnant les données à conserver, et rétrospectives, qui réalisent la correction *a posteriori*, après l'acquisition des données. Actuellement, les techniques les plus prometteuses sont rétrospectives, car elles permettent d'utiliser l'ensemble des données du cycle respiratoire.

6.2 Synchronisation respiratoire

La synchronisation respiratoire correspond à un découpage du cycle respiratoire selon la phase (voir figure 6.1) ou l'amplitude (voir figure 6.2). Une seule des phases ou amplitude sera sélectionnée pour la reconstruction. En théorie cela permet d'avoir le meilleur résultat, car il est possible de sélectionner les événements correspondants à la phase ou l'amplitude où a été acquise la carte d'atténuation.

Cette technique est notamment présentée dans [Nehmeh et al., 2002], où le signal respiratoire est estimé par une caméra qui suit un marqueur placé sur le torse du patient (système RPM de Varian Medical Systems). L'étude a été réalisée sur 5 patients volontaires comme suit : un scan de transmission de 3 minutes, suivi d'une acquisition avec correction de 10 minutes, puis d'une acquisition témoin non corrigée de 3 minutes. La décomposition du cycle s'effectue en fonction de la phase. L'auteur annonce une amélioration de l'estimation du volume des tumeurs pouvant aller jusqu'à 34%, avec une augmentation du SUV_{max} de 160%.

Une autre publication [Boucher et al., 2004] utilise un thermomètre détectant l'air chaud émis en début de cycle respiratoire pour réaliser la synchronisation. Les

CHAPITRE 6. CORRECTION DU MOUVEMENT RESPIRATOIRE

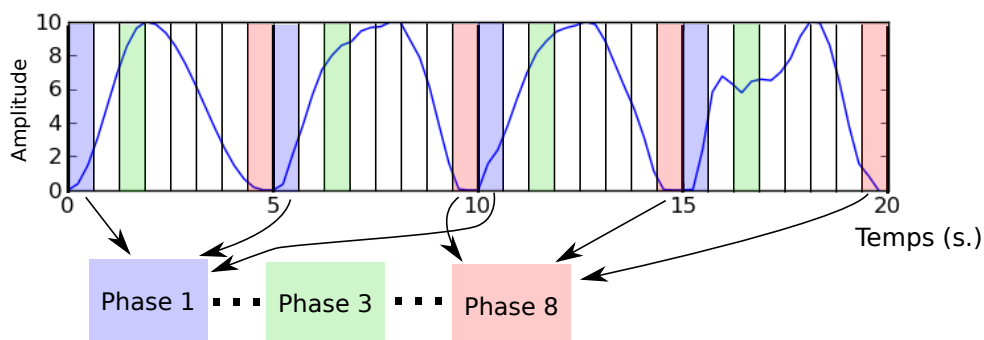


FIG. 6.1: Illustration de la synchronisation respiratoire en phase : Le cycle respiratoire acquis est découpé selon la position du signal acquis dans le cycle. Le signal est analysé pour déterminer les débuts et fins de cycles. Chaque cycle est découpé en un nombre déterminé de phases égales.

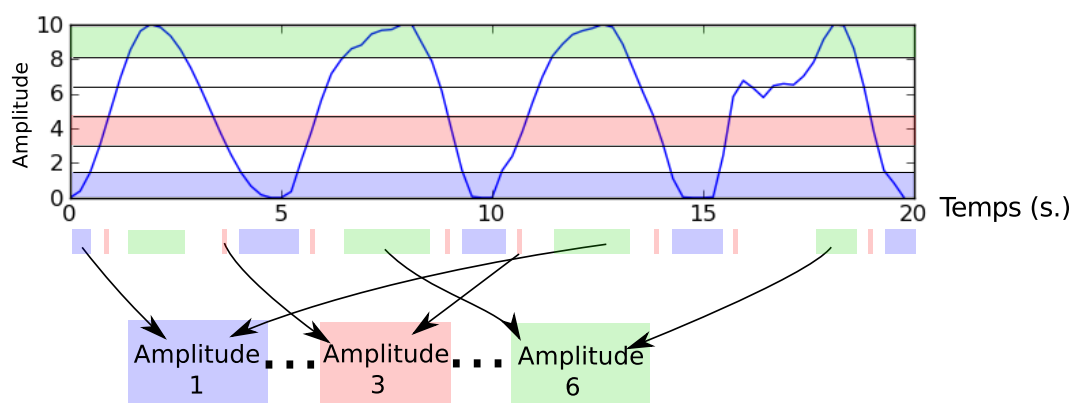


FIG. 6.2: Illustration de la synchronisation respiratoire en amplitude : Le cycle respiratoire acquis est découpé selon son amplitude.

différentes reconstructions issues de l'expérience sont visibles sur la figure 6.3. Il faut noter que la partie clinique de cette étude a été réalisée sur 10 patients sains, et qu'il n'y a donc pas de mesures de performance de la correction du mouvement sur les lésions.

Une variante de cette technique ne nécessitant pas de capteur est décrite dans [Nehmeh et al., 2003]. Un point faiblement radioactif est fixé au-dessus du torse du patient. Les acquisitions de l'imageur sont ensuite enregistrées par blocs temporels de 1 seconde, et une zone d'intérêt est reconstruite dans chacune des images. Les données reconstruites montrant le point source dans cette zone d'intérêt sont sommées et l'image finale reconstruite.

Cette technique a été comparée avec celle présentée précédemment [Nehmeh et al., 2002] basée sur le système RPM de Varian. Ces deux techniques n'ont été testées cliniquement que sur un patient mais ont montré des performances tout à fait semblables

6.2. SYNCHRONISATION RESPIRATOIRE

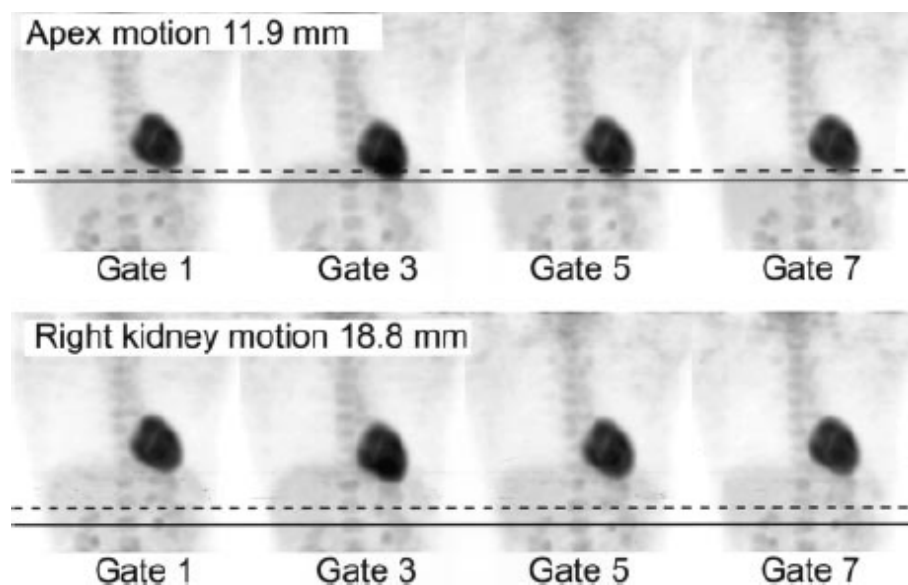


FIG. 6.3: Illustration de l'étendue du mouvement respiratoire sur des images reconstruites après synchronisation respiratoire [Boucher et al., 2004]. La rangée du haut montre l'étendue du mouvement de l'apex du coeur, et celle du bas l'étendue du mouvement du rein.

(6% de différence dans les activités et 2% pour le volume de la lésion).

Cependant, ces techniques n'utilisent pas forcément une carte d'atténuation optimisée pour la position du cycle correspondant aux acquisitions TEP. L'article [Boucher et al., 2004] utilise par exemple une carte d'atténuation obtenue par transmission, qui est une moyenne acquise sur plusieurs cycles respiratoires. Chang et al. [Chang et al., 2010] estiment la carte d'atténuation à partir d'une image TDM réalisée en respiration libre synchronisée. La phase du cycle respiratoire où l'acquisition TDM a été réalisée est enregistrée, et les données TEP acquises sur tous les cycles à cette phase sont reconstruites (voir exemples figure 6.4). Les résultats présentés sur 13 patients (21 tumeurs) montrent une amélioration du rapport signal sur bruit pouvant aller de -3.4 à 81% suivant les tumeurs, avec une amélioration moyenne de 26.3%.

L'inconvénient majeur de ces techniques est qu'elles demandent un temps d'acquisition beaucoup plus long à qualité d'image égale. Si l'on ne conserve que 20% des événements détectés, cela signifie qu'il faut augmenter le temps d'acquisition d'un facteur 5 pour obtenir une image de qualité égale. Il n'est donc pas envisageable de mettre en place ces protocoles en routine clinique, car le temps disponible n'est pas suffisant. C'est pour cela que de nombreuses équipes se sont mises à travailler sur une évolution de cette technique, où les images sont acquises en mode synchronisé puis recalées et sommées pour prendre en compte toute la statistique.

CHAPITRE 6. CORRECTION DU MOUVEMENT RESPIRATOIRE

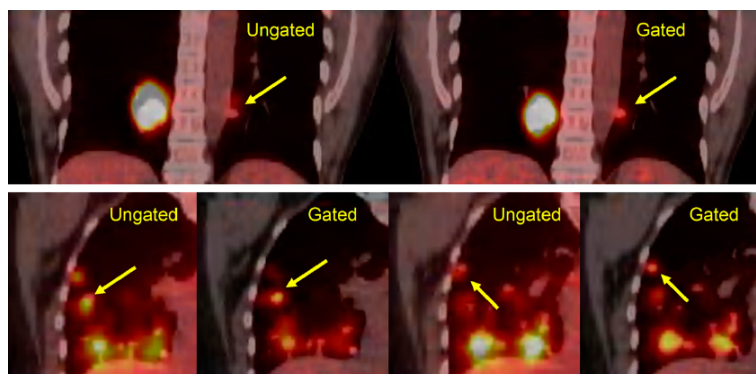


FIG. 6.4: Images TEP/TDM superposées du poumon reconstruites avec et sans synchronisation respiratoire en utilisant la méthode décrite dans [Chang et al., 2010]. On peut observer que les tumeurs sont mieux définies et correspondent à l'image TDM qui sert de référence.

6.3 Synchronisation respiratoire avec recalage

Dans cette section, nous allons détailler la technique consistant à corriger le mouvement en recalant les images reconstruites de chaque phase du cycle respiratoire.

Ces techniques se basent toutes sur une estimation préalable du mouvement respiratoire. Les images de chaque phase sont reconstruites indépendamment, puis recalées sur une phase de référence grâce au champ de mouvement. Enfin, les images déformées sont sommées. La difficulté principale se situe dans l'estimation du champ de mouvement interne lors de la respiration, car ce mouvement est complexe.

Les premières publications décrivant cette technique l'utilisaient notamment pour réaliser de l'imagerie cardiaque en TEP [Klein and Huesman, 1997]. Cette publication démontre la faisabilité du procédé sur un animal en utilisant des techniques de flux optique pour estimer le champ de mouvement. En effet le coeur a l'avantage d'avoir une activité métabolique intense, ce qui rend l'estimation de son mouvement relativement aisée même sur des images avec une faible statistique.

L'estimation du champ de mouvement interne se fait à l'aide d'une des méthodes présentées précédemment en 5.3 et rappelée brièvement :

Estimation du mouvement respiratoire basé sur l'Imagerie TEP

L'acquisition TEP est réalisée en même temps que l'acquisition du signal respiratoire. Une image est reconstruite par phase du signal respiratoire, puis un algorithme d'estimation de mouvement est utilisé pour calculer le champ de mouvement entre les instants du cycle.

Les premiers algorithmes étaient utilisés en imagerie cardiaque [Klein et al., 2001] avec des transformations simples (affines), puis d'autres algorithmes plus adaptés aux images corps entier ont été utilisés, comme les flux optiques [Dawood et al., 2006,

6.4. CORRECTION PRÉ-RECONSTRUCTION

Dawood et al., 2006], ou l'interpolation par B-spline [Bai and Brady, 2009].

Estimation du mouvement respiratoire basé sur l'Imagerie TDM 4D

Les images TDM 4D peuvent être utilisées pour réaliser l'estimation du mouvement respiratoire au lieu des données TEP. Cela nécessite un temps d'acquisition plus long et augmente la dose d'irradiation reçue par le patient. Dawood a réalisé plusieurs publications sur le sujet en utilisant le flux optique pour l'estimation du champ de mouvement [Dawood et al., 2006, Dawood et al., 2008]. L'algorithme a été étudié sur des images de patients réels. Une autre publication [Thorndyke et al., 2006] indique une amélioration du rapport de contraste sur bruit (CNR) d'un facteur 3 grâce à la correction.

6.4 Correction pré-reconstruction

Les méthodes de correction du mouvement pré-reconstruction modifient les positions des lignes de réponse (LDR) estimée par l'imageur TEP. Ce recalage des LDR correspond à un déplacement des lignes de réponse dans l'espace du détecteur (voir figure 6.5) en fonction du mouvement respiratoire. La limitation principale de ce type de méthode est que le champ de mouvement ne peut pas être élastique.

Cependant, il a été étudié en imagerie du cerveau [Bloomfield et al., 2003], où il permettait de corriger les mouvements de la tête. Il a été aussi utilisé en imagerie cardiaque TEP [Livieratos et al., 2005] en utilisant un champ de mouvement rigide (rotation suivie d'une translation).

Dans les deux cas, les résultats ont montré une nette amélioration des images (voir figure 6.6).

Dans le cadre du mouvement respiratoire du thorax, l'approche de recalage par LDR a été expérimentée notamment par Frédéric Lamare [Lamare et al., 2007a], mais avec des résultats mitigés. En effet, le champ était approximé par une transformation affine, qui peut difficilement modéliser le mouvement du thorax dans son ensemble. La figure 6.7 montre un profil d'image réalisé pour différents types de corrections. On peut voir que, bien que la correction pré-reconstruction estime correctement la position de la lésion, son activité est beaucoup plus faible que l'activité réelle.

En effet, [Lamare et al., 2007a] réalise une estimation du mouvement, en recalant les images TDM de chaque instant du cycle sur l'image de référence à l'aide d'une transformation affine par maximisation de l'information mutuelle normalisée. Les données sont des simulations réalisées à partir du logiciel GATE [Jan et al., 2004] et du fantôme NCAT. Des lésions de 7, 11, 15 et 21mm de diamètre ont été placées dans le poumon, avec un contraste de 8. Le champ de mouvement est calculé séparément pour le poumon, le coeur et trois organes sous le diaphragme (foie, estomac, rate).

Ces résultats ont été améliorés par l'utilisation de la technique suivante qui permet la prise en compte d'un mouvement élastique. Les performances comparées des deux

CHAPITRE 6. CORRECTION DU MOUVEMENT RESPIRATOIRE

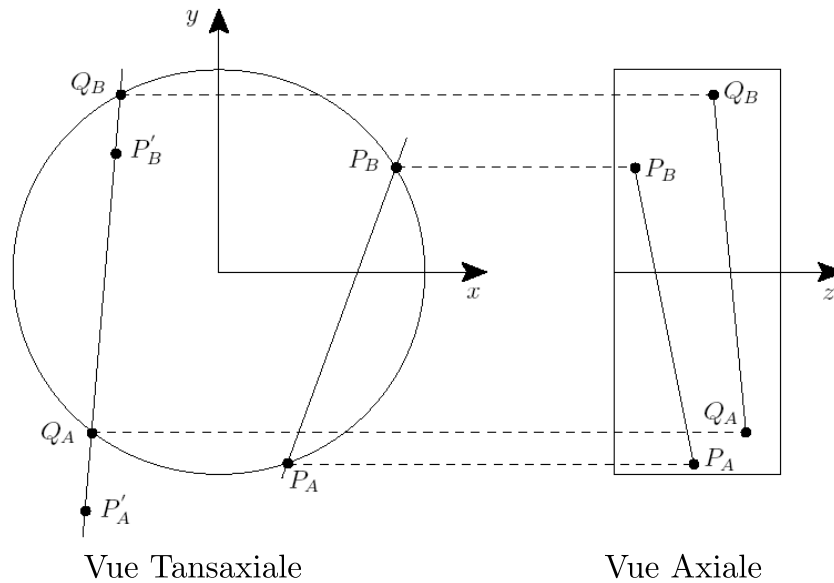


FIG. 6.5: Illustration du recalage des lignes de réponse dans l'espace du détecteur : P_A et P_B représentent les positions des détections, P'_A et P'_B les positions des points corrigés et Q_A et Q_B les détections correspondantes

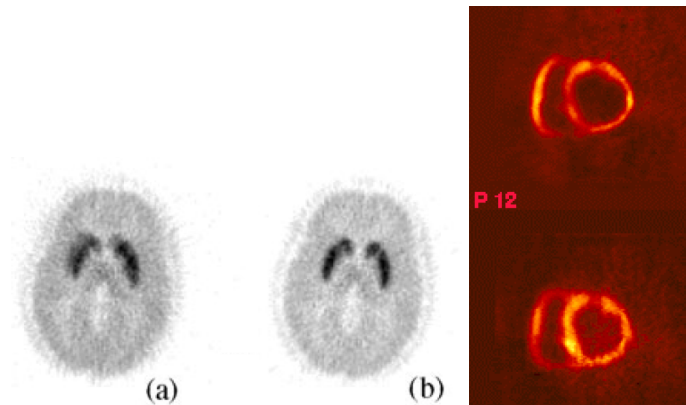


FIG. 6.6: Résultats de l'algorithme de recalage des LDR sur des images de patients utilisant le radio-traceur $[^{11}\text{C}]\text{raclopride}$. (a) montre une image non corrigée du mouvement et (b) une image corrigée. On peut noter que les structures internes du cerveau sont beaucoup mieux définies. (c) représente une coupe du cœur petit axe non corrigée (en haut) et corrigée (en bas). On peut voir une amélioration de la définition de l'image.

méthodes sont présentées dans la section suivante.

6.5. CORRECTION PENDANT LA RECONSTRUCTION

6.5 Correction pendant la reconstruction

Plusieurs auteurs ont présenté des méthodes permettant de réaliser la correction de mouvement pendant la reconstruction. Qiao et al. [Qiao et al., 2006] et Lamare et al. [Lamare et al., 2007b] ont proposé une méthode de correction du mouvement respiratoire basé sur une modification de la matrice de sensibilité lors de la reconstruction pour prendre en compte le mouvement. Tous les deux utilisent un champ de mouvement élastique estimé en utilisant un champ interpolé par B-splines.

L'algorithme original utilisé est basé sur OPL-EM [Reader et al., 2002] qui organise les données en "sous-ensemble" de la même manière que OS-EM [Hudson, 1994] mais en utilisant les informations séquentielles présentées en 2.6.1. Le principe de la reconstruction avec correction du mouvement respiratoire est décrit par la formule suivante :

$$f^{k+1} = \frac{f^k}{S} \sum_{t=1}^N P_t^T \frac{1}{P_t} f^k \quad (6.1)$$

f^k est l'image à l'itération k ,

T est l'opérateur de transposition

P_t représente la matrice système à l'instant t . Chaque élément p_{ij}^t de cette matrice indique la probabilité de détecter à la ligne de réponse i un évènement généré au voxel j au temps t . Cette équation est basée sur l'algorithme OPL-EM présenté en 3.1.3. Les informations du champ de mouvement sont contenues dans la matrice P_t . $P_t f$ correspond à la projection de l'image f à l'instant t .

N correspond au nombre d'instants temporels considérés.

S est la matrice de sensibilité :

$$S = \frac{1}{N} \sum_{t=1}^N P_t^T M A_t \quad (6.2)$$

A_t est la matrice permettant de corriger les effets de l'atténuation au temps t et M est la matrice de normalisation qui compense l'inhomogénéité spatiale de la sensibilité. A_t est estimée à partir des cartes d'atténuation de chaque instant t .

L'étude [Lamare et al., 2007b] a été réalisée sur des données simulées semblables à celles de l'étude précédente [Lamare et al., 2007a]. Il s'agit de simulations réalisées à l'aide du logiciel GATE, d'un cycle respiratoire du modèle NCAT, constitué de 8 instants temporels ainsi qu'une acquisition statique de référence de statistique équivalente à la somme des 8 acquisitions précédentes. Des lésions de 7, 11, 15 et 21 mm de diamètre et d'activité 8 fois plus importante que le fond ont été insérées dans le poumon.

Dans la publication, deux variantes de la technique de correction du mouvement respiratoire sont comparées avec la correction par synchronisation respiratoire avec

CHAPITRE 6. CORRECTION DU MOUVEMENT RESPIRATOIRE

recalage présentée ci-dessus ainsi que la correction pré-reconstruction précisé en 6.4. Des variantes de la méthode nommées “Reconstruction 1” et “Reconstruction 2” sur la figure correspondent à des différences d’interpolation du champ de mouvement dans la matrice système. Les résultats présentés montrent un clair avantage pour la correction pendant la reconstruction, avec des performances couramment améliorées d’un facteur 2.

Par exemple, le PRD sur la métrique contraste (équation 6.3) pour une lésion de 7mm présente dans la partie haute du poumon est de 28% pour les images non corrigées, contre 4.4% pour les images corrigées par la méthode de correction pré-reconstruction et de 1.2% pour la méthode de reconstruction pendant la reconstruction. De la même manière, pour des lésions de 7mm présentes dans le bas du poumon, les images non corrigées montrent une différence relative de contraste de 32%, contre 2.63% pour la correction pré-reconstruction et 1.66% pour la correction pendant la reconstruction.

$$\%Amélioration = \left| \frac{\text{Contraste sur image Évaluée} - \text{Contraste sur Référence}}{\text{Contraste sur Référence}} \right| \quad (6.3)$$

La figure 6.7 montre un profil de l’interface poumon/foie avec une tumeur pour les différentes techniques de correction. On voit clairement que l’image non corrigée montre un retard dû au flou de mouvement. Ce retard est partiellement corrigé par la correction de mouvement pré-reconstruction, mais le profil de courbe de la méthode de correction du mouvement pendant la reconstruction est celui qui s’approche le plus de la référence.

La publication [Thielemans et al., 2006] présente un système hybride de synchronisation respiratoire avec recalage et de correction de mouvement pendant la reconstruction. Ce système a été développé pour pouvoir isoler l’effet de la sommation des images recalées sur le résultat. Pour cela, les données de chaque instant temporel sont corrigées et reconstruites séparément à l’aide de la formule 6.1. Les N images sont ensuite directement sommées pour obtenir l’image corrigée. Les résultats de cette étude, dont la méthode est décrite au paragraphe 4.5.1, sont reportés sur la figure 6.8. Ils comparent les performances de détection estimées par un observateur numérique (présenté en 4.5.1) sur des images non corrigées du mouvement (symbole étoile), puis sur des image corrigées par les deux modèles de corrections présentés en 6.3 (symbole rond) et 6.5 (symbole hexagone) et enfin sur sa méthode hybride (symbole triangle). La figure montre que cette technique hybride améliore de manière sensible la détection par rapport aux images non corrigées mais reste inférieure aux deux autres techniques mentionnées pour la mesure de détectabilité, comme présenté dans la figure 6.8

6.6 Déconvolution de l’image

Cette technique décrite dans [El Naqa et al., 2006] utilise une connaissance du mouvement respiratoire acquise à partir d’une image TDM 4D pour déduire un filtrage ap-

6.6. DÉCONVOLUTION DE L'IMAGE

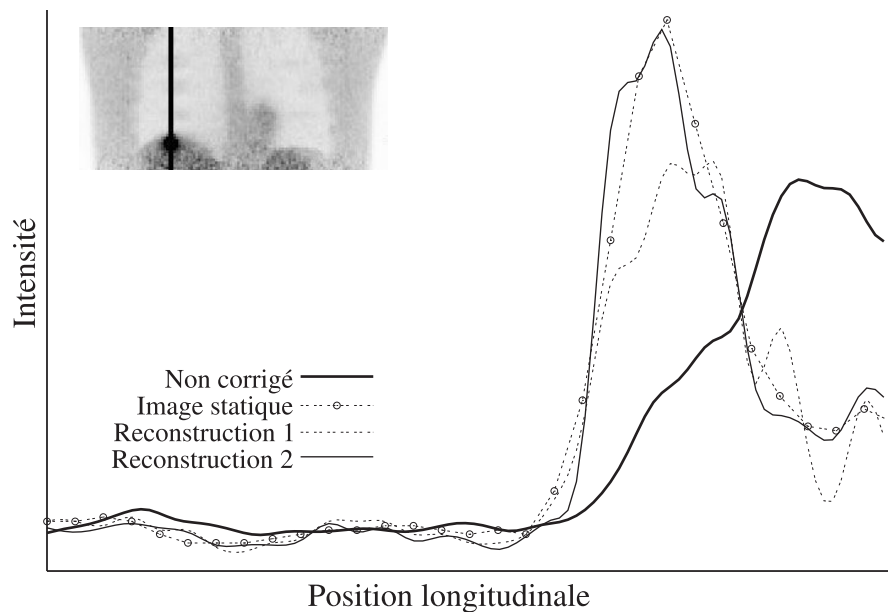


FIG. 6.7: comparaison des performances des différentes techniques de correction du mouvement sur un profil d'image TEP contenant une tumeur placée au niveau du diaphragme. La référence est l'image statique, "Reconstruction 1" correspond à la correction pré-reconstruction, et "Reconstruction 2" correspond à la correction pendant la reconstruction.

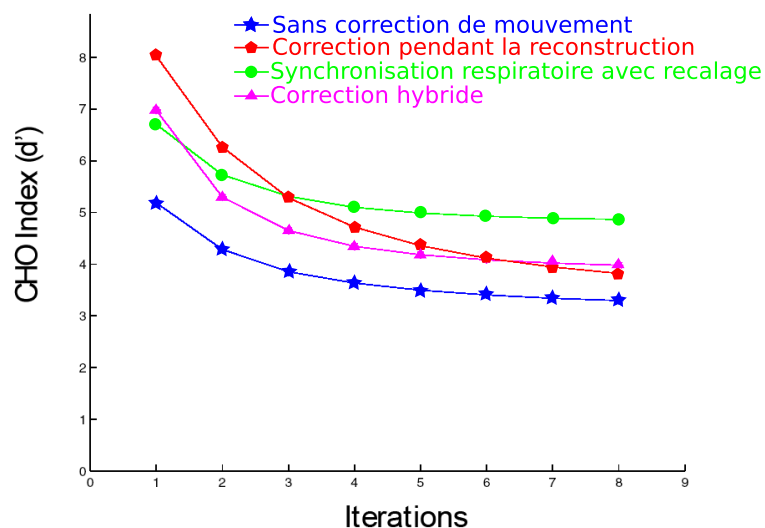


FIG. 6.8: Index de CHO pour différentes techniques de correction du mouvement respiratoire en fonction du nombre d'itérations de l'algorithme de reconstruction. Plus l'indice est important meilleure est la détectabilité. La méthode utilisant la correction post-reconstruction est comparée avec une méthode hybride présentée dans le document en 6.5.

pelé TLP (*Tumor Location Probability*) qui correspond à la dégradation due au mouvement respiratoire. Un exemple de filtre est représenté sous forme 3D dans la figure 6.9.

CHAPITRE 6. CORRECTION DU MOUVEMENT RESPIRATOIRE

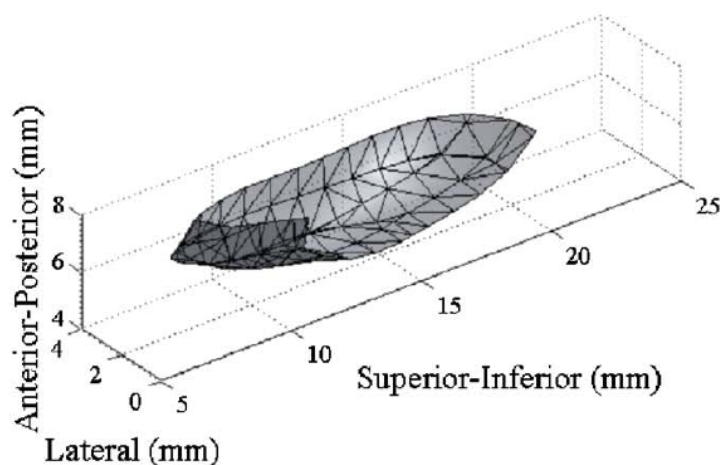


FIG. 6.9: Exemple de filtres TLP calculé dans l'article [El Naqa et al., 2006] pour la correction par déconvolution, rendu en 3D après tessellation, lissage et normalisation. Il représente la probabilité de présence du centre de la lésion.

L'image est déconvoluée pixel par pixel pour corriger les effets du mouvement respiratoire. Cette méthode a été évaluée sur un fantôme physique et des patients réels à l'aide d'un grand nombre de critères provenant pour partie de la TEP (sous-estimation de l'activité de la tumeur, exemples d'images), et pour partie du domaine de la déconvolution (entropie, "rugosité").

Les résultats présentés sur la figure 6.10 montrent une nette amélioration des performances sur des fantômes, pour un déplacement axial simple de 20 mm. L'activité des lésions de fort diamètre (> 20 mm) est correctement récupérée quelque soit l'algorithme utilisé, mais il n'a pratiquement pas d'effets sur les lésions de 1 cm de diamètre.

Ce type de méthode est peu présent dans la littérature par rapport aux autres présentés précédemment.

Un autre article utilisant aussi des algorithmes de déconvolution a été présenté par Wiemker [Wiemker et al., 2008]. Cependant il ne cherche pas à corriger le mouvement respiratoire sur l'ensemble de l'image mais principalement à améliorer la mesure du SUV sur une lésion. Pour cela, il réalise une estimation de la fonction d'étalement du point (FEP) de l'imageur TEP au niveau de la lésion à l'aide d'un contourage de la lésion réalisé préalablement sur une image TDM. L'estimation de la FEP permet de prendre en compte à la fois les effets du mouvement respiratoire et ceux de la FEP intrinsèque à l'imageur TEP.

Cependant, cette technique est inapplicable dans notre cas car les lésions doivent être de taille suffisamment importante et homogène pour pouvoir les délimiter de manière fiable sur les images TDM.

6.6. DÉCONVOLUTION DE L'IMAGE

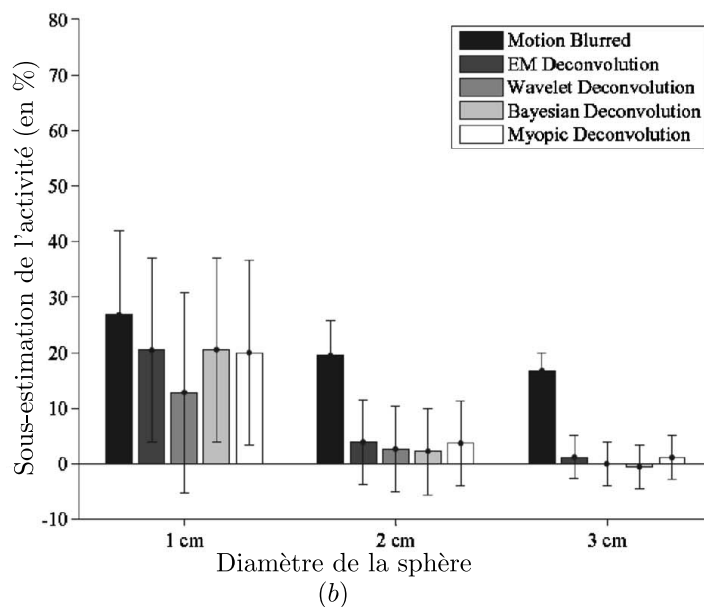
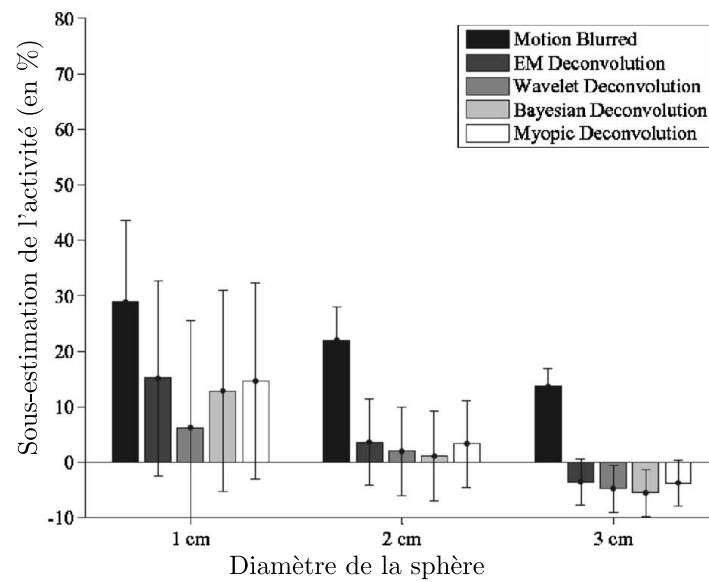


FIG. 6.10: Comparaison de l'erreur de sous-estimation de l'activité des lésions en fonction de l'algorithme de déconvolution utilisé sur des fantômes. En a) la lésion a une activité moyenne, tandis qu'en b) l'activité de la lésion est faible par rapport au fond. Le déplacement de la lésion est le même dans les deux cas (2cm).

CHAPITRE 6. CORRECTION DU MOUVEMENT RESPIRATOIRE

Troisième partie

Evaluation des performances de détection

Cette partie va présenter tout d'abord les techniques mises en place dans la littérature pour réaliser des comparaisons de performances entre plusieurs observateurs, qu'ils soient humains ou numériques. Puis nous présenterons les systèmes numériques de détection, avec notamment un état de l'art sur la détection des lésions en imagerie TEP.

Enfin, le dernier chapitre récapitulera les choix méthodologiques que nous avons effectués pour répondre à la problématique.

Chapitre 7

Performance des outils de détection

7.1 Généralités

En oncologie, la détection des sites tumoraux est une étape capitale dans la prise en charge des patients. Dans cette partie, je vais détailler les techniques qui permettent de comparer les performances de plusieurs observateurs (médecins ou algorithmes) face aux mêmes images, ou alors du même observateur face à plusieurs types d'images différentes.

Pour cette section, le problème va être simplifié au cas où un observateur doit classer un signal en "Sain" (normal, HO) ou "Pathologique" (anormal, H1).

Les performances de l'observateur sont indiquées par la matrice de confusion (table 7.1), qui recense les signaux correctement et incorrectement classés.

		Classe estimée	
		Sain	Pathologique
Classe réelle	Sain	VN (<i>Vrai Négatif</i>)	FP (<i>Faux Positif</i>)
	Pathologique	FN (<i>Faux Négatif</i>)	VP (<i>Vrai Positif</i>)

TAB. 7.1: Matrice de confusion : donne une vue d'ensemble des performances du classifieur. Elle indique le résultat de la classification de signaux connus.

On utilise habituellement deux grandeurs pour mesurer les performances de l'observateur :

La *sensibilité* (voir l'équation 7.1) correspond à la proportion d'images correctement évaluées pathologiques par l'observateur par rapport au nombre total d'images réellement pathologiques. Elle donne une information sur la capacité de l'observateur à détecter les cas pathologiques.

$$\text{Sensibilité} = \frac{VP}{VP + FN} \quad (7.1)$$

La *spécificité* (voir équation 7.2) représente le même type de grandeur, mais cette-

CHAPITRE 7. PERFORMANCE DES OUTILS DE DÉTECTION

fois ci appliquée aux cas non pathologiques : elle correspond à la capacité de l'observateur à donner un résultat négatif lorsque l'image est non pathologique.

$$\text{Spécificité} = \frac{VN}{VN + FP} \quad (7.2)$$

Ces deux grandeurs sont complémentaires mais ne permettent pas à elle seules de comparer les performances de différents observateurs. En effet, un utilisateur va souvent donner des notes, qui vont indiquer son niveau de confiance sur la présence de la pathologie (à ne pas confondre avec des notations sur la gravité des lésions, comme les techniques de gradation de [Genestie et al., 1998]).

Les techniques de comparaison de systèmes de décision comme l'analyse ROC (Receiver-Operating Curve) permettent de prendre en compte ces incertitudes. Elles proviennent à l'origine du domaine des télécommunications pendant la seconde guerre mondiale, où il fallait une métrique permettant de tester les performances des systèmes RADAR [Zou et al., 2007] pour la détection des avions ennemis. Les courbes ROC servent donc à évaluer la capacité d'un ou plusieurs "observateurs" à discriminer des signaux entre deux classes "normal" et "anormal". Les informations de sensibilité et de spécificité se limitent à comparer les performances pour un niveau de confiance donné.

7.1.1 Définition d'un Vrai positif / Faux positif

En général, les systèmes CAD, que nous présentons dans le chapitre 8 génèrent des cartes paramétriques, sur lesquels les voxels ou des groupements de voxels reçoivent une valeur correspondant à leur classe.

Afin de déterminer si l'élément (voxel ou groupement de voxels) noté "anormal" fait effectivement partie d'une lésion, sa position est généralement comparée à celle des lésions de la vérité terrain. Une distance d'acceptation est utilisée pour prendre en compte une éventuelle imprécision du système CAD. Par exemple, l'auteur de [Paik et al., 2004] considère que l'élément est un vrai positif si il est contenu dans le volume de la tumeur de la vérité terrain.

Cependant, il existe plusieurs stratégies pour compter les FP :

- Par voxel ou groupe de voxel (3D).
- Par coupe (2D).

Ces deux techniques vont très fortement influencer sur le nombre de faux positifs pour le même système CAD. En effet, dans le cas où le comptage des FP est réalisé sur l'ensemble de l'image, chaque groupement ou voxel ne correspondant pas à une lésion est noté FP. L'auteur de [Paik et al., 2004] indique par exemple 165 faux positifs par images TDM pour une sensibilité de 100%. De même, [Zhao, 2003] montre une sensibilité de 94% avec 906 faux positifs par image TDM pour la détection du cancer du poumon. Ces résultats montrent des nombre de faux positifs très importants, mais ne peuvent pas être comparés avec ceux utilisant un chiffrage des FP par coupe. En effet, dans ce cas, "FP" correspond au nombre de coupes contenant au moins un élément

7.2. MÉTHODOLOGIE ROC - RECEIVER-OPERATING CURVE

n'étant pas une tumeur. Le nombre de faux positifs est donc borné par le nombre de coupes, et ne fait pas de différences entre des coupes contenant un grand nombre de lésions et celles n'en contenant qu'une.

Dans le reste du document, nous considèrerons le premier système de comptage des FP : tout élément du volume à observer n'étant pas en contact avec une tumeur sera compté comme un faux positif.

7.2 Méthodologie ROC - Receiver-Operating Curve

Les courbes ROC [Swets and Pickett, 1982, Metz, 1986] sont des courbes indiquant la spécificité et la sensibilité de l'observateur (humain ou numérique) pour différents niveaux de certitude. Elles fournissent une mesure objective des performances d'un observateur dans une tâche de discrimination entre deux classes.

