



HAL
open science

Détermination automatique des volumes fonctionnels en imagerie d'émission pour les applications en oncologie

Mathieu Hatt

► **To cite this version:**

Mathieu Hatt. Détermination automatique des volumes fonctionnels en imagerie d'émission pour les applications en oncologie. Informatique [cs]. Université de Bretagne occidentale - Brest, 2008. Français. NNT: . tel-00373581

HAL Id: tel-00373581

<https://theses.hal.science/tel-00373581>

Submitted on 6 Apr 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Bretagne Occidentale - Telecom Bretagne

DETERMINATION AUTOMATIQUE DES VOLUMES
FONCTIONNELS EN IMAGERIE D'EMISSION POUR LES
APPLICATIONS EN ONCOLOGIE

THESE

Soutenue le 3 décembre 2008 pour l'obtention du

Doctorat de l'Université de Bretagne Occidentale et de Telecom Bretagne
(Santé, Information, Communication, Matière et Mathématiques)

par

Mathieu Hatt

Directeur de thèse : Prof. Christian Roux

Encadrant : Dr. Dimitris Visvikis

Financement : ce projet a bénéficié du soutien financier de la Région Bretagne

Rapporteurs : Professeurs Wojciech Pieczynski et Roland Hustinx

Membres : Professeurs Olivier Pradier et Christian Roux

Invités : Docteurs Dimitris Visvikis et Catherine Cheze Le Rest

INSERM U650 - Laboratoire de Traitement de l'Information Médicale (LaTIM)

IFR 148 - Science et Ingénierie en Biologie-Santé (ScInBioS)

Université de Bretagne Occidentale - école doctorale SICMA, Institut Telecom - Telecom Bretagne



Ce n'est pas une image juste, c'est juste une image

- Jean-Luc Godard dans "sauve qui peut la vie"

Remerciements

Ce doctorat m'a permis de travailler au sein d'une équipe composée de collègues motivés et dynamiques, engagés et chaleureux. Je tiens à remercier tout particulièrement Dimitris Visvikis pour son encadrement exceptionnel, sa disponibilité et ses conseils. Dimitris a été un mentor présent, encourageant et motivant, prompt à partager son expérience. Je le remercie pour ses efforts, notamment en ce qui concerne la valorisation scientifique de ce travail et son inclination à m'associer à la mise en place de projets intéressants et valorisants.

Merci à Christian Roux pour sa direction, ses conseils et notamment pour avoir encouragé et facilité ma participation à l'école d'été en imagerie Biomedicale¹ sur l'île de Berder, qui a été une expérience enrichissante.

Merci aux professeurs Wojciech Pieczynski et Roland Hustinx pour avoir accepté la lourde tâche d'être les rapporteurs de cette thèse.

Merci également au docteur Nicolas BouSSION et au doctorant Adrien Le Pogam pour leur enthousiasme et leur disponibilité pour les travaux effectués en collaboration au sein de l'équipe, dans une atmosphère cordiale permettant des échanges stimulants.

Merci au médecin nucléaire et docteur Catherine Cheze Le Rest pour ses nombreuses explications pédagogiques concernant les aspects médicaux de l'imagerie TEP, et pour m'aider à replacer mon travail dans ce contexte précis. Merci à Alexandre Turzo pour son aide et ses explications lors des acquisitions de fantôme.

Je voudrais également remercier les docteurs Frédéric Lamare et Thierry Cresson, qui ont été présents durant la première partie de ma thèse et dont le partage d'expérience et l'aide volontiers offerte ont été précieux, notamment pour l'utilisation des outils de simulation. Le docteur Patrice Descourt a également été d'une grande aide pour la compréhension des simulations.

Les collaborations nombreuses avec des scientifiques extérieurs au LaTIM ont été nécessaires, notamment pour obtenir des données sur lesquelles travailler. Je voudrais donc remercier en particulier le docteur Pascal Bailly du service de médecine nucléaire du CHU Sud Amiens, l'équipe de la clinique MAASTRO de Maastricht aux Pays-bas avec en particulier Andre

¹7th IEEE EMBS International Summer School on Biomedical Imaging

Dekker, et enfin l'équipe du service de médecine nucléaire du Royal Hospital de Belfast, en Irlande, en particulier Kathryn Carson et Ashraf Zadari. Merci également à Paul Segars pour sa collaboration et son aide sur l'utilisation du fantôme NCAT pour les simulations réalistes d'images corps entier de patients.

Merci plus généralement à tous les membres du LaTIM, techniciens, doctorants, post-docs, ingénieurs et chercheurs pour leur sympathie et l'ambiance cordiale qu'ils instaurent sur le lieu de travail.

Merci à ma famille et à mes amis, tant à Strasbourg qu'à Brest, pour leur soutien constant et leurs encouragements réguliers.

Un grand merci au Professeur Patrick Lutz et tout son service d'onco-hématologie de l'enfant de l'hôpital de Hautepierre, sans qui je n'aurais pas pu faire cette thèse.

Enfin, merci à tous les développeurs de Mario Kart Wii, Far Cry, Flat out, Unreal Tournament et Urban Terror pour avoir créé des jeux amusants pour se défouler entre collègues pendant les pauses.

Table des matières

Abstract	11
Résumé	13
Détermination des volumes fonctionnels : enjeux et objectifs	25
1 Imagerie TEP / TDM	31
1.1 Historique	31
1.2 Physique de l'imagerie	32
1.2.1 Fluor 18	32
1.2.2 Annihilation	33
1.2.3 Radiopharmaceutique ou radiotracteur	34
1.2.4 Chaîne de détection	36
1.2.5 Atténuation, LORs diffusées et aléatoires	38
1.3 Imagerie multimodalités	41
1.4 L'examen 18F-TEP	43
1.5 De l'acquisition à l'image	44
1.5.1 Acquisition	44
1.5.2 Reconstruction tomographique	47
1.5.3 Temps de vol	49
1.6 Principales limitations	52
1.6.1 Bruit	53

1.6.2	Résolution spatiale	53
1.6.3	Mouvements et respiration	57
1.7	Contexte clinique, besoins et objectifs	58
1.7.1	Analyse semi-quantitative et SUV	59
1.7.2	Utilisation diagnostique de la TEP en oncologie	60
1.7.3	Suivi thérapeutique	61
1.7.4	Radiothérapie	61
2	Etat de l’art de la segmentation en TEP	65
2.1	Introduction	65
2.1.1	Objectif	65
2.1.2	Définitions	66
2.2	Approches manuelles	70
2.3	Approches semi-automatiques : seuillages	71
2.3.1	Seuillages fixes	72
2.3.2	Seuillages adaptatifs	73
2.4	Approches automatiques	78
2.4.1	Clustering	78
2.4.2	Watersheds	81
2.4.3	Filtrages et détection de contours	82
2.4.4	Modèles déformables	82
2.4.5	Inférence bayésienne et segmentation statistique d’image	84
2.5	Conclusion	85
3	Méthodologies développées	89
3.1	Approche choisie : combinaison du modèle statistique et de la mesure floue	89
3.1.1	Inférence bayésienne et segmentation statistique	89
3.1.2	Travail sur l’image	92
3.1.3	Segmentation et détectabilité	93
3.1.4	Délimitation du volume et quantification	94