Elles peuvent être utilisées pour comparer les performances relatives de différents observateurs avec la même modalité d'imagerie ou pour déterminer les paramètres optimaux de ces images. L'évaluation d'un observateur par la méthode ROC nécessite la création d'un jeu de données labellisées en deux classes : Normale (H0) et Pathologiques (H1). L'observateur va se voir présenter l'ensemble des images et devra les noter individuellement selon un barème défini à l'avance (par exemple 0 : définitivement normal, 1 : potentiellement normal, 2 : équivoque, 3 : potentiellement pathologique, 4 : définitivement pathologique). Par convention, plus la note (notée λ) sera élevée, plus l'observateur va considérer qu'il est en présence d'un cas pathologique. A l'inverse, une note basse va indiquer un cas présumé sain ou normal.

L'observateur peut être un humain ou un algorithme, et les notes peuvent être discrètes ou continues.

Le tracé de la courbe ROC se fait en reportant la sensibilité et la valeur "1-spécificité" du classifieur pour différents seuils de décision. Par construction, la courbe va commencer au point (0,0) (tous les points sont diagnostiqués négatifs) et se terminer au point de coordonnée (1,1) (tous les points sont diagnostiqués positifs). Deux exemples de courbes ROC sont présentés sur la figure 7.1.

Pour analyser les courbes ROC, on considère que les distributions de probabilité des notes des cas H0 et H1 suivent une loi gaussienne (voir figure 7.2). Ce modèle de décision suppose que l'ensemble des valeurs de λ évaluées sur des cas H0 (sain) suit une distribution de probabilité $P(\lambda_0, \sigma_0)$ de valeur moyenne λ_0 et d'écart-type σ_0 . De même, les valeurs de λ évaluées sur des cas H1 (pathologiques), suivent une distribution de probabilité $P(\lambda_1, \sigma_1)$. Le mécanisme de décision se base sur le choix d'une valeur de seuil λ_s au-delà de laquelle les observations sont considérées comme pathologiques.

Ce seuil permet de modifier de manière dynamique la répartition des observations dans la matrice de confusion. Cela permet d'enrichir la comparaison des observateurs par rapport au couple (sensibilité/spécificité) seul.

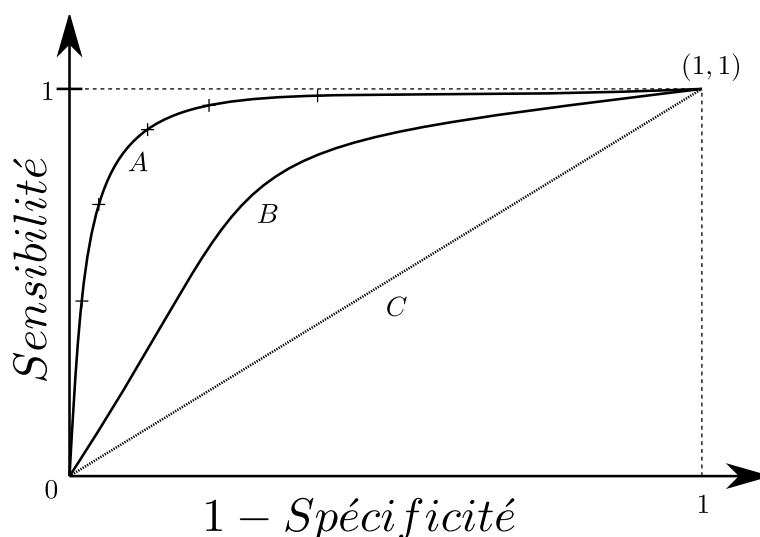


FIG. 7.1: Exemple de courbe ROC. La courbe A représente le résultat d'une évaluation ROC avec un barème à 6 niveaux. Pour chaque note du barème, le point correspondant est affiché en reportant la spécificité et la sensibilité (représentés par les croix). La courbe B représente une autre évaluation, avec des performances inférieures : Pour chaque niveau de "1-Spécificité", la sensibilité de la courbe B est inférieure à celle de la courbe A. La courbe C représente le d'un classifieur qui donne ses réponses de manière aléatoire.

Selon cette hypothèse gaussienne, on peut montrer que la distance d , appelée indice de détectabilité, correspond à l'aire sous la courbe ROC :

$$d = \frac{\lambda_1 - \lambda_0}{\sqrt{\sigma_1 + \sigma_0}} \quad (7.3)$$

Un ensemble d'indicateurs et de figures de mérite (FDM) permettent de comparer les performances de détection à partir des courbes ROC. La FDM la plus simple consiste à choisir un niveau de spécificité (noté α) et à comparer les sensibilités des différents observateurs. L'avantage de ce système est qu'il permet de comparer les performances dans des conditions proches de la réalité, où l'on cherche à rester dans un taux de spécificité donné. Cependant, les résultats vont dépendre du paramètre α . Une métrique plus globale est l'aire sous la courbe ROC. Étant donné que la courbe est nécessairement comprise dans un carré unitaire, la valeur de l'aire sera comprise entre 0 (le classifieur donne systématiquement les mauvaises réponses), 0.5 (le classifieur donne des réponses aléatoires) et 1 (le classifieur donne toujours la bonne réponse) [Nie et al., 2006].

Il est important, lors du calcul de la FDM, d'avoir une estimation de l'erreur. Il est possible de l'estimer en ajustant une courbe théorique (répondant à la loi théorique de la figure 7.2). Plusieurs logiciels ont été développés pour estimer les paramètres, qui ont été comparés dans la publication [Stephan et al., 2003] (AccuROC, Analyse-It, CMDT, GraphROC, MedCalc, mROC, ROCKIT, and SPSS).

7.2. MÉTHODOLOGIE ROC - RECEIVER-OPERATING CURVE

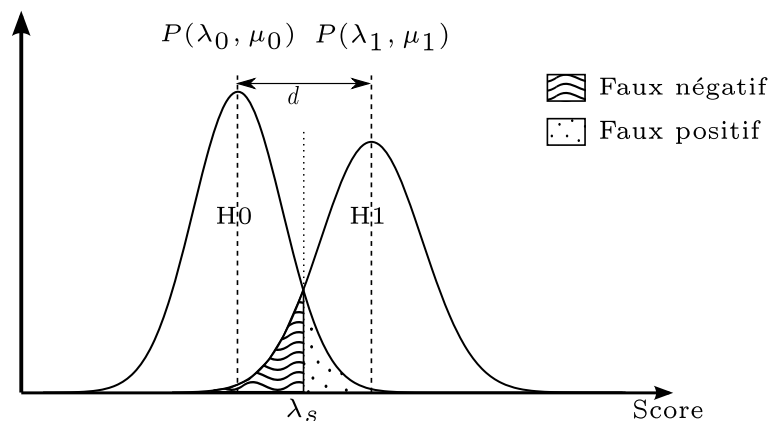


FIG. 7.2: Modèle de la distribution de probabilité de la variable de décision dans pour les populations H_0 ($P(\lambda_0, \mu_0)$) et H_1 ($P(\lambda_1, \mu_1)$) dans les études ROC. λ_s représente le seuil à partir duquel une observation sera catégorisée H_0 ou H_1 .

Une grandeur souvent utilisée dans la littérature pour évaluer la pertinence d'un résultat est la *p-valeur*. Elle représente la probabilité d'obtenir des différences entre deux courbes au moins aussi extrême que la différence observée (dans notre cas, les courbes ROC), en prenant en compte l'hypothèse selon laquelle le classifieur est aléatoire. Elle permet de vérifier si le test est statistiquement significatif.

Le problème de la méthodologie ROC est que l'observateur ne donne pas d'information de localisation de la pathologie dans l'image. On lui présente une image et il doit la noter sans indiquer la localisation de la zone suspecte. Dans notre cas, nous voulons comparer des observateurs qui réalisent une tâche de détection et de localisation de la zone suspecte. Il faut non seulement savoir si des lésions sont présentes, mais aussi avoir leur nombre et leur localisation. Cela est proche du travail en routine clinique, qui consiste à évaluer l'étendue et le nombre des lésions pour déterminer l'efficacité d'un traitement par exemple.

Pour éviter cette limitation, plusieurs extensions à la méthodologie ROC sont décrites dans la littérature : L-ROC, AF-ROC ou encore F-ROC. Les L-ROC sont décrites ci-après, tandis que les AF-ROC et F-ROC seront décrites dans la section suivante.

7.2.1 Courbes "Localization ROC" (L-ROC)

L'analyse L-ROC [Farquhar et al., 1999] ajoute l'information de localisation lors de la décision. L'observateur doit indiquer sur l'image qu'il considère comme pathologique la localisation de la lésion la plus probable. Elle est considérée comme un vrai positif si la distance entre la localisation indiquée et la localisation réelle de la lésion est inférieure à une certaine distance.

Cependant, bien que cette technique prenne en compte l'information de localisation, elle ne permet pas de traiter de manière satisfaisante les cas de lésions multiples.

7.3 Courbes Free ROC

Les courbes Free-ROC permettent de prendre en compte plus d'informations que les courbes vues précédemment car elles utilisent non seulement les informations de localisation, mais prennent aussi en compte le concept de faux positifs de manière plus approfondie, avec la possibilité d'avoir à la fois des faux positifs et des vrais positifs dans une même image.

7.3.1 Courbes Free-ROC

Les courbes F-ROC [Bunch et al., 1978] sont une généralisation des courbes ROC aux cas où l'on évalue la capacité de l'observateur à détecter un ensemble de lésions dans une série d'images. Chaque image pouvant contenir un nombre indéfini de lésions. L'observateur va donc devoir pointer sur l'image l'ensemble des sites suspects et y associer une note.

Dans ce cas, on ne peut pas utiliser le formalisme ROC car le terme de spécificité n'est pas directement calculable pour chaque niveau de confiance. On utilise à sa place le nombre moyen de faux positifs par image pour un seuil donné (voir figure 7.3).

On utilise les termes de LL (Lésion localisée) et NL (Non-Lésion) en lieu et place des informations de vrai positifs et faux positifs sur les courbes ROC. De la même manière, la sensibilité et la spécificité sont respectivement FLL (Fraction de lésions correctement détectées et localisées) et NFM (Nombre de faux positifs moyens par image).

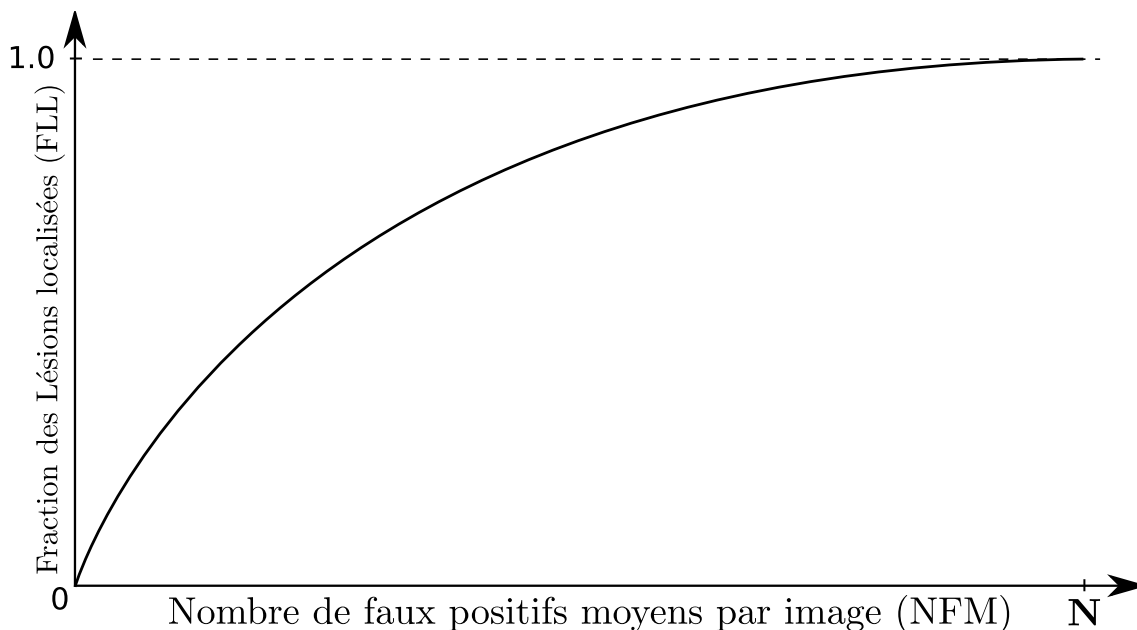


FIG. 7.3: Exemple de courbe Free ROC

Les courbes F-ROC n'ayant pas de bornes sur l'axe des abscisses, il est impossible de comparer plusieurs courbes à partir de l'aire sous la courbe. Il reste cependant

7.3. COURBES FREE ROC

possible de comparer la sensibilité pour un nombre de faux positifs donnés, mais on retrouve les mêmes problèmes que pour l'analyse ROC : il faut choisir un paramètre.

7.3.2 Courbes Alternative Free-ROC

Les courbes A-FROC [Chakraborty and Winter, 1990] sont des extensions des courbes Free-ROC présentées précédemment mais qui ne prennent en compte que le faux positif de plus haut score par image, ce qui ne pénalise pas le cas où un observateur indique un grand nombre de fausses détections (NL). On se retrouve alors dans la situation de courbes bornées sur l'axe des abscisses, ce qui permet leur comparaison.

7.3.3 Comparaison des courbes

Plusieurs techniques ont été développées pour permettre de réaliser des comparaisons de courbes dérivées de la méthodologie F-ROC. De la même manière que pour les courbes ROC, il est possible de comparer les courbes F-ROC en fonction de la FLL pour un nombre de faux positifs donnés. Cependant, étant donné que les courbes F-ROC n'ont pas de fin déterminée, il n'est pas possible d'utiliser l'aire sous la courbe.

JAFROC [Chakraborty and Berbaum, 2004] (JACKknife Free Receiver Operating Curve) est un algorithme et un logiciel développé par Chakraborty et se base sur une FDM non liée directement à la courbe. Cette mesure de performance utilise un algorithme dérivé des études A-FROC, ce qui signifie qu'il n'utilise pas l'ensemble des informations disponibles dans les courbes F-ROC mais se borne à comparer les scores des faux positifs de plus forte note pour chaque image avec les notes des vrais positifs. La figure de mérite θ_j proposée mesure la probabilité d'avoir un score de vrai positif supérieur à celui d'un faux positif (de n'importe quelle image).

Soit θ_j la valeur de la FDM, N_T le nombre total d'images, indexées par i , N_A le nombre total de cas pathologiques, indexées par j . n_j est le nombre total de lésions dans le cas anormal j .

$$\theta_j = \frac{1}{N_T N_A} \sum_{i=1}^{N_T} \sum_{j=1}^{N_A} \sum_{k=1}^{n_j} W_{jk} \psi(X_i, Y_{jk})$$

$$\psi(X, Y) = \begin{cases} 1.0 & \text{si } Y > X \\ 0.5 & \text{si } Y = X \\ 0.0 & \text{si } Y < X \end{cases} \quad (7.4)$$

$$\text{avec } \sum_{k=1}^{n_j} W_{jk} = 1$$

X_i le score du plus haut faux positif de l'image i , Y_{jk} est la note de la lésion k de l'image j . Si une lésion n'a pas été détectée, alors sa note sera par défaut de "0".

CHAPITRE 7. PERFORMANCE DES OUTILS DE DÉTECTION

Les poids W_{jk} correspondent à l'importance relative de détecter la lésion k dans l'image j pour le diagnostic. Pour chaque image, la somme des poids doit être égale à 1.

Une seconde version de JAFROC existe avec une puissance statistique plus importante, mais elle nécessite de disposer d'un grand nombre de cas non pathologiques. La formule est la même que la précédente (équation 7.4). La seule différence est que la première sommation se fait sur l'ensemble des cas non pathologiques N_N (équation 7.5) au lieu de l'ensemble des cas N_T .

$$\theta_J = \frac{1}{N_T N_A} \sum_{i=1}^{N_N} \sum_{j=1}^{N_A} \sum_{k=1}^{n_j} W_{jk} \psi(X_i, Y_{jk}) \quad (7.5)$$

Chapitre 8

Systemes de detection

8.1 Principe des CAD

8.1.1 Introduction

Les systemes CAD (Computer-Aided-Detection) sont des algorithmes permettant d'assister le praticien dans la detection des lesions ou le classement des images medicales. Dans le cadre de l'imagerie TEP oncologique, le besoin principal est celui du suivi therapeutique, pour lequel, il est important de detecter d'eventuelles lesions residuelles. Pour cela, il faut que le systeme CAD soit particulierement adapte a la recherche de petites lesions de faible contraste qui pourraient echapper au medecin. Cependant, le diagnostic, qui consiste a evaluer la dangerosite des lesions, et leur caractere pathologique est une tache plus complexe qui relve plus des systemes d'aide au diagnostic, qui ne seront pas traites ici.

Le developpement des systemes CAD a debuté dans les années 1980 [Chan et al., 1987], notamment pour detecter les micro-calcifications en mammographie. Bien qu'il existe plusieurs systemes CAD commerciaux pour l'imagerie TDM (xLNA pour Philips par exemple), aucun CAD commercial pour la TEP n'existe a notre connaissance.

Étant donné que les systemes d'aide a la detection dédiées a l'imagerie TEP restent limitées, le travail de cette these s'est largement inspire de la methodologie suivie lors du developpement de systemes CAD dans d'autres modalites d'imagerie medecale et particulierement la mammographie et l'imagerie TDM. La classification supervisee permet de profiter de la disponibilite *a priori* d'images prealablement classees pour fournir un 'superviseur' efficace et paramétrable. Le paragraphe suivant definit les differentes etapes standardisees d'un systeme CAD supervise.

8.1.2 Les differentes etapes d'un systeme CAD supervise

Comme rappele dans la publication [Li, 2007], un systeme CAD est constitue de deux etapes majeures : l'identification initiale des tumeurs et la reduction de faux po-

CHAPITRE 8. SYSTÈMES DE DÉTECTION

sitifs.

L'objectif de l'identification initiale des tumeurs (figure 8.1.1),2) et 3)) est de localiser les structures potentiellement suspectes de l'image avec la plus grande sensibilité possible au coût d'un grand nombre de fausses détections. Différents choix méthodologiques doivent être précisés pour cette étape :

- Les régions ou organes d'intérêt du système, lorsque celui-ci n'est pas dédié au corps entier, peuvent être séparés. L'image ou les régions d'intérêt de l'image peuvent ensuite être améliorées lors d'une étape de prétraitement, reposant par exemple sur du rehaussement du contraste ou du débruitage (figure 8.1.1)).
- Le processus d'extraction des tumeurs suspectes passe ensuite par l'utilisation d'un système de classification (figure 8.1.3)), optimisé par un apprentissage, s'appliquant sur un ensemble de caractéristiques de l'image ou de la structure cible (figure 8.1.2)). Ce dernier a pour but de déterminer la classe, « tumeur présente » ou « tumeur absente », de l'image ou de la structure cible.

L'objectif de l'étape de réduction de faux positifs (figure 8.1.4)) est de fortement réduire les fausses détections extraites de l'étape précédente, tout en conservant une sensibilité élevée. Cette étape rassemble un ensemble de méthodes permettant de trier les tumeurs candidates précédemment extraites afin de ne conserver que les candidats les plus 'suspects'.

8.2 Les CAD en TEP

Quelques systèmes CAD sur des images TEP ont été décrits dans des travaux académiques, tels que [Guan et al., 2006], qui recherche des anomalies dans le corps entier. Cependant, ils ne donnent pas de mesures de performances quantitatives. La publication de [Kanakatte et al., 2008] utilise des systèmes de classification supervisés (K plus proches voisins et SVM décrits en 8.5). Un seuillage est réalisé, puis les candidats sont classés en utilisant des caractéristiques fréquentielles et les moments statistiques.

Certaines publications utilisent seulement l'imagerie fonctionnelle TEP [Ying et al., 2004, Kanakatte et al., 2007, Kanakatte et al., 2008, Saradhi et al., 2009, El Naqa et al., 2009, Mhd Saeed et al., 2010], tandis que d'autres combinent cette information avec l'imagerie TDM [Jafar et al., 2006, Nie et al., 2006, Potesil et al., 2007].

La plupart des travaux réalisés dans le domaine ont portés sur la segmentation de lésions présélectionnées pour réaliser de l'estimation de volume pour de la planification de radiothérapie ou du suivi thérapeutique [Kanakatte et al., 2007, Potesil et al., 2007, Kanakatte et al., 2008, Yu et al., 2009]. Des publications récentes se sont focalisées sur la problématique du diagnostic pour la classification de nodules [Nie et al., 2006] ou la prédiction d'évolution des lésions suite à un traitement [El Naqa et al., 2009]. Cependant, très peu d'études se sont focalisées sur la problématique de la localisation des lésions [Guan et al., 2006, Kanakatte et al., 2007, Kanakatte et al., 2008, Mhd Saeed et al., 2010], surtout dans le cas de lésions de faible diamètre et faible contraste [Ying

8.3. EXTRACTION INITIALE DES LÉSIONS

et al., 2004, Jafar et al., 2006, Saradhi et al., 2009].

Les systèmes de classifications existants se concentrent en général sur la délimitation de lésions de fort contraste, ce qui ne correspond pas à notre problématique de détection.

8.3 Extraction initiale des lésions

8.3.1 Pré-traitements de l'image

Cette étape (figure 8.1.1) a pour objectif de préparer et améliorer l'image initiale à l'extraction des tumeurs candidates.

Cette étape n'est pas obligatoire et certaines études utilisent simplement l'image originale en imagerie TDM [Mougiakakou et al., 2007]. D'autres modifient l'image initiale, pour améliorer les performances finales de classification. En imagerie TDM, des méthodes de rehaussement de contraste sont mises en place à partir de filtres sélectifs [Li et al., 2003]. Certains prétraitements visent à éliminer des structures gênantes (vaisseaux sanguins dans les poumons...). En mammographie, il est également courant de débruiter les images par des techniques de filtrage [Bazzani et al., 2001]. Cependant, le filtrage doit être correctement optimisé pour les images à traiter, car il peut être à l'origine d'une baisse de sensibilité [Ge et al., 2005].

8.3.2 Extraction des caractéristiques

Un système de classification ne traite que très rarement directement les voxels de l'image brute. Différentes caractéristiques descriptives de l'image ou du cas clinique (patient) (figure 8.1.2) sont utilisées pour réaliser la classification. Celles-ci sont en effet porteuses d'informations plus pertinentes facilitant leur interprétation. Ces caractéristiques sont assimilées à l'analyse réalisée par l'oeil humain avant de réaliser le diagnostic sur la présence ou non de tumeurs. Elles sont l'interprétation de l'image d'un point de vue numérique. Elles doivent permettre au classifieur de discriminer les éléments pathologiques des éléments sains et être caractéristiques de chacune des classes.

Un état de l'art des différentes caractéristiques a été présenté par [Cheng et al., 2006], classées selon leurs catégories. On peut différencier plusieurs familles de caractéristiques selon leur type :

les régions d'intérêt de l'image : Que ce soit en imagerie anatomique ou fonctionnelle, les tumeurs apparaissent visuellement avec un degré de distinction variable par rapport au fond de l'organe d'appartenance. Dans le cas de la mammographie et de la TEP au 18F-FDG, les tumeurs sont de captation généralement plus élevée que le tissu environnant. L'imagerie TEP associée à d'autres radio-traceurs que le 18F-FDG peut aussi être caractérisée par un hyposignal dans les

CHAPITRE 8. SYSTÈMES DE DÉTECTION

tumeurs. Dans le cas de la TDM, les tumeurs correspondent à des nodules de forme spécifique (ellipsoïde) et d'atténuation plus élevée. Etant donné que les tumeurs peuvent, dans certains cas, être caractérisées par la valeur même des pixels ou voxels les constituant, certaines études utilisent directement les voxels de la région d'intérêt comme caractéristiques d'image.

les caractéristiques statistiques : Les caractéristiques statistiques reflètent généralement les propriétés de bruit et de texture des zones étudiées. Les statistiques du premier ordre caractérisent les niveaux de gris de l'image sans prendre en compte leur distribution relative. Des mesures de moyennes ou d'histogrammes peuvent être extraites. Les statistiques d'ordre deux représentent généralement des propriétés de texture. Celle-ci est assimilée à l'apparence locale de l'image, c'est à dire l'organisation des détails d'une petite partie de l'image telle que le système visuel humain l'aperçoit.

les caractéristiques cliniques : Ce sont les données contenues dans le dossier médical du patient qui peuvent être utilisées comme critère par le système de détection.

les caractéristiques fréquentielles : Selon [Sachs et al., 1971], le système visuel humain traite l'information à travers des canaux sélectifs en fréquence. Il est ainsi possible de se rapprocher d'une modélisation assez fiable du système visuel humain en se basant sur les caractéristiques fréquentielles des images cibles.

les caractéristiques géométriques : Les caractéristiques géométriques des tumeurs, exprimées par exemple par des critères de convexité, peuvent aussi être discriminantes. Elles sont généralement simples à interpréter car elles correspondent aux détails visuels sur lesquels se focalisent consciemment les cliniciens. Cette famille de caractéristiques est généralement dédiée aux imageries de type anatomique étant donné qu'elle repose sur une délimitation précise des structures étudiées. Le principal inconvénient des caractéristiques de forme est d'être fortement corrélé à la segmentation initiale des tumeurs étudiées.

8.4 Méthodes de classification

Dans cette section je vais décrire les deux principales classes de techniques de classification : supervisée et non supervisée.

8.4.1 Méthodes non supervisées

Dans le système de classification non supervisée, le classifieur reçoit directement l'ensemble des données à traiter, sans informations supplémentaires. Il devra de lui-même les classer par similitude en groupes. On utilise ce type de classifieur si on ne connaît pas *a priori* les classes, comme présenté dans la figure 8.2. Le nombre de classes peut être un des paramètres d'entrée notamment pour les k-moyennes, ou être déterminé de manière automatique.

8.5. MACHINES À VECTEUR DE SUPPORT (SVM)

La classification non supervisée repose sur une méthode statistique utilisant une fonction de proximité. Toutes les observations d'une même classe devront être proches au sens de cette fonction. Le partitionnement idéal est obtenu lorsque la covariance intraclasse est minimisée, ce qui implique que les classes sont le plus homogènes possible, et que la distance entre les classes est maximisée, pour que les classes soient le plus différenciées possible.

Plusieurs algorithmes sont décrits dans la littérature pour réaliser des classifications non supervisées :

K-moyennes (K-means) : c'est un algorithme itératif qui associe à chaque point la classe dont le barycentre est le plus proche, puis remet à jour le centre des classes à la prochaine itération [Herwig et al., 1999]. Cet algorithme nécessite de fixer le nombre de classes *a priori*.

Cartes auto-adaptatives : Cet algorithme utilise des techniques dérivées des graphes pour partitionner les données [Kohonen, 1982].

Regroupement ascendant hiérarchique : Cet algorithme utilise une mesure de dissimilarité pour regrouper les observations. A la première itération, chaque observation dispose de sa propre classe. A chaque itération suivante, les deux classes contenant les observations les plus proches vont être regroupées, jusqu'à obtenir la classification finale [Ward Jr, 1963].

8.4.2 Méthodes supervisées

La classification supervisée vise à étiqueter des observations à partir de connaissances acquises *a priori*.

Les classifieurs supervisés nécessitent une connaissance *a priori* des classes. On entraîne le classifieur en lui fournissant des *exemples* de cas avec l'étiquette associée. A partir de cette base de données d'entraînement, le classifieur va générer un *modèle* prédictif permettant de classer de futurs exemples non encore connus, comme présenté dans la figure 8.3.

8.5 Machines à vecteur de support (SVM)

La "Machine à Vecteur de Support", aussi appelée "Séparateur à Vaste Marge", ou "Support Vector Machine" en Anglais, est un classifieur qui comme son nom l'indique vise à maximiser la marge [Boser et al., 1992], qui est la distance minimale entre les points des données et la surface séparatrice (voir figure 8.4).

Les SVM sont des classifieurs supervisés qui cherchent à maximiser la séparation entre chaque classe. L'idée sous-jacente aux SVM est de trouver l'hyperplan optimal de séparation, et non pas n'importe quel hyperplan qui permettrait de séparer correctement les données d'apprentissage (figure 8.5).

CHAPITRE 8. SYSTÈMES DE DÉTECTION

La définition de la marge et de l'hyperplan se fait à partir des vecteurs de support. Ils correspondent aux cas extrêmes des deux classes qui pourraient éventuellement poser des problèmes de classification.

Soit $\{(x_i, y_i) | x_i \in \mathbb{R}^p, y_i \in \{-1, 1\}\}$ un ensemble de points étiquetés correspondant à la base d'apprentissage. x_i correspond au vecteur caractéristique, tandis que y_i correspond à la classe du point.

Tout hyperplan dans l'espace des caractéristiques \mathbb{R}^p peut être écrit $w \cdot x - b = 0$, avec w la normale à l'hyperplan et $\frac{b}{\|w\|}$ le décalage de l'hyperplan par rapport à l'origine.

La contrainte principale sur w et b est de classer correctement les données. Pour cela, les points vecteur de support devront respecter la contrainte $y_i(w \cdot x_i - b) \geq 1$. Par construction, les vecteurs de support correspondent aux points où l'égalité stricte est observée : $y_i(w \cdot x_i - b) = 1$, comme indiqué sur la figure 8.4.

La seconde contrainte est la maximisation de la marge, qui correspond à $\frac{2}{\|w\|}$, ce qui revient à minimiser $\|w\|$ comme indiqué sur la figure 8.5.

La solution de ce problème d'optimisation est apportée par l'algorithme des multiplicateurs de Lagrange.

Cependant, il peut arriver que les classes ne soient pas séparables. Ce problème est résolu par l'autorisation d'une erreur [Cortes and Vapnik, 1995] ξ_i que l'on cherchera à minimiser : $(w \cdot x_i - b) \geq 1 - \xi_i$. Cela engendre l'ajout d'un paramètre noté C , qui sera utilisé dans le lagrangien pour pénaliser plus ou moins les erreurs.

Dans le cas où l'hyperplan de séparation ne pourrait pas être défini par une équation linéaire, des fonctions à noyaux sont utilisées pour projeter les points dans un espace de plus grande dimension où le problème devient linéaire (voir figure 8.6). Cependant, le passage des points dans l'espace de dimensions supérieure n'est jamais réalisé explicitement, car en pratique, le changement de dimension est utilisé uniquement lors de la comparaison entre deux points. On utilise donc des fonctions de comparaison modifiées, appelées noyaux, qui réalisent cette transformation de manière implicite. Le noyau le plus couramment utilisé est le RBF (fonctions à base radiale) qui utilise un espace d'arrivée de dimensions infinies, ce qui est rendu possible par le fait que l'on ne réalise par la transformation directement, mais qu'elle est réalisée implicitement lors de la comparaison des points x_i et x_j à l'aide de l'équation suivante :

$$k(x_i, x_j) = e^{-\gamma \|x_i - x_j\|^2} \quad (8.1)$$

Grâce à la maximisation de la marge, les SVM sont performants dans le cas où l'on ne dispose que de faibles quantités de données. D'autres algorithmes tels que les forêts aléatoires parviennent à obtenir des performances similaires [Statnikov et al., 2008].

8.5.1 Autres classifieurs supervisés

De nombreux classifieurs supervisés ont été développés :

8.5. MACHINES À VECTEUR DE SUPPORT (SVM)

Le classifieur LDA [Fisher, 1936] approxime les deux classes en supposant qu'elles ont une distribution gaussienne de covariance égale. Seule la moyenne de chaque distribution varie, et tout nouveau point est classé en fonction de sa distance à l'une ou l'autre des classes. Bien qu'il soit très rapide, ce classifieur ne fonctionne que dans le cas linéaire. L'approximation gaussienne donne des résultats suffisamment bons dans de nombreux cas, ce qui explique son utilisation pour la détection des lésions en mammographie [Baydush et al., 2007] ou en oncologie TDM [Gurcan et al., 2002] par exemple.

Les réseaux de neurones [Haykin, 1999] sont des classifieurs non linéaire qui se basent sur un fonctionnement simplifié des neurones du cerveau humain. Ils sont constitués de neurones artificiels qui miment les réactions des neurones biologiques. Un neurone applique une fonction à la combinaison linéaire de ses stimuli d'entrée. Chaque élément du vecteur de caractéristique est entré dans un neurone, qui ira stimuler un ou plusieurs autres neurones. L'information se propage jusqu'à un neurone terminal qui donnera le résultat. Les réseaux de neurones sont très performants, mais souffrent de difficultés de mise en place à cause de leur grand nombre de paramètres et de temps d'apprentissage très importants.

Une forêt de décision est un classifieur relativement récent basé sur plusieurs arbres de décisions initialisés différemment [Ho, 1998]. Le résultat de ce classifieur est le mode des résultats de tous les arbres. La combinaison de plusieurs arbres différents permet d'améliorer la stabilité du classifieur, qui est un gros désavantage des arbres de décision pris indépendamment. Ce classifieur est très rapide et très performant.

CHAPITRE 8. SYSTÈMES DE DÉTECTION

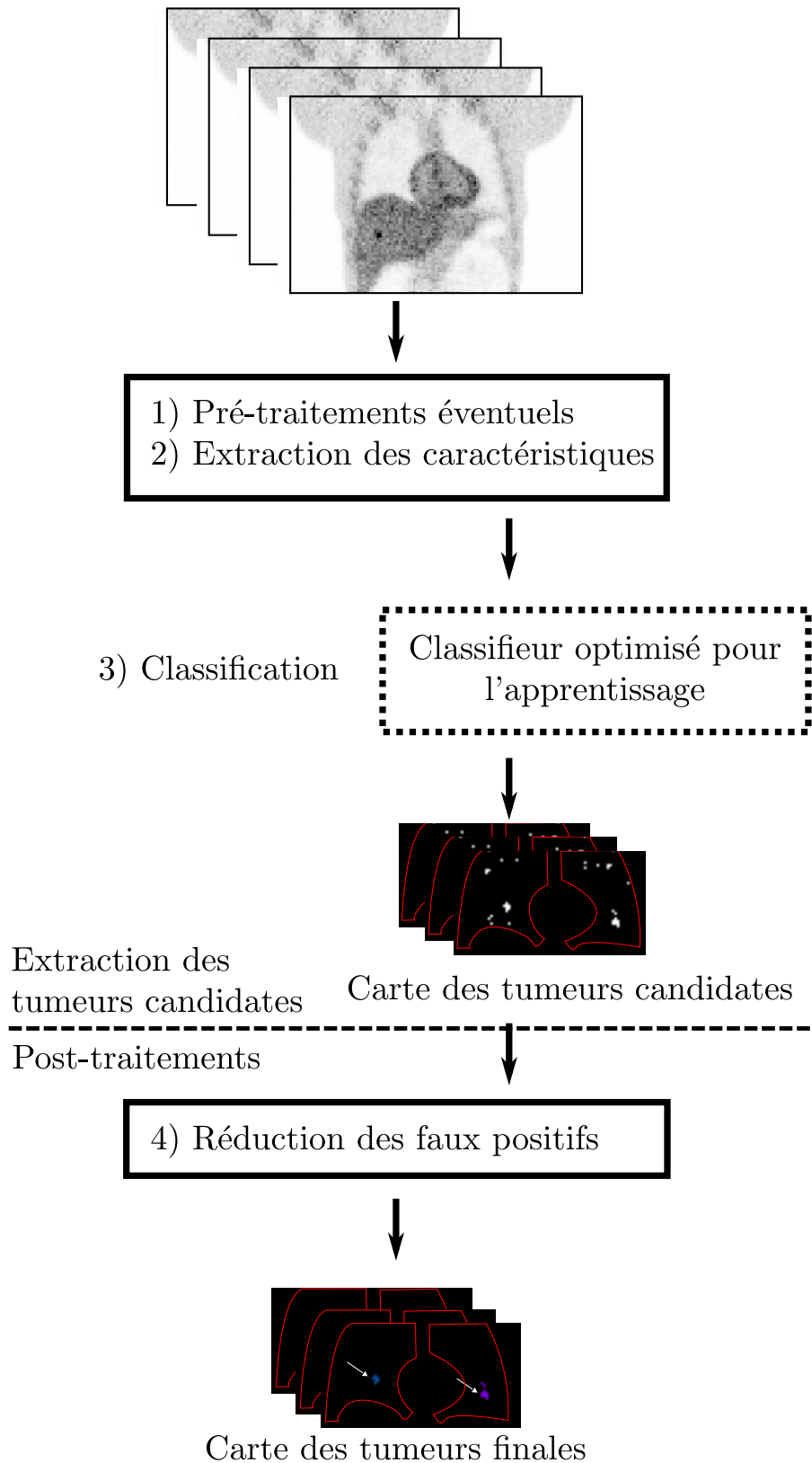


FIG. 8.1: Les différentes étapes d'un système CAD.

8.5. MACHINES À VECTEUR DE SUPPORT (SVM)

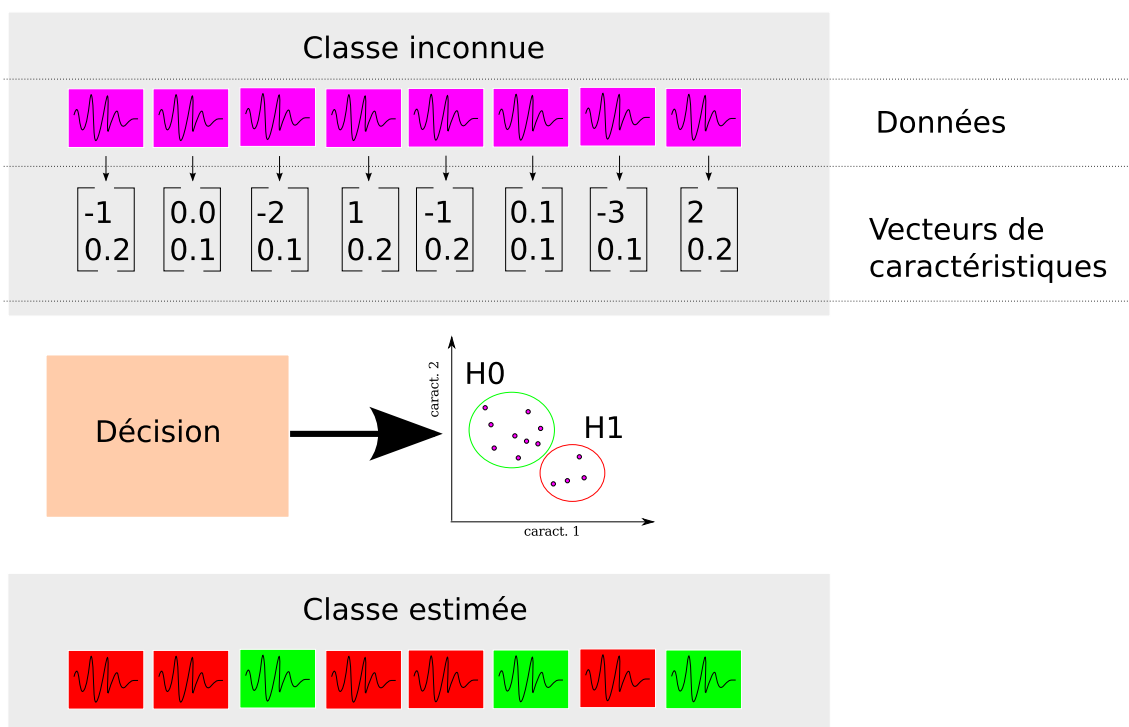


FIG. 8.2: Fonctionnement d'un classifieur non supervisé : Les données brutes sont envoyées au classifieur qui va les regrouper en classes en fonction de leur répartition dans l'espace des caractéristiques.