3.1.5	Prise en compte du bruit et du flou	94
3.1.6	Idée générale et historique	95
3.1.7	Implémentation pratique	96
3.1.8	Limites et extensions possibles	96
3.2	Approche globale : chaînes de Markov	96
3.2.1	Approche globale : hypothèse de Markov	97
3.2.2	Modèle <i>a priori</i> : chaîne de Markov	97
3.2.3	Chaîne de Markov floue	99
3.2.4	Lois des observations	100
3.2.5	Estimation des paramètres	100
3.2.6	Segmentation	102
3.3	Approche locale adaptative	102
3.3.1	Approche locale sans hypothèse markovienne	102
3.3.2	Estimation adaptative	103
3.3.3	Loi <i>a priori</i>	104
3.3.4	Loi des observations	104
3.3.5	Estimation des paramètres	105
3.3.6	Segmentation	105
3.4	Approche locale étendue à trois classes	106
3.4.1	Problématique	106
3.4.2	Extension à trois classes dures	109
3.4.3	Modélisation des transitions floues	110
3.4.4	Loi des observations	111
3.4.5	Estimation des paramètres	111
3.4.6	Segmentation	113
3.5	Modélisation de la loi des observations	113
3.5.1	Système de Pearson	114
4	Validation et optimisation des méthodes	119

4.1	Données de validation et d'évaluation	119
4.1.1	Fantôme	119
4.1.2	Protocole d'acquisition commun	120
4.1.3	Acquisitions du fantôme IEC obtenues avec différents scanners	121
4.1.4	Images synthétiques d'objets hétérogènes	127
4.2	Analyse et mesure de performance	128
4.3	Optimisation sur sphères homogènes	132
4.3.1	Optimisation du nombre de niveaux de flou et valeurs associées	133
4.3.2	Optimisation de la loi des observations	136
4.3.3	Optimisation de la taille du cube d'estimation dans le cas de FLAB	138
4.3.4	Importance de l'estimation adaptative dans FLAB	140
4.3.5	Reproductibilité de FLAB	141
4.4	Erreurs de classification et de volume sur les acquisitions de fantôme	142
4.4.1	Resultats	142
4.4.2	Conclusions de l'étude sur fantôme	145
4.5	Validation sur objets non sphériques et non homogènes	147
4.6	Résultats de l'étude sur la robustesse	150
5	Performances des méthodes dans les applications cliniques	155
5.1	Images de tumeurs réalistes	155
5.1.1	Intérêt, avantages et inconvénients	155
5.1.2	Images cliniques	156
5.1.3	Images simulées de tumeurs	157
5.2	Précision de segmentation sur tumeurs simulées et réelles pour la radiothérapie	160
5.2.1	Tumeurs simulées	160
5.2.2	Tumeur réelles avec résultats d'histologie	166
5.3	Suivi du volume tumoral au cours du temps pour le suivi thérapeutique :	
	résultats sur tumeurs simulées	168
5.3.1	Cas considérés	168

5.3.2	Résultats pour chaque cas	171
6	Conclusion et perspectives	183
6.1	Méthode développée	183
6.1.1	Mérites et performances	183
6.1.2	Limites	184
6.2	Perspectives et applications	185
6.2.1	Application : radiothérapie guidée par l'image (IGRT)	186
6.2.2	Application correction d'effets de volume partiel et quantification pour le suivi thérapeutique et le diagnostic	192
6.2.3	Modèles de Markov couples et triplets	197
6.2.4	Etendre le champ d'application de FLAB	197
7	Annexes	199
7.1	Glossaire des termes et abréviations	199
7.2	Communications et publications dans le cadre de cette thèse	202
7.2.1	Publications sur le travail de thèse	202
7.2.2	Publications annexes en co-auteur	204
7.3	Algorithme FLAB : schéma d'ensemble	206
7.4	Résultats complémentaires	206
7.4.1	Optimisation et validation	206
	Bibliographie	206

Résumé

Une des principales causes d'erreur en analyse semi-quantitative en imagerie par émission de positons (TEP) est la méthode utilisée pour déterminer les volumes d'intérêt sur les images fonctionnelles. Ceci concerne le diagnostic et le suivi thérapeutique en oncologie ainsi que la nouvelle application en plein développement qu'est la radiothérapie guidée par l'image. La faible qualité des images d'émission, liée notamment au bruit et au flou induits par les effets de volume partiels et la variabilité des protocoles d'acquisition et de reconstruction des images, ainsi que le grand nombre de procédures proposées pour définir les volumes, en sont la cause. La plupart des méthodes proposées jusqu'alors sont basées sur des seuillages déterministes, peu robustes au bruit et aux variations de contraste et incapables de gérer les hétérogénéités dans la distribution d'activité des tumeurs.

L'objectif de cette thèse est de proposer une approche de segmentation des images 3D, automatique, robuste, précise et reproductible pour déterminer le volume fonctionnel des tumeurs de toutes tailles dont la distribution d'activité peut être très hétérogène. L'approche proposée est fondée sur la segmentation statistique des images, couplée à une modélisation floue, permettant de prendre en compte à la fois l'aspect bruité et l'aspect flou des images de médecine nucléaire. Elle fait appel à une étape d'estimation itérative des paramètres et une modélisation locale du voxel et de son voisinage pour l'estimation et la segmentation.

Les méthodes développées ont été évaluées sur de nombreuses données simulées et réelles, tant pour des images de fantômes que pour des images de tumeurs. Les résultats sur fantôme ont permis de valider les performances de l'approche proposée en terme de taille d'objet d'intérêt, jusqu'à 13 mm de diamètre (environ deux fois la résolution spatiale en TEP), ainsi que de

confirmer un comportement plus robuste par rapport au bruit, aux variations de contraste ou des paramètres d'acquisition et de reconstruction, que les méthodes de référence basées sur des seuillages.

Les résultats obtenus sur différents ensemble de données d'images cliniques de tumeurs, fournies par différents services de médecine nucléaire dans le cadre de multiples collaborations, ont montré la capacité de l'approche à segmenter avec précision des tumeurs complexes, tant en terme de forme que de distribution d'activité, pour lesquelles les méthodes de référence échouent à produire des segmentation cohérentes. La méthode de segmentation développée est également capable de définir des régions d'intérêt au sein même de la tumeur grâce à sa gestion de l'hétérogénéité de l'activité de la tumeur, là où les méthodes de références sont strictement binaires.

Les résultats concernant la robustesse et la précision de l'approche sur tumeurs amènent à penser que son utilisation peut être envisagée tant dans le cadre du diagnostic et du suivi thérapeutique, que pour la définition des volumes cibles en radiothérapie, avec le potentiel d'augmenter les doses délivrées aux tumeurs tout en réduisant dans le même temps les doses délivrées aux tissus sains et aux organes à risque environnants, conformément au principe de "dose painting". Des travaux pour évaluer l'impact effectif de la méthodologie dans le contexte de la radiothérapie ont commencé dans le cadre d'un projet ANR, avec un dépôt de brevet associé.

Mots clés : TEP, oncologie, radiothérapie, segmentation floue, volumes d'intérêt.

Abstract

One of the main factors of error for semi-quantitative analysis in positron emission tomography (PET) imaging for diagnosis and patient follow up, as well as new flourishing applications like image guided radiotherapy, is the methodology used to define the volumes of interest in the functional images. This is explained by poor image quality in emission tomography resulting from noise and partial volume effects induced blurring, as well as the variability of acquisition protocols, scanner models and image reconstruction procedures.

The large number of proposed methodologies for the definition of a PET volume of interest does not help either. The majority of such proposed approaches are based on deterministic binary thresholding that are not robust to contrast variation and noise. In addition, these methodologies are usually unable to correctly handle heterogeneous uptake inside tumours.

The objective of this thesis is to develop an automatic, robust, accurate and reproducible 3D image segmentation approach for the functional volumes determination of tumours of all sizes and shapes, and whose activity distribution may be strongly heterogeneous. The approach we have developed is based on a statistical image segmentation framework, combined with a fuzzy measure, which allows to take into account both noisy and blurry properties of nuclear medicine images. It uses a stochastic iterative parameters estimation and a locally adaptive model of the voxel and its neighbours for the estimation and segmentation.

The developed approaches have been evaluated using a large array of datasets, comprising both simulated and real acquisitions of phantoms and tumours. The results obtained on phantom acquisitions allowed to validate the accuracy of the segmentation with respect to the size of considered structures, down to 13 mm in diameter (about twice the spatial resolution of a

typical PET scanner), as well as its robustness with respect to noise, contrast variation, acquisition parameters, scanner models or reconstruction algorithms. The performance of the developed algorithm is shown to be superior to thresholding reference methodologies.

The results demonstrate the ability of the developed approach to accurately delineate tumours with complex shapes and activity distributions, for which the reference methodologies fail to generate coherent segmentation maps. The algorithm is also able to delineate multiples regions inside the tumour, whereas reference methodologies are usually binary only.

Both robustness and accuracy results demonstrate that the proposed methodology may be used in clinical context for diagnosis and patients follow up, as well as for radiotherapy treatment planning and "dose painting", facilitating optimized dosimetry and potentially reduced doses delivered to healthy tissues around the tumour and nearby organs. Such studies to evaluate the impact of the methodology in radiotherapy treatment planning have already started in a project which aims to explore the potential of the algorithm which has been successfully patented.

Keywords : PET, oncology, radiotherapy, fuzzy segmentation, volumes of interest.

Table des figures

1.1	Premier dispositif clinique d'imagerie par positon (1953) [22].	32
1.2	Trois scanners TEP/TDM commerciaux actuels. A gauche le Philips Gemini, au centre le Siemens Biograph et à droite le GE Discovery LS. Photos : source constructeurs. . .	32
1.3	Illustration de l'annihilation du positon et de l'électron.	34
1.4	Formule du 18F-FDG.	35
1.5	Photo d'un détecteur avec cristal <i>lutetium oxyorthosilicate</i> (LSO). Source Siemens. . .	37
1.6	Illustration de la détection d'une LOR.	39
1.7	Images TEP corps entier, sans correction d'atténuation (colonne de gauche) et avec correction (colonne de droite), en utilisant une image TDM (en haut) ou une image de transmission (en bas).	40
1.8	Illustration des LORs erronées : les fortuites (b) et les diffusées (c) par rapport à une LOR vraie (a).	41
1.9	Illustration d'un examen TEP/TDM classique et les images obtenues, avec leur fusion.	44
1.10	Comparaison entre les deux modes de fonctionnement, 2D et 3D.	45
1.11	Une projection $p(s, \phi)$ représente l'intégrale selon toutes les LORs ayant une même direction ϕ . Les projections sont organisées en sinogramme de manière à ce que chaque projection dans une direction donnée ϕ corresponde à une ligne du sinogramme. . . .	46
1.12	Images TEP cliniques reconstruites par rétroprojection filtrée (FBP) ou un algorithme itératif (OSEM), les deux avec correction d'atténuation (AC).	49
1.13	Illustration du principe de fonctionnement du temps de vol (image issue de la brochure Philips GEMINI TF).	50

1.14	Comparaison de la qualité d'image entre un système TEP classique au centre, et un système TEP TOF à droite (image CT à gauche), pour des conditions d'acquisitions et de reconstruction identiques. Les images du haut illustrent un cas de cancer du colon bien mieux visible avec le TOF. En bas, illustration de l'amélioration sur la structure de l'aorte, mieux visible en TOF. Image issue de [88].	52
1.15	Illustration de la largeur à mi-hauteur ou FWHM.	54
1.16	Illustration des effets de volume partiel. En haut à gauche, un objet idéal contenant des sphères. A droite, l'image acquise par un scanner TEP de cet objet (simulation). En bas, les profils correspondants aux barres rouges sur chaque image illustre la perte d'intensité et l'étalement du signal dus aux effets de volume partiel.	56
1.17	Illustration des effets de la respiration sur des images simulées, et de la performance des méthodes de correction. A gauche, une image sans respiration, au centre l'image incorporant les effets de respiration et à droite l'image corrigée par la correction élastique, au sein de la reconstruction [93]. D'un point de vue quantitatif, la récupération d'activité dans les tumeurs simulées est de l'ordre de 72-90%, l'information de position est corrigée à hauteur de 78-100% et l'étalement (FWHM) de 67-92%.	58
2.1	La fameuse photo de Léna, rendue populaire par la communauté scientifique de l'imagerie.	67
2.2	Illustration des résultats obtenus sur des sphères homogènes avec un seuillage fixe. La notation de chaque configuration est : taille (mm) / durée d'acquisition (min) / contraste sphère :fond. Ces résultats illustrent bien la faible robustesse de la méthode à des changements de taille, de contraste ou de bruit.	73
2.3	Figure issue de [99]. La tumeur utilisée pour tester l'approche sur un cas clinique a été choisie suffisamment grande et surtout avec une activité totalement homogène.	77
2.4	Illustration des résultats obtenus sur un objet hétérogène avec les deux seuillages adaptatifs décrits précédemment. Ces résultats montrent qu'une approche binaire, même adaptative par rapport au fond, est insuffisante pour rendre compte correctement de la complexité de la tumeur à segmenter, et que les résultats peuvent varier de façon très significative suivant les configurations de bruit et de contrastes.	78