CHAPITRE 8. SYSTÈMES DE DÉTECTION

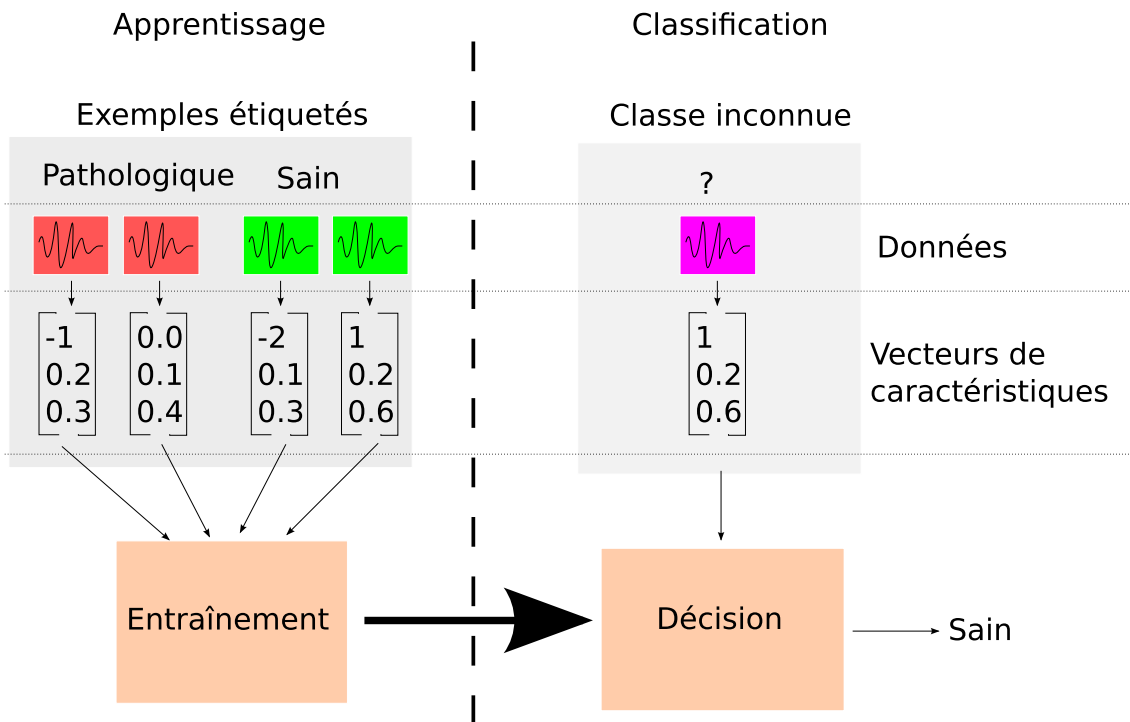


FIG. 8.3: Fonctionnement d'un classifieur supervisé : Les données d'apprentissage servent à entraîner le classifieur pour générer un modèle. Ce modèle permettra de rattacher des observations aux classes apprises.

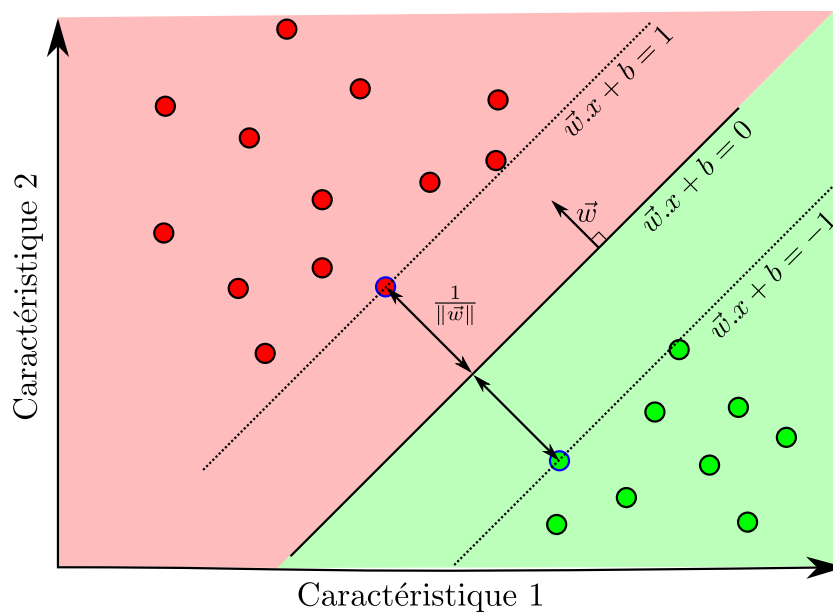


FIG. 8.4: Machine à Vecteur de Support : les points vecteur de support (entourés de bleu) sont les seuls utilisés pour calculer la surface de séparation d'équation $w \cdot x + b = 0$. Le vecteur w est normal à la surface de séparation et permet de calculer la marge $\frac{1}{\|w\|}$.

8.5. MACHINES À VECTEUR DE SUPPORT (SVM)

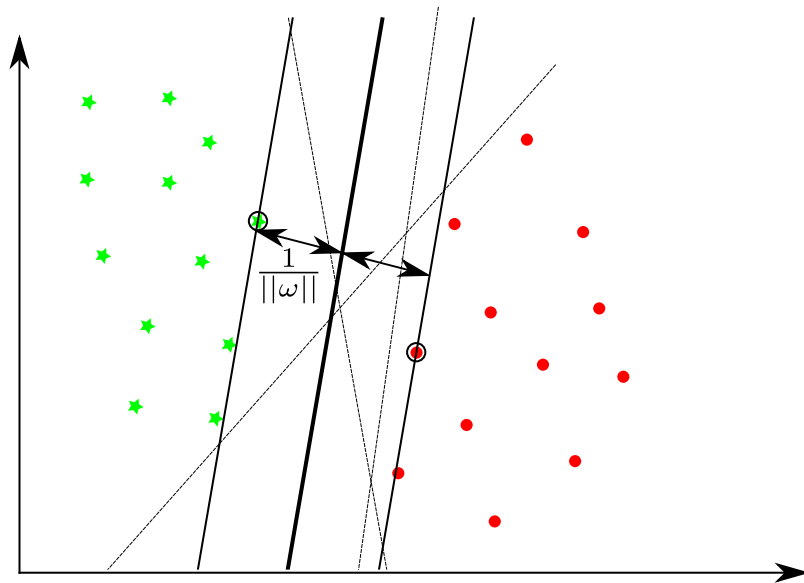


FIG. 8.5: Illustration du principe d'optimisation de la marge sur un cas à deux classes et deux caractéristiques : Il existe une infinité d'hyperplans qui séparent correctement les classes "étoile vertes" et "ronds rouges", représentés par les traits fins en pointillé. Cependant, il existe seulement une seule solution qui maximise la marge. Cette solution est représentée en trait plein, avec la marge $\frac{1}{\|w\|}$. Les points entourés en noir sont les points support utilisés pour contraindre l'hyperplan.

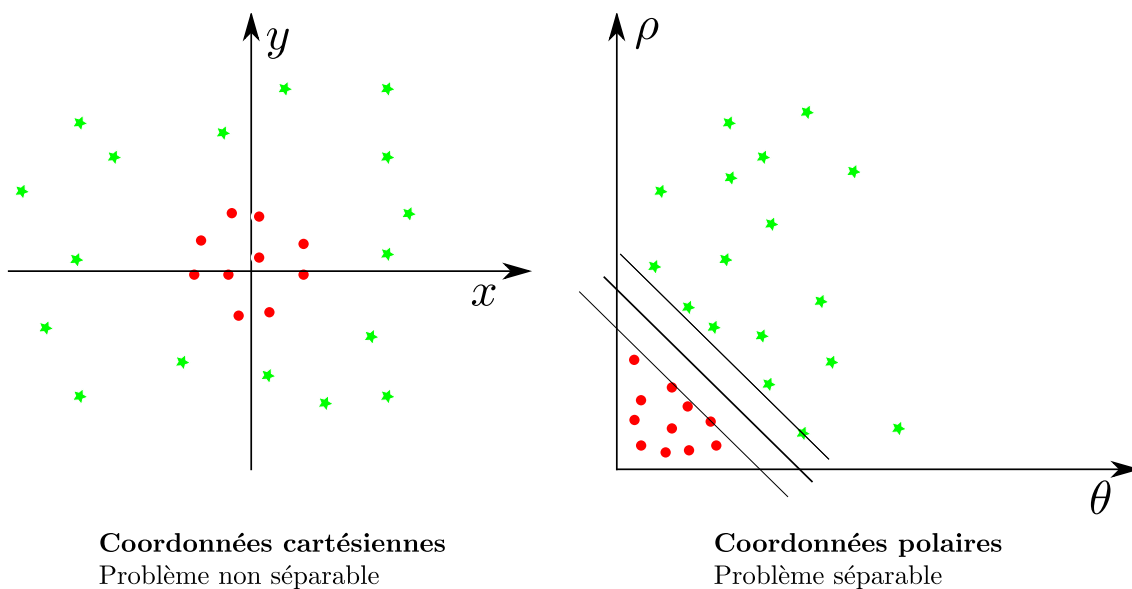


FIG. 8.6: Illustration du principe du changement de base : Des données non séparables dans le repère original peuvent devenir séparables en utilisant les noyaux. Dans ce cas, le passage des coordonnées cartésiennes aux coordonnées polaires permet de rendre le problème original linéairement séparable.

CHAPITRE 8. SYSTÈMES DE DÉTECTION

Chapitre 9

Choix méthodologiques

La problématique de ce travail de thèse relève du domaine de l'évaluation en imagerie médicale. Cette discipline nécessite :

- De bien définir la tâche.
- De choisir les bonnes métriques.
- De disposer d'une référence (ou vérité terrain).

Pour nous, la tâche diagnostique est d'améliorer la détection. Dans ce court chapitre, nous justifierons les choix méthodologiques que nous avons dû faire, d'une part pour la construction de la vérité terrain et d'autre part pour définir la méthode d'évaluation des performances de détection.

9.1 Construction de la vérité terrain

Le choix des types de données à utiliser pour une étude relevant du domaine de l'évaluation est important. Pour notre problème, il est important de disposer de données TEP oncologiques dynamiques et synchronisées à la respiration contenant des lésions parfaitement annotées. La collection d'images cliniques annotées est un travail de longue haleine nécessitant le concours de différentes personnes, et principalement des cliniciens.

Quelques bases de cas cliniques sont disponibles en imagerie TEP oncologiques. Celles-ci ne sont cependant pas distribuées à la communauté scientifique, sont rarement annotées et ne présentent pas forcément assez d'images de chaque classe. La littérature des CAD en imagerie TEP (section 8.2) confirme ce manque de bases d'images cliniques distribuées car les auteurs utilisent généralement quelques cas fournis par les hôpitaux partenaires de leur étude.

Une alternative à l'utilisation d'images cliniques repose sur la simulation d'exams. La quatrième partie de ce manuscrit s'attache à décrire la méthode utilisée pour créer notre jeu de données TEP 3D simulées avec la respiration, puis à décrire les caractéristiques de cette base de données.

CHAPITRE 9. CHOIX MÉTHODOLOGIQUES

Nous avons cherché à générer une base de données réaliste prenant en compte la variabilité rencontrée dans les acquisitions réelles. Pour cela, nous avons utilisé un modèle de patient déformable incorporant un modèle de respiration pour générer les images. Cela permet d'incorporer une variabilité importante dans les formes des organes patients de la base de données. Un mouvement respiratoire irrégulier a été utilisé pour générer les données, pour prendre en compte les imprécisions des systèmes d'acquisition du signal respiratoire.

Les contrastes des lésions ont été calibrés pour obtenir des taux de détection par un observateur humain de 10% à 90%. Les activités des organes sont calculées à partir de zones d'intérêts extraites de 70 images TEP. La simulation a été réalisée avec le logiciel PET-SORTEO, utilisant des algorithmes Monte-Carlo accélérés. Nous avons simulé deux jeux de données : le premier correspond à une acquisition dynamique d'un patient respirant, avec 4 cycles respiratoires différents. Les cycles ont été coupés en 8 instants temporels simulés séparément pour représenter une acquisition synchronisée. Le second jeu de données correspond à une acquisition statique de même durée que le précédent (224 secondes), mais sans mouvement respiratoire. Il va être utilisé comme "témoin", correspondant à une hypothétique correction parfaite du mouvement respiratoire.

La quatrième partie de ce manuscrit s'attarde à décrire la méthode utilisée pour créer notre jeu de données TEP 3D simulées avec la respiration, puis à décrire les caractéristiques de cette base de données.

9.2 Mesure des performances de détection

Nous avons montré dans le chapitre 7 que les analyses ROC ou F-ROC sont les méthodes de référence pour évaluer les performances de détection. Ces études sont néanmoins lourdes à mettre en oeuvre lorsque l'on cherche à optimiser des paramètres et requièrent la participation active d'observateurs, souvent difficiles à mobiliser lorsque les tâches de lecture sont longues et répétitives.

Nous avons donc cherché à développer une stratégie qui permette de respecter les conditions de l'analyse F-ROC tout en s'affranchissant d'observateurs humains.

Le chapitre 8 a présenté le principe des systèmes d'aide à la détection, développés pour fournir une seconde lecture au médecin lors de son diagnostic. Nous avons proposé d'utiliser ces systèmes CAD comme observateurs numériques.

Les systèmes CAD permettent de surmonter les inconvénients des observateurs humains lors des évaluations de méthodes. Ils permettent de limiter la variabilité inter-observateurs en permettant de réaliser l'ensemble des détections avec un seul observateur, et garantissent que toutes les observations sont réalisées dans les mêmes conditions. Il est aussi possible de reproduire à l'identique les conditions d'évaluation pour traiter de nouvelles données.

Nous avons développé une méthode d'évaluation des performances de correction basée sur l'utilisation d'un système de détection des lésions automatisé. Les perfor-

9.2. MESURE DES PERFORMANCES DE DÉTECTION

mances de ce système de détection sur les lésions présentes dans les modèles vont définir la qualité et la pertinence de la correction. Les différentes étapes de cette méthode d'évaluation sont présentées dans le chapitre 12

Nous utilisons un classifieur basé sur les Machines à Vecteurs de Support (SVM) utilisant les coefficients d'ondelettes non décimées comme caractéristiques fréquentielles. La classification est réalisée sur chaque voxel des images, puis une étape de réduction des faux positifs est appliquée, en agrégeant les voxels classés positifs, et en supprimant les agrégats de trop petite taille.

Les performances sont mesurées par l'analyse psychophysique des courbes free-receiver operating characteristics (FROC) que nous avons adaptée aux spécificités de l'imagerie TEP. L'évaluation des performances est réalisée sur deux méthodes prometteuses de correction du mouvement respiratoire, en les comparant avec les performances obtenues sur les images non corrigées ainsi que sur les images idéales sans mouvement respiratoire.

CHAPITRE 9. CHOIX MÉTHODOLOGIQUES

Quatrième partie

Construction d'une base d'images TEP simulées respirante

Chapitre 10

Simulations

Simuler des images réalistes est un enjeu important en imagerie médicale. Le fait d'avoir accès à la vérité terrain, contrairement à la routine clinique, permet de valider des hypothèses de manière beaucoup plus fine. De plus, avec l'augmentation des performances des ordinateurs et l'accès à des centres de calculs, il devient possible de créer des bases de données de patients de plus en plus importantes et réalistes.

Nous allons tout d'abord présenter les différents types d'algorithme de simulations disponibles. Ensuite, nous ferons un état de l'art des simulateurs développés par la communauté. Enfin, nous décrirons le simulateur retenu ainsi que les contributions réalisées pour adapter ce simulateur à nos besoins.

10.1 Principe des simulations

10.1.1 Simulations analytiques

Les simulateurs analytiques utilisent des approximations fortes pour résoudre les problèmes de manière analytique. Ils ne simulent pas de manière réaliste le déplacement des particules (positon, photon) comme les simulateurs Monte-Carlo.

Dans le cas de l'imagerie TEP, la simulation analytique revient à réaliser des projections du volume TEP dans l'espace du sinogramme. Ce sinogramme est ensuite bruité et modifié pour prendre en compte les effets des coïncidences aléatoires et des photons diffusés. L'efficacité des détecteurs est aussi prise en compte, tout comme les effets de l'atténuation.

La reconstruction est ensuite réalisée de manière classique à l'aide des algorithmes de reconstruction TEP.

Cependant, bien que ce type de simulateur soit extrêmement rapide, les images générées ne parviennent pas à reproduire de manière parfaite les différents phénomènes physiques à l'origine du bruit.

10.1.2 Simulation Monte-Carlo

Les simulations Monte-Carlo¹ (MC) ont une approche probabiliste en modélisant la trajectoire de chaque photon indépendamment.

Dans le cadre de l'imagerie TEP, le modèle probabiliste est appliqué à l'émission, puis au parcours du positon, à son annihilation, ainsi que les interactions et la probabilité de détection des photons dans les détecteurs.

La génération et le suivi de chaque photon ainsi que de chaque désintégration prennent un temps extrêmement important, amplifié par le fait que deux désintégrations successives dans des organes différents peuvent amener à une coïncidence fortuite. Il faut aussi prendre en compte les limitations de l'électronique (incapacité à séparer des événements trop proches en temps comme en énergie, saturation des capteurs, ...), ce qui amène à des temps de simulation prohibitifs de l'ordre de plusieurs mois, même sur des centres de calculs. Le grand Challenge TERA 10 créé en 2006 a mobilisé un supercalculateur comprenant 7000 processeurs pendant 3h pour simuler une image corps entier à l'aide du simulateur GATE [Jan et al., 2004], avant la mise en place des procédures d'accélération. Cela revient à plus 10 jours de calculs ininterrompus pour une soumission sur un centre de calculs avec 100 processeurs disponibles en permanence.

10.1.3 Simulation Monte-Carlo accélérées

Les simulateurs Monte-Carlo accélérés utilisent des heuristiques pour augmenter la vitesse des calculs, notamment en réalisant des calculs préliminaires afin de simplifier les simulations futures.

La séparation des simulations entre les coïncidences directes et diffusées, à l'aide des statistiques créées précédemment, permet encore d'accélérer les calculs. Par exemple le projet ANR fGATE (2006-2009) a permis de réaliser une amélioration des temps de calcul d'un facteur 20 par rapport à ceux présentés ci-dessus pour le logiciel GATE.

10.2 État de l'art des simulateurs TEP

De nombreux simulateurs ont été développés par la communauté. Ils sont souvent développés à partir d'une librairie dédiée aux calculs de trajectoires de particules, comme GATE [Jan et al., 2004], qui se base sur le jeu d'outils de simulation d'interactions particule/matière Geant4 [Allison et al., 2006], le simulateur Penelo-PET [España et al., 2009] basé sur la librairie de simulation PENELOPE [Salvat et al., 2006], ou encore EGS-PET. L'université de Washington développe le simulateur Monte-Carlo simSET ainsi que le simulateur analytique ASIM.

¹Le nom de Monte-Carlo fait allusion aux jeux de hasard pratiqués dans la ville du même nom.

10.3. PROCESSUS DE SIMULATION AVEC SORTEO

De très nombreux simulateurs ont été développés en interne par de petites équipes pour les besoins des laboratoires, mais une grande partie de la communauté scientifique intéressée par cette activité s'orientent actuellement vers GATE, qui bénéficie d'un développement rapide et d'une grande communauté d'utilisateurs et de développeurs.

Une étude décrit l'ensemble des codes disponibles en 2002 [Buvat et al., 2002], et a été mise à jour en 2006 [Buvat and Lazaro, 2006] pour prendre en compte les avancées techniques réalisées.

Nous avons utilisé le logiciel PET-SORTEO [Reilhac et al., 2004], car il a été développé pour augmenter la vitesse de simulations et profiter de la parallélisation offerte par les centres de calculs. De plus, nous avons réalisé à CREATIS une première version de la base de données OncoPET_BD [Tomei et al., 2010] (http://www.creatis.insa-lyon.fr/oncoPET_DB) sur ce simulateur, ce qui permet de capitaliser sur l'expérience déjà acquise.

10.3 Processus de simulation avec SORTEO

Pour reproduire les données fournies par les systèmes cliniques TEP, les simulateurs Monte-Carlo simulent les désintégrations une par une et suivent les sous-produits dans les tissus jusqu'aux détecteurs. Étant donné qu'un examen TEP génère plusieurs millions de désintégrations, les temps de simulations deviennent très rapidement prohibitifs. PET-SORTEO dispose de plusieurs heuristiques qui permettent d'accélérer les simulations, notamment en séparant les simulations des différents types de coïncidences en fonction de statistiques estimées au début de la simulation. Cependant il faut tout de même encore plusieurs dizaines d'heures pour simuler une image correspondant à une acquisition clinique TEP corps entier au ^{18}FDG .

a) Le logiciel prend en entrée deux cartes de labels voxelisées correspondant respectivement au type de traceur et à sa concentration (en Bq/cm^3), ainsi qu'à l'atténuation pour chaque voxel. A chaque label (nombre entier définissant une région) correspond un organe, qui peut être différent pour les deux cartes. Le type de scanner doit être précisé, ainsi que les paramètres d'acquisition (2D ou 3D), et le format de sortie (mode séquence ou sinogramme).

Toutes ces informations sont stockées dans un fichier texte, appelé fichier de protocole. Les cartes d'émission et d'atténuation voxeliques doivent être fournies au format ECAT (développé par Siemens).

b) La première partie de la simulation est utilisée pour déterminer un ensemble de probabilités permettant d'accélérer les calculs futurs. Pour cela, SORTEO réalise des simulations Monte-Carlo complètes pour chaque label avec un faible nombre de photons. Cela permet de calculer plusieurs grandeurs telle que la probabilité de détection des photons pour chaque couple (label, détecteur) en prenant en compte les phénomènes physiques, ou encore la probabilité de détection en coïncidence des deux photons émis lors d'une désintégration pour chaque label.

c) La seconde partie de la simulation correspond à la réalisation de l'examen proprement dit. C'est à cette étape que sont simulées les coïncidences vraies, qu'elles soient diffusées ou non. Chaque coïncidence de chaque label est simulée une par une, et le trajet de chaque photon est calculé séparément. Si un des photons est absorbé ou si son énergie n'est pas dans la fenêtre énergétique déterminée par les paramètres d'acquisition, alors la paire de photons est considérée comme perdue. Le programme passe à la désintégration suivante. À cette étape, le programme prend en compte les temps morts (limitation de l'électronique qui ne peut pas prendre en compte un trop grand nombre de photons en même temps), qui dépend de l'activité. Il est intéressant de noter que grâce aux statistiques estimées à la première étape, la trajectoire de certains photons n'est pas calculée car le simulateur considère qu'ils font partie des "pertes". C'est ce qui permet de réaliser les calculs très rapidement par rapport aux simulations Monte-Carlo classiques.

d) La dernière étape consiste à simuler les coïncidences aléatoires, qui sont ajoutées aux données précédentes. Ces dernières sont calculées à l'aide d'un modèle utilisant les paramètres mesurés à l'étape b), notamment le taux de photons simples arrivant sur chaque détecteur.

Le modèle utilisé par SORTEO est un compromis intéressant entre le réalisme apporté par les méthodes de Monte-Carlo et les performances apportées par les techniques de pré-calculs de statistiques ainsi que de séparation du calcul des coïncidences directes et fortuites. L'auteur de PET-SORTEO annonce que toutes les sources majeures de bruit et de biais sont prises en compte.

Le simulateur PET-SORTEO a été validé pour le simulateur Ecat Exact HR+ [Reilhac et al., 2004] ainsi que pour la plateforme PET Philips Allegro, qui est à la base du scanner TEP/TDM Gemini [Lemaître, 2008].

10.4 Contribution à PET-SORTEO

Au cours de cette thèse, nous avons effectué plusieurs contributions au code de PET-SORTEO, notamment au niveau de la partie découpage des tâches pour l'exécution en centre de calcul, et au niveau de l'adaptation du programme aux données séquences.

10.4.1 Adaptation pour l'exécution sur centre de calcul

Le code original de SORTEO était adapté à une exécution sur des centres de calculs de petite taille, où la communication entre les processus n'est pas limitée. Cependant, étant donné les volumes de calculs représentés par nos simulations, nous avons fait appel au centre de calcul de l'in2p3 (Institut National de Physique Nucléaire et de Physique des Particules).

Ce centre de calcul regroupe plus de 1300 machines totalisant plus de 17000 cœurs, ainsi que 13 Petaoctets de stockage sur disques en 2011. Les technologies mises en

10.4. CONTRIBUTION À PET-SORTEO

places par les administrateurs du centre de calcul sont spécialisées pour gérer cette quantité de données, ce qui représente des contraintes particulières quand aux techniques employées par les logiciels.

Par exemple, les différents processus du simulateur PET-SORTEO dialoguaient à travers des fichiers partagés. Cela engendrait des problèmes de saturation de la bande passante entre les nœuds. J'ai donc réalisé des modifications en profondeur du code pour séparer le simulateur en plusieurs entités, chacune réalisant une seule partie du travail :

1. Estimation des paramètres nécessaires à la simulation accélérée par simulation Monte-Carlo pur (lancé pour chaque processus).
2. Combinaison des résultats Monte-Carlo.
3. Simulation simplifiée des désintégrations (lancé pour chaque processus).
4. Combinaison des désintégrations détectées pour chaque processus dans un seul fichier de données.

Ensuite, un ensemble de scripts a été réalisé pour automatiser les opérations de combinaison des résultats et de calcul des statistiques, puis pour relancer les simulations des coïncidences vraies et fortuites. Une dernière étape consiste à réassembler les détections pour générer les données séquence.

10.4.2 Sortie en mode Séquence

Bien que le code original permettait de spécifier un format de sortie, en pratique seul le format sinogramme était pris en compte.

En effet, le code original ne permettait pas la sauvegarde de l'information temporelle de chaque évènement détecté. Or cette information est nécessaire aux méthodes de correction du mouvement respiratoire.

Nous avons donc adapté PET-SORTEO au format séquence. Nous avons repris le format LMF développé lors de la collaboration CrystalClear pour générer les données, car ce format de données est simple à utiliser. Il se compose d'un fichier texte contenant les informations sur l'acquisition ainsi que d'un fichier binaire contenant les évènements organisés de manière séquentielle.

L'adaptation de SORTEO a nécessité entre autres ces adaptations :

- Intégrer la spécification du format de sortie dans le code.
- Intégrer dans le code de PET-SORTEO des enregistreur d'évènements pour chaque désintégration détectée, car le code original se contentait de modifier un sinogramme. Cela a nécessité une modification en profondeur du code original.
- Adapter la géométrie du modèle simulé pour correspondre aux conventions du LMF.
- Ajouter une information temporelle non présente originellement à chaque évènement.

CHAPITRE 10. SIMULATIONS

- Réaliser le code permettant l'assemblage et le tri de toutes les émissions, qui sont réalisées séparément pour chaque organe.

Chapitre 11

Base de données : OncoPET_Respi

11.1 Présentation

Pour répondre à notre problématique de détection des lésions sur les images TEP respirantes, nous avons choisi d'utiliser une base de données d'images simulées. En effet, notre problématique repose sur une connaissance très précise de la position des lésions pour l'évaluation des performances, ainsi que sur l'acquisition de données en mode séquence. Ces deux objectifs sont relativement difficiles à réaliser à partir d'images cliniques, notamment parce que l'acquisition en mode séquence ne fait pas partie de la routine clinique en oncologie.

11.1.1 Avantages des bases de données simulées

Le principal avantage des bases de données simulées est le contrôle qu'elles permettent d'exercer sur les modèles. Il est possible de spécifier les conditions des acquisitions de manière très précise, et de garantir une homogénéité difficile à atteindre en utilisant des données patient.

Dans le cadre de nos travaux sur la détection, la vérité terrain est particulièrement appréciable, car elle permet de savoir, avec une certitude impossible à atteindre en clinique, la position, le contraste et le nombre des lésions. En effet, dans le cas de lésions de petit diamètre et de petit contraste, le diagnostic ne peut être donné avec une précision totale par le praticien.

De plus, les simulations permettent d'obtenir des images avec des paramètres qui ne sont pas forcément disponibles en routine clinique. Dans notre cas, la génération des images en mode séquence est indispensable, et très peu d'exams sont réalisés de cette façon, notamment à cause de l'espace disque nécessaire et des temps de reconstruction qu'ils impliquent.

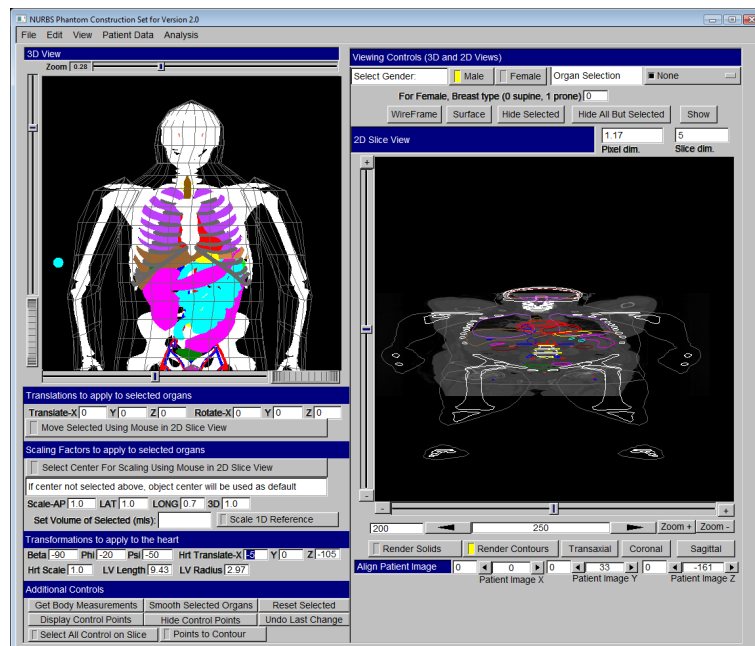


FIG. 11.1: Adaptation du modèle XCAT sur une image TDM de patient : le logiciel ne gère pas correctement l'anisotropie des voxels, ce qui explique la déformation du modèle sur la vue de droite. Chaque organe est représenté par une surface paramétrique qu'il est possible d'adapter sur un modèle de patient.

11.2 Modèles

11.2.1 Modèle anatomique

Nous avons utilisé une base de données d'images TDM fournie par le Centre Hospitalier Lyon-Sud (CHLS) pour créer 15 modèles de patients. Le fantôme paramétrique XCAT [Segars and Tsui, 2009] développé par Paul Segars a été déformé de manière manuelle sur chacune des images TDM à l'aide d'un logiciel fourni avec le modèle, présenté dans la figure 11.1. Ce modèle est basé sur des surfaces paramétriques NURB (Non-Uniform Rational B-Splines) que l'on peut ajuster manuellement sur les organes d'intérêt du patient en déplaçant les points de contrôle. L'adaptation n'est pas parfaite, du fait des limitations du logiciel, mais permet d'obtenir une variabilité suffisante pour notre étude. Deux images TDM associées aux fantômes adaptés sont présentées dans la figure 11.2.

Nous avons choisi le XCAT car il intègre directement un modèle de mouvement respiratoire, et il est conçu pour pouvoir être déformé et ainsi engendrer différents patients virtuels.

11.2. MODÈLES

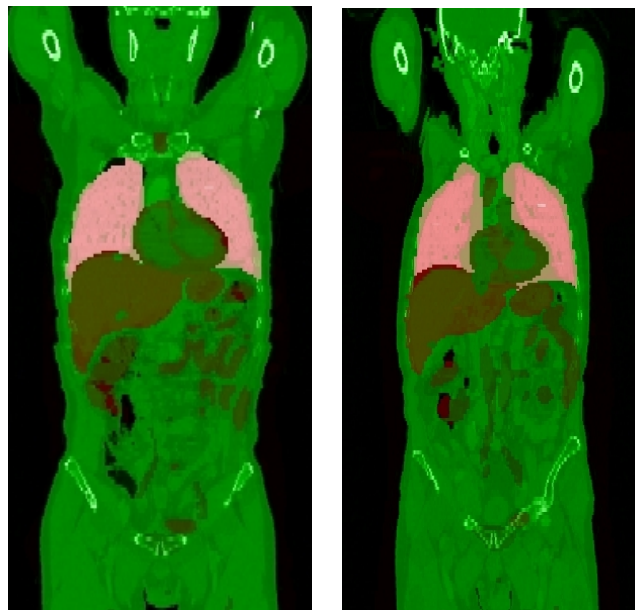


FIG. 11.2: Exemple d'adaptation d'images du modèle XCAT sur des images TDM : deux modèles XCAT représentés en niveaux de rouge sont adaptés sur des images TDM représentées en vert.

11.2.2 Activités des organes

Les activités des organes sont décrites dans la table 11.5. Elles ont été estimées à partir d'une étude réalisée précédemment lors des travaux de thèse de Sandrine Tomeï [Tomei et al., 2008b], en mesurant l'activité dans des zones d'intérêt tracées dans une série de 70 patients [Tomei et al., 2010].

11.2.3 Modèle de Respiration

Pour prendre en compte la variabilité du cycle respiratoire (voir figure 11.3), nous avons utilisé 4 cycles différents pour modéliser la respiration du patient. Un signal respiratoire complet a été acquis sur une durée de plusieurs minutes à l'aide d'un spiromètre (voir chapitre 5.2.1). Nous avons extrait 3 cycles semblables correspondant à la phase de respiration normale, ainsi qu'un cycle "anormal" pour prendre en compte les effets d'une respiration irrégulière du patient pendant une partie du cycle.

Le signal respiratoire obtenu est présenté dans la figure 11.4. Il est constitué de 4 cycles de 5.6 secondes qui se répètent 10 fois pour former un signal total de 224 secondes, soit un peu moins de 4 minutes.

Chacun de ces 4 cycles est ensuite discrétisé en 8 parties qui seront utilisées pour générer les modèles de la simulation.

CHAPITRE 11. BASE DE DONNÉES : ONCOPET_RESPI

Organe	Activité (Bq/cm^3)
Foie	7740
Myocarde	11610
Os	3863
Poumon	1338
Prostate	2575
Rate	4939
Reins	8220
Sang	5340
Tissus mous	2575
Urètre	33815
Vésicule biliaire	2575
Vessie	50973

TAB. 11.1: Activités définies pour les organes des patients virtuels.

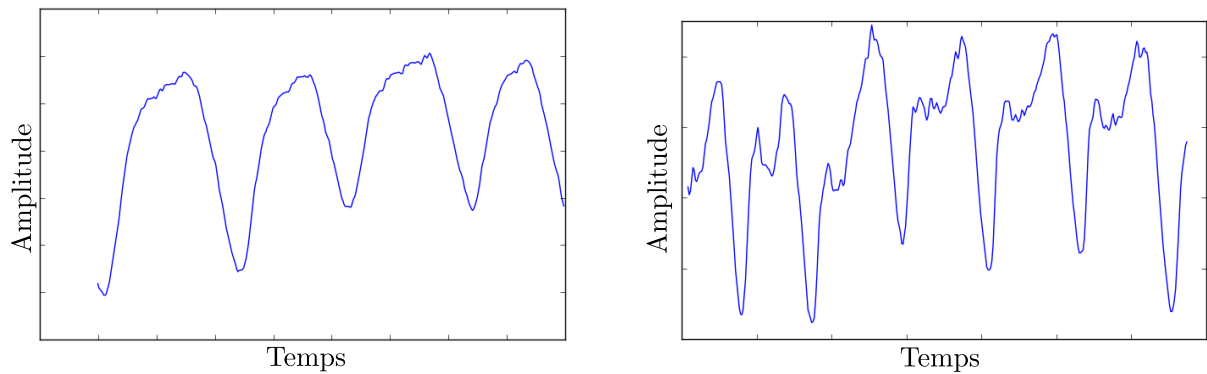


FIG. 11.3: Deux exemples de courbes respiratoires acquises par un spiromètre : celle de gauche montre une respiration régulière, tandis que celle de droite est irrégulière.

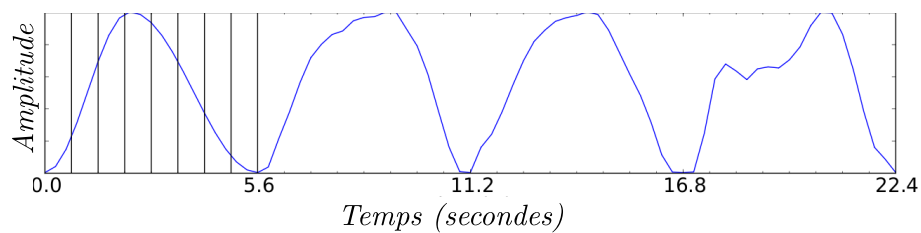


FIG. 11.4: Courbe respiratoire utilisée pour générer les modèles respirants. Les traits verticaux du premier cycle montrent les 8 instants sélectionnés pour discrétiser le mouvement respiratoire de ce cycle.

11.2. MODÈLES

11.2.4 Modèle de tumeur

Nous avons choisi de focaliser cette étude sur la détection des lésions situées dans le foie et les poumons. Ces organes, en effet, sont les plus influencés par le mouvement respiratoire.

Nous avons utilisé un modèle de lésion correspondant adapté aux lésions de faible intensité. Nous avons utilisé des lésions sphériques [Bai and Brady, 2009], avec des diamètres de l'ordre de la résolution de l'imageur TEP.

Nous avons inséré 280 lésions dans les 15 modèles de patients que nous avons créés. 173 lésions sont présentes dans le poumon, contre 103 dans le foie. La différence de taille entre les organes du XCAT explique la plus petite taille de la base de lésions du foie.

Les lésions ont été placées manuellement dans tout le volume de l'organe, en privilégiant les parties à fort mouvement proche du diaphragme.

Les contrastes ont été calibrés pour correspondre à des lésions de faible détectabilité par l'étude suivante :

Calibration des contrastes des lésions

Notre but est de simuler des images avec un ensemble de tumeurs qui ne soient ni trop évidentes ni impossibles à détecter. Nous avons donc cherché à obtenir une échelle de contrastes qui permette un taux de détection par un observateur humain de 10%, 30%, 50%, 70% et 90%.

Pour réaliser cela, un ensemble d'images a été simulé sans mouvement respiratoire, avec différentes tailles de tumeurs et en utilisant, en premier lieu, des valeurs de contrastes présentés dans la base de données créée dans notre laboratoire et présentée dans [Tomei et al., 2010]. Les valeurs de contraste de la base OncoPET_DB sont présentées dans la table 11.2 pour les lésions du poumon et du foie.

Ensuite, une étude ROC avec deux observateurs humains non médecin a été réalisée afin d'estimer la détectabilité de ces couples activité/taille de tumeurs (fig.11.5).