2.5	Figure issue de [41]. La segmentation est faite sur l'image corps entier et la segmentation des zones marquées est clairement trop grossière. Cela est particulièrement flagrant sur les images de fantôme.	86
3.1	Le parcours fractal d'Hilbert-Peano en 2D et à plusieurs niveaux (en haut) et en 3D pour un volume 4x4x4 (en bas), utilisé pour transformer les images 2D et 3D en chaîne unidimensionnelle pour les traitement par chaîne de Markov.	98
3.2	Illustration d'une tumeur pulmonaire dont la distribution d'activité est fortement hétérogène. Noter que l'imagerie anatomique TDM ne permet pas d'apprécier les variations au sein de la tumeur, qui ne sont mises en évidence que sur l'imagerie fonctionnelle.	106
3.3	Différents cas de figure (en haut, la vérité terrain, en bas l'image TEP simulée correspondante) : tumeur homogène (à gauche), légèrement hétérogène (au centre) et fortement hétérogène (à droite).	107
3.4	Exemple de segmentation en deux phases, appliquée à une tumeur faiblement homogène pour différencier deux régions distinctes en son sein.	108
3.5	Illustration des différentes zones envisageables dans le cadre d'une tumeur fortement hétérogène.	109
3.6	Représentation de la modélisation adoptée pour la segmentation à trois classes dures et trois transitions floues.	110
3.7	Le graphe de Pearson permet de localiser une distribution en fonction des paramètres γ_1 et γ_2	115
4.1	Illustration du fantôme IEC. (a) Acquisition TDM et (b) TEP (vues axiales, sagittales et coronales) du fantôme réalisées sur le TEP/TDM Philips Gemini.	121
4.2	Paramètres considérés pour chaque ensemble de données acquis avec différents types de scanner TEP/TDM.	122
4.3	Images obtenues sur le Philips Gemini. Contrastes (A) 8 :1 et (B) 4 :1. Voxels de (a) 4x4 mm et (b) 2x2 mm. Acquisitions de (1) 5 min et (2) 1 min.	125

4.4	Images obtenues sur le Philips Gemini TF. Contrastes (A) 10 :1 et (B) 4 :1. Voxels de (a) 4x4 mm et (b) 2x2 mm. Acquisitions de (1) 5 min et (2) 1 min..	125
4.5	Images obtenues sur le Siemens Biograph. Contrastes (A) 8 :1 et (B) 4 :1. Voxels de (a) 5.3x5.3 mm et (b) 2x2 mm. Acquisitions de (1) 5 min et (2) 1 min.	126
4.6	Images obtenues sur le GE Discovery LS. Contrastes (A) 9 :1 et (B) 5 :1. Voxels de (b) 2x2 mm. Acquisitions de (1) 5 min et (2) 1 min.	126
4.7	Objets synthétiques générés avec différentes conditions pour étudier la robustesse de l'algorithme FLAB.	128
4.8	Illustration des différences de résultats suivant la mesure utilisée pour évaluer les performances des méthodes de segmentation : (a) vérité terrain ; (b) image TEP simulée ; (c) segmentation ; (d) différences entre segmentation et vérité terrain. En rouge, les voxels de l'objet classés comme fond (entraînent une sous-évaluation du volume), en vert, ceux du fond classés comme objet (entraînent une sur-évaluation du volume). EC est l'erreur de classification telle que définie dans ce travail. EC globale est la même erreur mais calculée par rapport à l'ensemble de l'image et non l'objet. EV est l'erreur sur le volume.	130
4.9	Comment lire les graphiques d'erreurs présentés dans la suite.	131
4.10	Différentes représentations d'une sphère de 10 mm sur les acquisitions réelles du fantôme IEC (5 minutes d'acquisition, ce qui correspond au bruit le plus faible considéré dans cette étude) sur Philips Gemini et reconstruites par Ramla 3D. (a) contraste de 8 :1, voxels de 2x2x2 mm ; (b) contraste de 8 :1, voxels de 4x4x4 mm ; (c) contraste de 4 :1, voxels de 2x2x2 mm ; (d) contraste de 4 :1, voxels de 4x4x4 mm.	133
4.11	Erreurs de classification obtenus par les chaînes de Markov floues (FHMC) sur les images du fantôme en considérant différentes configurations de niveaux de flou. Pour comparaison, les résultats obtenus par une approche par chaînes dures (HMC), c'est à dire avec seulement deux classes dures sans modélisation floue, sont également présentés.	134

4.12 Erreurs de classification obtenues par l'approche locale adaptative (FLAB) sur les images du fantôme en considérant une implémentation avec 2 niveaux de flou et différentes valeurs pour chaque niveau de flou. 136

4.13 Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini (avec voxels de 2x2x2 mm) en considérant des lois gaussiennes ou bêta I dans l'implémentation de FLAB. 137

4.14 Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini en considérant un cube d'estimation de taille 3x3x3 ou 5x5x5 voxels dans l'implémentation de FLAB. Images avec voxels de 4x4x4 mm. 138

4.15 Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini en considérant un cube d'estimation de taille 3x3x3 ou 5x5x5 voxels dans l'implémentation de FLAB. Images avec voxels de 2x2x2 mm. 139

4.16 Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini (avec voxels de 2x2x2 mm) en considérant l'approche adaptative (FLAB) avec un cube d'estimation de taille 3x3x3 et l'approche "aveugle" sans estimation adaptative (FLB). 140

4.17 Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini (avec voxels de 2x2x2 mm) en considérant l'approche adaptative (FLAB) optimisée, sur cinq acquisitions indépendantes de 1 min. Contrairement aux autres graphiques, les barres noires de variation indiquent ici l'écart type des erreurs obtenues sur les cinq réalisations indépendantes de une minute chacune. 141

4.18 Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm. 143

4.19 Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm. 144

4.20 Résultats de FLAB et FCM obtenus sur l'objet synthétique à trois classes, avec deux configurations de contrastes et deux niveaux de bruit différents (c à f). 148

4.21 Résultats de FLAB et FCM obtenus sur l'objet synthétique binarisé, convolué par des PSF de largeur 4 (g) à 10 (j) mm. 149

4.22 Comment lire les résultats de l'étude de robustesse sur les différents scanners considérés. 151

4.23 Erreur de volume moyenne et ecart type pour chaque méthode considérée, sur l'ensemble des données disponibles pour l'étude de robustesse. 152

4.24 Erreur de classification moyenne et ecart type pour chaque méthode considérée, sur l'ensemble des données disponibles pour l'étude de robustesse. 153

5.1 Procédure suivie pour obtenir des images simulées réalistes de tumeurs, et évaluer la qualité de la segmentation. 159

5.2 Illustration des performances des méthodes considérées sur une petite tumeur homogène au fond légèrement hétérogène. 161

5.3 Illustration des performances des méthodes considérées sur une tumeur pulmonaire légèrement hétérogène et de taille moyenne (environ 4,5 cm). 162

5.4 Illustration des performances des méthodes considérées sur une tumeur fortement hétérogène et de grande taille (environ 6 cm). 163

5.5 Erreurs de classifications obtenues par quatre méthodes de segmentation (FLAB, FCM, TSBR et T42) sur l'ensemble des 20 tumeurs simulées. Les barres bleues représentent l'erreur moyenne sur les 20 tumeurs, et la barre de variation indique l'erreur minimale et maximale obtenue. 164

5.6 Erreurs de classifications obtenues par quatre méthodes de segmentation (FLAB, FCM, TSBR et T42) sur l'ensemble des 20 tumeurs simulées. Les barres bleues représentent l'erreur moyenne sur les 20 tumeurs, et la barre de variation indique l'écart type. . . . 165

5.7 Erreurs de classifications obtenues par FLAB et FCM sur les tumeurs fortement hétérogènes pour les classes 2 et 3 définissant les tumeurs. Les barres bleues représentent l'erreur moyenne, et la barre de variation indique les erreurs minimales et maximales atteintes. 166