On présente à chaque observateur les images les unes après les autres. Ils les annotent en recherchant toutes les lésions et en leur attribuant une note à l'aide du logiciel Amide [Loening and Gambhir, 2003]. La localisation des lésions n'est pas connue à l'avance, et ils doivent donner pour chaque lésion une note entre 1 et 5 correspondant au barème suivant :

1. Possible.
2. Probable.
3. Très probable.
4. Pratiquement certain.
5. Certain.

CHAPITRE 11. BASE DE DONNÉES : ONCOPET_RESPI

Poumon :

Poumon	10%	30%	50%	70%	90%
8mm	2	5	8	10	13
12mm	2	5	8	10	13
16mm	2	3.5	4.5	7.5	10

Foie :

Poumon	10%	30%	50%	70%	90%
8mm	2.5	4	4.5	6	9
12mm	2	3	4	4.5	7.5
16mm	2	3	4	4.5	7.5

TAB. 11.2: Contraste de la base originale OncoPET_DB des lésions du poumon et du foie, avec les pourcentages de détection associés.

Nous avons simulé 5 réalisations d'un même modèle basé sur le XCAT avec 12 lésions différentes implantées dans les poumons et le foie de chaque réalisation. Les données ont été reconstruites avec l'algorithme OPL-EM présenté dans le chapitre 3.1.3 en utilisant 10 itérations et un seul sous-ensemble. Le choix de ces paramètres a été fixé *a priori* car cette étude est antérieure à la recherche des meilleurs paramètres présentés dans le chapitre 12.

Les résultats de cette étude sont reportés sur les figures 11.5 et 11.6, qui représentent respectivement le taux de détection des lésions du foie et du poumon moyennée sur 2 observateurs en fonction du contraste.

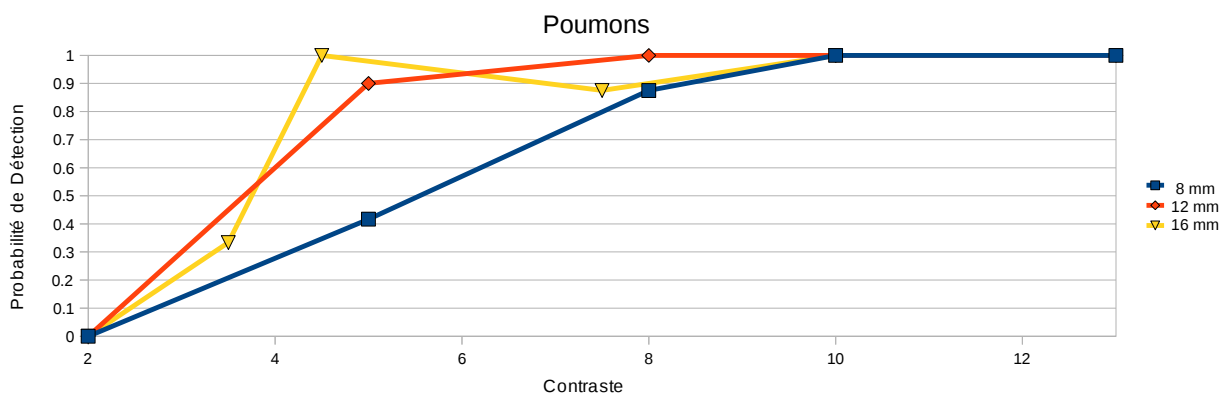


FIG. 11.5: Détectabilité des tumeurs du poumon moyennée sur 2 observateurs en fonction du contraste et de la taille des tumeurs. Chaque courbe représente une taille de lésion différente.

Ces résultats sont également résumés dans la table 11.3. Ils montrent une détectabilité largement supérieure à celle que nous avons estimée initialement à partir des résultats obtenus sur la première base de données OncoPET_DB. Les écarts peuvent s'expliquer par le fait que nous n'avons pas simulé le même scanner TEP (Caméra HR+ pour la base OncoPET_BD, Philips Gemini pour la nouvelle base), et que les données ont été reconstruites avec un algorithme différent. Il s'avère également que les lésions de 16 mm sont trop visibles pour être intégrées dans l'étude. Nous

11.3. SIMULATION

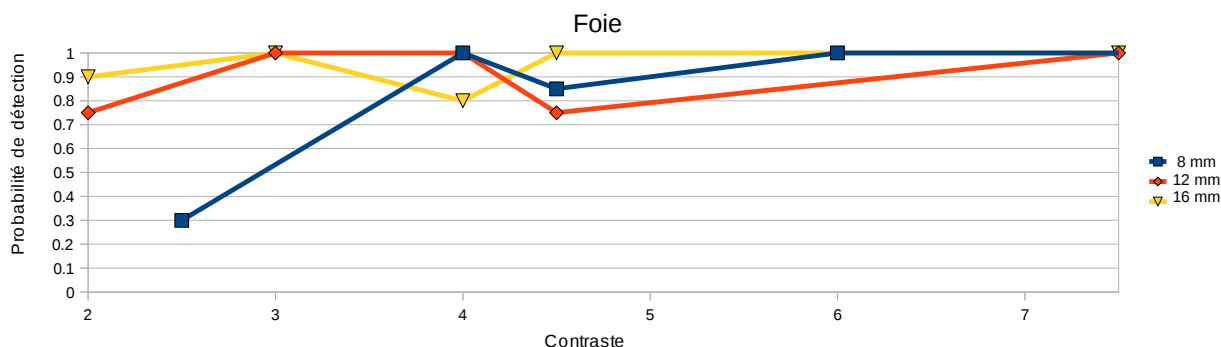


FIG. 11.6: Détectabilité des tumeurs du foie moyennée sur 2 observateurs en fonction du contraste et de la taille des tumeurs. Chaque courbe représente une taille de lésion différente.

Poumon :

Détection voulue	8mm	12mm	16mm
90 %	100 %	100 %	100 %
70 %	100 %	100 %	90 %
50 %	90 %	100 %	100 %
30 %	44 %	90 %	30 %
10 %	0 %	0 %	0 %

Foie :

Détection voulue	8mm	12mm	16mm
90 %	100 %	100 %	100 %
70 %	100 %	75 %	100 %
50 %	85 %	100 %	75 %
30 %	100 %	100 %	100 %
10 %	30 %	75 %	90 %

TAB. 11.3: Détectabilité des lésions réalisée à l'aide d'une étude par des humains sur 5 patients totalisant 60 lésions. La colonne de gauche indique la détectabilité des lésions calibrée pour OncoPET-BD.

avons fait des interpolations linéaires entre les statistiques obtenues pour définir les nouvelles gammes d'activités. Les valeurs de contraste utilisées dans notre étude sont finalement reportées dans la table 11.4 pour les lésions du foie et du poumon.

11.3 Simulation

Nous avons utilisé le logiciel SORTEO pour simuler la caméra TEP de l'imageur TEP/TDM Philips Gemini. La simulation a été réalisée sur 3 lits, avec un recouvrement de 9 cm par lits. Pour référence, le champ de vue axial de l'imageur est de 18 cm.

Nous avons réalisé les simulations de deux acquisitions :

Dynamique : Les 8 modèles de chacun des 4 cycles respiratoires ont été simulés

CHAPITRE 11. BASE DE DONNÉES : ONCOPET_RESPI

	8mm Foie	12mm Foie	8mm Poumon	12mm Poumon
10%	1.8	1.3	3.0	2.5
30%	2.0	1.5	4.0	3.0
50%	2.5	1.8	5.0	3.5
70%	3.0	2.0	6.5	4.0
90%	3.5	2.3	8.0	5.0

TAB. 11.4: Contrastes appliqués aux lésions de la base de données en fonction du taux de détectabilité désiré.

séparément pour chaque lit. La durée de chaque examen simulé est de 7 secondes, ce qui représente une simulation totale de 224 secondes. Les données simulées de chaque cycle sont concaténées pour simuler le résultat d'une acquisition avec synchronisation respiratoire, et former 8 fichiers de données séquence par lit.

Statique : Le modèle de fin d'expiration a été utilisé pour simuler un examen de 224 secondes par lit pour correspondre à la durée de la simulation dynamique.

Le temps de simulation nécessaire pour générer une image statique est de 110 h par processeur environ, contre 130 heures pour générer une image dynamique. La différence de temps s'explique par le processus de simulation. En effet, chaque simulation est réalisée en deux temps (voir 10.3), dont le premier a une durée fixe, qui est réalisée pour chacune des 8 simulations des 4 cycles dans le cadre de la simulation dynamique, et seulement une fois pour la simulation statique.

Le volume de données généré est d'environ 6 Go par patient sans prendre en compte les données intermédiaires générées lors de la simulation.

11.3.1 Exemples de données reconstruites

La base de données générée contient les données de simulation de 15 patients, reconstruits avec une correction parfaite (image statique), sans correction, et avec deux méthodes de correction du mouvement respiratoire présentées dans les sections 6.3 et 6.5. Deux exemples d'images reconstruites complètes sont visibles sur la figure 11.7. La figure 11.8 illustre le mouvement respiratoire dans les images reconstruites.

11.4 Données clef

Cette section correspond à un résumé des caractéristiques de la base de données.

11.4.1 Modèles

Nous avons créé 15 modèles, qui ont été adaptés depuis autant d'images TDM fournies par le centre hospitalier Lyon-SUD. Pour cela nous avons utilisé les outils

11.4. DONNÉES CLEF

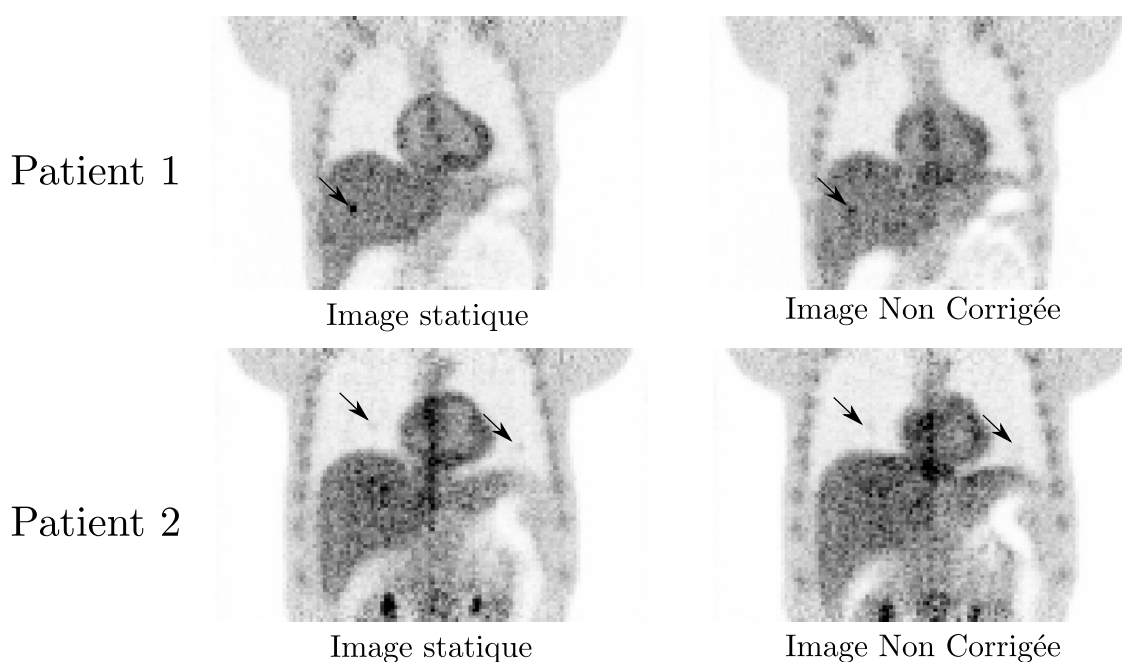


FIG. 11.7: Images reconstruites (8 itérations et 5 sous-ensembles) à partir des simulations de la base de données. Les flèches représentent une lésion du foie de diamètre 8 mm avec un contraste de 3.5 : patient 1, ainsi que deux lésions du poumon de 8 mm et de contraste 6.5 : patient 2.

Niveau de confiance		1	2	3	4	5
Poumon (173)	8 mm (90)	3 (19)	4 (18)	5 (18)	6.5 (18)	8 (17)
	12 mm (83)	2.5 (16)	3 (16)	3.5 (18)	4 (17)	5 (16)
Foie (107)	8 mm (54)	1.8 (11)	2 (11)	2.5 (10)	3 (11)	3.5 (11)
	12 mm (53)	1.3 (10)	1.5 (10)	1.8 (11)	2 (11)	2.3 (11)

TAB. 11.5: Tableau récapitulatif des contrastes des lésions présentes dans le foie et les poumon des modèles. le nombre de lésions est indiqué entre parenthèses.

fournis par Paul Segars [Segars, 2001].

11.4.2 Lésions

Nous avons inséré 280 lésions dans les 15 modèles de patients que nous avons créés. 173 lésions sont présentes dans le poumon, contre 107 dans le foie. La différence de taille entre les organes explique la plus petite taille de la base de lésions du foie.

Les lésions ont été placées manuellement dans tout le volume de l'organe, en préférant les parties à fort mouvement proche du diaphragme.

Le tableau 11.5 récapitule les contrastes, ainsi que les tailles des lésions implantées dans les organes.

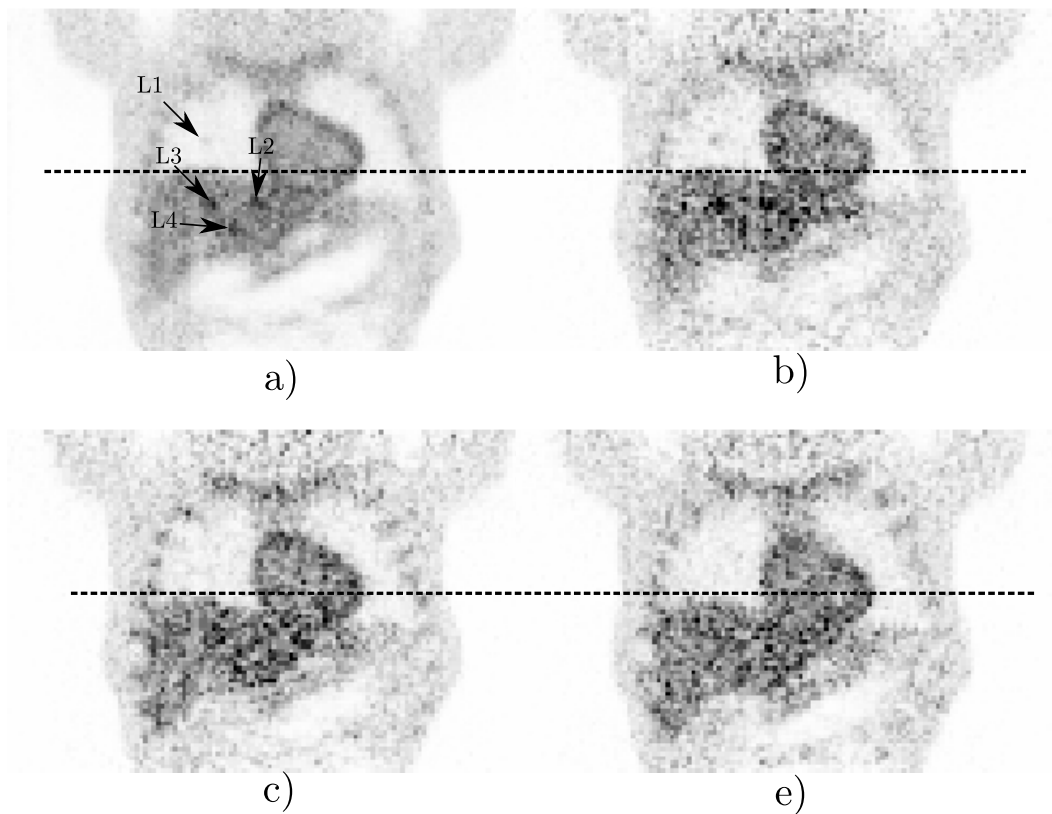


FIG. 11.8: Images reconstruites (8 itérations et 5 sous-ensembles) à partir des simulations statiques a) et dynamiques b),c) et d) de la base de données. b), c) et d) représentent respectivement les données du premier instant temporel (fin d'expiration), du troisième et du quatrième. L1 représente une lésion pulmonaire de 8 mm et de contraste 8, L2 une lésion hépatique de 12 mm et de contraste 1.5, L3 et L4 des lésions hépatiques de diamètre 12 mm et de contraste de 2.3.

Cinquième partie

Évaluation de la correction du mouvement respiratoire sur la détection des lésions pulmonaires et hépatiques

Chapitre 12

Méthodes

Nous avons utilisé la base de données présentée dans la partie 11 pour évaluer les performances des techniques de correction du mouvement respiratoire présentées dans le chapitre 6.

Les techniques de correction du mouvement implémentées sont les suivantes :

1. Correction pendant la reconstruction par modification de la matrice système (voir section 6.5). Elle sera désignée par l'acronyme **TE-MS** (Transformation Élastique de la Matrice Système).
2. Correction post-reconstruction par recalage des images prises à différents instants du cycle (voir section 6.3). Elle sera désignée par l'acronyme **TE-IM** (Transformation Élastique des IMages reconstruites).

Les images TEP corrigées à l'aide de ces deux méthodes sont comparées dans cette partie avec les images non corrigées et des images statiques (qui représentent une correction parfaite). Leurs acronymes sont respectivement **NoCorr** et **Statique**.

L'objectif est d'évaluer les performances des techniques de correction du mouvement sur la détection des lésions de faible contraste et faible diamètre. Pour cela, les performances d'un système de détection automatique seront mesurées à l'aide des courbes F-ROC.

Ce chapitre va tout d'abord présenter la démarche réalisée pour sélectionner les paramètres de reconstruction des images, puis la méthode d'estimation du mouvement utilisée.

La suite de ce chapitre présente les caractéristiques principales du système d'aide à la détection (CAD) que nous avons proposé (voir chapitre 7), l'étape d'optimisation des paramètres de ce CAD que nous avons utilisé, puis la méthode que nous avons développée pour mesurer les performances de détection.

12.1 Reconstruction des images

12.1.1 Principe de la reconstruction

Nous avons utilisé le logiciel de reconstruction fourni par le Laboratoire de Traitement de l'Information Médicale (LaTIM) dans le cadre d'un partenariat, créé et utilisé par Frédéric Lamare pour ses travaux sur la correction du mouvement respiratoire [Lamare et al., 2007b].

Ce logiciel est capable de reconstruire les images acquises en données séquentielles à l'aide de l'algorithme OPL-EM décrit en 3.1.3. La reconstruction permet de prendre en compte la correction de l'atténuation, mais ne permet pas la correction des coïncidences aléatoires et des coïncidences diffusées. Nous avons donc choisi de simplifier le problème en supposant une correction parfaite de ces deux effets. Pour cela, nous n'avons pas inclus les photons diffusés ou les coïncidences aléatoires dans les données séquentielles.

L'algorithme utilisé peut intégrer les corrections de mouvement à partir d'une estimation du champ de déplacement qui doit être calculé séparément. La qualité de ce champ conditionne la qualité de la correction. De plus, OPL-EM nécessite de choisir le nombre de sous-ensembles d'itération. Cette section présente l'étude réalisée pour fixer ces différents paramètres.

12.1.2 Paramètres de reconstruction

Lors de l'utilisation des méthodes itératives, il faut définir le nombre d'itérations à appliquer lors de la reconstruction. La figure 12.1 présente des reconstructions réalisées avec différentes valeurs du nombre d'itérations. L'algorithme OPL-EM utilisé pour réaliser les reconstructions emploie la technique des sous-ensembles pour accélérer la reconstruction.

Le nombre d'itérations total (nombre d'itérations \times nombre de sous-ensembles) va déterminer la qualité du résultat. Nous avons choisi de travailler avec 5 sous-ensembles et de faire varier le nombre d'itérations de 1 à 7, ce qui correspond à une évaluation de 5 à 35 itérations totales.

L'optimisation du nombre totale d'itérations est réalisée en maximisant le rapport contraste sur bruit [Takahara et al., 2004] (CSB) des lésions, défini comme suit :

$$CSB = \frac{\mu_{signal} - \mu_{fond}}{\sqrt{\sigma_0}} \quad (12.1)$$

Où μ_{signal} et μ_{fond} représentent, respectivement, l'activité moyenne d'une zone d'intérêt de taille $3 \times 3 \times 3$ voxels centrée sur une lésion, et l'activité moyenne d'une zone saine. Cette zone saine correspond à un volume de 200 voxels dans le poumon et de 64 voxels dans le foie. σ_0 représente la variance du bruit. Dans notre cas, cette variance est évaluée sur la zone saine définie précédemment. Appliquée aux lésions,

12.1. RECONSTRUCTION DES IMAGES

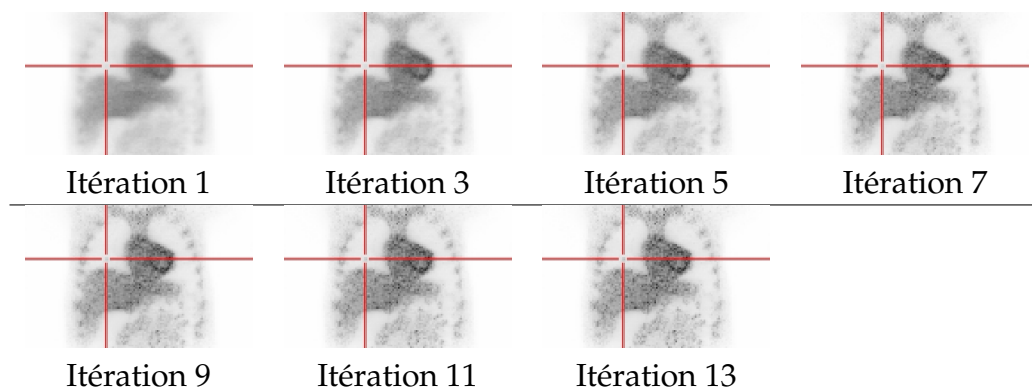


FIG. 12.1: Illustration de l'évolution des images en fonction du nombre d'itérations utilisées pour la reconstruction. Une lésion est présente au centre de la croix rouge.

cette métrique a l'avantage de permettre une représentation rapide de l'évolution du contraste entre la lésion et le fond, tout en pénalisant le bruit.

L'évaluation du rapport contraste sur bruit est réalisée sur une acquisition statique d'un modèle de patient contenant 12 lésions (6 dans le foie et 6 dans le poumon), de diamètre et de contraste échantillonnés à partir des valeurs calibrées dans la table 11.5 présentée à la page 121.

Les résultats sont présentés dans la figure 12.2 pour le foie et 12.3 pour le poumon.

La première observation est que certaines courbes sont strictement descendantes, surtout dans le foie. Cela s'explique par le fait que le niveau de bruit augmente beaucoup plus rapidement que la différence d'activité entre le signal et le fond. Cependant, les lésions concernées ont toutes un contraste très élevé, ce qui les rend plus facilement détectables. Par contre, pour les autres, on peut voir que l'optimum du CSB des lésions du foie est situé autour de 4 itérations, suivi par un plateau, tandis que l'optimum du CSB des lésions du poumon est plus proche de 8 itérations complètes.

La lésion du foie ayant un contraste de 2.5 et un diamètre de 8mm a une valeur de rapport contraste sur bruit négative. Cela s'explique par le fait que pour les premières itérations, une faible différence d'activité en faveur du fond peut être fortement accentuée par l'écart type qui est très faible. De la même manière, pour les itérations suivantes, l'activité de cette lésion redevient supérieure à celle du fond, mais l'écart-type du bruit devient plus important, ce qui fait que le rapport est proche de zéro. Lorsque le nombre d'itérations est très faible, on observe une diminution de l'activité des petites structures par effet d'étalement (spill-out). Pour des nombres d'itérations plus importants, l'effet de volume partiel diminue, ce qui signifie que l'activité du centre de la lésion augmente.

Nous avons choisi de réaliser les reconstructions avec 8 itérations de 5 sous-ensembles car nous avons montré que ce nombre d'itérations est optimal pour les lésions du poumon, et ne pénalise pas trop la détectabilité des lésions du foie.

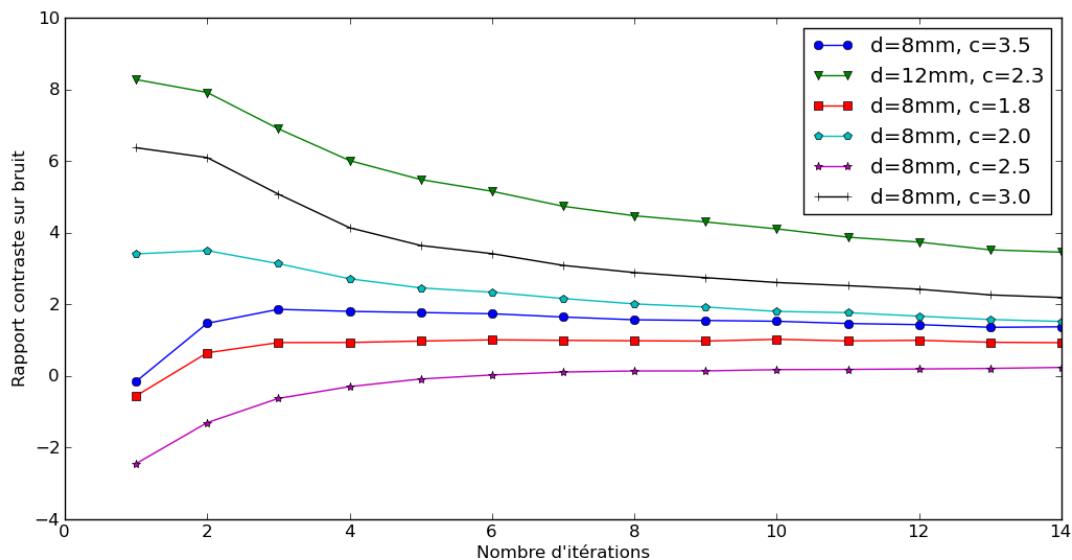


FIG. 12.2: Evaluation du rapport "contraste sur bruit" des lésions du foie en fonction du nombre d'itérations : sont indiqués, pour chaque lésion, le niveau de contraste réel ainsi que le diamètre de la lésion

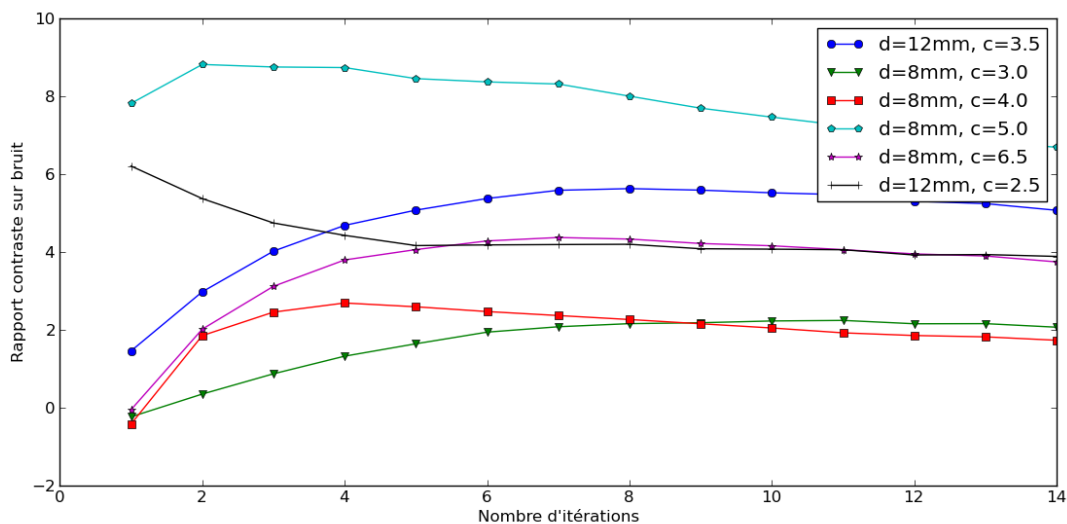


FIG. 12.3: Evaluation du rapport Contraste sur bruit des lésions du poumon en fonction du nombre d'itérations : sont indiqués, pour chaque lésion, le niveau de contraste réel ainsi que le diamètre de la lésion

12.2. ESTIMATION DU MOUVEMENT

12.2 Estimation du mouvement

L'estimation du mouvement est réalisée uniquement à partir des données TEP et des informations de synchronisation respiratoire.

12.2.1 Principe de l'estimation de mouvement

L'estimation du mouvement est réalisée selon le principe énoncé en 5.3.1 :

1. Les données simulées de chaque partie de cycle sont additionnées pour les N instants du cycle respiratoire.
2. Les N images reconstruites à partir des données précédentes sont utilisées pour calculer le mouvement respiratoire : les images des temps 2 à N sont recalées sur l'image de référence (numéro 1), ce qui associe à chaque instant respiratoire un champ de déformation.

Le champ de mouvement est estimé en utilisant une méthode de recalage par B-splines. Il est représenté par un nombre réduit de coefficients c de fonctions β^n placées sur une grille dans le volume, comme présenté dans la figure 12.4. L'implémentation de cet algorithme a été fournie par Philips France (laboratoire Medisys). La transformation $g_t(x)$ utilisée pour associer les images temporelle $f(x, t)$ (pour t allant de 2 à N) avec l'image de référence $f(x, 1)$, est représentée de la manière suivante :

$$g_t(x) = x + \sum_{j \in \mathbb{Z}^N} c_j \beta^n \left(\frac{x}{h} - j \right) \quad (12.2)$$

Où $\beta^n(x)$ est la valeur de la fonction B-spline de degré n au point x , et j représente les indices des positions de la grille. Le paramètre h correspond à l'espacement entre les points de la grille. A chaque fonction B-spline de la grille correspond un coefficient de pondération associé nommé c_j qui représente l'apport de la B-spline correspondante au signal final.

Les coefficients c_j sont obtenus à l'aide d'une optimisation multi-échelle sur 3 niveaux : L'estimation est réalisée sur des images de résolution différente, en se basant sur l'estimation à la résolution $r - 1$ pour réaliser celle de l'estimation à la résolution r . L'optimisation est réalisée par gradient conjugué selon la méthode de Polak-Ribière [Polak and Ribiere, 1969]. Nous avons utilisé l'erreur quadratique moyenne pour comme métrique de recalage. Elle a l'avantage d'être simple, rapide, et de ne pas montrer de discontinuités [Ledesma-Carbayo et al., 2005]. Nous avons choisi d'utiliser des splines d'ordre 1.

12.2.2 Paramètres de l'estimation de mouvement

L'estimation du champ de mouvement est réalisée à partir des 8 images TEP correspondant aux 8 positions temporelles du cycle respiratoire. Ces données contiennent

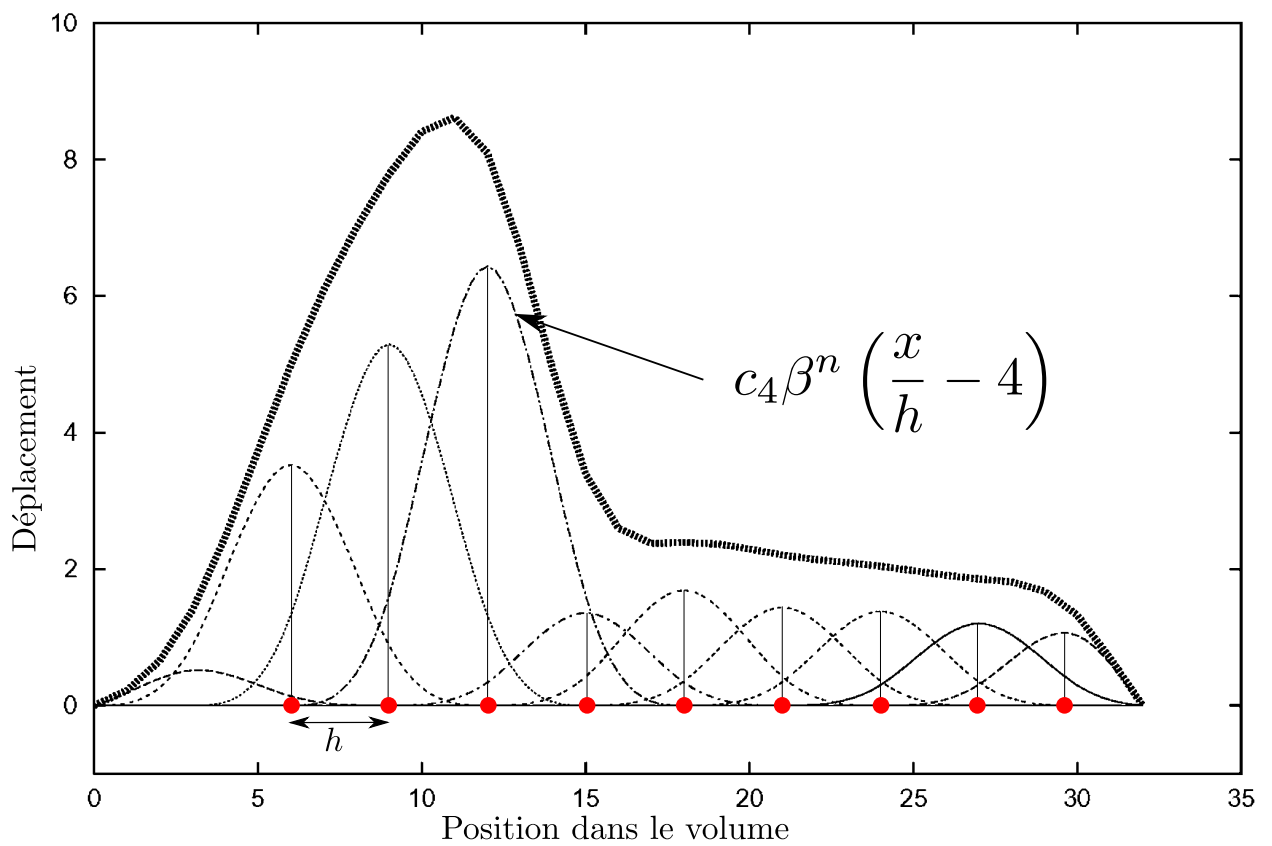


FIG. 12.4: Exemple d'interpolation par B-spline en 1 dimension : La courbe en trait pointillé épais est représentée par la somme des contributions des courbes B-spline représentées en traits pointillés fins. Les points de la grille utilisée pour le placement des courbes sont représentés par les ronds rouges.

12.2. ESTIMATION DU MOUVEMENT

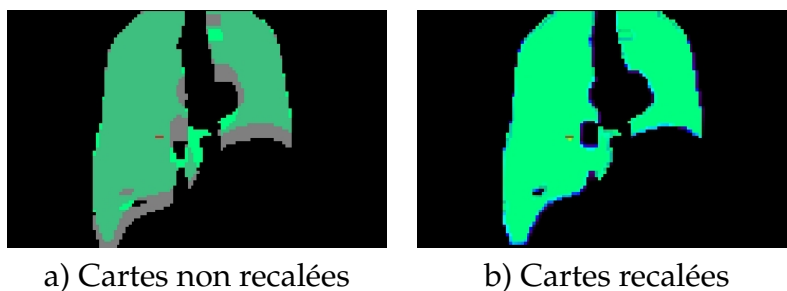


FIG. 12.5: Illustration de la pertinence du recalage obtenu à partir de l'estimation de mouvement réalisée sur les images. En vert l'image de référence et en gris l'image du temps correspondant à la différence la plus importante.

donc 8 fois moins d'évènements que les images statiques. Ainsi, les paramètres de reconstruction déterminés en 12.1.2 ne sont pas optimaux pour ces images, d'autant que nous cherchons, à ce stade, à optimiser la qualité de l'estimation de mouvement. C'est pour ces raisons que nous avons réalisé une autre étude pour estimer les paramètres de reconstruction pour cette problématique spécifique.

Évaluation de l'estimation de mouvement

Pour évaluer la performance de chaque jeu de paramètres, nous avons utilisé la vérité terrain fournie par la carte de labels pour créer une image d'évaluation correspondant à l'image de référence, ainsi qu'une autre correspondant à l'image "respirante". Les images d'évaluation sont créées à partir des fantômes de référence en assignant aux organes étudiés (foie et poumon) une valeur de 1, et une valeur plus importante pour les lésions. Le reste a une valeur de 0. Une illustration est présentée dans la figure 12.5.a).

L'image d'évaluation correspondant à l'image "respirante" est alors recalée sur l'image de référence à l'aide du champ de mouvement élastique calculé précédemment. Nous évaluons ensuite la qualité du recalage en calculant l'écart quadratique moyen entre les deux images d'évaluation 12.5.b).

Optimisation des paramètres de l'estimateur de mouvement

Nous avons évalué l'influence de l'échantillonnage spatial des B-spline sur la qualité de l'estimation de mouvement. En effet, si l'espacement entre les nœuds de contrôle est trop important, les mouvements locaux vont interférer entre eux et le résultat ne sera pas valide. De la même manière, en cas d'espacement trop faible, la nature extrêmement bruitée des images va engendrer des micro-mouvements parasites locaux. Pour réduire le bruit, nous appliquons un filtrage gaussien de largeur à mi-hauteur égale à 6 mm à chaque itération.

Nous avons évalué le recalage des reconstructions pour les paramètres suivants :

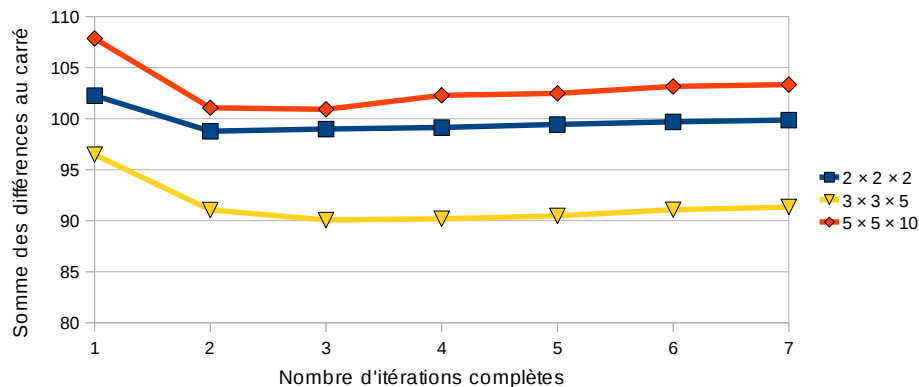


FIG. 12.6: Performances de l'estimation de mouvement en fonction du nombre d'itérations de la reconstruction selon la taille de la grille utilisée pour l'estimation de mouvement

Nombre d'itérations : Le nombre de sous-ensembles est toujours de 5, mais nous faisons varier le nombre d'itérations de 1 à 7, ce qui représente un nombre d'itérations totales de 5 à 35.

Précision de la grille : nous avons évalué les performances de la détection pour 3 tailles différentes, allant d'une grille très peu précise de 2 noeuds en x, y et z, à une grille plus précise avec 5 noeuds en x et y, et 10 noeuds en z. Une grille intermédiaire a été évaluée avec 3 noeuds en x et y, ainsi que 5 noeuds en z.

Les résultats sont présentés dans la figure 12.6. Les meilleures performances sont obtenues pour la configuration $3 \times 3 \times 5$ avec 3 itérations et 5 sous-ensembles.

Optimisation des paramètres de la reconstruction

Nous avons ensuite souhaité vérifier que la régularisation avait un impact positif sur la qualité du recalage. Nous avons utilisé les paramètres présentés précédemment pour la valider :

Nombre d'itérations : Le nombre de sous-ensembles est de 5, mais nous faisons varier le nombre d'itérations de 1 à 7, ce qui représente un nombre d'itérations totales de 5 à 35.

Présence ou absence de régularisation pendant la reconstruction : Une régularisation par filtrage gaussien de largeur à mi-hauteur de 6 mm est appliquée ou non à chaque itération.

Les résultats sont présentés dans la figure 12.7. Ils montrent clairement que la régularisation entraîne une amélioration des performances, et que la meilleure estimation de mouvement est réalisée pour une reconstruction de 3 itérations avec 5 sous-ensembles. Nous avons par conséquent utilisé ces paramètres avec une grille de 3 noeuds en x et en y, soit un espacement de 17 cm, et de 5 noeuds en z, soit un espacement de 7.7 cm.

12.3. DESCRIPTION DES DIFFÉRENTS TYPES D'IMAGES À ÉVALUER

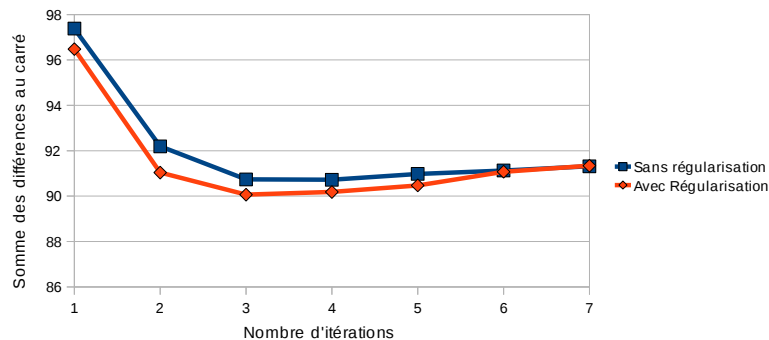


FIG. 12.7: Performances de l'estimation de mouvement en fonction du nombre d'itérations complètes de la reconstruction selon la présence ou l'absence de régularisation pendant la reconstruction. Plus la valeur en ordonnée (mesure de différence) est faible, meilleure est la performance

12.3 Description des différents types d'images à évaluer

Nous avons choisi de corriger les images TEP en utilisant la méthode de correction post reconstruction décrite en 6.3 ainsi que la correction par modification de la matrice système décrite en 6.5. La première est déjà implémentée sur les imageur GE (technologie motionFree), tandis que la seconde est activement étudiée en recherche et intégrée dans le logiciel de reconstruction fourni par le LATIM. Des images générées sur deux patients sont présentées dans la figure 12.8.

12.3.1 Images Statiques

Les images statiques sont reconstruites à partir des données de simulation statiques, réalisées à partir d'une acquisition complète du modèle au premier instant du cycle respiratoire (fin d'expiration).

Chaque image est donc en phase avec la carte d'atténuation qui est elle aussi dérivée du modèle en fin d'expiration. Les images ont été reconstruites à l'aide de l'algorithme OPL-EM avec 8 itérations comprenant 5 sous-ensembles chacune.

12.3.2 Images Non corrigées

Les images non corrigées sont produites de la même manière que les images statiques, mais à partir des fichiers séquence dynamiques. Ces fichiers correspondent à une acquisition en respiration libre de 224 secondes.

La carte d'atténuation utilisée est celle correspondant au premier instant du cycle respiratoire.

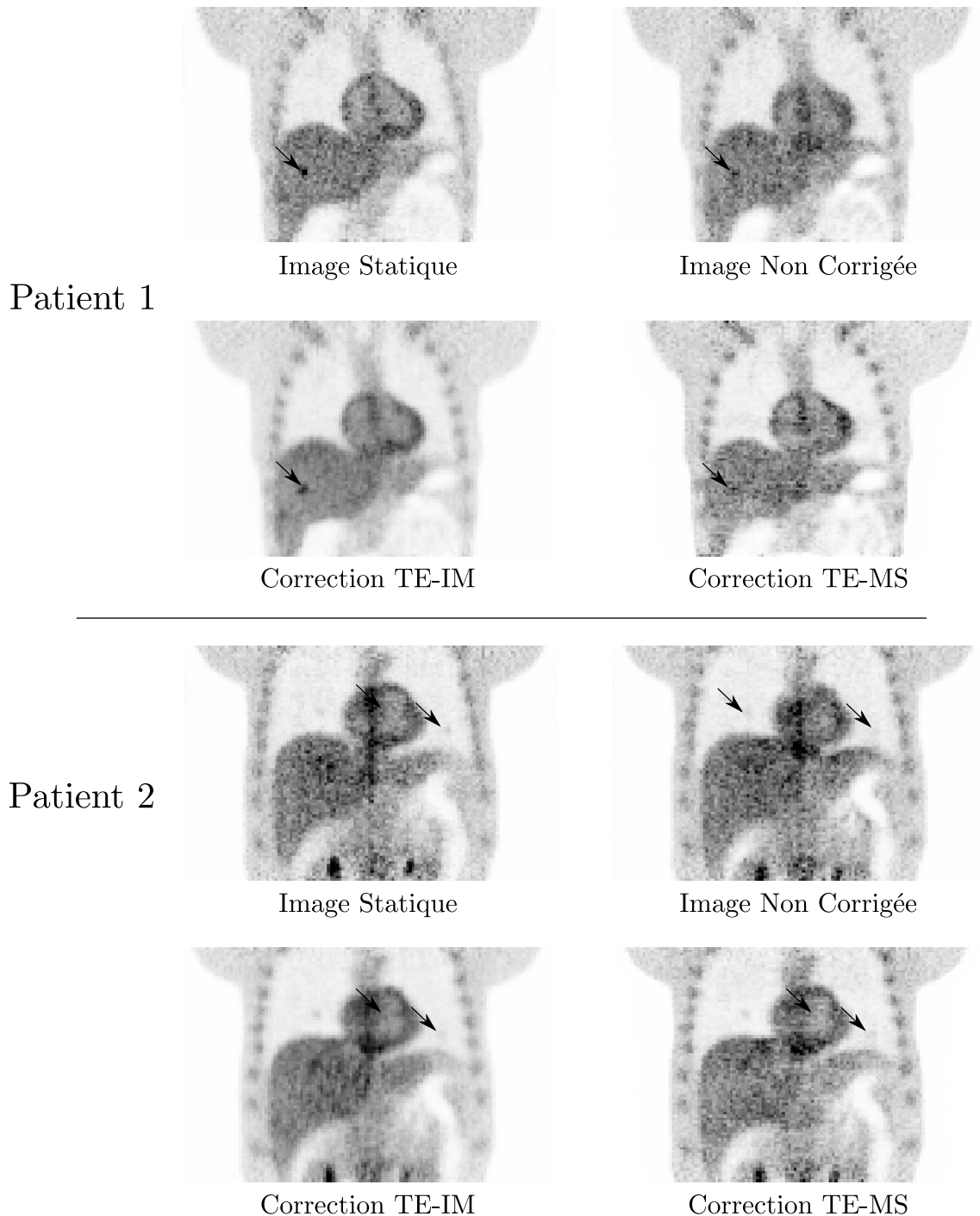


FIG. 12.8: Images reconstruites avec et sans correction du mouvement respiratoire (8 itérations et 5 sous-ensemble). Les flèches représentent une lésion du foie de diamètre 8 mm avec un contraste de 3.5 dans le patient 1, ainsi que deux lésions du poumon de 8 mm et de contraste 6.5 pour le patient 2.

12.4. SYSTÈME CAD

12.3.3 Images TE-MS

Ces images sont reconstruites en utilisant la méthode décrite en 6.5, qui prend en compte l'estimation de mouvement réalisée précédemment ainsi que les fichiers séquence générés par la simulation dynamique pour reconstruire directement les images corrigées.

La carte d'atténuation de l'instant de référence est déformée à l'aide des champs de mouvement pour générer 8 cartes d'atténuation. Les données de chaque instant du cycle respiratoire seront corrigées avec la carte d'atténuation correspondante.

Ces images sont reconstruites avec OPL-EM en utilisant 8 itérations et 5 sous-ensembles.

12.3.4 Images TE-IM

La méthode TE-IM est présentée en 6.3. Les 8 images correspondant aux 8 instants du cycle moyen sont reconstruites séparément à l'aide de la carte d'atténuation recalée correspondante. Chacune de ces images est reconstruite avec les mêmes paramètres que pour les autres types d'images, soit 8 itérations et 5 sous-ensembles. Ces 8 images sont ensuite déformées sur l'image correspondant au premier instant du cycle, à l'aide de la carte de mouvement estimée précédemment. Ces 8 cartes recalées sont ensuite sommées pour obtenir les images corrigées.

Bien que chaque image prise indépendamment soit très bruitée à cause du faible nombre d'évènements, la déformation suivie par la sommation des différentes images apporte une régularisation importante, comme on peut le voir dans la figure 12.8.

12.4 Système CAD

Le système CAD que nous utilisons a été développé à l'origine pendant les travaux de thèse de Sandrine Tomei ainsi que mes travaux de master [Tomei et al., 2008a, Lartizien et al., 2010]. Nous l'avons amélioré et adapté aux besoins de cette étude, notamment en développant les mesures de performances.

Le CAD utilise des informations fréquentielles obtenues par décomposition des images en ondelettes biorthogonale 4/4 non décimée (figure 12.9). Ces données sont utilisées par le système de classification basé sur un SVM travaillant voxel par voxel. Une étape de réduction des faux positifs est ajoutée par la suite.

Le nombre de tumeurs utilisées pour générer la base d'apprentissage est de 107 pour la base d'apprentissage poumon et 173 pour la base d'apprentissage foie. Ce nombre est relativement faible [Hua et al., 2005] par rapport à la taille du vecteur de caractéristiques utilisé pour le CAD (entre 24 et 32), ce qui nécessite de prendre en compte un maximum de données pour l'apprentissage et le test pour éviter le sous-apprentissage. Nous avons donc utilisé une méthode d'évaluation des performances

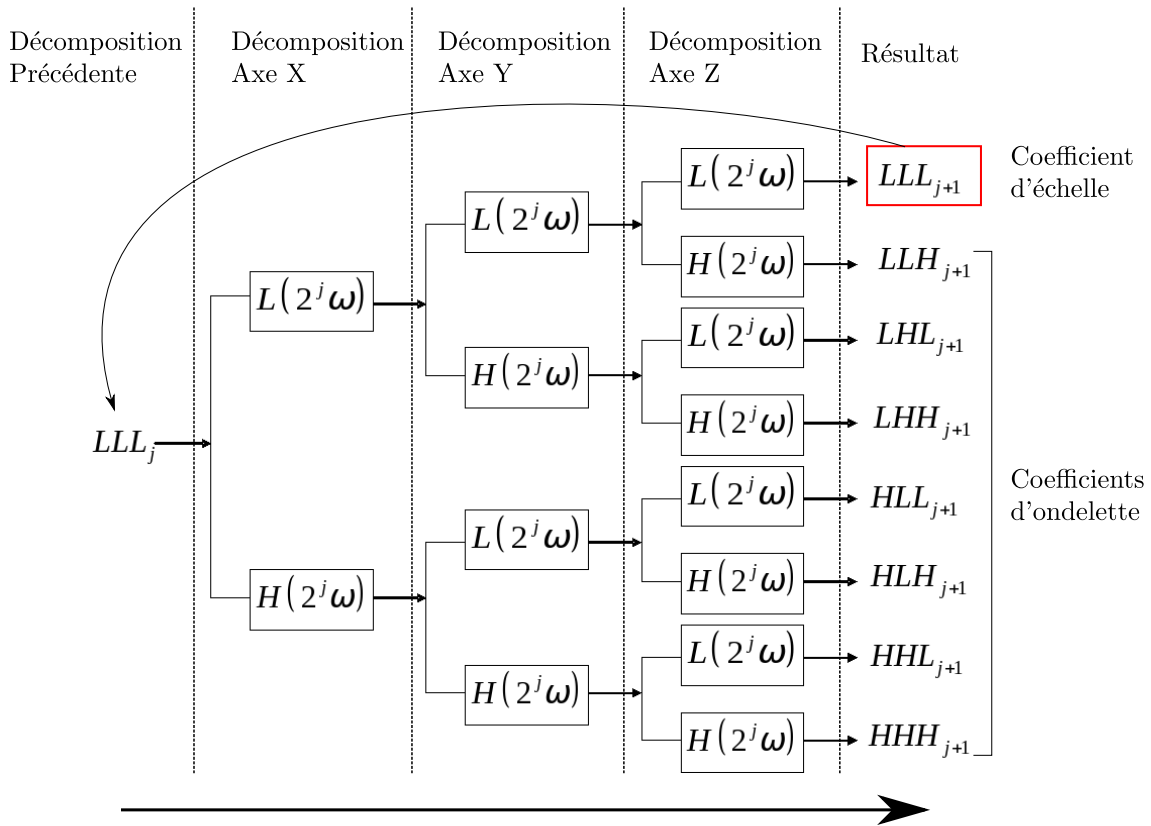


FIG. 12.9: Décomposition en ondelettes par banc de filtres : Chaque image est filtrée selon les 3 dimensions pour obtenir les coefficients d'ondelettes et d'échelle de chaque voxel de l'image. L correspond à un filtrage passe-bas tandis que H correspond à un filtrage passe-haut. Les coefficients d'échelle sont contenus dans l'image LLL_{j+1} tandis que les coefficients d'ondelettes sont présents dans les sept images de détail $LLH_{j+1}, LHL_{j+1} \dots HHH_{j+1}$

par resubstitution. Cela signifie que les apprentissages et les tests sont réalisés sur la même base de données constituée de 15 patients présentés dans le chapitre 10.

12.4.1 Vecteur de caractéristiques

Nous avons choisi d'utiliser une décomposition en ondelettes 3D non décimées par banc de filtres. Dans le cas tridimensionnel, la décomposition par banc de filtres est résumée par la figure 12.9. L'image de départ est traitée successivement dans les trois directions de l'espace par un filtre fréquentiel passe-haut correspondant à la fonction d'ondelettes (noté H) et passe-bas correspondant à la fonction d'échelle (noté L).

Ainsi sept images de détails (HHH, LLH, \dots) et une image d'approximation (LLL) sont produites pour chaque niveau j de décomposition. Les huit images du niveau suivant $j + 1$ sont générées de la même manière, mais en considérant l'image d'approximation LLL_j du niveau précédent comme image de départ. Les caractéristiques des images, rassemblées dans un vecteur descripteur de taille $8 \times j$, correspondent ici à l'ensemble de ces coefficients pour chaque voxel de l'image. Des exemples d'images

12.4. SYSTÈME CAD

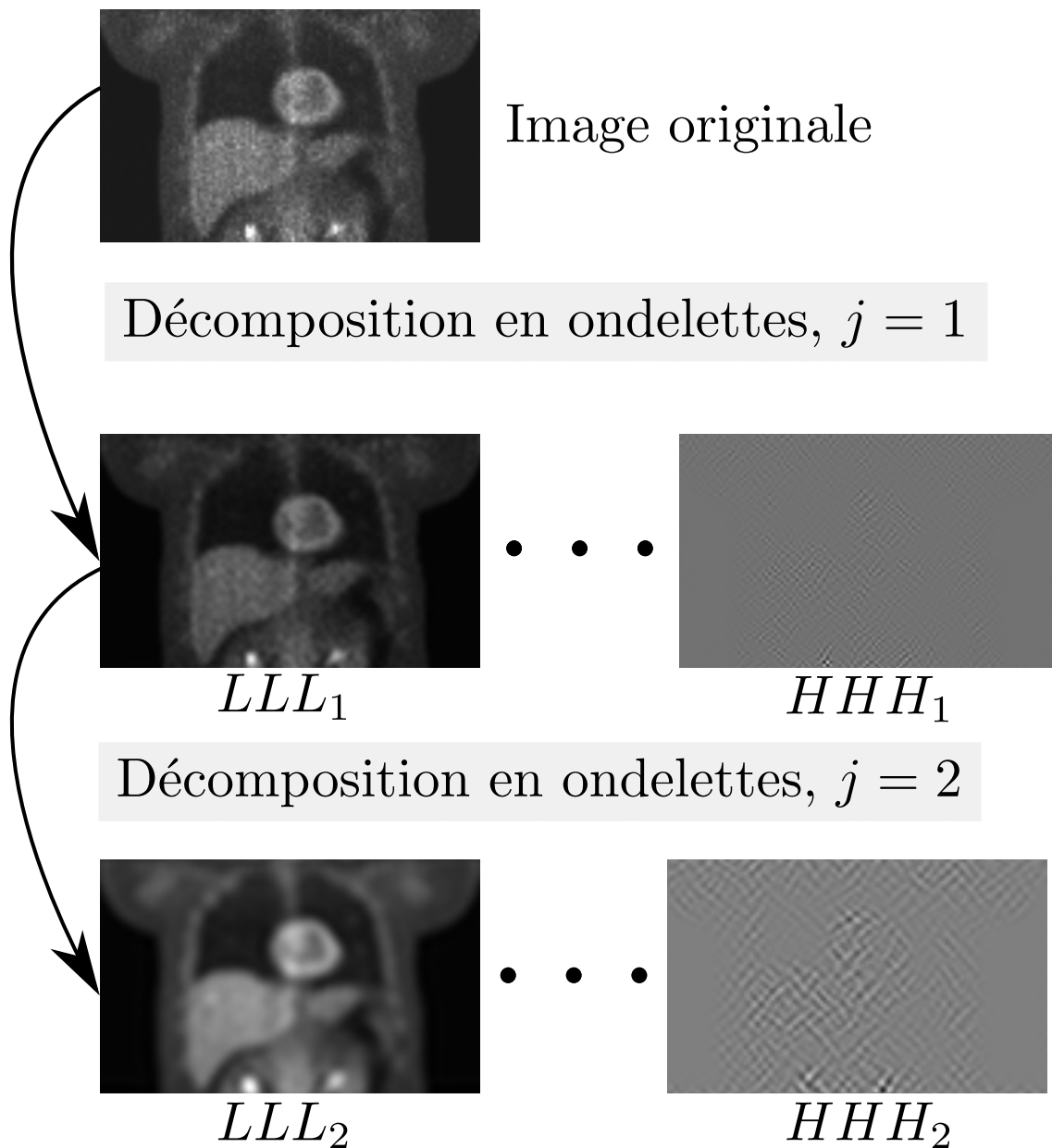


FIG. 12.10: Exemples d'images TEP décomposées en ondelettes. Sont représentées l'image originale, puis l'image des coefficients d'échelle notée LLL_j et une image des coefficients d'ondelettes notée HHH_j pour deux niveaux de décomposition. L'ondelette utilisée pour la décomposition est la biorthogonale 4/4

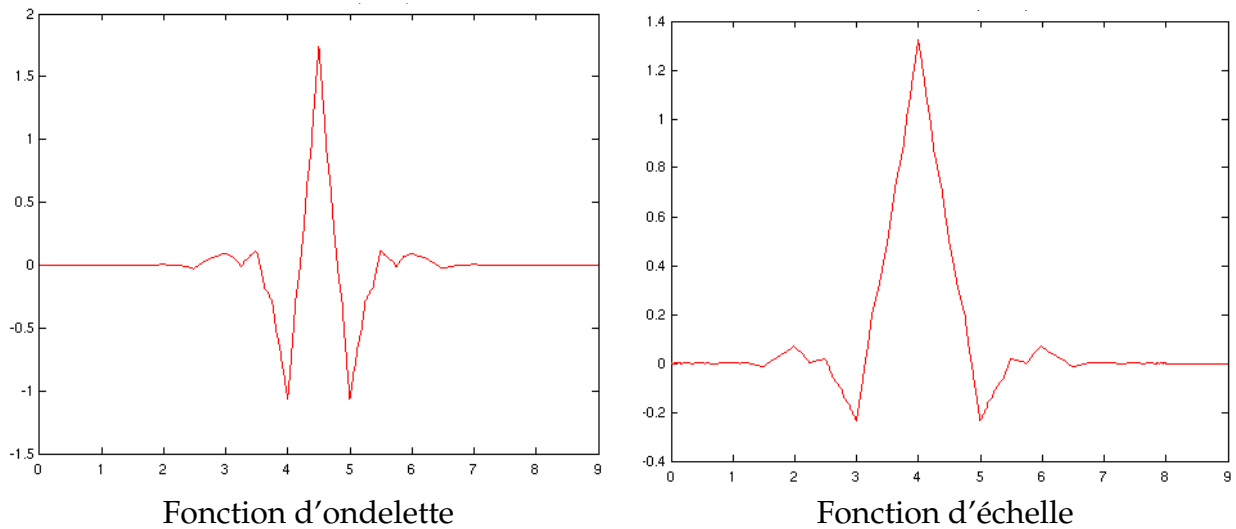


FIG. 12.11: *Fonction d'ondelette et d'échelle biorthogonale 4/4*

obtenues par les filtrages sont présentés dans la figure 12.10.

Nous utilisons l'ondelette biorthogonale 4/4 dont la fonction est présentée sur la figure 12.11.

12.4.2 Génération de la base d'apprentissage

Les données générées par la décomposition en ondelettes de chaque image se présentent sous la forme de volumes 4D, indiquant pour chaque voxel de l'image 3D l'ensemble des $8 \times j$ coefficients associés.

La base d'apprentissage sert à entraîner le classifieur dans un processus de classification supervisée, en lui fournissant un ensemble d'exemples avec leur classes associées (voir figure 8.3) page 98).

La génération de cette base demande l'extraction de points de la classe "pathologique", et de la classe "sain". Pour chaque tumeur de la vérité terrain, nous extrayons de l'image correspondante le vecteur de caractéristiques du voxel placée au centre de la tumeur. Ce vecteur de caractéristique correspond aux coefficients de la décomposition en ondelette du voxel de l'image ($8 \times j$).

Les points de la classe "sain" sont extraits de manière aléatoire dans les volumes de toutes les images (hors tumeurs). Pour des raisons de simplicité d'implémentation, un nombre fixe de points est extrait de chaque image.

12.4.3 Apprentissage de la Machine à Vecteur de Support (SVM)

La base de données d'apprentissage est utilisée par le classifieur SVM (Machine à Vecteur de Support) pour l'apprentissage.

12.4. SYSTÈME CAD

Le principe des SVM est de trouver l'hyperplan optimal de séparation dans l'espace des caractéristiques, qui va maximiser la marge de séparation entre les deux classes "pathologique" et "sain" (le SVM est défini plus en détail en 8.5). Cette marge correspond à la distance entre les plus proches vecteurs de caractéristiques appartenant à chacune des classes et l'hyperplan. La définition de cette marge et donc de l'hyperplan, se fait uniquement à partir de vecteurs de support, qui correspondent à l'enveloppe du groupement de point de chacune des deux classes.

Le SVM va calculer un modèle décrivant l'hyperplan permettant de séparer les données.

12.4.4 Génération des sites présumés

Une fois le classifieur entraîné, il est capable de classer rapidement les nouveaux vecteurs de caractéristiques. Ainsi, on lui soumet les données correspondants aux voxels des images pour qu'il les classe. On obtient en sortie un ensemble de cartes de score (une par image 3D), dans lesquelles chaque voxel correspond au score indiqué par le SVM. Ce score est une mesure de distance polarisée par rapport à l'hyperplan de séparation dans l'espace des points d'apprentissage, et donne, en plus de la classe, une mesure assimilable à la confiance du SVM vis-à-vis de sa classification.

Cette carte de score est ensuite seuillée pour fournir une carte binaire, indiquant pour chaque voxel la classe sélectionnée par le SVM pour ce niveau de seuil. Le mécanisme de sélection du seuil est défini en 12.5.3. Une exemple de carte de score est visible dans la figure 12.12.a.

À cette étape, le résultat est sous forme d'une carte de voxels étiquetés comme "sain" ou "pathologiques". Nous avons choisi de travailler sur des agrégats de points plutôt que directement sur les voxels car les cartes binarisées sont relativement bruitées (voir 12.12.b), et donc non directement exploitables. Les voxels "pathologiques" sont donc regroupés selon une 26-connexité (3x3x3). Un score est associé à chaque agrégat, correspondant au score le plus important observé dans l'agrégat.

12.4.5 Règles d'évaluation du résultat

Pour évaluer le résultat, nous avons élaboré des règles permettant de classer les agrégats à l'aide de la vérité terrain. Ce sont ces algorithmes qui sont utilisés dans les évaluateurs de performances présentés ci-après.

Les agrégats sont classés en LL (Lésion localisée) et NL (Non Lésion) selon qu'ils peuvent être considérés comme des vrais positifs ou des faux positifs :

Soit L l'ensemble des points de la lésion, A les points correspondant à l'amas candidat.

Les agrégats seront considérés comme des vrais positifs si ils intersectent une tumeur selon les règles décrites ci-après, ou comme des faux positifs dans le cas

contraire. Cependant, si leur taille est inférieure à la taille minimale définie par la première règle, l'amas n'est pas considéré.

Règles de classification :

1. $\text{card}(\mathbf{L} \cap \mathbf{A}) > \alpha \times \text{card}(\mathbf{L})$: où α définit la proportion minimale de la tumeur qui doit être présente dans l'amas. Elle permet d'éviter les amas qui intersecteraient la tumeur par accident.
2. $\text{card}(\mathbf{L} \cap \mathbf{A}) > \beta \times \text{card}(\mathbf{A})$: où β limite l'étendue de l'amas en dehors de la tumeur.

α et β sont des constantes empiriques fixées respectivement à 0.05 et 0.20 dans nos travaux. Nous avons choisi ces valeurs en visualisant les cartes de score pour obtenir des résultats corrects sans pénaliser les performances.

12.4.6 Résumé

Ce pseudo-code décrit les différentes étapes du CAD, depuis l'importation des images jusqu'à l'extraction des lésions potentielles. Il est illustré par la figure 12.12 :

1. Décomposition des images en ondelettes : pour chaque voxel de l'image d'origine, on obtient entre 8 et 32 coefficients suivant le niveau de décomposition de l'image, qui correspondent au vecteur de caractéristiques utilisé par le classifieur.
2. Extraction de la base d'apprentissage : les coefficients des centres de toutes les tumeurs sont extraits des volumes décomposés, et vont former la base d'apprentissage pathologique. Un certain nombre de voxels sont tirés aléatoirement dans les zones normales de chaque image et leurs coefficients sont ajoutés à la base saine.
3. Apprentissage : le classifieur SVM est entraîné sur cette base d'apprentissage pour générer le modèle qui sera utilisé pour le test.
4. Tests : le SVM entraîné est utilisé pour classer chaque voxel contenu dans les organes à évaluer (poumon et foie).
5. Réduction des faux-positifs : les points sont agrégés en composantes connexes (26-connexité en 3 dimensions).
6. Évaluation : lorsque nous avons accès à la vérité terrain, il est possible d'évaluer les performances du CAD en classant les agrégats en LL (Lésion Localisée) et LN (Lésion Non Localisée).

12.5 Optimisation des paramètres du système CAD

Les différentes étapes de l'évaluation des performances nécessitent de fixer un grand nombre de paramètres que nous allons détailler dans cette partie.

12.5. OPTIMISATION DES PARAMÈTRES DU SYSTÈME CAD

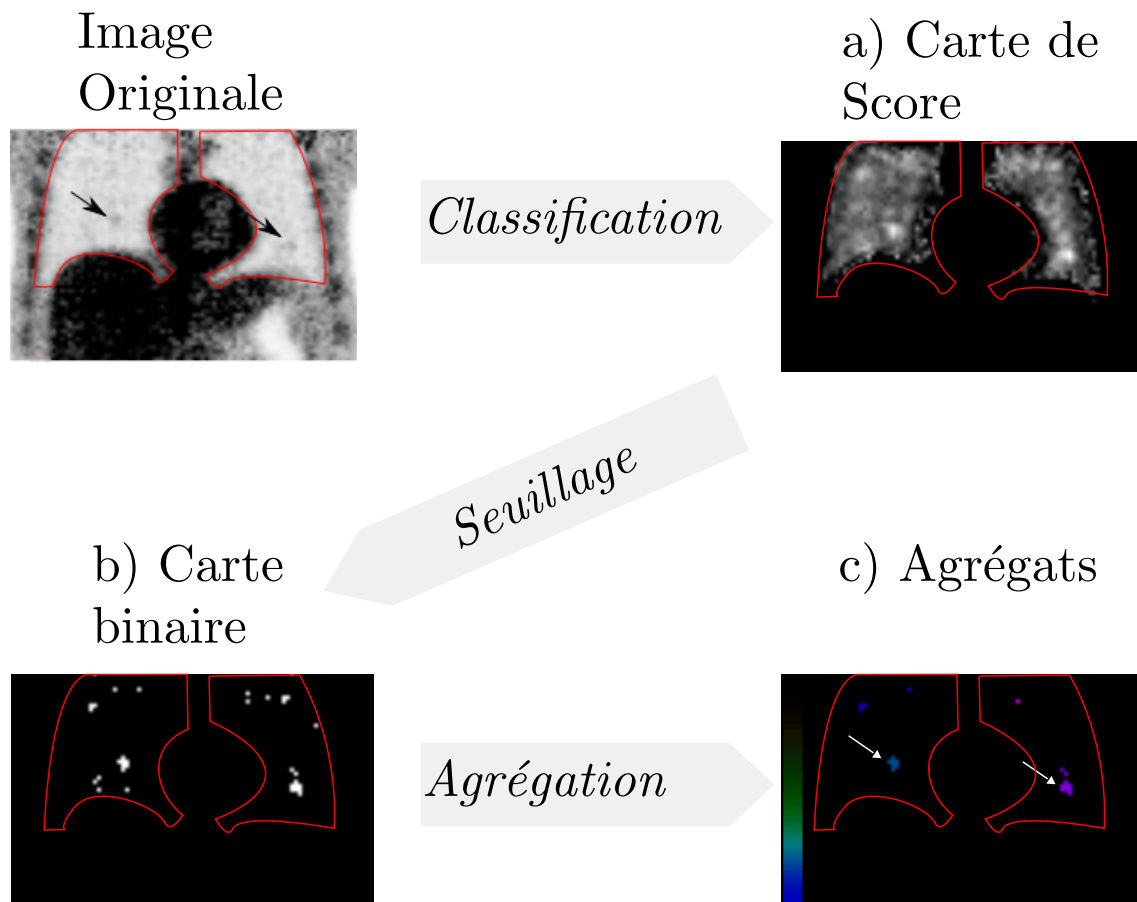


FIG. 12.12: Schéma du système CAD : l'image d'origine en haut à gauche est utilisée par le classifieur pour générer la carte de score. Cette carte est ensuite seuillée par l'étape b) pour générer une carte binaire correspondant aux sites qui dépassent un certain score s . L'étape c) correspond à la formation des agrégats qui seront le résultat du système CAD. Cette étape s'accompagne d'une suppression des sites de trop petite taille. Les flèches représentent des lésions dans l'image d'origine

12.5.1 Paramètres à optimiser

Paramètres de la base d'apprentissage

Nous avons identifié plusieurs autres paramètres qui, sont susceptibles d'influencer les performances de détection :

1. **Normalisation** : Le but de la normalisation est d'homogénéiser les plages de valeurs des différentes caractéristiques pour faciliter le travail du classifieur, dont les paramètres C et γ (définis ci-après) dépendent de la distance entre les points et ne permettent pas de gérer des différences trop importantes d'étendues dans les caractéristiques. Nous avons proposé de comparer deux méthodes de normalisation. La première consiste à modifier les données pour que la moyenne et l'écart-type aient une valeur respectivement de 0 et 1 ($(\mu, \sigma) = (0, 1)$). Cette méthode est notée "*moyenne*" dans la suite du texte. La seconde méthode consiste à adapter des données pour que l'ensemble des valeurs soit comprises entre -1 et +1. Cette méthode est notée "*écart*". La première méthode (*moyenne*) a l'avantage d'être relativement peu sensible aux valeurs extrêmes, contrairement à la seconde (*écart*).
2. **Nombre de points de la base d'apprentissage** : Le nombre de point de la base d'apprentissage détermine directement la qualité du modèle. Le nombre de caractéristiques de chaque exemple de la base est de $8 \times j$, avec j le niveau de décomposition des images. Il n'existe pas de règle définitive pour choisir le nombre d'exemples nécessaires en fonction du nombre de caractéristiques, mais l'on considère cependant que le nombre d'échantillons de la base d'apprentissage doit être largement supérieur au nombre de caractéristiques pour éviter le sur-apprentissage. Cependant, les SVM sont relativement efficaces pour éviter ce problème. Dans notre étude, le nombre de cas pathologiques est de 173 pour le poumon et de 107 pour le foie. Il faudrait idéalement sélectionner autant de cas "sains" que de cas pathologiques pour chaque organe. Cependant le nombre total de cas risque d'être trop faible par rapport au nombre de caractéristiques. Nous avons par conséquent analysé l'influence du nombre de cas sur les performances de détection en considérant trois valeurs : 100 points normaux par images (pts/im.) (soit 1500 pts. négatifs), 200 pts/im. (soit 3000 pts. négatifs) et 1000 pts/im. (soit 15000 pts. négatifs).
3. **Positions des points extraits** : Les points normaux extraits des images pour alimenter la base vont avoir une influence directe sur la qualité des résultats. Idéalement ils devraient être représentatifs de l'ensemble des cas rencontrés dans la base de tests, néanmoins les bords de certains organes ont un profil proche de celui des tumeurs, et rendent l'estimation de la surface de séparation plus difficile. Nous avons donc voulu évaluer la performance du CAD sur une base dépourvue de ces données ambiguës. Pour cela nous avons réalisé une érosion de 2 voxels sur les masques des volumes sains à extraire, afin de ne pas prendre en compte les frontières des organes.

12.5. OPTIMISATION DES PARAMÈTRES DU SYSTÈME CAD

Paramètres du classifieur

Le classifieur (SVM) utilise les données d'apprentissage formatées selon les choix opérés précédemment (normalisation, volumes utilisés pour extraire les points de la base de points "sain" et nombre de ces points) pour générer un modèle de prédiction. Cependant, cet algorithme dispose de paramètres intrinsèques, qui sont beaucoup plus dépendants des données d'apprentissage (voir 8.5).

Niveau de la décomposition en ondelettes j : Ce paramètre détermine la taille du vecteur de caractéristiques. Il correspond au nombre de niveaux de décomposition pris en compte par le SVM. Les niveaux de décomposition élevés (j grand) correspondant à des informations de très basse fréquence, les informations qu'ils apportent ne sont pas forcément pertinentes pour la détection des lésions. Cependant, les caractéristiques fréquentielles des images générées par les différents jeux d'images étant différentes, il est nécessaire d'adapter ce paramètre à chaque type d'images.

Coefficient de pénalisation C : lors du calcul de l'hyperplan de séparation des données, chaque point mal classé va pénaliser la surface selon un facteur proportionnel à C , comme indiqué dans le paragraphe 8.5.

Largeur de bande γ : cette valeur influe directement la largeur de bande du noyau utilisé par le classifieur (Fonction de Base Radiale gaussienne, ou RBF), comme indiqué dans la description du SVM présentée au paragraphe 8.5.

12.5.2 Méthodes d'optimisation

Pour chaque combinaison des paramètres de la base d'apprentissage définis en 12.5.1, nous avons déterminé le triplet optimal de paramètres (C, γ, j) à l'aide d'une recherche exhaustive par grille. Cela correspond à évaluer les performances sur le produit cartésien d'une discrétisation des valeurs de chaque paramètre. Cette étude a été réalisée sur le jeu d'images statiques afin de ne pas être influencé par les éventuels artefacts des autres types d'images.

Par exemple, si nous disposons de deux paramètres A et B , et que nous recherchons le meilleur jeu de paramètres, il faut tout d'abord sélectionner pour chaque paramètre la taille de la zone de recherche. Supposons par exemple que les valeurs de A soient des valeurs entières comprises entre -1 et 2 , et que B puisse prendre les valeurs 100 , 10000 et 50000 . Les critères de performance seront évalués pour toutes les combinaisons des valeurs de A et B , tels que $(-1, 100)$, $(-1, 10000)$, $(-1, 50000)$, $(0, 100)$, ..., $(2, 50000)$. Le jeu de paramètres ayant obtenu les meilleures performances sera donc sélectionné.

Les critères de performance que nous avons sélectionnés sont la sensibilité et la spécificité obtenue pour une validation croisée à 5 éléments réalisée sur la base d'apprentissage comme présenté sur la figure 12.13. La validation croisée permet de réduire le biais sur les performances [Varma and Simon, 2006].

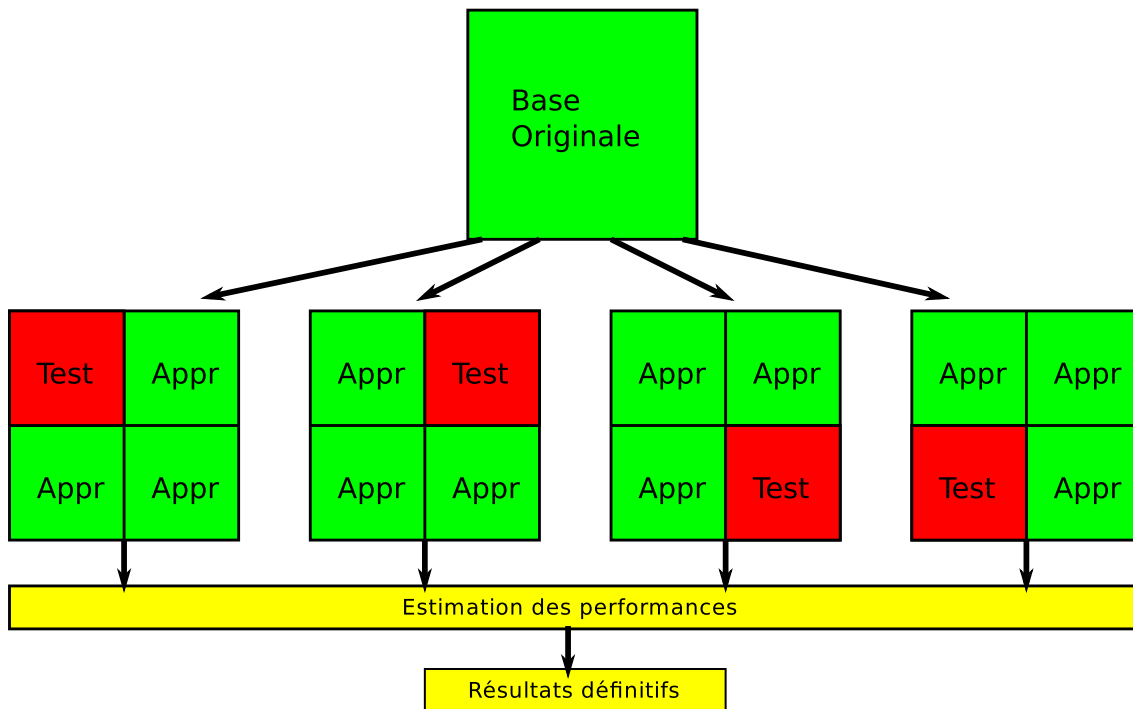


FIG. 12.13: Réalisation d'une validation croisée à n éléments (ici 4) : La base d'exemples d'origine est décomposée en n parts égales. La mesure de performance se fait n fois, avec à chaque fois un apprentissage sur $n - 1$ éléments et un test sur l'élément restant. La mesure de performance est réalisée sur les résultats des n tests.