5.8 Illustration des segmentations obtenues sur une tumeur réelle des données de l'étude histologique. 167

5.9 Erreur moyenne et écart type obtenu par chaque méthode par rapport au diamètre maximal de la tumeur mesuré par l'étude histologique. 167

5.10 Les huit cas considérés pour l'évaluation de l'utilisation des méthodes pour le suivi thérapeutique. 169

5.11 Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 1. 172

5.12 Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 2. 173

5.13 Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 3. 174

5.14 Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 4. 175

5.15 Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 5. 176

5.16 Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 6. 177

5.17 Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 7. 178

5.18 Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 8. 179

5.19 Erreurs moyennes et écart types calculés sur chacun des 8 cas, pour les méthodes considérées. 180

6.1 Illustration du fonctionnement de la radiothérapie externe conformationnelle. 188

6.2 Illustration d'un cas où utiliser le volume anatomique pour irradier la tumeur ne prend pas en compte la réalité du métabolisme, que l'on peut observer sur l'image fonctionnelle. 189

6.3 Positionnement de l'algorithme FLAB dans la planification de traitement par radiothérapie. 191

6.4 Principe de l'approche multirésolution développée par N. Bousson pour une image cérébrale. 193

6.5 Image correctrice générée et résultat de correction sur la même image de cerveau. . . 194

6.6 Résultats de correction de volume partiel en combinant la carte de segmentation de FLAB avec l'approche multirésolution développée par N. Bousson, dans le cas de deux classes (en haut) et trois classes (en bas). 196

7.1 Schéma d'ensemble du fonctionnement de l'algorithme FLAB. Les cases à traits continus constituent les états des paramètres, et les cases à traits en pointillés indiquent des opérations de transitions entre états. 207

7.2 Résultats obtenus (erreur de classification) sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm. 208

7.3	Résultats obtenus (erreur de classification) sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.	209
7.4	Résultats obtenus (erreur de volume) sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.	210
7.5	Résultats obtenus (erreur de volume) sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.	211
7.6	Comparaison des performances (erreur de classification) entre FLAB et T42 sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.	212
7.7	Comparaison des performances (erreur de classification) entre FLAB et T42 sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.	213
7.8	Comparaison des performances (erreur de volume) entre FLAB et T42 sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.	214
7.9	Comparaison des performances (erreur de volume) entre FLAB et T42 sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.	215
7.10	Comparaison des performances (erreur de classification) entre FLAB et FCM sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.	216
7.11	Comparaison des performances (erreur de classification) entre FLAB et FCM sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.	217
7.12	Comparaison des performances (erreur de volume) entre FLAB et FCM sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.	218

7.13 Comparaison des performances (erreur de volume) entre FLAB et FCM sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.	219
7.14 Comparaison des performances (erreur de classification) entre FLAB et FHMC sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.	220
7.15 Comparaison des performances (erreur de classification) entre FLAB et FHMC sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.	221
7.16 Comparaison des performances (erreur de volume) entre FLAB et FHMC sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.	222
7.17 Comparaison des performances (erreur de volume) entre FLAB et FHMC sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.	223

Détermination des volumes fonctionnels : enjeux et objectifs

L'imagerie de tomographie par émission de positons (TEP) ou tomographie par émission mono-photonique (TEMP) est une imagerie fonctionnelle qui permet de visualiser le métabolisme d'une fonction de l'organisme, telle que la consommation de sucre des cellules ou la synthèse de dopamine, entre autres. Le principe sous-jacent est d'injecter au patient une molécule biologique appropriée (par exemple du glucose pour observer la consommation de sucre) marquée par une substance radioactive, puis on procède à la détection des particules émises par cette substance. Dans le cas de la TEMP, il s'agit d'émetteur de photons (qui sont émis dans toutes les directions) tandis que dans le cas de la TEP, il s'agit d'émetteurs de positons, particules s'annihilant avec les électrons dans les tissus du patient, cette annihilation résultant en l'émission de deux photons à 180° l'un de l'autre.

Bien que cette modalité offre des informations complémentaires cruciales aux médecins pour établir ou conforter un diagnostic, notamment par rapport aux imageries anatomiques classiques comme le scanner à rayons X, elle souffre de limitations importantes en terme de qualité d'image et de biais déterministes et aléatoires concernant l'information contenue dans les images reconstruites à partir des acquisitions brutes par le principe de tomographie. La principale limite est la résolution spatiale, qui est au mieux d'environ 5 à 6 mm en TEP et le double en TEMP, pour l'imagerie corps entier.

Malgré ses nombreuses limitations, l'imagerie d'émission est aujourd'hui considérée comme la méthode de référence et l'outil principal pour le diagnostic et le suivi thérapeutique des

patients en oncologie [76] [84]. De plus, de nouvelles applications telles que la radiothérapie guidée par l'image fonctionnelle connaissent actuellement un fort développement et un intérêt grandissant [81] [13]. L'intérêt porté aux développements méthodologiques concernant le pré-traitement des données et leur acquisition (avec la correction des mouvements respiratoires par exemple), la reconstruction des données en images (reconstructions plus rapides, l'incorporation de la modélisation de la résolution spatiale, l'incorporation des mouvements respiratoires, l'utilisation d'informations *a priori*, etc.) ou les traitements *a posteriori* (correction des effets de volume partiel, segmentation, débruitage...) s'est renforcé, en particulier depuis que l'imagerie multimodalités s'est imposée en routine clinique. L'imagerie multimodalités consiste en la combinaison dans un même appareil de deux scanners de modalité d'imagerie différentes, comme un scanner TEP ou TEMP associé à un scanner à rayons X (TEP/TDM ou TEMP/TDM), ou avec une imagerie par résonance magnétique (IRM) avec les TEP/IRM, ces derniers étant toujours en développement. L'imagerie multimodalités offre la complémentarité de l'information et la possibilité d'utiliser l'une pour améliorer l'autre, mais elle introduit également de nouveaux défis à relever. La demande des cliniciens pour des outils performants d'analyse et de traitement des images est donc d'autant plus forte.

Bien que de nombreux aspects de l'imagerie multimodalités soient aujourd'hui bien maîtrisés, comme la correction d'atténuation, l'acquisition et la reconstruction 3D optimisée ou la fusion et le recalage des données anatomiques et fonctionnelles, il reste plusieurs problématiques importantes qui font l'objet de travaux. La correction des effets de volume partiel provenant de la résolution spatiale limitée, la prise en compte et la correction des mouvements respiratoires, et la segmentation et quantification automatique des volumes fonctionnels font encore l'objet de nombreuses études. Ce dernier sujet, qui fait l'objet de ce travail, est également directement lié aux deux autres, car une quantification correcte nécessite une correction efficace du flou résultant de la résolution spatiale et des mouvements respiratoires, qui entraînent des biais importants par rapport à l'activité réelle d'un objet d'intérêt.

D'une part, l'analyse des images TEP en routine clinique, notamment pour le diagnostic et le suivi thérapeutique, se fait encore largement sur la base d'analyse strictement visuelle ou à l'aide d'outils d'analyse d'image implémentant des méthodes à base de seuillages très ru-

dimentaires, dont le manque de robustesse, de reproductibilité et de précision est désormais largement souligné dans de récentes études [9] [75] [118] [70] [115]. Quant à la radiothérapie guidée par l'image, bien qu'il ait été montré que l'utilisation de l'imagerie fonctionnelle permet d'améliorer la délimitation des régions cibles et de réduire la variabilité inter utilisateurs [3] [56] [1] [158], cette dernière reste encore largement réalisée à la main sur les images anatomiques [52], plus faciles à segmenter, faute d'outils précis, automatisés et faciles d'utilisation pour la segmentation des images fonctionnelles.

D'autre part, une grande majorité des travaux visant à développer une approche semi-automatique ou automatique pour segmenter les volumes d'intérêt sur les images fonctionnelles s'appuie le plus souvent sur des méthodes trop simplistes à base de seuillages, souvent utilisateurs- et systèmes- dépendantes et dont l'application à des tumeurs réelles se révèle difficile (voir le chapitre 2 pour l'état de l'art), en particulier dans le cas de tumeurs dont l'activité est hétérogène. En conséquence, l'objectif de ce travail est de proposer une méthode automatique, précise, robuste et reproductible de segmentation des images d'émission en 3D, pour améliorer la détermination des volumes fonctionnels, facilitant ainsi l'évolution vers une utilisation plus quantitative de l'imagerie d'émission dans des applications telles que la radiothérapie guidée par l'image, le suivi thérapeutique et le diagnostic. Pour ce faire, nous proposons d'explorer une voie nouvelle de recherche pour la segmentation automatique de volumes fonctionnels en imagerie d'émission, en utilisant des méthodes basées sur la segmentation statistique d'image, avec modélisation floue et prise en compte du bruit, deux aspects importants lorsque l'on considère cette modalité d'imagerie.

Cette thèse est organisée de la façon suivante : le premier chapitre décrit le fonctionnement et la raison d'être de l'imagerie fonctionnelle TEP et des aspects spécifiques de l'imagerie multimodalités, en commençant par un rapide historique et les principes physiques sur lesquels elle se fonde, puis les méthodes utilisées pour acquérir les données et les reconstruire en images exploitables par les cliniciens. La suite du chapitre se concentre sur les problèmes existants et les limites de l'imagerie fonctionnelle, tant matérielles que logicielles. La fin du chapitre est consacrée à son utilisation en routine clinique et les applications susceptibles de bénéficier d'une méthode améliorée de segmentation des images d'émission, notamment le diagnostic, le

suiwi thérapeutique et la radiothérapie guidée par l'image.

Le deuxième chapitre propose un état de l'art de la segmentation semi-automatique et automatique des images d'émission. Ce dernier présente les approches ayant été proposées pour segmenter et définir des volumes d'intérêt sur les images fonctionnelles, ainsi que certaines méthodes n'ayant que peu été utilisées dans ce contexte. Le chapitre commence par une introduction générale sur la segmentation d'image, les caractéristiques recherchées dans une approche de segmentation d'images d'émission, puis passe en revue les différentes approches, d'abord manuelles et semi-automatiques, puis les méthodes automatiques.

Le chapitre 3 détaille les méthodologies développées dans le cadre de cette thèse, en commençant par la justification de notre approche, puis détaillant la structure commune de modélisation statistique et floue considérée. La partie suivante est consacrée à la première approche développée et testée : les chaînes de Markov floues. L'approche locale adaptative, développée pour pallier à certaines lacunes des chaînes est ensuite détaillée, suivie de son extension à trois classes dures et trois transitions floues de façon à pouvoir prendre en compte des structures fortement hétérogènes tant en forme qu'en fixation, que présentent parfois les tumeurs réelles. Enfin, le chapitre se conclue sur l'utilisation du système de Pearson, envisagée pour rendre l'algorithme plus robuste face à la variabilité des différents modèles de scanners et les algorithmes de reconstruction associés.

Le chapitre 4 présente les études d'optimisation et de validation des méthodes développées et les résultats obtenus, c'est à dire les performances de chaque méthode développée au cours de cette thèse, comparées à d'autres méthodes de référence, sur des images de fantômes et des simulations d'objets hétérogènes. Le chapitre détaille tout d'abord les données sur lesquelles les méthodes ont été validées, puis les résultats obtenus sur chaque type de données par les différentes méthodes sont présentés et discutés. Le chapitre se conclue sur les résultats de robustesse par rapport à différents modèles de scanners et algorithmes de reconstruction associés.

Le cinquième chapitre est consacré aux résultats obtenus par les méthodes dans le cas des tumeurs simulées réalistes et réelles avec données histologiques. L'approche a été testée sur des images cliniques, ainsi que sur les simulations générées à partir d'images cliniques, de façon

à bénéficier d'une rigoureuse vérité terrain rigoureuse, difficile à obtenir dans le cas d'images de patients. Comme dans le chapitre 4, nous commençons par présenter les données, puis les résultats obtenus.

Le dernier chapitre conclut cette thèse en rappelant les développements effectués, puis examine les perspectives ouvertes et les travaux futurs envisagés pour poursuivre les études entamées au cours de cette thèse en présentant notamment les applications susceptibles de bénéficier des méthodes développées dans ce doctorat et les études préliminaires qui ont été menées pour montrer l'intérêt de l'utilisation de nos méthodes.

Enfin, l'annexe contient le détail des algorithmes, suivi d'un glossaire des termes et abréviations, quelques résultats complémentaires, et une liste des communications et publications réalisées dans le cadre de ce doctorat.

Chapitre 1

Imagerie TEP / TDM

Ce travail porte sur l'automatisation de la détermination des volumes en imagerie fonctionnelle pour les applications d'oncologie. Pourquoi est-elle qualifiée de fonctionnelle ? Qu'apporte-t-elle par rapport aux imageries dites anatomiques, par exemple le scanner à rayons X ou tomodynamométrie (TDM) ? Quels sont les principes qui la régissent ? En quoi est-il intéressant de combiner ces différentes modalités d'imagerie médicale ?

1.1 Historique

Le principe de la TEP en imagerie médicale voit le jour dès les années 50, avec la première machine construite au Laboratoire de Recherche en Physique à l'Hôpital Général du Massachusetts (MGH), utilisée dans le cadre de la localisation de tumeurs cérébrales [151]. Il s'agit alors d'un simple dispositif utilisant deux détecteurs opposés d'iodure de sodium [24]. La figure 1.1 montre cette première machine en action. Après plusieurs versions de ce premier modèle, un scanner "hybride" fut développé dans les années 60 [23]. Ce modèle possédait deux rangées de neuf détecteurs, chacun en coïncidence avec trois détecteurs de la rangée opposée, permettant d'obtenir une image bidimensionnelle. Puis les premiers modèles commerciaux à anneaux furent développés avec par exemple le ECATI d'ORTEC en 1978. Parallèlement, les radiopharmaceutiques furent également développés. Initialement utilisés avec l'oxygène 15, l'utilisation du carbone 11, du fluor 18 ou de l'azote 13 se répandit à partir des années 70.