Pour choisir les meilleurs paramètres du classifieur, j'ai effectué une recherche exhaustive par grille avec les paramètres suivants :

C : de 1 à 10000 en 15 pas logarithmiques.

γ : de 0.0001 à 1 en 15 pas logarithmiques.

j : de 1 à 4, soit de 8 à 32 caractéristiques.

L'optimisation a été réalisée à l'aide du logiciel rapid-i [Mierswa et al., 2006] pour chaque type d'image. Les indicateurs de performance calculés par validation croisée sont les suivants : sensibilité, spécificité et précision définis dans le chapitre 7 (section 7.1). Le triplet de paramètres retenu est celui qui maximise la sensibilité.

Nous avons représenté pour chaque jeu d'image le nuage de points correspondant à la répartition de chaque triplet (C, γ, j) dans un espace à deux dimensions ("Sensibilité", "Spécificité"). De ce nuage de points nous pouvons voir le front de Pareto. Ce type de diagramme permet de rechercher un optimum selon plusieurs critères antagonistes. Dans notre cas, nous voulons à la fois une sensibilité et une spécificité importante, sachant qu'il n'existe pas de jeu de paramètres "parfaits" qui permettent d'avoir 100% aux deux. Dans notre cas, le front de Pareto va permettre de vérifier que le choix par maximisation de la sensibilité ne se fait pas au détriment de la spécificité.

12.6. COMPARAISON DES PERFORMANCES DE DÉTECTION DES DIFFÉRENTES MÉTHODES DE CORRECTION DU MOUVEMENT RESPIRATOIRE

12.5.3 Mesure des performances de détection

Les performances de détection obtenues par la méthode de validation croisée considèrent uniquement le voxel central de la tumeur pour déterminer si la lésion est correctement détectée. Ces performances ne traduisent donc pas forcément le comportement d'un observateur humain, qui détecte plutôt des agrégats. Nous avons donc proposé d'estimer les performances en considérant la notion d'agrégats. Pour un seuil donné, le CAD génère un ensemble d'agrégats associés à un score, comme présenté en 12.4.4. Il est ensuite possible d'évaluer la performance de ce CAD à l'aide de courbes F-ROC décrites en 7.3.1 construites à l'aide de la vérité terrain.

Le seuil sélectionné s pour générer les agrégats est celui qui maximise la Fraction de Localisation de Lésion (FLL : nombre de lésions localisées divisé par le nombre total de lésions), c'est à dire celui qui va permettre de détecter un maximum de lésions.

Cette sélection est réalisée selon l'algorithme suivant :

Le processus d'estimation de la sensibilité pour un seuil s se fait de la manière suivante. Les cartes de scores sont binarisées en fonction de s . Les agrégats sont estimés sur ces cartes binaires, en prenant en compte les informations provenant des cartes de score correspondantes pour attribuer un score à chaque agrégat. Ils sont ensuite classés en "Lésion localisée" (LL) ou "Non Lésions" (NL) à l'aide de la vérité terrain, selon les règles présentées en 12.4.5. Une valeur de sensibilité est calculée à partir de ces informations pour le seuil s .

Le seuil optimal retenu est celui qui maximise la sensibilité. Nous recherchons ce maximum pour 40 valeurs réparties de manière uniforme entre -2 et +2. Nous utilisons une recherche exhaustive car le résultat possède de nombreux minima locaux et le temps de calcul est faible (moins d'une heure pour les 40 valeurs).

Ce critère de sélection va naturellement engendrer un grand nombre de faux positifs, mais il faut garder à l'esprit qu'il sera utilisé pour réaliser des courbes F-ROC, qui indiquent une spécificité pour chaque nombre de faux positifs en jouant sur un second seuil, toujours supérieur à celui retenu pour extraire les agrégats.

Le second critère utilisé pour comparer les bases d'apprentissage est la figure de mérite extraite de l'analyse JAFROC décrite en 7.3.3. Elle va comparer pour chaque image le score du faux positif le plus haut avec les scores des vrais positifs de l'image. Pour cette étude, nous avons utilisé la version 4.0 du logiciel JAFROC.

12.6 Comparaison des performances de détection des différentes méthodes de correction du mouvement respiratoire

L'optimisation des paramètres décrits dans la section 12.5.1 nous a permis de fixer les paramètres de la base d'apprentissage (normalisation, nombre de cas sains, localisation des cas sains).

CHAPITRE 12. MÉTHODES

Afin de comparer les performances des quatre séries d'images (statique, non corrigée, TE-IM et TE-MS) nous avons dans un premier temps réalisé une nouvelle optimisation des paramètres (C, γ, j) pour chaque type d'images en suivant la méthode décrite en 12.5.2. Puis nous avons comparé les courbes F-ROC et réalisé une étude JAFROC comme décrit dans la section 12.5.3

Chapitre 13

Analyse des résultats

Dans ce chapitre, nous détaillons les performances de détection obtenue par les quatre séries d'images que nous avons comparées : TE-IM, TE-MS, Statiques et Non Corrigées. Nous commençons par présenter les courbes obtenues lors de l'étape d'optimisation et d'adaptation du CAD aux données, puis nous parlons des performances de détection obtenues par ce CAD sur les lésions hépatiques et pulmonaires.

Dans tous les cas, l'estimation des performances se fait de la manière suivante :

- Optimisation des paramètres du classifieur au jeu de données.
- Comparaison des courbes Free-ROC.
- Comparaison des Figures de Mérite JAFROC.
- Conclusion.

13.1 Optimisation des paramètres

Nous avons réalisé des mesures de performances pour les différentes valeurs des paramètres suivants :

- Nombre de points de la base d'apprentissage (100, **200**, 1000).
- Normalisation des données (**moyenne**, écart).
- Position des points de la base d'apprentissage (**organe complet**, organe avec érosion).

La recherche des paramètres optimaux a été réalisée sur le foie, pour la série d'images statiques, afin de ne pas être influencé par les artefacts des méthodes de correction du mouvement respiratoire. Pour sélectionner le jeu de paramètre optimal, nous avons créé une base **Témoin** comprenant les valeurs en gras des paramètres présentés ci-dessus. Toutes les autres bases reprennent les valeurs de la base Témoin en modifiant un seul paramètre. Nous utilisons par la suite les termes définis ci-dessous pour qualifier les 5 jeux de paramètres :

Base Témoin : contient des données normalisées par la méthode "moyenne" avec 200 points sains extraits de chaque ensemble du volume des organes.

CHAPITRE 13. ANALYSE DES RÉSULTATS

Base Érodée : contient des données normalisées par la méthode "moyenne" avec 200 points sains extraits du volume de chaque organe érodé (érosion morphologique de 2 voxels).

Base Appauvrie : contient des données normalisées par la méthode "moyenne" avec 100 points sains extraits de l'ensemble du volume de chaque organe.

Base Enrichie : contient des données normalisées par la méthode "moyenne" avec 1000 points sains extraits de l'ensemble du volume de chaque organe.

Base Normalisée Écart : contient des données normalisées par la méthode "écart" avec 200 points sains extraits de l'ensemble du volume de chaque organe.

Pour vérifier que les résultats de la sélection des paramètres du classifieur sont stables, une seconde base témoin (**Témoin 2**) a été générée avec les données de seulement 14 images sur les 15 que nous avons simulées. Les points "sains" ne sont pas non plus extraits aux mêmes endroits que pour la base **Témoin**. Nous allons donc vérifier que les paramètres optimaux du CAD pour un jeu de données ne sont pas dépendants de la base.

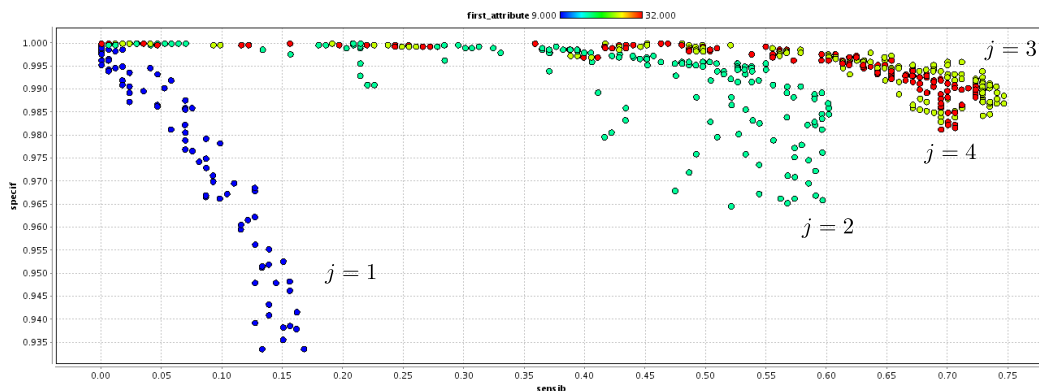
13.1.1 Sélection des meilleurs paramètres du classifieur

Les paramètres du classifieur sont déterminés par une recherche par grille. Elle consiste à rechercher l'optimum en évaluant la performance de chaque jeu de paramètre dans un ensemble déterminé à l'avance. La performance de chaque triplet (C, γ, j) est estimée en réalisant une validation croisée à 5 sous-ensembles sur l'ensemble de la base d'apprentissage.

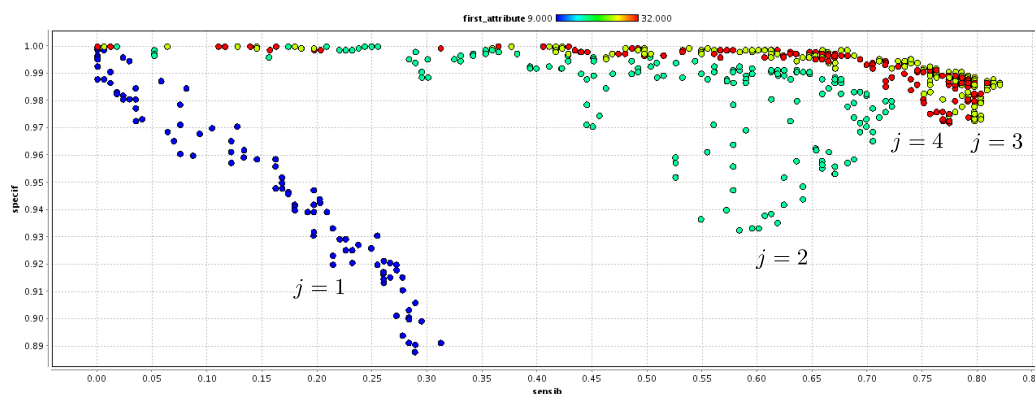
Les résultats sont reportés sur les figures 13.1 et 13.2 qui tracent, pour chaque base, les variations de spécificité en fonction de la sensibilité pour chaque triplet (C, γ, j) évalué sur le poumon. Les paramètres optimaux sont choisis à partir du front de Pareto des figures 13.1 et 13.2 en maximisant la sensibilité.

Dans l'ensemble, on peut voir clairement que les points correspondant au premier niveau de décomposition (bleu foncé) ont une performance systématiquement inférieure aux autres. Pour la base **Témoin**, la performance maximale est atteinte pour environ 15% de sensibilité et une spécificité de 93.5%, ce qui correspond à la valeur de sensibilité la plus faible de tous les points de la base **Témoin**. On observe cependant un front de Pareto marqué, bien qu'en fort retrait par rapport aux autres niveaux de décomposition. Ce même constat se retrouve pour les bases **Appauvrie** (1500 points d'apprentissage) et **Enrichie** (15000 points d'apprentissage). Les trois autres bases montrent des comportements différents. Pour la base **Normalisée Écart**, les performances pour $j = 1$ s'effondrent, de ce fait le classifieur est quasiment incapable de discerner les classes. Cela indique qu'il ne parvient pas à trouver une surface de séparation des données avec les paramètres indiqués. Puisque la normalisation moyenne parvient à mieux séparer les données pour ce niveau de décomposition, il est probable que la normalisation **Écart** ne parvienne pas à homogénéiser les valeurs des différentes dimensions de manière satisfaisante. Dans le cas de la base **Érodée**, la simplification du problème de classification fait que les performances sont nettement

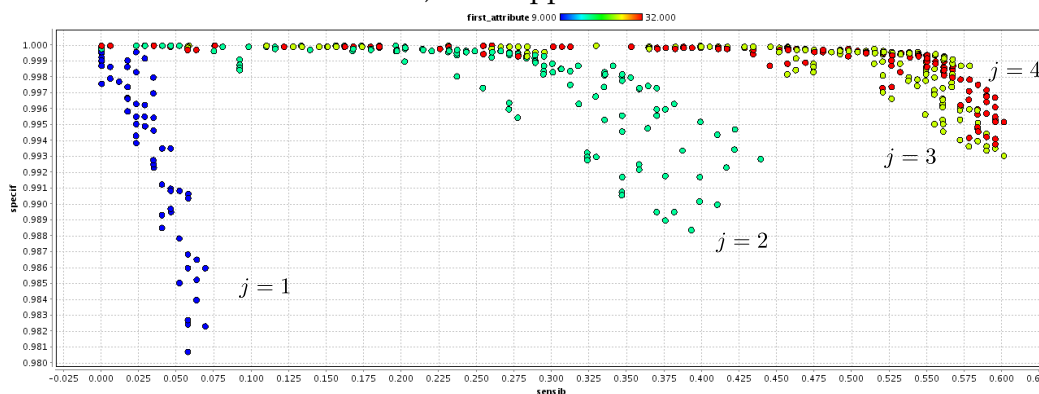
13.1. OPTIMISATION DES PARAMÈTRES



a) Base Témoin



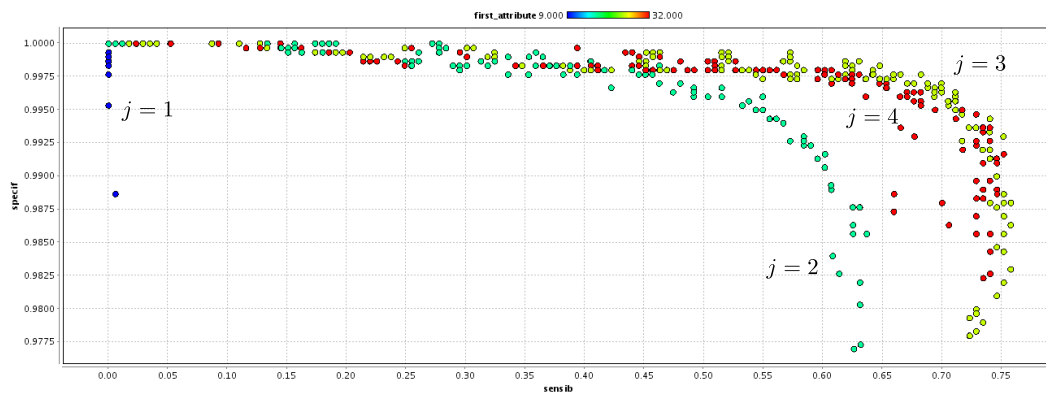
b) Base Appauvrie



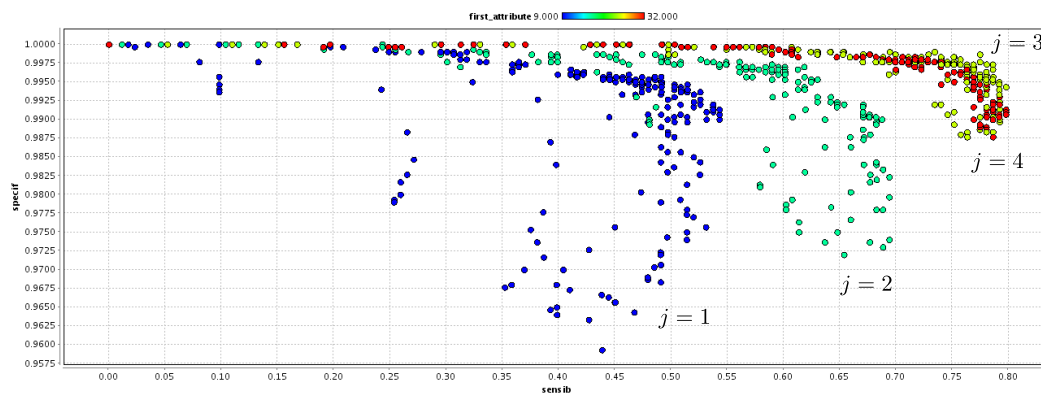
c) Base Enrichie

FIG. 13.1: (1/2) Fronts de Pareto des résultats de la recherche des meilleurs paramètres du classifieur. Pour chaque triplet de paramètres (C, γ, j) , la sensibilité et la spécificité sont reportées sur le graphique. Le code de couleurs correspond à la valeur de j . Bleu correspond à $j = 1$, turquoise à $j = 2$, vert à $j = 3$ et rouge à $j = 4$. En a), la base **Témoin**, avec 200 points négatifs par image et une normalisation moyenne, en b) la base **Appauvrie** avec 100 points négatifs par image et une normalisation moyenne, et en c) la base **Enrichie** avec 1000 points négatifs par image et une normalisation moyenne.

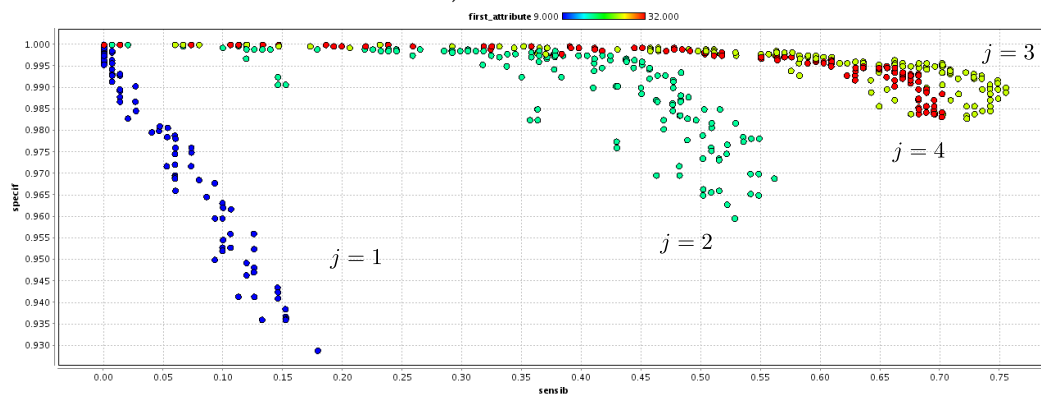
CHAPITRE 13. ANALYSE DES RÉSULTATS



a) Base Normalisée Écart



b) Base Érodée



c) Base Témoin 2

FIG. 13.2: (2/2) Fronts de Pareto des résultats de la recherche des meilleurs paramètres du classifieur. Pour chaque triplet de paramètres (C, γ, j) , la sensibilité et la spécificité sont reportées sur le graphique. Le code couleur correspond à la valeur de j . En a) la base **Normalisée Écart** avec 200 points négatifs par image et une normalisation écart. En b), la base **Érodée**, avec 200 points négatifs par image et une normalisation moyenne, en c) la base **Témoin 2**, réalisée de la même manière que la base **Témoin** mais en retirant une image.

améliorées : 55% de sensibilité pour 99.2% de spécificité, ce qui est le score le plus élevé

13.1. OPTIMISATION DES PARAMÈTRES

de toutes les bases.

Les performances du second niveau de décomposition (bleu ciel) sont toujours situées environ à mi-chemin entre les performances du premier niveau et celles des niveaux 3 et 4.

Les performances des décompositions de niveaux 3 (vert) et 4 (rouge) sont systématiquement meilleures que les deux premières mais sont très variables. La base **Témoin** et la base **Appauvrie** montrent une bonne avance tant en terme de sensibilité que de spécificité pour 3 niveaux de décomposition. La base **Érodée** ne montre, quand à elle, aucune différence de spécificité entre ces deux niveaux de décomposition, mais montre cependant une faible amélioration de la sensibilité pour les 4 niveaux (0.5%). Quant à la base *Normalisée Écart*, on n'observe pas de réelles différences de performance entre les deux niveaux de décomposition.

La table 13.1 référence tous les paramètres sélectionnés à partir des courbes de Pareto. On peut observer que la sensibilité la plus importante est atteinte pour la base **Appauvrie**, avec 82% de bonne détection. Ce taux est sensiblement le même que celui de la base **Érodée** (80%). Ces valeurs importantes, par rapport aux autres, peuvent s'expliquer par le fait que ces deux bases proposent un problème simplifié. Dans le cas de la base **Érodée**, les cas litigieux (discontinuités proches des bords des organes) ont été retirés de la base, ce qui simplifie le problème, tandis que pour la base **Appauvrie**, c'est le nombre de point "sain" plus faible (1500 contre 3000) qui permet de rendre le problème plus simple à traiter par le classifieur au détriment de la généralisation du résultat obtenu.

Il est intéressant de constater que les performances en sensibilité pour les bases **Appauvrie**, **Témoin** et **Enrichie** sont inversement proportionnelles à la taille de la base. En effet, plus la complexité de la base est importante, plus il devient difficile de trouver une surface de séparation efficace. Cependant, une base trop simpliste va engendrer une solution qui sera sans rapport avec la réalité, comme nous le verrons plus loin avec les courbes F-ROC.

Les valeurs de spécificité sont toutes très supérieures à 99%, ce qui montre que le classifieur n'a pas de problème pour classer correctement les points sains. En effet, la base étant très déséquilibrée, avec un rapport de 17 points sains pour 1 point "lésion" dans la base **Témoin**, il est normal que le classifieur favorise la classification des points "sains". Il existe des techniques classiques permettant de compenser ce déséquilibre lors de l'apprentissage, notamment le choix d'un paramètre de pénalisation C différent pour chacune des classes, mais tous les tests que nous avons réalisés n'ont pas montré d'amélioration du résultat. En effet, l'étape de sélection du seuil (voir section suivante) permet de compenser ces différences.

Il est intéressant d'observer que les fronts de Pareto de la base **Témoin 2** sont très semblables à ceux de la base **Témoin** pour une décomposition au troisième niveau. Des disparités apparaissent pour le quatrième niveau, mais les performances optimales sont obtenues pour le même jeu de paramètre. Cela semble indiquer que la base d'apprentissage est suffisamment complète en exemples de lésions.

CHAPITRE 13. ANALYSE DES RÉSULTATS

	Base Témoin	Base Érodée	Base Appauvrie	Base Enrichie	Base Normalisée Écart
C	464	74	5412	5412	10000
γ	0.0053	0.0094	0.00031	0.0017	0.052
j	3	3	3	4	3
Sensibilité	0.75	0.80	0.82	0.60	0.76
Spécificité	0.99	0.99	0.99	0.99	0.99
Précision	0.98	0.98	0.97	0.99	0.98

TAB. 13.1: Paramètres (C, γ, j) sélectionnés pour l'optimisation des performances sur les différentes bases. Sont indiqués pour chaque base le triplet de paramètres sélectionnés ainsi que sa position sur le front de Pareto.

13.1.2 Courbe Free-ROC

Les courbes Free-ROC de la figure 13.3 permettent de comparer les performances du CAD sur les différentes bases d'apprentissage. Les courbes ont volontairement été tronquées à 40 faux positifs par image, car ce nombre est déjà trop important pour un système CAD.

On peut observer que les courbes correspondant aux bases **Témoin** et **Enrichie** atteignent leur maximum de sensibilité pour un nombre de faux positifs moyen par image (NFM) relativement faible par rapport aux autres bases : entre 17 et 20 faux positifs pour ces bases, contre plus de 40 pour les bases **Appauvrie** et **Érodée**. Cela tend à montrer que le système CAD est plus performant pour ces bases car il crée moins d'agrégats là où il n'y a pas de lésions.

En ce qui concerne la sensibilité maximale obtenue sur les courbes, elle est atteinte pour la base **Témoin** avec environ 62% de sensibilité, suivie par la base **Appauvrie** avec 60%, mais pour un nombre de faux positifs beaucoup plus important (38 contre 18 pour la base **Témoin**). La troisième courbe est la courbe **Normalisation Écart**, suivie par la base enrichie puis la base **Érodée**. Il est important de noter que les sensibilités maximales observées sont très proches, entre 55% et 62%, ce qui indique que la qualité de la base d'apprentissage n'a pas d'impact réel sur la sensibilité maximale atteinte par le CAD, mais qu'il pourra être plus ou moins difficile pour ce dernier de différencier les lésions du bruit de fond.

En pratique, on choisira le seuil pour avoir la certitude d'avoir un nombre de faux positifs "raisonnable" par image. Dans notre cas, les images contiennent environ 10 lésions par image. Il peut être intéressant de comparer les performances des bases pour un ratio de 1 faux positif par lésion, soit 10 faux positifs. Dans ce cas, la base **Témoin** a des performances très semblables avec celles de la base **Enrichie**, à environ 55%, ce qui est déjà très proche de leurs performances maximales. La base **Normalisation Écart** et la base **Appauvrie** sont, quant à elles, à 40% de sensibilité, tandis que la base **Érodée** atteint 35%.

13.1. OPTIMISATION DES PARAMÈTRES

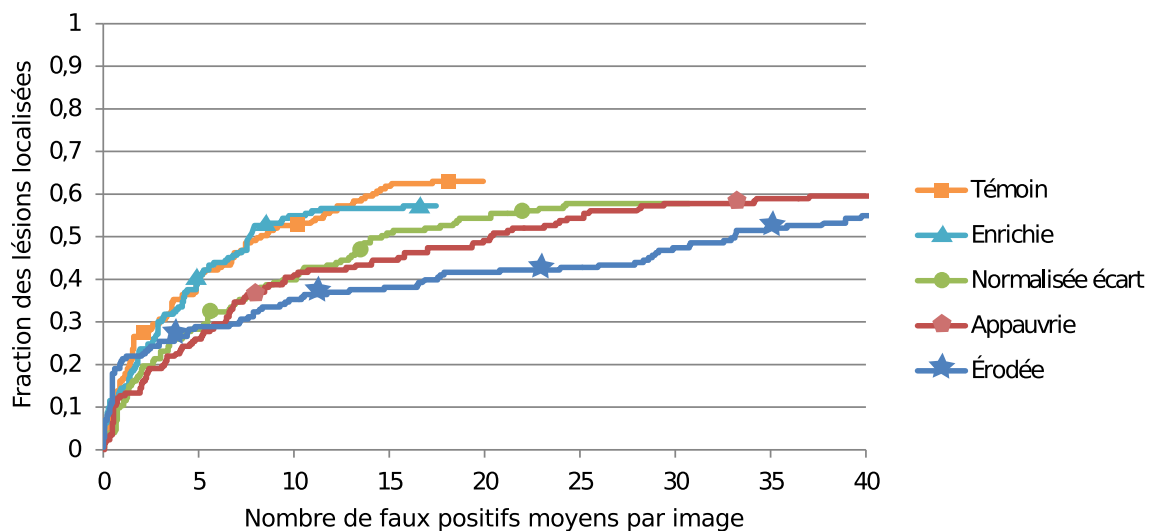


FIG. 13.3: Courbes Free-ROC comparant les performances du CAD sur une base **Témoïn** (normalisation moyenne et 200 points négatifs par image), sur une base **Enrichie** (1000 points négatifs par image), sur une base **Appauvrie** (100 points négatifs par image), sur une base **Normalisée Écart** (normalisation entre -1 et +1 et 200 points négatifs par image) et enfin sur une base de 100 points négatifs par image mais dont les volumes ont été érodés de 2 voxels.

13.1.3 Comparaison des performances JAFROC

La comparaison des performances obtenues par l'algorithme JAFROC [Chakraborty and Winter, 1990] de la figure 13.4 nous montre les FDM (Figure de Mérite) obtenues pour les différentes bases par l'algorithme JAFROC. Les FDM des bases **Témoïn**, **Érodée** et **Enrichie** sont quasiment au même niveau (0.18), mais les barres d'erreurs semblent montrer un léger avantage pour la base **Témoïn**.

Les bases **Normalisée Écart** et **Appauvrie** quand à elles ont une FDM de 0.1 environ, ce qui indique une performance plus faible que les autres.

D'un point de vue statistique, la p-valeur (voir 7.2) calculée par le logiciel est de 0.049, ce qui ne permet pas de pouvoir annoncer avec une fiabilité de 95% que le test est significatif, c'est à dire que les FDM sont effectivement toutes différentes. Cela se vérifie aisément en regardant l'étendue des barres d'erreur. Mais il faut noter que cette FDM est basée sur une méthode avec une puissance statistique faible [Chakraborty and Berbaum, 2004], ce qui signifie qu'elle sous-estime la p-valeur.

13.1.4 Conclusion

Nous avons vu que tous les indicateurs montrent que le maximum de performance est apporté par la base **Témoïn**. De plus, les performances relativement proches de la base **Témoïn 2** semblent indiquer que les performances sont stables. Nous allons donc conserver les paramètres de cette base pour la comparaison des différents types

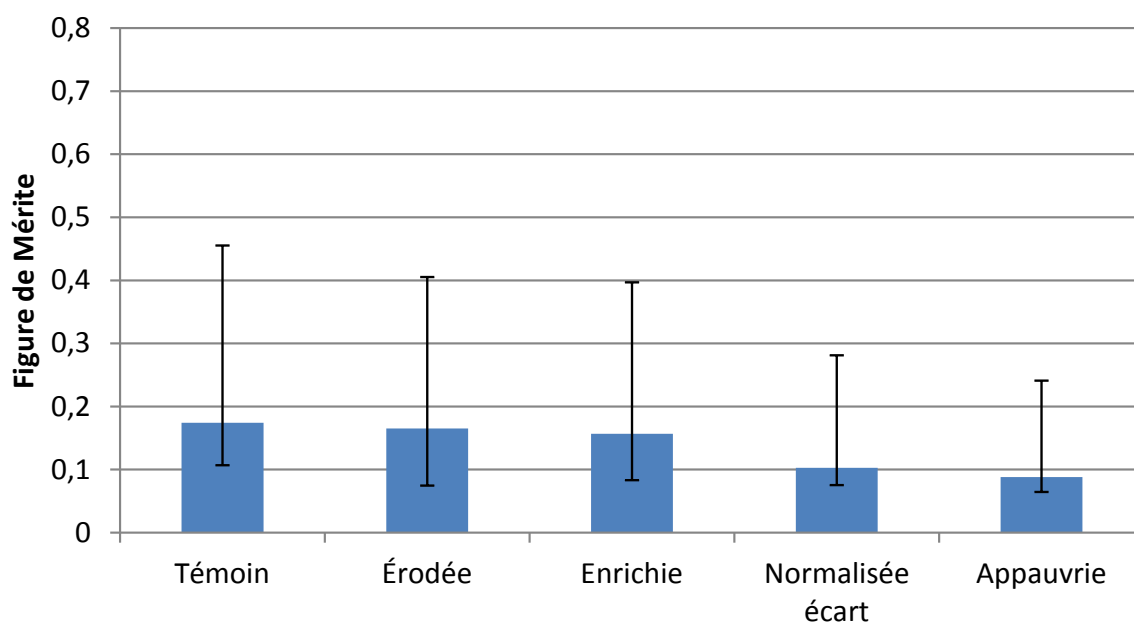


FIG. 13.4: Les FDM (Figure de Mérite) JAFROC obtenues pour les différents paramètres de génération de la base d'apprentissage. Les bases **Témoin**, **Érodée** et **Enrichie** montrent les meilleures performances, avec un faible avantage pour la base **Témoin**. Les bases **Normalisée écart** et **Appauvrie** ont une figure de mérite nettement plus faible.

d'images :

Pas d'érosion, une normalisation visant à ramener la moyenne et l'écart-type sur les caractéristiques à 1, et 200 points extraits de chaque image de la base d'apprentissage.

13.2 Comparaison des performances de la détection des tumeurs pulmonaires

Les paramètres de génération de la base d'apprentissage retenus sont ceux de la base **Témoin**. Ils correspondent aux choix suivants :

- 200 points tirés aléatoirement dans le volume complet du poumon de chaque image (hors tumeurs).
- normalisation par neutralisation de la moyenne et de la variance, tels que $(\mu, \sigma) = (0, 1)$.

Quatre jeux d'images seront comparés :

Statique : correspond aux images de la "vérité terrain", à savoir des images sans mouvement respiratoire. Ce jeu doit donner la performance haute.

NoCorr : représente les images simulées avec mouvement respiratoire mais reconstruites sans aucune correction de mouvement. Ce jeu représente le cas le plus défavorable.

13.2. COMPARAISON DES PERFORMANCES DE LA DÉTECTION DES TUMEURS PULMONAIRES

TE-IM (Transformation Élastique images) : correspond aux images reconstruites avec la correction de mouvement post-reconstruction.

TE-MS (Transformation Élastique Matrice Système) : correspond aux images reconstruites avec correction de mouvement pendant la reconstruction.

a	Base Statique	Base TE-IM	Base TE-MS	Base NoCorr
C	464	10000	10000	10000
γ	0.0053	0.00097	0.00031	0.00055
j	3	3	4	3
Sensibilité	0.75	0.81	0.82	0.83
Spécificité	0.99	0.99	0.99	0.99
Précision	0.98	0.98	0.98	0.98

TAB. 13.2: Paramètres sélectionnés pour l'optimisation des performances de détection des tumeurs pulmonaires. Chaque triplet de paramètres sélectionné (C, γ, j) est indiqué ainsi que sa valeur de sensibilité et spécificité.

13.2.1 Sélection des meilleurs paramètres du classifieur

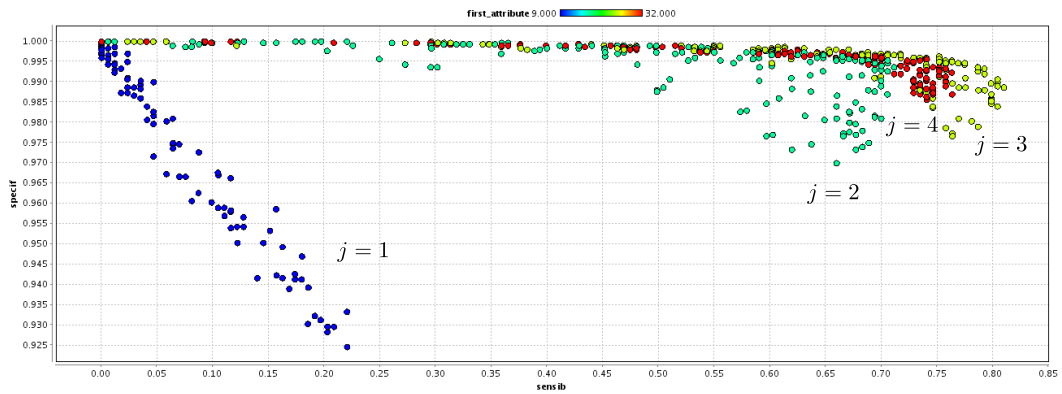
De la même manière que pour la sélection de la meilleure base d'apprentissage, il faut adapter les paramètres du classifieur (C, γ, j) aux bases des quatre jeux d'images que nous souhaitons comparer.

La figure 13.5 montre les nuages de points associés aux différents triplets de paramètres (C, γ, j) pour chaque type d'images. Les performances des images statiques sont celles présentées sur la base **Témoin** de la figure 13.1.a. La table 13.2 résume les triplets (C, γ, j) correspondant aux meilleures performances mesurées à l'aide des nuages de points des figures 13.1.a et 13.5.

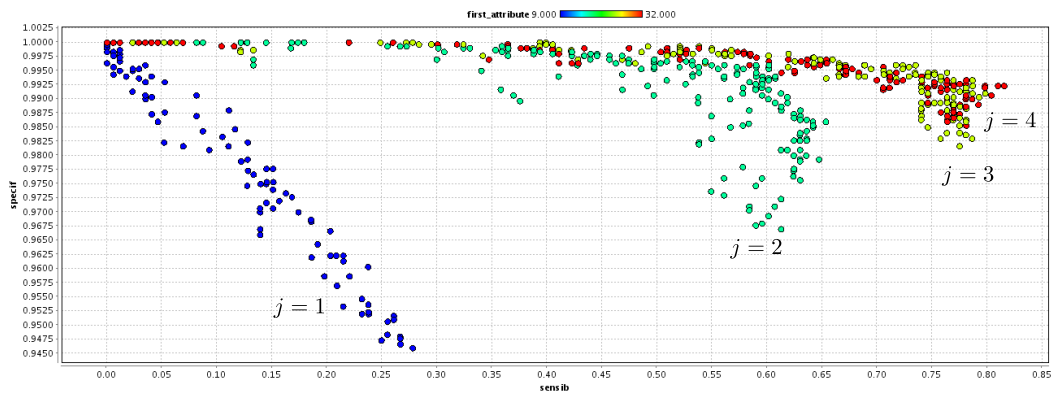
Il est étonnant de constater que la base 13.1.a, qui correspond aux images **Statique**, offre les plus mauvaises performances pour la décomposition de niveau 1 (17% de sensibilité). La mauvaise performance de cette base d'apprentissage se retrouve pour tous les niveaux de décomposition. Le résultat est également souligné par le tableau 13.2 (sensibilité de 75% contre des valeurs supérieures à 80% pour les autres types d'images). On peut cependant observer que les nuages de points des autres jeux d'images ont la même distribution spatiale que ceux de la base **Statique**, par opposition aux distributions des bases **Érodée** ou **Normalisée Écart** vues dans la partie précédente. Par exemple, pour $j = 1$, on observe une forte corrélation de la sensibilité et de la spécificité pour les bases **Statique**, **NoCorr**, **TE-IM** et **TE-MS** que l'on ne retrouve pas pour les bases **Érodée** et **Normalisée Écart**.

On peut observer qu'il y a une forte corrélation entre la sensibilité et la spécificité des points pour tous les types d'images au premier niveau de décomposition.

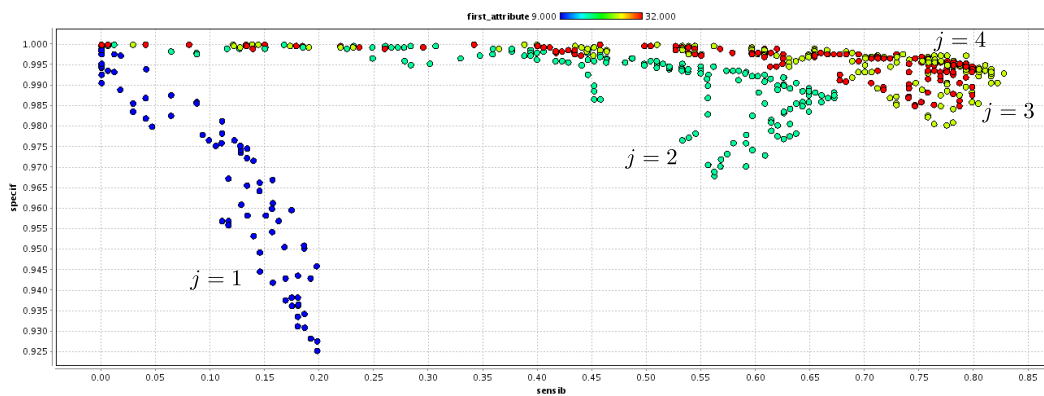
CHAPITRE 13. ANALYSE DES RÉSULTATS



a) TE-IM



b) TE-MS



c) NoCorr

FIG. 13.5: Fronts de Pareto des résultats de la recherche des meilleurs paramètres du classifieur pour les différents jeux d'images pulmonaire, avec 200 points négatifs par image. Pour chaque triplet de paramètres (C, γ, j) , la sensibilité et la spécificité sont reportées sur le graphique. Le code couleur correspond à la valeur de j . a) représente la correction d'image TE-IM, b) les images corrigées par la méthode TE-MS, et c) les images NoCorr.