FIG. 1.1 – Premier dispositif clinique d'imagerie par positon (1953) [22].

Plus de détails sur l'histoire de l'imagerie par positons sont disponibles dans [22]. Aujourd'hui, trois grand fabricants commercialisent des systèmes TEP : Siemens, GE Healthcare et Philips (voir figure 1.2).



FIG. 1.2 – Trois scanners TEP/TDM commerciaux actuels. A gauche le Philips Gemini, au centre le Siemens Biograph et à droite le GE Discovery LS. Photos : source constructeurs.

1.2 Physique de l'imagerie

1.2.1 Fluor 18

Le fluor 18 ($^{18}_9F$) est un radio-isotope de l'atome de fluor de masse atomique 18 (9 protons et 9 neutrons) dont la demi-vie est de 110 minutes. Il n'existe pas à l'état naturel et doit être fabriqué grâce à un cyclotron. Lors de la désintégration, un proton est converti en neutron

par l'intermédiaire de la force nucléaire faible et un positon ainsi qu'un neutrino sont émis :



Le positon est l'anti-particule associée à l'électron. Il possède une charge électrique de +1 (contre -1 pour l'électron), le même spin et la même masse que l'électron.

Nous nous concentrerons sur le fluor 18 (F18) car c'est le principal radio-isotope utilisé en oncologie avec la TEP. Il en existe d'autres utilisés en oncologie : l'I-124, le CU-64 et le GA-68. Aucun de ces isotopes n'existent à l'état naturel et un cyclotron est nécessaire pour les produire. Etant donnée la courte durée de demi-vie de ces isotopes, les centres d'examens doivent se trouver soit à distance raisonnable du cyclotron en question (dans le cas du fluor 18), soit sur site pour les éléments à la durée de demi-vie plus courtes.

1.2.2 Annihilation

Lorsqu'un positon émis par le ${}^{18}\text{F}$ se propage dans la matière environnante, il perd de l'énergie à cause des passages successifs près des atomes de la matière et des interactions résultantes. Après un parcours qui varie de 1 à 3 mm environ en fonction de l'énergie moyenne du positon émis, il perd suffisamment d'énergie pour s'annihiler avec un électron, qui porte une charge opposée à la sienne. Cette annihilation produit l'émission de deux photons gamma, émis à 180° ($\pm 0.25^\circ$) l'un de l'autre, du fait de la loi de la conservation de quantité de mouvement. La figure 1.3 illustre ce phénomène. C'est grâce à cette particularité précise de colinéarité que cette annihilation va pouvoir être indirectement détectée, de façon à aboutir à une reconstruction en trois dimensions de la distribution radioactive au sein du patient imagé. Toutefois, l'isotope n'est qu'une partie du dispositif de l'examen. En effet, il faut l'associer à une molécule biologique permettant de visualiser un processus physiologique spécifique lié à la pathologie que l'on souhaite détecter. Le composé résultant de l'association de cette molécule avec l'isotope radioactif utilisé pour la marquer est appelé radiopharmaceutique ou radiotraceur.

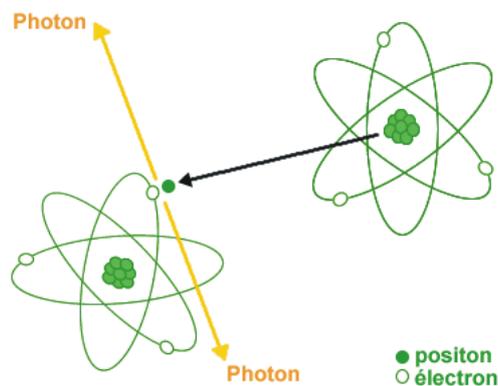


FIG. 1.3 – Illustration de l’annihilation du positon et de l’électron.

1.2.3 Radiopharmaceutique ou radiotraceur

18F-fluorodéoxyglucose

En oncologie, un des radiotraceurs les plus utilisés est le fluorodéoxyglucose (abrégé 18F-FDG ou FDG). Ce traceur est semblable au glucose mais il possède un atome d’oxygène de moins que la molécule de glucose, sur le deuxième carbone. Cette modification a un impact très important et est nécessaire à l’examen TEP, dans le sens où ce composé n’est pas assimilé de la même façon que le glucose normal et se fixe en conséquence plus longtemps dans les cellules concernées, ce qui entraîne une accumulation du traceur au sein des cellules. Cette accumulation permet la visualisation ultérieure du processus grâce aux quantités mises en jeu et c’est la raison pour laquelle l’examen est effectué environ une heure après l’injection, pour laisser le temps à l’accumulation de se réaliser.

Fixations du FDG

Il est important de noter que ce traceur n’est pas spécifique du cancer, mais seulement de la consommation de sucre, ce qui explique que certaines tumeurs ne fixent pas le 18F-FDG. Le cerveau fixe de façon systématique et le coeur de façon variable. Le foie fixe également mais de façon moins intense. On retrouve également parfois une fixation au niveau du tube digestif. Le FDG étant éliminé par voie urinaire, la fixation au niveau de la vessie

est particulièrement importante. Les muscles ont également tendance à fixer le FDG, c'est pourquoi on demande aux patients de rester calmes et immobiles avant l'examen, pour éviter des fixations musculaires importantes et gênantes pour le diagnostic. Ces fixations sont toutes considérées comme normales, ou physiologiques. Les fixations anormales, ou pathologiques, sont celles que l'examen TEP a pour objectif de détecter.

La consommation anormalement élevée de certaines cellules tumorales est connue depuis les années 1950 [166], ainsi que la fixation liée à des processus inflammatoires, eux mêmes pouvant être indépendants ou liés au processus tumoral. Même si ce n'est pas le cas pour tous les types de cancers, une corrélation a été établie entre la fixation du FDG et la prolifération cellulaire [117].

L'image obtenue en TEP est donc une représentation tri-dimensionnelle de la distribution du glucose marqué, ce qui permet de détecter et d'analyser les anomalies fonctionnelles comme une tumeur, à condition que sa consommation de glucose soit différente des tissus sains environnants. Logiquement, ce traceur n'est pas adapté pour détecter et analyser tous les types de tumeurs, notamment cérébrales (étant donné l'accumulation normale du cerveau).

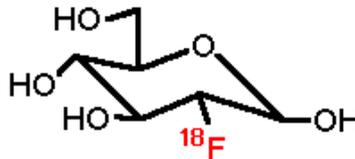


FIG. 1.4 – Formule du ^{18}F -FDG.