Le type d'images ayant les meilleures performances pour la décomposition de ni-

13.2. COMPARAISON DES PERFORMANCES DE LA DÉTECTION DES TUMEURS PULMONAIRES

veau 2 est **TE-IM**, avec plus de 70% de sensibilité, comparés aux 60% de la base **Statique**.

Les meilleures performances du CAD sont atteintes pour le troisième niveau de décomposition pour tous les types d'images sauf **TE-MS**, avec une sensibilité d'environ 81 et 83% pour **TE-IM** et **NoCorr**, et 75% pour la base **Statique**. Les performances maximales sur **TE-MS** sont atteintes pour $j = 4$.

La spécificité observée pour les meilleurs jeux de paramètres est sensiblement la même quelque soit le type d'image, entre 99% et 99.5%.

13.2.2 Courbes Free-ROC

Les courbes F-ROC obtenues sur les différents types d'images à l'aide des triplets (C, γ, j) reportés dans la table 13.2 sont présentées dans la figure 13.6.

Toutes les courbes ont un NFM (nombre de faux positifs moyen par image) à la sensibilité maximale à peu près équivalent entre 20 et 24, excepté **TE-MS** avec un NFM à la sensibilité maximale de 34. Cependant, les performances en FLL (fraction des lésions détectées et localisées) de **TE-MS** sont constantes à partir d'une NFM de 20 environ.

L'ordre des courbes est constant pour tous les NFM à partir de 5 faux positifs par image. Au delà de cette limite, les performances des images **statique** sont supérieures à celles de **TE-IM** de 1 à 5%, tandis que **TE-IM** a des performances supérieures de 3 à 10% à celles des images **TE-MS** et des images **NoCorr**. Ces deux dernières ont des performances quasiment identiques.

En dessous de 5 faux positifs par image, les courbes sont trop proches pour pouvoir en déduire une tendance.

Globalement, le graphique montre que le maximum de performance est apporté par la base **Statique**, suivi par la base **TE-IM**. Les bases **TE-MS** et **NoCorr** sont très proches l'une de l'autre mais nettement en dessous des deux premières en terme de FLL, à NFM égale.

13.2.3 Comparaison des performances JAFROC

Les figures de mérite obtenues par l'algorithme de JAFROC sont présentées dans la figure 13.7. On observe une tendance proche de celle observée sur les courbes F-ROC. Les images **Statique** montrent les meilleures performances, suivies par **TE-IM**. Il est intéressant de noter que les valeurs de **TE-IM** et **TE-MS** sont relativement proches, y compris leurs mesures d'erreur, ce qui indique que l'algorithme JAFROC les distingue difficilement. **TE-MS** est quand à lui nettement en retrait.

CHAPITRE 13. ANALYSE DES RÉSULTATS

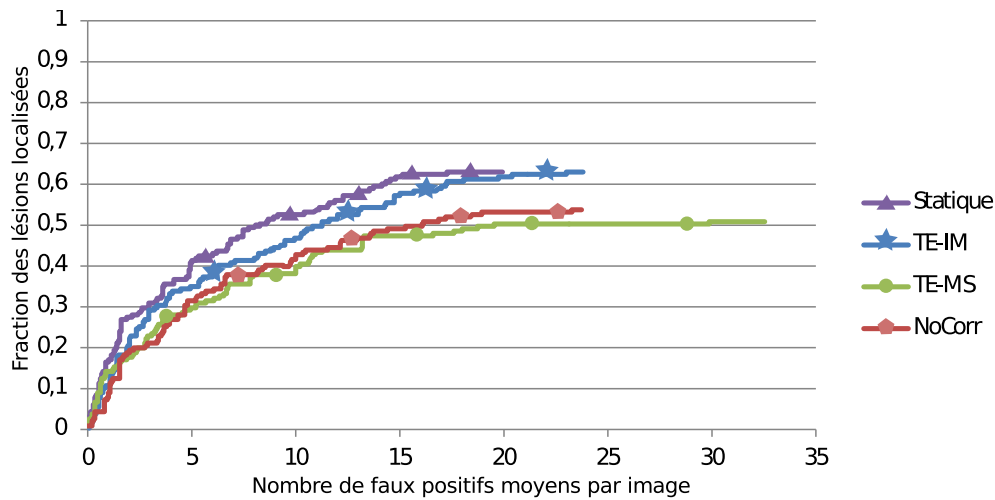


FIG. 13.6: Courbes Free-ROC comparant les performances du CAD sur les lésions pulmonaires selon la technique de correction du mouvement respiratoire.

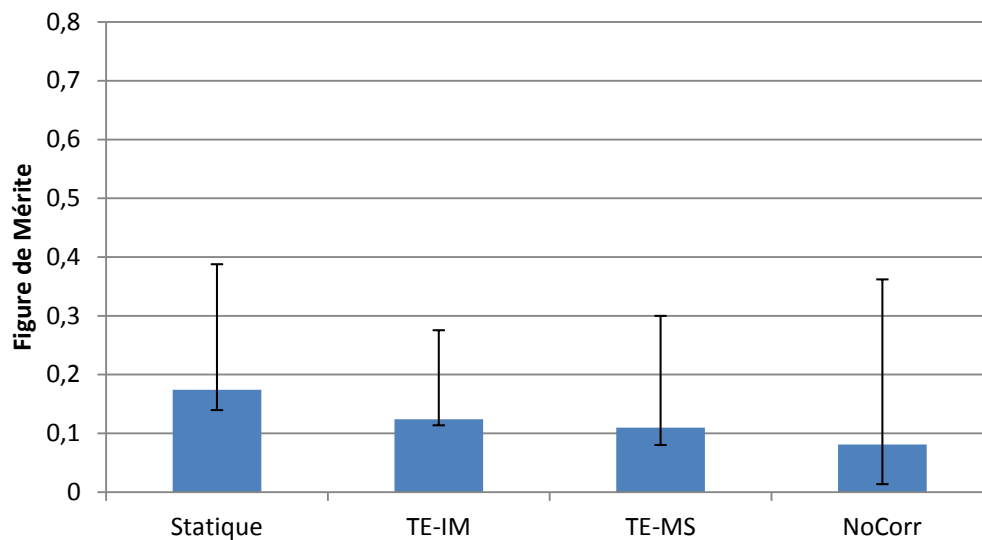


FIG. 13.7: Les FDM (Figure de Mérite) obtenues pour les différentes techniques de correction du mouvement respiratoire sur la détectabilité des lésions du poumon.

13.2.4 Conclusion

Les résultats présentés précédemment indiquent que les techniques de correction du mouvement respiratoire améliorent les performances de détection par rapport aux images **NoCorr** pour les lésions pulmonaires. Le résultat est net pour **TE-IM**, surtout lorsque l'on regarde les courbes F-ROC. Il est plus difficile de statuer sur les performances de la méthode **TE-MS**, dont les performances selon JAFROC sont au même niveau que celles de la méthode **TE-IM**, mais qui est nettement moins performante que cette dernière sur les courbes F-ROC.

13.3. COMPARAISON DES PERFORMANCES DES DIFFÉRENTES MÉTHODES POUR LA DÉTECTION DES TUMEURS HÉPATIQUES

13.3 Comparaison des performances des différentes méthodes pour la détection des tumeurs hépatiques

13.3.1 Sélection des meilleurs paramètres du classifieur

Les figures 13.8 et 13.9 nous montrent une répartition des performances très différente de celle observée précédemment pour les travaux sur les tumeurs pulmonaires.

Pour le premier niveau de décomposition ($j = 1$), les images **Statique**, **TE-IM** et **TE-MS** montrent une très grande dispersion des valeurs de performances. **NoCorr**, par contre, montre des points plus proches, mais avec des sensibilités très faibles (inférieures à 15%). Cette dispersion montre une rupture par rapport à la corrélation observée sur les bases du poumon.

De la même manière que pour $j = 1$, les performances obtenues pour $j = 2$ montrent une dispersion très importante des points pour les images **Statique** et **TE-IM**. Le nuage de points correspondant à **TE-MS** a une forme plus proche de celles observées précédemment. La base **NoCorr** est clairement en difficulté car l'apport des informations du second niveau de décomposition diminue les performances maximales du CAD en sensibilité.

En ajoutant les informations des niveaux de décomposition supérieurs, on observe une amélioration nette des performances, surtout pour la base **NoCorr**. Tous les types d'images montrent une amélioration des performances lorsque l'on prend en compte le quatrième niveau de décomposition, par opposition aux tumeurs du poumon où l'ajout de ces informations faisait baisser les performances du CAD. Étant donné que les informations fournies par le quatrième niveau de décomposition des ondelettes sont de très basse fréquence, elles ne donnent pas d'information sur la lésion elle-même mais sur son environnement. Cela semble indiquer que le classifieur est mal adapté pour gérer ces données.

Les résultats des couples (C, γ, j) pour les quatre types d'images sont reportés dans le tableau 13.3. Comme pour les lésions pulmonaires, **TE-IM** a la sensibilité la plus forte avec 68%. Et cette fois ci, **Statique** est seconde avec 62%.

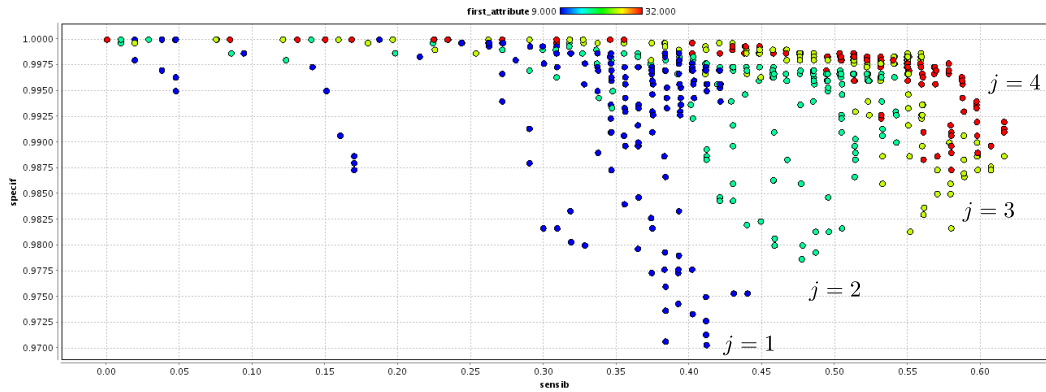
13.3.2 Courbes Free-ROC

Les courbes Free-ROC de la figure 13.10 montrent un ordre équivalent à celles de la figure 13.6 pour le poumon.

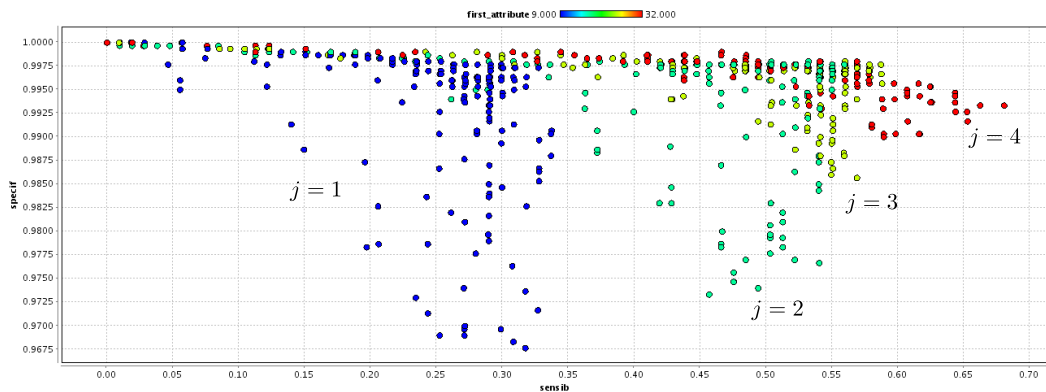
Le maximum de performances est apporté par les images statiques, suivi par les images **TE-IM**, puis **TE-MS** et enfin **NoCorr**. Contrairement au poumon, les courbes sont relativement bien séparées avec une différence de sensibilité de 20% entre les images non **NoCorr** et **Statique**.

Il est étonnant d'observer que les images **Static** et **TE-MS** ont toutes les deux un NFM maximum d'environ 17, très supérieur à celui de **TE-IM** et des images **NoCorr**

CHAPITRE 13. ANALYSE DES RÉSULTATS



a) Statique



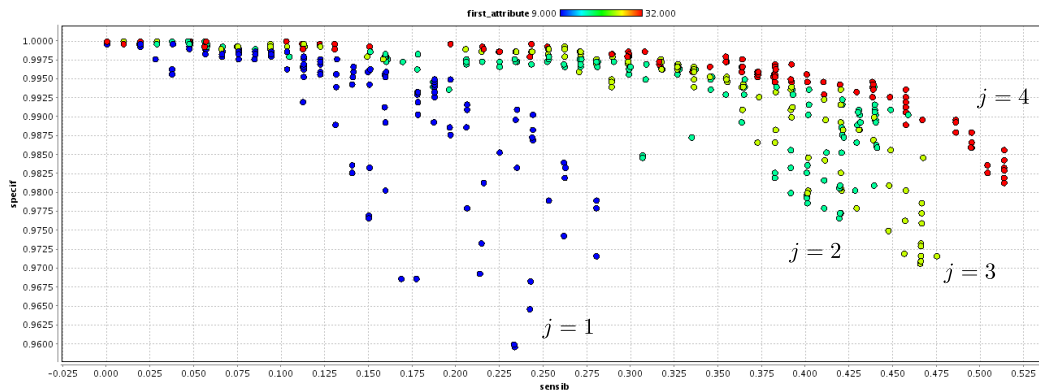
b) TE-IM

FIG. 13.8: Fronts de Pareto des résultats de la recherche des meilleurs paramètres du classifieur pour les lésions du foie, avec 200 points négatifs par image. Pour chaque triplet de paramètres (C, γ, j) , la sensibilité et la spécificité sont reportées sur le graphique. Le code couleur correspond à la valeur de j . a) représente les images statiques, b) les images corrigées du mouvement post-reconstruction.

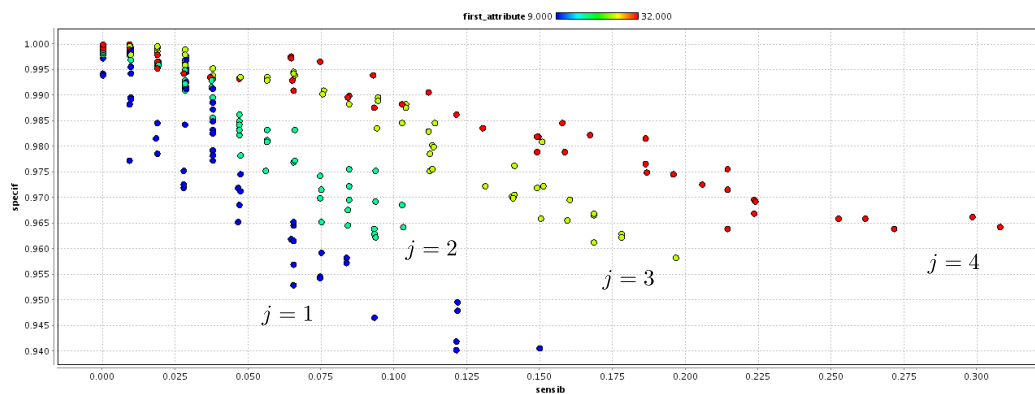
a	Base Statique	Base TE-IM	Base TE-MS	Base NoCorr
C	858	5412	251	5412
γ	0.002	0.00055	0.0053	0.0017
j	4	4	4	4
Sensibilité	0.62	0.68	0.51	0.31
Spécificité	0.99	0.99	0.99	0.96
Précision	0.98	0.98	0.97	0.94

TAB. 13.3: Paramètres sélectionnés pour l'optimisation des performances du CAD sur les tumeurs hépatiques. Sont indiqués pour chaque base le triplet de paramètres sélectionné ainsi que sa position sur le front de Pareto.

13.3. COMPARAISON DES PERFORMANCES DES DIFFÉRENTES MÉTHODES POUR LA DÉTECTION DES TUMEURS HÉPATIQUES



c) TE-MS



d) NoCorr

FIG. 13.9: Fronts de Pareto des résultats de la recherche des meilleurs paramètres du classifieur pour les lésions du foie, avec 200 points négatifs par image. Pour chaque triplet de paramètres (C, γ, j) , la sensibilité et la spécificité sont reportées sur le graphique. Le code couleur correspond à la valeur de j . c) représente la correction d'image **TE-MS** et d) les images **NoCorr**.

qui ne dépassent pas 9 faux positifs par image. Mais comme dans les cas précédents, ce nombre important de faux positifs n'apporte qu'une amélioration très faible de la sensibilité pour les images de la base **Statique**, au contraire de **TE-MS** où un important bond de sensibilité (de 50% à 60%) apparaît pour un NFM de 11.5 .

Si l'on se contente de comparer les courbes F-ROC pour un NFM donné de 9, la base d'images **Statique** a la meilleure sensibilité avec 58%, suivie par **TE-IM** avec 53% et **TE-MS** avec 48%. Les images **NoCorr** sont en net retrait avec seulement 37% de sensibilité.

Les performances sont globalement en retrait par rapport à celles du poumon, mais l'ordre des courbes reste cohérent avec celui observé sur le poumon.

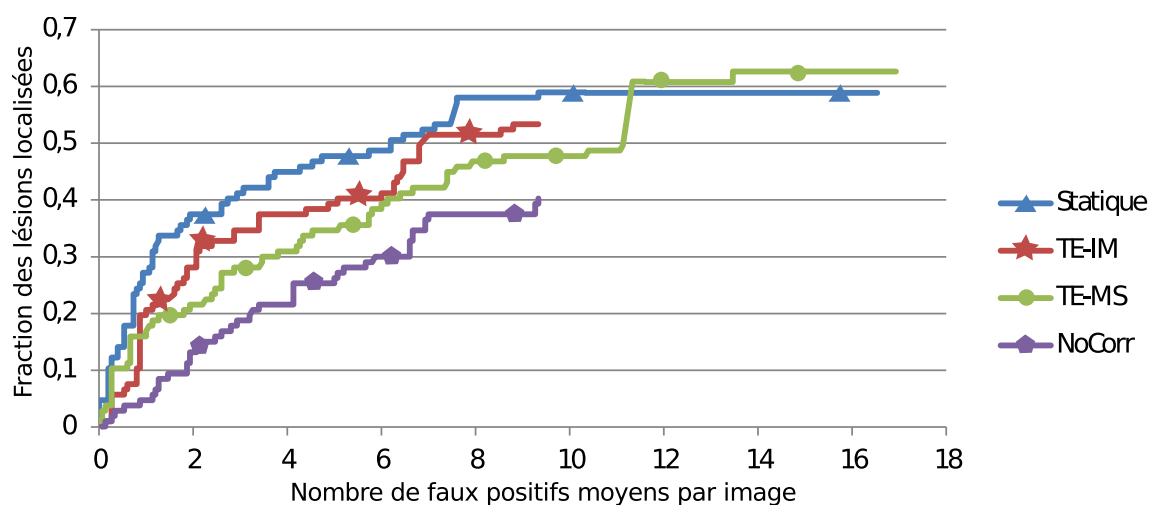


FIG. 13.10: Courbes Free-ROC comparant les performances du CAD sur les lésions hépatiques selon la technique de correction du mouvement respiratoire.

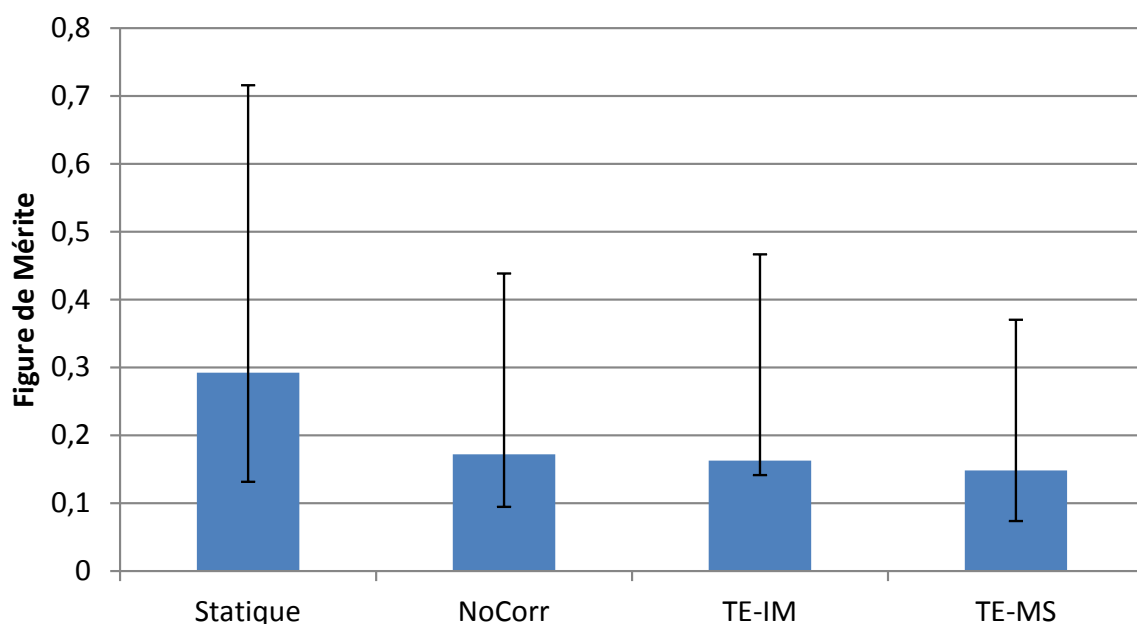


FIG. 13.11: Les FDM (Figure de Mérite) JAFROC obtenues pour les différents types d'images pour la détectabilité des lésions du foie.

13.3.3 Comparaison des performances JAFROC

Les Figures de mérite obtenues par l'algorithme de JAFROC sont présentées dans la figure 13.11. La p-valeur est de 0,1, ce qui ne permet pas de déclarer que statistiquement les données sont différentes. Il n'est pas étonnant de constater que les FDM sont plus élevées que précédemment car le nombre de faux positifs est pratiquement deux fois inférieur à celui obtenu pour le poumon, ce qui est le signe que le classifieur réussit mieux à discerner les vrais positifs.

13.3. COMPARAISON DES PERFORMANCES DES DIFFÉRENTES MÉTHODES POUR LA DÉTECTION DES TUMEURS HÉPATIQUES

On observe que les images **Statique** ont un score supérieur de 60% environ à celui des images **NoCorr**. Par contre, il est surprenant de constater que les deux techniques de correction du mouvement respiratoire ont un score égal ou plus faible que les images **NoCorr**, en contradiction avec les résultats de l'analyse F-ROC. Cependant, la proximité des valeurs ne permet pas de montrer une réelle hiérarchie entre les jeux d'images. Tout au plus pourrait-on observer que l'étendue des barres d'erreur semble indiquer que **TE-IM** aurait de meilleures performances que les images **NoCorr**, elles-mêmes supérieures aux images **TE-MS**, mais il n'est pas possible de se prononcer de manière définitive à partir de ces données.

13.3.4 Conclusion

Les résultats des mesures de performance sur les tumeurs hépatiques sont contrastés. Les courbes F-ROC montrent un clair avantage des techniques de correction du mouvement respiratoire, surtout pour **TE-IM** si l'on compare la sensibilité pour un nombre moyen de fausses localisations. Par contre, les figures de mérite obtenues par JAFROC montrent à l'inverse que **TE-IM** et **TE-MS** sont au même niveau que les images **NoCorr**, en net retrait par rapport aux images **Statique**.

CHAPITRE 13. ANALYSE DES RÉSULTATS

Sixième partie

Discussion

Chapitre 14

Discussion

Les contributions de cette thèse sont doubles. Elles portent, à la fois sur la création d'une base de données de patients simulés respirant, en TEP dynamique, et sur un ensemble d'outils d'évaluation des performances de détection des lésions.

Nous allons tout d'abord présenter une synthèse des résultats, puis les limites et perspectives de notre système de détection ainsi que de la base de données.

14.1 Synthèse des résultats

Les résultats que nous avons présentés montrent une différence importante de détectabilité des lésions dans les images non corrigées et dans les images statiques avec tous les critères utilisés. Ce résultat conforte la fiabilité du système de détection que nous avons développé. Les détections réalisées dans les images corrigées montrent des performances de détectabilité intermédiaires entre celles obtenues dans les images statiques et celles avec obtenues dans les images dynamiques non corrigées.

Pour le poumon, les courbes Free-ROC montrent que la détectabilité des lésions est bien retrouvée pour la correction du mouvement respiratoire TE-IM, avec une sensibilité de 62% contre 63% sur les images statiques, pour 20 faux positifs en moyenne par image. La méthode de correction du mouvement TE-MS apporte une sensibilité de seulement 50% dans les mêmes conditions. En revanche, les FDM issues de l'analyse JAFROC montrent une amélioration semblable pour les deux techniques de correction du mouvement. Pour le foie, les courbes F-ROC montrent que les techniques de correction du mouvement respiratoire ont des performances intermédiaires par rapport à celles des images statiques et non corrigées avec, par exemple, pour 9 faux positifs par image une sensibilité de 58% pour les images statiques, 53% pour la méthode TE-IM, 48% pour la méthode TE-MS et 37% pour les images non corrigées. Cependant, les FDM issues des images corrigées du foie ne montrent pas d'améliorations notables de la détectabilité apportée par la correction du mouvement par rapport à celle des images non corrigées.

Les performances observées à l'aide des courbes Free-ROC sont très différentes

entre les deux organes, notamment le nombre de faux positifs moyen qui est beaucoup plus faible pour le foie que pour le poumon. Cette différence s'explique par les différences de volumes des organes. En effet, le volume des poumons est approximativement deux fois plus important que celui du foie dans le modèle XCAT que nous avons utilisé.

Les courbes Free-ROC sont plus pertinentes que les FDM pour la détection, mais il n'existe pas de métrique permettant de synthétiser l'ensemble des performances, comme peut le faire l'aire sous la courbe pour les étude ROC. Il est intéressant de constater que les deux métriques que nous avons utilisées (F-ROC et FDM JAFROC) montrent des résultats parfois contradictoires. Cela peut s'expliquer par la différence de l'information qu'elles mesurent. L'analyse JAFROC ne prend en compte que le faux positif de plus haut score pour chaque image, tandis les courbes F-ROC utilisent le nombre de faux positifs dépassant un certain seuil. Les FDM sont naturellement plus difficiles à interpréter et sont, par nature, plus sensibles au bruit car elles utilisent des maxima de distributions, et sont donc très dépendantes des valeurs extrêmes. Cependant, cette métrique renseigne sur la différence de détectabilité entre les vrais positifs et le faux positifs de plus haut score. C'est une métrique qui est handicapée par la faible quantité d'exemples à disposition, comme le montrent les barres d'erreur de la figure 13.7.

Les résultats que nous avons obtenus montrent néanmoins, sans ambiguïté, que la détectabilité des lésions est améliorée par la correction du mouvement respiratoire, notamment en utilisant la correction TE-IM. La détectabilité des lésions indiquée par la courbe F-ROC est systématiquement supérieure sur les images corrigées par rapport à celle des images non corrigées. Cependant la correction TE-MS a des performances inférieures à celles observées pour TE-IM, ce qui est en contradiction avec les résultats présentés dans la littérature sur la quantification des lésions [Lamare et al., 2007b, Qiao et al., 2006]. Il faut noter que nous n'avons pas réalisé notre étude sur les mêmes principes, en privilégiant une approche détection sur la quantification. Ces deux problématiques sont différentes et demandent des paramétrages spécifiques. De plus, nous n'avons pas utilisé la même implémentation de l'estimateur de mouvement et du correcteur d'image que les auteurs, ce qui pourrait expliquer les différences observées. La précision de l'estimation de mouvement est aussi différente, car nous utilisons un espacement entre les noeuds de la grille B-spline de 17 cm en X et en Y contre 2.5 cm pour l'évaluation de [Lamare et al., 2007b].

14.2 Limites et perspectives

14.2.1 Système de détection

La métrique utilisée pour choisir les paramètres du système CAD (voir 12.5.1) est très différente de celle utilisée pour mesurer les performances de détection sur les différents jeux d'images de la base (voir 12.5.3). L'étude de l'optimisation des paramètres du système CAD est basée sur une validation croisée qui ne considère que

14.2. LIMITES ET PERSPECTIVES

le voxel central des lésions. Si le score de ce voxel est supérieur à 0, alors la lésion est considérée comme correctement détectée. Les performances de détectabilité sont, quant à elles, basées sur une approche région, en travaillant sur des agrégats. De plus, les métriques FDM et Free-ROC utilisent des seuils variables pour déterminer les performances de chaque technique.

Il serait intéressant de travailler sur une métrique “basée région” pour la recherche des meilleurs paramètres de la base d’apprentissage mais, dans ce cas, l’évaluation des performances pour chaque jeu de paramètres nécessiterait de traiter toutes les images, ce qui représente plusieurs heures de calculs par jeu de paramètres. La recherche exhaustive que nous avons réalisée demande l’évaluation de 900 triplets (c, γ, j) . Il n’est donc pas possible d’utiliser la même métrique, à moins d’opérer une étude séparée sur les algorithmes d’optimisation à utiliser, et de réaliser les calculs sur une grille.

Il serait également intéressant d’augmenter la taille de la base de données pour pouvoir éviter le phénomène de sur-apprentissage (optimisation des paramètres pour le jeu de données et non pour les classes), qui est relativement important dans notre étude, où les données d’apprentissage et de tests sont extraites de la même base de données. Cela engendre un biais positif sur les résultats. Ce biais n’est cependant pas très grave dans notre cas car il affecte toutes les modalités, et nous ne sommes pas intéressés par les performances absolues, mais par les performances relatives. Pour éliminer ce biais, il faudrait entre autre, que la base de données utilisée pour le réglage des paramètres soit différente de celle utilisée pour réaliser l’estimation des performances.

Les performances que nous avons observées, avec des taux de détection de lésions de l’ordre de 60%, sont cohérentes avec la calibration réalisée sur les activités des lésions. Le nombre de faux positifs par image que nous observons pour ces niveaux de performances (entre 8 et 15) est trop important pour un système d’assistance à la détection clinique, mais ce n’était pas notre objectif.

14.2.2 Base de données

Nous avons réalisé notre base de données à partir du modèle anthropomorphique XCAT de Paul Segars, principalement pour sa capacité à s’adapter à des morphologies différentes et sa prise en charge native d’un modèle respiratoire complexe.

Cependant, cette approche par modèle a des limitations connues, notamment l’homogénéisation de l’activité des organes. En effet, contrairement à l’imagerie TDM, où les coefficients d’atténuation sont relativement constants dans les tissus et entre les patients, en TEP l’activité mesurée pour un même organe varie entre les patients, et n’est pas constante à l’intérieur d’un même organe. Les différences d’activités entre les patients s’expliquent par la quantité de traceur injecté, d’éventuelles inflammations, le taux de glycémie (pour le ^{18}FDG), ou encore le temps écoulé depuis l’injection.

Nous avons prévu, au début de la création de la base d’intégrer une variabilité de captation inter-patients, mais nous avons dû y renoncer par manque de temps. De plus, incorporer la variabilité de captation intra-organe pour les patients n’aurait pas

été possible car le modèle utilisé est trop simpliste.

Dans la base de données que nous avons simulée, nous nous sommes restreints à utiliser des lésions sphériques de petit diamètre (8 et 12 mm). Aux résolutions permises par la TEP, cette approximation est suffisante pour modéliser correctement les fixations observées en clinique. Cependant cette limite est uniquement liée à notre choix de sujets, et n'est pas représentative des possibilités du simulateur. Nous avons notamment travaillé sur des lésions de formes différentes dans le cadre de la publication de [Le Maitre et al., 2009].

Du point de vue des choix de paramètres réalisés, une étude de ce type demande la fixation d'un nombre très important de paramètres, qui ont tous un impact sur les résultats. Nous avons justifié le choix d'un maximum d'entre eux, mais ils sont souvent interdépendants, ce qui engendre des problèmes impossibles à optimiser de manière exhaustive.

Une évolution de nos travaux pourrait être de mettre en place une réflexion et une refonte du système d'optimisation des paramètres en se basant sur la théorie des plans d'expériences. Cela demanderait cependant des travaux supplémentaires pour évaluer de manière précise les interactions entre chaque type de paramètres et son implication dans le résultat final.

14.3 Imagerie TEP/TDM

En routine clinique, les acquisitions sont maintenant réalisées conjointement avec la TDM. Le rapprochement des scanners TEP et TDM sur le même appareil a permis d'éviter la réalisation de la carte de transmission et fournit au praticien des informations anatomiques qui permettent d'améliorer la précision du diagnostic. Nous avons prévu au début de la thèse d'inclure la simulation d'images TDM, et d'utiliser conjointement les deux modalités pour l'évaluation de la détectabilité des lésions.

Un partenariat avait été mis en place avec le laboratoire Leti du CEA de Grenoble portant sur l'utilisation de leur simulateur SINDBAD. Nous avons passé plusieurs mois à l'évaluer et à adapter nos données pour le valider. En effet, ce simulateur a été développé à l'origine pour le contrôle non destructif, et n'avait jamais été utilisé pour simuler des données cliniques où les géométries, ainsi que les énergies utilisées, sont totalement différentes. Nous avons embauché un stagiaire ingénieur pour valider les simulations par rapport à des mesures réelles sur des fantômes [Leduvehat, 2010]. Cependant, les images que nous sommes parvenues à simuler étaient trop "parfaites" pour pouvoir être utilisées conjointement avec les images TEP. En effet, le modèle que nous utilisons (XCAT) est trop peu détaillé pour une modalité comme la TDM, qui dispose d'une résolution inférieure au millimètre. Les inclusions étaient beaucoup trop visibles dans les tissus, en particulier au niveau pulmonaire, car ces derniers n'étaient pas assez complexes pour pouvoir les dissimuler malgré la modélisation des bronchioles et du système sanguin. En effet, les images TDM réelles montrent fréquemment des artefacts physiologiques qui doivent être éliminés par le radiologue, ou encore, notamment pour le poumon, des textures très complexes que nous ne sommes pas

14.3. IMAGERIE TEP/TDM

parvenus à reproduire. Malgré notre travail avec l'auteur du modèle, il n'a pas été possible d'améliorer suffisamment la complexité du modèle pour obtenir des résultats satisfaisants.

Malgré ces difficultés, et en prenant la suite de nos travaux, le simulateur est actuellement en cours d'intégration dans le projet VIP (Virtual Imaging Platform) porté par le laboratoire CREATIS (<http://www.creatis.insa-lyon.fr/vip>), où il devrait être utilisé pour la simulation médicale. Le but de ce projet est de réaliser une interface commune pour plusieurs simulateurs et les modèles associés. Une interface en cours de développement permettra de réaliser les simulations sur une grille de calculs (actuellement la grille Égée). Les simulateurs que nous avons utilisés (PET-SORTEO et SINDBAD) sont en cours d'intégration ainsi que le modèle XCAT. Des travaux sont en cours pour améliorer le réalisme des modèles, notamment grâce à l'ajout d'informations de texture.

Conclusion

L'étude que nous avons réalisée a montré une amélioration globale de la détectabilité des lésions, apportée par les techniques de correction du mouvement respiratoire. Cependant, les résultats sont parfois contrastés selon les deux techniques que nous avons évaluées, notamment pour le poumon, ce qui semble indiquer que les organes réagissent différemment à ces corrections.

L'approche région que nous avons proposée et intégrée à notre système d'évaluation est originale, dans le sens où nous avons appliqué un estimateur région qui se rapproche du modèle de détection humain. Cependant, il reste de nombreuses possibilités d'amélioration, notamment au niveau de la complexité du modèle utilisé et de la sélection des paramètres.

De plus, la référence reste toujours l'observateur humain, et les résultats que nous avons observés doivent être comparés avec les performances d'un observateur humain, et validés sur des données cliniques. Nous avons montré qu'il était possible de répondre à la problématique uniquement à partir de données simulées et d'observateurs informatiques.

Bibliographie

- [Allison et al., 2006] Allison, J., Amako, K., Apostolakis, J., Araujo, H., Dubois, P., Asai, M., Barrand, G., Capra, R., Chauvie, S., Chytracsek, R., et al. (2006). Geant4 developments and applications. *IEEE Transactions on Nuclear Science*, 53(1) :270–278.
- [Bai and Brady, 2009] Bai, W. and Brady, M. (2009). Regularized B-spline deformable registration for respiratory motion correction in PET images. *Physics in Medicine and Biology*, 54 :2719.
- [Bailey, 2005] Bailey, D. (2005). *Positron emission tomography : basic sciences*. Springer Verlag.
- [Barrett et al., 1993] Barrett, H., Yao, J., Rolland, J., and Myers, K. (1993). Model observers for assessment of image quality. *Proceedings of the National Academy of Sciences of the United States of America*, 90(21) :9758.
- [Baydush et al., 2007] Baydush, A., Catarious, D., Lo, J., and Floyd, C. (2007). Incorporation of a Laguerre–Gauss Channelized Hotelling Observer for False-Positive Reduction in a Mammographic Mass CAD System. *Journal of Digital Imaging*, 20(2) :196–202.
- [Bazzani et al., 2001] Bazzani, A., Bevilacqua, A., Bollini, D., Brancaccio, R., Campanini, R., Lanconelli, N., Riccardi, A., and Romani, D. (2001). An SVM classifier to separate false signals from microcalcifications in digital mammograms. *Physics in Medicine and Biology*, 46 :1651.
- [Beach et al., 2004] Beach, R., Pretorius, P., Boening, G., Bruyant, P., Feng, B., Fulton, R., Gennert, M., Nadella, S., and King, M. (2004). Feasibility of stereo-infrared tracking to monitor patient motion during cardiac SPECT imaging. *IEEE Transactions on Nuclear Science*, 51(5) :2693–2698.
- [Bissantz et al., 2006] Bissantz, N., Mair, B., and Munk, A. (2006). A multi-scale stopping criterion for MLEM reconstructions in PET. In *Nuclear Science Symposium Conference Record, 2006. IEEE*, volume 6, pages 3376–3379. IEEE.
- [Bloomfield et al., 2003] Bloomfield, P., Spinks, T., Reed, J., Schnorr, L., Westrip, A., Livieratos, L., Fulton, R., and Jones, T. (2003). The design and implementation of a motion correction scheme for neurological PET. *Physics in Medicine and Biology*, 48 :959.
- [Boser et al., 1992] Boser, B., Guyon, I., and Vapnik, V. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM.

BIBLIOGRAPHIE

- [Boucher et al., 2004] Boucher, L., Rodrigue, S., Lecomte, R., and Benard, F. (2004). Respiratory gating for 3-dimensional PET of the thorax : feasibility and initial results. *Journal of Nuclear Medicine*, 45(2) :214.
- [Brasse et al., 2005] Brasse, D., Kinahan, P., Lartizien, C., Comtat, C., Casey, M., and Michel, C. (2005). Correction methods for random coincidences in fully 3D whole-body PET : impact on data and image quality. *Journal of nuclear medicine*, 46(5) :859.
- [Browne and De Pierro, 1996] Browne, J. and De Pierro, A. (1996). A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography. *IEEE Transactions on Medical Imaging*, 15(5) :687–699.
- [Bunch et al., 1978] Bunch, P., Hamilton, J., Sanderson, G., and Simmons, A. (1978). A free-response approach to the measurement and characterization of radiographic-observer performance. *J. Appl. Photogr. Eng.*, 4(4) :166–171.
- [Bundsuh et al., 2007] Bundschuh, R., Martinez-Moeller, A., Essler, M., Martinez, M., Nekolla, S., Ziegler, S., and Schwaiger, M. (2007). Postacquisition detection of tumor motion in the lung and upper abdomen using list-mode PET data : a feasibility study. *Journal of Nuclear Medicine*, 48(5) :758.
- [Buvat et al., 2002] Buvat, I., Castiglioni, I., et al. (2002). Monte Carlo simulations in SPET and PET. *Journal of Nuclear Medicine*, 46(1) :48–61.
- [Buvat and Lazaro, 2006] Buvat, I. and Lazaro, D. (2006). Monte Carlo simulations in emission tomography and GATE : an overview. *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*, 569(2) :323–329.
- [Cachin et al., 2006] Cachin, F., Kelly, A., and Maublant, J. (2006). Évaluation de la réponse thérapeutique en cancérologie : place de l'imagerie isotopique. *Bulletin du cancer*, 93(12) :1191–1199.
- [Chakraborty and Berbaum, 2004] Chakraborty, D. and Berbaum, K. (2004). Observer studies involving detection and localization : modeling, analysis, and validation. *Medical physics*, 31 :2313.
- [Chakraborty and Winter, 1990] Chakraborty, D. and Winter, L. (1990). Free-response methodology : alternate analysis and a new observer-performance experiment. *Radiology*, 174(3) :873.
- [Chan et al., 1987] Chan, H., Doi, K., Galhotra, S., Vyborny, C., MacMahon, H., and Jokich, P. (1987). Image feature analysis and computer-aided diagnosis in digital radiography. Automated detection of microcalcifications in mammography. *Medical Physics*, 14 :538.
- [Chang et al., 2010] Chang, G., Chang, T., Pan, T., Clark, John W., J., and Mawlawi, O. R. (2010). Implementation of an automated respiratory amplitude gating technique for PET/CT : clinical evaluation. *J Nucl Med*, 51(1) :16–24.
- [Cheng et al., 2006] Cheng, H., Shi, X., Min, R., Hu, L., Cai, X., and Du, H. (2006). Approaches for automated detection and classification of masses in mammograms. *Pattern recognition*, 39(4) :646–668.
- [Cherry and Dahlbom, 2006] Cherry, S. and Dahlbom, M. (2006). *PET : physics, instrumentation, and scanners*.

BIBLIOGRAPHIE

- [Chiang et al., 2004] Chiang, S., Cardi, C., Matej, S., Zhuang, H., Newberg, A., Alavi, A., and Karp, J. (2004). Clinical validation of fully 3D versus 2.5-D RAMLA reconstruction on the Philips-ADAC CPET PET scanner. *Nuclear medicine communications*, 25(11) :1103.
- [Chornoboy et al., 1990] Chornoboy, E., Chen, C., Miller, M., Miller, T., and Snyder, D. (1990). An evaluation of maximum likelihood reconstruction for SPECT. *IEEE Transactions on Medical Imaging*, 9(1) :99–110.
- [Cortes and Vapnik, 1995] Cortes, C. and Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20 :273–297. 10.1007/BF00994018.
- [Daube-Witherspoon et al., 2001] Daube-Witherspoon, M., Matej, S., Karp, J., and Lewitt, R. (2001). Application of the row action maximum likelihood algorithm with spherical basis functions to clinical PET imaging. *IEEE Transactions on Nuclear Science*, 48(1) :24–30.
- [Dawood et al., 2008] Dawood, M., Buther, F., Jiang, X., and Schafers, K. (2008). Respiratory Motion Correction in 3D PET Data With Advanced Optical Flow Algorithms. *IEEE Transactions on Medical Imaging*, 27(8) :1164 –1175.
- [Dawood et al., 2006] Dawood, M., Lang, N., Jiang, X., and Schafers, K. (2006). Lung motion correction on respiratory gated 3D PET/CT images. *IEEE transactions on medical imaging*, 25(4) :476.
- [Detorie and Dahlbom, 2008] Detorie, N. C. and Dahlbom, M. (2008). Quantitative evaluation of reconstruction algorithms for motion compensated PET. In *Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE*, pages 5366 –5372.
- [Dimitrakopoulou-Strauss et al., 2002] Dimitrakopoulou-Strauss, A., Strauss, L., Heichel, T., Wu, H., Burger, C., Bernd, L., and Ewerbeck, V. (2002). The role of quantitative ^{18}F -FDG PET studies for the differentiation of malignant and benign bone lesions. *Journal of Nuclear Medicine*, 43(4) :510.
- [El Naqa et al., 2009] El Naqa, I., Grigsby, P., Apte, A., Kidd, E., Donnelly, E., Khullar, D., Chaudhari, S., Yang, D., Schmitt, M., Laforest, R., et al. (2009). Exploring feature-based approaches in PET images for predicting cancer treatment outcomes. *Pattern recognition*, 42(6) :1162–1171.
- [El Naqa et al., 2006] El Naqa, I., Low, D., Bradley, J., Vicic, M., and Deasy, J. (2006). Deblurring of breathing motion artifacts in thoracic PET images by deconvolution methods. *Medical physics*, 33 :3587.
- [Erdi et al., 2004] Erdi, Y., Nehmeh, S., Pan, T., Pevsner, A., Rosenzweig, K., Mageras, G., Yorke, E., Schoder, H., Hsiao, W., Squire, O., et al. (2004). The CT motion quantitation of lung lesions and its impact on PET-measured SUVs. *Journal of Nuclear Medicine*, 45(8) :1287.
- [España et al., 2009] España, S., Herraiz, J., Vicente, E., Vaquero, J., Desco, M., and Udias, J. (2009). PeneloPET, a Monte Carlo PET simulation tool based on PENELOPE : features and validation. *Physics in Medicine and Biology*, 54 :1723.
- [Evans, 1955] Evans, R. (1955). *The atomic nucleus*. International series in pure and applied physics. McGraw-Hill.

BIBLIOGRAPHIE

- [Fahey, 2002] Fahey, F. (2002). Data acquisition in PET imaging. *Journal of nuclear medicine technology*, 30(2) :39–49.
- [Farquhar et al., 1999] Farquhar, T., Llacer, J., Hoh, C., Czernin, J., Gambhir, S., Seltzer, M., Silverman, D., Qi, J., Hsu, C., and Hoffman, E. (1999). ROC and localization ROC analyses of lesion detection in whole-body FDG PET : effects of acquisition mode, attenuation correction and reconstruction algorithm. *Journal of Nuclear Medicine*, 40(12) :2043.
- [Fayad et al., 2009] Fayad, H., Pan, T., Roux, C., Le Rest, C., Pradier, O., Clement, J., and Visvikis, D. (2009). A patient specific respiratory model based on 4D CT data and a time of flight camera (TOF). In *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, pages 2594–2598. IEEE.
- [Fayad et al., 2010] Fayad, H., Pan, T., Roux, C., and Visvikis, D. (2010). Application of a generic motion model for pet respiratory motion correction. *Journal of Nuclear Medicine*, 51(2) :524.
- [Fisher, 1936] Fisher, R. (1936). The use of multiple measurements in taxonomic problems. *Annals of Human Genetics*, 7(2) :179–188.
- [Ge et al., 2005] Ge, Z., Sahiner, B., Chan, H., Hadjiiski, L., Cascade, P., Bogot, N., Kazerooni, E., Wei, J., and Zhou, C. (2005). Computer-aided detection of lung nodules : False positive reduction using a 3D gradient field method and 3D ellipsoid fitting. *Medical physics*, 32 :2443.
- [Genestie et al., 1998] Genestie, C., Zafrani, B., Asselain, B., Fourquet, A., Rozan, S., Validire, P., Vincent-Salomon, A., and Sastre-Garau, X. (1998). Comparison of the prognostic value of Scarff-Bloom-Richardson and Nottingham histological grades in a series of 825 cases of breast cancer : major importance of the mitotic count as a component of both grading systems. *Anticancer research*, 18(1B) :571–576.
- [Gould et al., 2001] Gould, M., Maclean, C., Kuschner, W., Rydzak, C., and Owens, D. (2001). Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions. *JAMA : the journal of the American Medical Association*, 285(7) :914.
- [Guan et al., 2006] Guan, H., Kubota, T., Huang, X., Zhou, X., and Turk, M. (2006). Automatic hot spot detection and segmentation in whole body fdg-pet images. In *Image Processing, 2006 IEEE International Conference on*, pages 85–88. IEEE.
- [Guivarc’h et al., 2004a] Guivarc’h, O., Turzo, A., Visvikis, D., and Bizais, Y. (2004a). Synchronization of pulmonary scintigraphy by respiratory flow and by impedance plethysmography. In *Proceedings of SPIE*, volume 5370, page 1166.
- [Guivarc’h et al., 2004b] Guivarc’h, O., Turzo, A., Visvikis, D., and Bizais, Y. (2004b). Synchronization of pulmonary scintigraphy by respiratory flow and by impedance plethysmography. 5370(1) :1166–1175.
- [Gurcan et al., 2002] Gurcan, M., Sahiner, B., Petrick, N., Chan, H., Kazerooni, E., Cascade, P., and Hadjiiski, L. (2002). Lung nodule detection on thoracic computed tomography images : preliminary evaluation of a computer-aided diagnosis system. *Medical Physics*, 29 :2552.

BIBLIOGRAPHIE

- [Haykin, 1999] Haykin, S. (1999). *Neural networks : a comprehensive foundation*. Prentice hall.
- [Herwig et al., 1999] Herwig, R., Poustka, A., Müller, C., Bull, C., Lehrach, H., and O'Brien, J. (1999). Large-scale clustering of cDNA-fingerprinting data. *Genome research*, 9(11) :1093.
- [Ho, 1998] Ho, T. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8) :832–844.
- [Horn and Schunck, 1981] Horn, B. and Schunck, B. (1981). Determining optical flow. *Artificial intelligence*, 17(1-3) :185–203.
- [Hua et al., 2005] Hua, J., Xiong, Z., Lowey, J., Suh, E., and Dougherty, E. (2005). Optimal number of features as a function of sample size for various classification rules. *Bioinformatics*, 21(8) :1509.
- [Hudson, 1994] Hudson, H. (1994). Accelerated Image Reconstruction Using Ordered Subsets of Projection Data. *IEEE Transactions On Medical Imaging*, 13(4).
- [Jafar et al., 2006] Jafar, I., Ying, H., Shields, A., and Muzik, O. (2006). Computerized detection of lung tumors in PET/CT images. In *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, pages 2320–2323. IEEE.
- [Jan et al., 2004] Jan, S., Santin, G., Strul, D., Staelens, S., Assie, K., Autret, D., Avner, S., Barbier, R., Bardies, M., Bloomfield, P., et al. (2004). GATE : a simulation toolkit for PET and SPECT. *Physics in medicine and biology*, 49 :4543.
- [Kanakatte et al., 2007] Kanakatte, A., Gubbi, J., Mani, N., Kron, T., and Binns, D. (2007). A pilot study of automatic lung tumor segmentation from positron emission tomography images using standard uptake values. In *Computational Intelligence in Image and Signal Processing, 2007. CIISP 2007. IEEE Symposium on*, pages 363–368. IEEE.
- [Kanakatte et al., 2008] Kanakatte, A., Mani, N., Srinivasan, B., and Giibbi, J. (2008). Pulmonary tumor volume detection from positron emission tomography images. In *BioMedical Engineering and Informatics, 2008. BMEI 2008. International Conference on*, volume 2, pages 213–217. IEEE.
- [Kinahan et al., 1988] Kinahan, P., Rogers, J., Harrop, R., and Johnson, R. (1988). Three-dimensional image reconstruction in object space. *IEEE Transactions on Nuclear Science*, 35(1) :635–638.
- [Klein and Huesman, 1997] Klein, G. and Huesman, R. (1997). A 3D optical flow approach to addition of deformable PET volumes. *nam*, page 0136.
- [Klein et al., 2001] Klein, G., Reutter, R., and Huesman, R. (2001). Four-dimensional affine registration models for respiratory-gated PET. *IEEE Transactions on Nuclear Science*, 48(3) :756–760.
- [Kohonen, 1982] Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1) :59–69.
- [Krak et al., 2005] Krak, N., Boellaard, R., Hoekstra, O., Twisk, J., Hoekstra, C., and Lammertsma, A. (2005). Effects of ROI definition and reconstruction method on

BIBLIOGRAPHIE

- quantitative outcome and applicability in a response monitoring trial. *European journal of nuclear medicine and molecular imaging*, 32(3) :294–301.
- [Lamare et al., 2007a] Lamare, F., Cresson, T., Savean, J., Cheze Le Rest, C., Reader, A., and Visvikis, D. (2007a). Respiratory motion correction using affine transformation of list mode data in PET. *Physics in medicine and biology*, 52 :121–140.
- [Lamare et al., 2007b] Lamare, F., Ledesma Carbayo, M., Cresson, T., Kontaxakis, G., Santos, A., Le Rest, C., Reader, A., and Visvikis, D. (2007b). List-mode-based reconstruction for respiratory motion correction in PET using non-rigid body transformations. *Physics in medicine and biology*, 52(17) :5187–5204.
- [Lamare et al., 2006] Lamare, F., Turzo, A., Bizais, Y., Rest, C., and Visvikis, D. (2006). Validation of a Monte Carlo simulation of the Philips Allegro/GEMINI PET systems using GATE. *Physics in medicine and biology*, 51 :943.
- [Langner, 2008] Langner, J. (2008). *Event-Driven Motion Compensation in Positron Emission Tomography : Development of a Clinically Applicable Method*. PhD thesis, Faculty of Medicine Carl Gustav Carus, University of Technology Dresden, Germany.
- [Lartizien et al., 2010] Lartizien, C., Marache-Francisco, S., Le Pogam, A., Tomei, S., and Visvikis, D. (2010). Impact of partial volume correction in whole-body PET imaging : A computer-aided detection study. In *Nuclear Science Symposium Conference Record (NSS/MIC), 2009 IEEE*, pages 3214–3217. IEEE.
- [Le Maitre et al., 2009] Le Maitre, A., Segars, W., Marache, S., Reilhac, A., Hatt, M., Tomei, S., Lartizien, C., and Visvikis, D. (2009). Incorporating Patient-Specific Variability in the Simulation of Realistic Whole-Body ^{18}F FDG Distributions for Oncology Applications. *Proceedings of the IEEE*, 97(12) :2026–2038.
- [Ledesma-Carbayo et al., 2005] Ledesma-Carbayo, M., Kybic, J., Desco, M., Santos, A., Suhling, M., Hunziker, P., and Unser, M. (2005). Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation. *IEEE transactions on medical imaging*, 24(9) :1113–1126.
- [Leduvehat, 2010] Leduvehat, L. (2010). Développement d'un modèle de simulation réaliste d'images TDM médicales avec SINDBAD. Master's thesis, ENSEIRB-MATMECA-Bordeaux.
- [Lemaître, 2008] Lemaître, A. (2008). Développement d'un modèle de simulation réaliste représentant la distribution de radioactivité chez un patient. Master's thesis, Telecom Bretagne.
- [Li, 2007] Li, Q. (2007). Recent progress in computer-aided diagnosis of lung nodules on thin-section CT. *Computerized Medical Imaging and Graphics*, 31(4-5) :248–257.
- [Li et al., 2003] Li, Q., Sone, S., and Doi, K. (2003). Selective enhancement filters for nodules, vessels, and airway walls in two-and three-dimensional CT scans. *Medical Physics*, 30 :2040.
- [Livieratos et al., 2005] Livieratos, L., Stegger, L., Bloomfield, P., Schafers, K., Bailey, D., and Camici, P. (2005). Rigid-body transformation of list-mode projection data for respiratory motion correction in cardiac PET. *Physics in Medicine and Biology*, 50 :3313.

BIBLIOGRAPHIE

- [Loening and Gambhir, 2003] Loening, A. and Gambhir, S. (2003). AMIDE : a free software tool for multimodality medical image analysis. *Molecular imaging*, 2(3) :131–137.
- [Metz, 1986] Metz, C. (1986). ROC methodology in radiologic imaging. *Investigative Radiology*, 21(9) :720.
- [Mhd Saeed et al., 2010] Mhd Saeed, S., Maysam, A., Abbes, A., and Habib, Z. (2010). Artificial Neural Network-Based System for PET Volume Segmentation. *International Journal of Biomedical Imaging*, 2010.
- [Mierswa et al., 2006] Mierswa, I., Wurst, M., Klinkenberg, R., Scholz, M., and Euler, T. (2006). Yale : Rapid prototyping for complex data mining tasks. In Ungar, L., Craven, M., Gunopulos, D., and Eliassi-Rad, T., editors, *KDD '06 : Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 935–940, New York, NY, USA. ACM.
- [Mougiakakou et al., 2007] Mougiakakou, S., Valavanis, I., Nikita, A., and Nikita, K. (2007). Differential diagnosis of CT focal liver lesions using texture features, feature selection and ensemble driven classifiers. *Artificial Intelligence in Medicine*, 41(1) :25–37.
- [Nehmeh et al., 2002] Nehmeh, S., Erdi, Y., Ling, C., Rosenzweig, K., Schoder, H., Larson, S., Macapinlac, H., Squire, O., and Humm, J. (2002). Effect of respiratory gating on quantifying PET images of lung cancer. *Journal of Nuclear Medicine*, 43(7) :876.
- [Nehmeh et al., 2003] Nehmeh, S., Erdi, Y., Rosenzweig, K., Schoder, H., Larson, S., Squire, O., and Humm, J. (2003). Reduction of respiratory motion artifacts in PET imaging of lung cancer by respiratory correlated dynamic PET : methodology and comparison with respiratory gated PET. *Journal of Nuclear Medicine*, 44(10) :1644.
- [Nie et al., 2006] Nie, Y., Li, Q., Li, F., Pu, Y., Appelbaum, D., and Doi, K. (2006). Integrating PET and CT information to improve diagnostic accuracy for lung nodules : A semiautomatic computer-aided method. *Journal of Nuclear Medicine*, 47(7) :1075–1080.
- [Paik et al., 2004] Paik, D., Beaulieu, C., Rubin, G., Acar, B., Jeffrey Jr, R., Yee, J., Dey, J., and Napel, S. (2004). Surface normal overlap : a computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT. *IEEE Transactions on Medical Imaging*, 23(6) :661–675.
- [Polak and Ribiere, 1969] Polak, E. and Ribiere, G. (1969). Note sur la convergence de méthodes de directions conjuguées. *Revue Fr Inf Rech Oper*, 3(1) :35–43.
- [Potesil et al., 2007] Potesil, V., Huang, X., and Zhou, X. (2007). Automated tumor delineation using joint PET/CT information. In *Proc. SPIE International Symposium on Medical Imaging : Computer-Aided Diagnosis*, volume 6514.
- [Qiao et al., 2006] Qiao, F., Pan, T., John Jr, W., and Mawlawi, O. (2006). A motion-incorporated reconstruction method for gated PET studies. *Physics in Medicine and Biology*, 51 :3769.
- [Rahmim et al., 2011] Rahmim, A., Tang, J., Ay, M., and Bengel, F. (2011). 4D Respiratory Motion-Corrected Rb-82 Myocardial Perfusion PET Image Reconstruction.

BIBLIOGRAPHIE

- [Reader et al., 2002] Reader, A., Ally, S., Bakatselos, F., Manavaki, R., Walledge, R., Jeavons, A., Julyan, P., Zhao, S., Hastings, D., and Zweit, J. (2002). One-pass list-mode EM algorithm for high-resolution 3D PET image reconstruction into large arrays. *IEEE Transactions on Nuclear Science*, 49(3) :693 – 699.
- [Reilhac et al., 2005] Reilhac, A., Batan, G., Michel, C., Grova, C., Tohka, J., Collins, D., Costes, N., and Evans, A. (2005). PET-SORTEO : validation and development of database of simulated PET volumes. *IEEE Transactions on Nuclear Science*, 52(5 Part 1) :1321–1328.
- [Reilhac et al., 2004] Reilhac, A., Lartizien, C., Costes, N., Sans, S., Comtat, C., Gunn, R., and Evans, A. (2004). PET-SORTEO : A Monte Carlo-based simulator with high count rate capabilities. *IEEE Transactions on Nuclear Science*, 51(1 Part 1) :46–52.
- [Rohren et al., 2004] Rohren, E., Turkington, T., and Coleman, R. (2004). Clinical applications of PET in oncology. *Radiology*, 231(2) :305.
- [Sachs et al., 1971] Sachs, M., Nachmias, J., and Robson, J. (1971). Spatial-frequency channels in human vision. *JOSA*, 61(9) :1176–1186.
- [Salvat et al., 2006] Salvat, F., Fernández-Varea, J., and Sempau, J. (2006). PENELOPE-2006 : A code system for Monte Carlo simulation of electron and photon transport. In *Workshop Proceedings*, volume 4, page 7.
- [Saradhi et al., 2009] Saradhi, G., Gopalakrishnan, G., Roy, A., Mullick, R., Manjeshwar, R., Thielemans, K., and Patil, U. (2009). A framework for automated tumor detection in thoracic FDG PET images using texture-based features. In *Biomedical Imaging : From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pages 97–100. IEEE.
- [Schöder et al., 2004] Schöder, H., Erdi, Y., Chao, K., Gonen, M., Larson, S., and Yeung, H. (2004). Clinical implications of different image reconstruction parameters for interpretation of whole-body PET studies in cancer patients. *Journal of Nuclear Medicine*, 45(4) :559.
- [Segars, 2001] Segars, W. (2001). *Development and Application of the new dynamic nurbs-based cardiac-torso (NCAT) phantom*. PhD thesis.
- [Segars and Tsui, 2009] Segars, W. and Tsui, B. (2009). The MCAT, NCAT, XCAT, and MOBY Computational Human and Mouse Phantoms. *Handbook of Anatomical Models for Radiation Dosimetry*, page 105.
- [Servant-Laval, 2007] Servant-Laval, A. (2007). Cours anatomie fonctionnelle. Disponible sur Internet à l'adresse <http://www.chups.jussieu.fr/polysPSM/anatfonctPSM2/poly/POLY.Chp.8.html>.
- [Shepp and Vardi, 1982] Shepp, L. and Vardi, Y. (1982). Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1(2) :113–122.
- [Sherwood, 2006] Sherwood, L. (2006). *Fundamentals of physiology : a human perspective*. Thomson Brooks/Cole.
- [Statnikov et al., 2008] Statnikov, A., Wang, L., and Aliferis, C. (2008). A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification. *BMC bioinformatics*, 9(1) :319.

BIBLIOGRAPHIE

- [Stephan et al., 2003] Stephan, C., Wesseling, S., Schink, T., and Jung, K. (2003). Comparison of eight computer programs for receiver-operating characteristic analysis. *Clin Chem*, 49(3) :433–439.
- [Swets and Pickett, 1982] Swets, J. and Pickett, R. (1982). *Evaluation of diagnostic systems : methods from signal detection theory*. Academic Press New York.
- [Takahara et al., 2004] Takahara, T., Imai, Y., Yamashita, T., Yasuda, S., Nasu, S., and Van Cauteren, M. (2004). Diffusion weighted whole body imaging with background body signal suppression (DWIBS) : technical improvement using free breathing, STIR and high resolution 3D display. *Matrix*, 160(160) :160.
- [Thévenaz and Unser, 2000] Thévenaz, P. and Unser, M. (2000). Optimization of mutual information for multiresolution image registration. *IEEE Transactions on Image Processing*, 9(12) :2083–2099.
- [Thielemans et al., 2006] Thielemans, K., Manjeshwar, R., Tao, X., and Asma, E. (2006). Lesion detectability in motion compensated image reconstruction of respiratory gated PET/CT. In *Nuclear Science Symposium Conference Record, 2006. IEEE*, volume 6, pages 3278–3282.
- [Thorndyke et al., 2006] Thorndyke, B., Schreibmann, E., Koong, A., and Xing, L. (2006). Reducing respiratory motion artifacts in positron emission tomography through retrospective stacking. *Medical physics*, 33 :2632.
- [Tomei et al., 2008a] Tomei, S., Marache-Francisco, S., Odet, C., and Lartizien, C. (2008a). Automatic detection of active nodules in 3D PET oncology imaging using the Hotelling Observer and the Support Vector Machines : a comparison study. In *Nuclear Science Symposium Conference Record, 2008. NSS'08. IEEE*, pages 5314–5319. IEEE.
- [Tomei et al., 2010] Tomei, S., Reilhac, A., Visvikis, D., Boussion, N., Odet, C., Giammarile, F., and Lartizien, C. (2010). OncoPET.DB : A Freely Distributed Database of Realistic Simulated Whole Body ^{18}F FDG PET Images for Oncology. *IEEE Transactions on Nuclear Science*, 57(1) :246–255.
- [Tomei et al., 2008b] Tomei, S., Reilhac, A., Visvikis, D., Odet, C., Giammarile, F., Mognetti, T., and Lartizien, C. (2008b). Development of a database of realistic simulated whole body ^{18}F -FDG PET images for lymphoma. In *Nuclear Science Symposium Conference Record, 2008. NSS'08. IEEE*, pages 4958–4963. IEEE.
- [Vandemeulebroucke et al., 2009] Vandemeulebroucke, J., Kybic, J., Clarysse, P., and Sarrut, D. (2009). Respiratory motion estimation from cone-beam projections using a prior model. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pages 365–372.
- [Varma and Simon, 2006] Varma, S. and Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC bioinformatics*, 7(1) :91.
- [Visvikis et al., 2004] Visvikis, D., Barret, O., Fryer, T., Lamare, F., Turzo, A., Bizais, Y., and Le Rest, C. (2004). Evaluation of respiratory motion effects in comparison with other parameters affecting PET image quality. In *Nuclear Science Symposium Conference Record, 2004 IEEE*, volume 6, pages 3668–3672. IEEE.

BIBLIOGRAPHIE

- [Ward Jr, 1963] Ward Jr, J. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, pages 236–244.
- [Wiemker et al., 2008] Wiemker, R., Paulus, T., Kabus, S., Bülow, T., Apostolova, I., Buchert, R., and Klutmann, S. (2008). Combined motion blur and partial volume correction for computer aided diagnosis of pulmonary nodules in PET/CT. *International Journal of Computer Assisted Radiology and Surgery*, 3(1) :105–113.
- [Ying et al., 2004] Ying, H., Zhou, F., Shields, A., Muzik, O., Wu, D., and Heath, E. (2004). A novel computerized approach to enhancing lung tumor detection in whole-body PET images. In *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, volume 1, pages 1589–1592. IEEE.
- [Yu et al., 2009] Yu, H., Caldwell, C., Mah, K., Poon, I., Balogh, J., MacKenzie, R., Khaouam, N., and Tirona, R. (2009). Automated radiation targeting in head-and-neck cancer using region-based texture analysis of PET and CT images. *International Journal of Radiation Oncology* Biology* Physics*, 75(2) :618–625.
- [Zhao, 2003] Zhao, B. (2003). Automatic detection of small lung nodules on CT utilizing a local density maximum algorithm. *journal of applied clinical medical physics*, 4(3) :248–260.
- [Zou et al., 2007] Zou, K., O'Malley, A., and Mauri, L. (2007). Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 115(5) :654.

BIBLIOGRAPHIE

Conférences premier auteur

S. Marache, R. Prost, J. M. Rouet, et C. Lartizien, (2011). Incorporating Patient-Specific Variability in the OncoPET_DB Database, *IEEE Nuclear Science Symposium and Medical Imaging Conference*, Valence, Espagne, sous presse, Octobre 2011

S. Marache, F. Lamare, H. Fayad, D. Visvikis, R. Prost, J. M. Rouet, and C. Lartizien, (2010), Impact of Respiratory Motion Correction on the Detection of Small Lesions in Whole-body PET Imaging : A Simulation Study, *IEEE Nuclear Science Symposium and Medical Imaging Conference*, Knoxville, Tennessee, USA., pp. 3531-3533, Octobre 2010.

Articles

C. Lartizien, S. Marache, R. Prost (2012), Automatic Detection of Lung and Liver Lesions in 3D Positron Emission Tomography Images : a Pilot Study, *IEEE Trans. on Nuclear Sciences*, sous presse.

A. Lemaitre, P. Segars, S. Marache, A. Reilhac, M. Hatt, S. Tomei, C. Lartizien, and D. Visvikis, (2009), Incorporating patient specific variability in the simulation of realistic whole body 18F-FDG distributions for oncology applications, *Proceedings of the IEEE*, vol. 97, no. 12, 2009.

Conférences

C. Lartizien, S. Marache, A. Le Pogam, S. Tomei, and D. Visvikis, (2009), Impact of Partial Volume Correction in Whole-Body PET Imaging : a Computer-Aided Detection Study, *Proc. of the 2009 IEEE NSS and MIC Conference*, Orlando, Floride, USA, pp. 3214-3217, Octobre 2009.

C. Lartizien, S. Tomei, and S. Marache, (2009), Improving performance of automated detection systems for 3D PET oncology imaging by use of resampled training images, *Proc. 23rd Int. CARS Conference*, Berlin, Allemagne, Springer, vol. 4, pp. (SS186-193), Juin 2009.

J. Tabary, S. Marache, S. Valette, W. Segars, and C. Lartizien, (2009), Realistic X-Ray CT Simulation of the XCAT Phantom with SINDBAD, *Proc. of the 2009 IEEE NSS and MIC Conference*, Orlando, Floride, USA, pp. 3980 - 3983, Octobre 2009.

S. Tomei, S. Marache, C. Odet, and C. Lartizien, (2008), Automatic detection of active nodules in 3D PET oncology imaging using the Hotelling Observer and the Support Vector Machines : a comparison study, *Proc of the 2008 IEEE NSS and MIC Conference*, Dresde, Allemagne, pp. 5314-5319, Octobre 2008.

INSA Direction de la Recherche - Ecoles Doctorales – Quinquennal 2011-2015

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	CHIMIE DE LYON http://www.edchimie-lyon.fr Insa : R. GOURDON	M. Jean Marc LANCELIN Université de Lyon – Collège Doctoral Bât ESCPE 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cedex Tél : 04.72.43 13 95 directeur@edchimie-lyon.fr
E.E.A.	ELECTRONIQUE, ELECTROTECHNIQUE, AUTOMATIQUE http://edeea.ec-lyon.fr Secrétariat : M.C. HAVGOUDOUKIAN eea@ec-lyon.fr	M. Gérard SCORLETTI Ecole Centrale de Lyon 36 avenue Guy de Collongue 69134 ECULLY Tél : 04.72.18 60 97 Fax : 04 78 43 37 17 Gerard.scorletti@ec-lyon.fr
E2M2	EVOLUTION, ECOSYSTEME, MICROBIOLOGIE, MODELISATION http://e2m2.universite-lyon.fr Insa : H. CHARLES	Mme Gudrun BORNETTE CNRS UMR 5023 LEHNA Université Claude Bernard Lyon 1 Bât Forel 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cédex Tél : 04.72.43.12.94 e2m2@biomserv.univ-lyon1.fr
EDISS	INTERDISCIPLINAIRE SCIENCES-SANTE http://ww2.ibcp.fr/ediss Sec : Safia AIT CHALAL Insa : M. LAGARDE	M. Didier REVEL Hôpital Louis Pradel Bâtiment Central 28 Avenue Doyen Lépine 69677 BRON Tél : 04.72.68 49 09 Fax :04 72 35 49 16 Didier.revel@creatis.uni-lyon1.fr
INFOMATHS	INFORMATIQUE ET MATHEMATIQUES http://infomaths.univ-lyon1.fr	M. Johannes KELLENDONK Université Claude Bernard Lyon 1 INFOMATHS Bâtiment Braconnier 43 bd du 11 novembre 1918 69622 VILLEURBANNE Cedex Tél : 04.72. 44.82.94 Fax 04 72 43 16 87 infomaths@univ-lyon1.fr
Matériaux	MATERIAUX DE LYON Secrétariat : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Saint Exupéry Ed.materiaux@insa-lyon.fr	M. Jean-Yves BUFFIERE INSA de Lyon MATEIS Bâtiment Saint Exupéry 7 avenue Jean Capelle 69621 VILLEURBANNE Cédex Tél : 04.72.43 83 18 Fax 04 72 43 85 28 Jean-yves.buffiere@insa-lyon.fr
MEGA	MECANIQUE, ENERGETIQUE, GENIE CIVIL, ACOUSTIQUE Secrétariat : M. LABOUNE PM : 71.70 –Fax : 87.12 Bat. Saint Exupéry mega@insa-lyon.fr	M. Philippe BOISSE INSA de Lyon Laboratoire LAMCOS Bâtiment Jacquard 25 bis avenue Jean Capelle 69621 VILLEURBANNE Cedex Tél :04.72.43.71.70 Fax : 04 72 43 72 37 Philippe.boisse@insa-lyon.fr
ScSo	ScSo* M. OBADIA Lionel Sec : Viviane POLSINELLI Insa : J.Y. TOUSSAINT	M. OBADIA Lionel Université Lyon 2 86 rue Pasteur 69365 LYON Cedex 07 Tél : 04.78.69.72.76 Fax : 04.37.28.04.48 Lionel.Obadia@univ-lyon2.fr

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

FOLIO ADMINISTRATIF

THESE SOUTENUE DEVANT L'INSTITUT NATIONAL DES SCIENCES APPLIQUEES DE LYON

NOM : MARACHE-FRANCISCO
(avec précision du nom de jeune fille, le cas échéant)

DATE de SOUTENANCE : 14/02/2012

Prénoms : Simon

TITRE : Évaluation de la correction du mouvement respiratoire sur la détection des lésions en oncologie TEP

NATURE : Doctorat

Numéro d'ordre : AAAAISALXXXX

Ecole doctorale : École doctorale: Électronique, Électrotechnique, Automatique

Spécialité : Images et Systèmes

RESUME :

La tomographie par émission de positons (TEP) est une méthode d'imagerie clinique en forte expansion dans le domaine de l'oncologie. De nombreuses études cliniques montrent que la TEP permet, d'une part de diagnostiquer et caractériser les lésions cancéreuses à des stades plus précoces que l'imagerie anatomique conventionnelle, et d'autre part d'évaluer plus rapidement la réponse au traitement. Le raccourcissement du cycle comprenant le diagnostic, la thérapie, le suivi et la réorientation thérapeutiques contribue à augmenter le pronostic vital du patient et maîtriser les coûts de santé.

La durée d'un examen TEP ne permet pas de réaliser une acquisition sous apnée. La qualité des images TEP est par conséquent affectée par les mouvements respiratoires du patient qui induisent un flou dans les images. Les effets du mouvement respiratoire sont particulièrement marqués au niveau du thorax et de l'abdomen.

Plusieurs types de méthode ont été proposés pour corriger les données de ce phénomène, mais elles demeurent lourdes à mettre en place en routine clinique. La problématique de la correction du mouvement respiratoire et le choix de la méthode appropriée sont des sujets d'actualité au sein de la communauté de médecine nucléaire. Des travaux récemment publiés proposent une évaluation de ces méthodes basée sur des critères de qualité tels que le rapport signal sur bruit ou le biais. Aucune étude à ce jour n'a évalué l'impact de ces corrections sur la qualité du diagnostic clinique. Ce problème pose des questions d'orientation stratégique et financière importantes.

Nous nous sommes focalisés sur la problématique de la détection des lésions du thorax et de l'abdomen de petit diamètre et faible contraste, qui sont les plus susceptibles de bénéficier de la correction du mouvement respiratoire en routine clinique.

Nos travaux ont consisté dans un premier temps à construire une base d'images TEP qui modélisent un mouvement respiratoire non-uniforme, une variabilité inter-individuelle et contiennent un échantillonnage de lésions de taille et de contraste variable. Ce cahier des charges nous a orientés vers les méthodes de simulation Monte Carlo qui permettent de contrôler l'ensemble des paramètres influençant la formation et la qualité de l'image. Une base de 15 modèles de patient a été créée en adaptant le modèle anthropomorphe XCAT sur des images tomodynamométriques (TDM) de patients.

Nous avons en parallèle développé une stratégie originale d'évaluation des performances de détection. Cette méthode comprend un système de détection des lésions automatisé basé sur l'utilisation de machines à vecteurs de support (SVM) utilisant des caractéristiques fréquentielles. Les performances sont mesurées par l'analyse psychophysique des courbes free-receiver operating characteristics (FROC) que nous avons adaptée aux spécificités de l'imagerie TEP.

L'évaluation des performances est réalisée sur deux techniques de correction du mouvement respiratoire, en les comparant avec les performances obtenues sur des images non corrigées ainsi que sur des images sans mouvement respiratoire.

Les résultats obtenus sont prometteurs et montrent une réelle amélioration de la détection des lésions après correction, qui approche les performances obtenues sur les images statiques.

MOTS-CLES : TEP, oncologie, mouvement respiratoire, reconnaissance de formes, Computer aided detection (CAD), Séparateurs à vaste marge (SVM), Simulation Monte-Carlo

Laboratoire (s) de recherche : CREATIS, INSA Lyon

Directeur de thèse: Rémy Prost, Professeur INSA Lyon, Université de Lyon, CREATIS ; CNRS UMR5220 ; Inserm U1044 ; INSA-Lyon ; Université Lyon 1, France.

Co-directrice de thèse : Carole Lartizien, Chargée de Recherche CNRS, Université de Lyon, CREATIS ; CNRS UMR5220 ; Inserm U1044 ; INSA-Lyon ; Université Lyon 1, France.

Président de jury :

Composition du jury :

Michel Desvignes
Su Ruan

Professeur, GIPSA-Lab, INP Grenoble, Rapporteur
Professeur, LITIS, Université de Rouen, Rapporteur

Frédéric Lamare
Christophe Odet
Jean-Michel Rouet

Docteur, Ingénieur de Recherche, INCIA, Examinateur
Professeur, CREATIS, INSA-lyon, Examinateur
Docteur, Ingénieur de Recherche Philips Healthcare Medisys, Examinateur

Carole Lartizien
Rémy Prost

CR CNRS, CREATIS, Lyon, Co-encadrante
Professeur, CREATIS, INSA-Lyon, Directeur de thèse