Autres traceurs

Notons qu'il existe différents radiopharmaceutiques utilisés dans différentes applications médicales de la TEP (cardiologie, neurologie,...). De nombreux autres traceurs sont en cours d'évaluation pour permettre le suivi de différents métabolismes, comme les caractéristiques de vascularisation, d'oxygénation, de radiosensibilité ou de radiorésistance. Les traceurs utilisant le fluor 18 ont un succès lié aux propriétés avantageuses de ce marqueur, notamment la faible énergie initiale de l'émission du positron, et sa durée de demie-vie de 110 minutes [35]. Citons

notamment le $[^{18}\text{F}]$ fluoromisonidazole (FMISO) pour l'hypoxie [132], définie comme une inadéquation entre les besoins cellulaires et les ressources en oxygène et connue pour être liée à une progression tumorale accrue, maligne et résistante aux thérapies [69]. Deux autres traceurs spécifiques à l'hypoxie ont également été proposés récemment : le méthylthiosemicarbazone marqué au cuivre 64 ($[^{64}\text{Cu}]$ -ATSM) [116] et le nitroimidazol-trifluoropropyl acétamide marqué au fluor 18 ($[^{18}\text{F}]$ -EF3) [30]. Le $[^{18}\text{F}]$ fluoro-3'-deoxythymidine (FLT) est quant à lui actuellement étudié pour suivre la prolifération cellulaire [145] et devrait améliorer la spécificité potentiellement limitée du FDG [34], liée à la fixation de l'inflammation, et qui n'avait pas été améliorée par l'utilisation du 2- $[^{18}\text{F}]$ -fluoro-L-tyrosine (TYR) [78]. La tendance actuelle pour l'avenir de l'oncologie est la volonté d'obtenir une sorte de carte d'identité biologique d'un objet d'intérêt en combinant les informations obtenues à l'aide de divers traceurs, ce qui a également un intérêt particulier dans le cas de la radiothérapie [59].

1.2.4 Chaîne de détection

Le principe de la TEP repose sur la détection de photons en coïncidences, c'est à dire dans deux détecteurs opposés, de façon à créer ce que l'on appelle une ligne de réponse (LOR). La LOR est une ligne virtuelle reliant les deux lieux de détections, c'est à dire les deux détecteurs. On sait alors que l'annihilation a eu lieu le long de cette ligne, mais on ne sait pas précisément où le long de cette ligne. Il est toutefois possible d'affiner la position de l'annihilation le long de la LOR en utilisant le principe du temps de vol (*Time Of Flight* ou TOF), ce principe est décrit plus loin en section 1.5.3. La reconstruction d'image optimisée par le principe de tomographie, ainsi que différentes méthodes de correction de divers effets, permettent d'affiner la qualité des images malgré l'incertitude concernant la localisation de l'annihilation.

Détecteurs

Les détecteurs sont formés en partie de blocs de matière possédant des propriétés spécifiques, qu'on appelle cristaux scintillateurs. Ces derniers sont couplés avec plusieurs photo-détecteurs,

chacun étant composé d'un tube photo-multiplicateur et d'une photo-diode à base de semi-conducteur, qui permettent la détection de la lumière visible. La figure 1.5 illustre un détecteur TEP. Il existe plusieurs types de cristaux utilisés en TEP, jusqu'il y a quelques années essentiel-

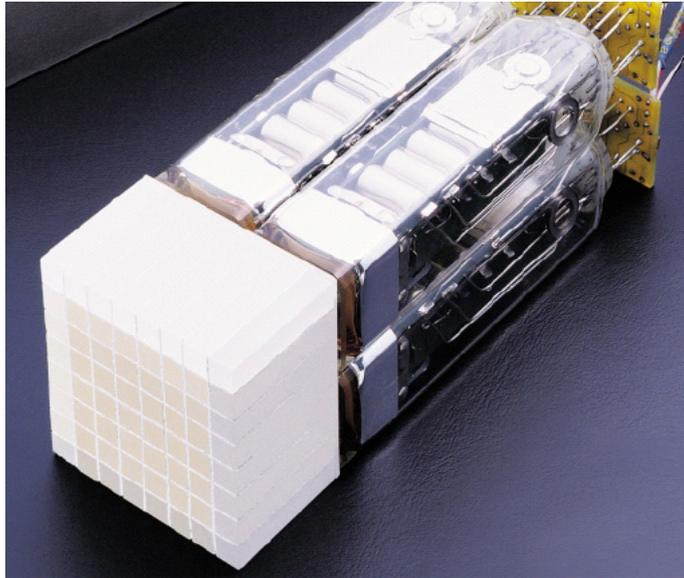


FIG. 1.5 – Photo d'un détecteur avec cristal *lutetium oxyorthosilicate* (LSO). Source Siemens.

lement de type *bismuth germanate* (BGO). Chacun d'entre eux a des propriétés différentes en terme de capacité d'arrêt d'un photon de 511 Kev (liée à la densité), d'une durée d'émission de lumière, d'intensité relative d'émission de lumière en sortie et de résolution intrinsèque d'énergie. Les propriétés désirées pour un cristal utilisé dans les détecteur TEP sont les suivantes [163] : une densité élevée et un numéro atomique suffisant pour une détection efficace des émissions gamma d'énergie élevée (511 Kev), ce qui permet au système d'atteindre une bonne sensibilité. Un grand nombre de photons émis par scintillation permettant d'améliorer la résolution spatiale, et, associé à une résolution énergétique appropriée, une réduction significative du nombre de coïncidences diffusées (voir section 1.2.5) détectées. Enfin, un délai de scintillation le plus court possible permet d'améliorer directement la résolution temporelle, facilitant la réduction du nombre de coïncidences aléatoires (voir section 1.2.5) détectées. Cette propriété a aussi le mérite d'intégrer plus de lumière émise pour chaque interaction entre un photon et le cristal, ce qui réduit le temps mort du détecteur, et mène donc à une sensibilité

supérieure et la capacité d'enregistrer plus d'évènements.

La recherche de matériaux mieux adaptés a entraîné l'utilisation récente de *lutetium oxyorthosilicate* (LSO) et *gadolinium oxyorthosilicate* (GSO). Ces deux cristaux possèdent des propriétés qui permettent une réduction importante des détections de coïncidences aléatoires et de diffusées par rapport aux cristaux BGO. LSO a une meilleure capacité d'arrêt des photons et une plus courte durée d'émission que GSO, mais sa résolution énergétique est inférieure [163]. Les recherches actuelles de nouveaux matériaux continuent, avec des candidats potentiels tels que le *lutetium aluminium perovskite* (LuAP) et *lutetium pyro-silicate* (LPS) par exemple, qui permettent une meilleure résolution temporelle. L'utilisation de semi-conducteurs à la place des détecteurs actuels qui sont formés par l'association entre cristaux et photomultiplicateurs, est également envisagée [112].

Ligne de réponse (LOR)

Pour qu'une LOR soit enregistrée, plusieurs paramètres doivent être vérifiés. D'abord, il existe une fenêtre d'énergie et une fenêtre temporelle. La fenêtre d'énergie a pour rôle d'éliminer les particules dont l'énergie est trop faible pour correspondre à celle d'un photon issu de l'annihilation d'un positon et d'un électron n'ayant pas été dévié (diffusé) et / ou atténué. En effet, les deux photons émis à 180° l'un de l'autre le sont avec une énergie de 511 Kev, mais ils peuvent perdre une partie de cette énergie en traversant les tissus du patients et en étant déviés de leur trajectoire initiale, voir être complètement absorbés. La fenêtre d'énergie est usuellement située entre 350-400 et 700 Kev, mais dépend du type de cristal utilisé et de sa résolution énergétique intrinsèque. Le rôle de la fenêtre temporelle est de réduire l'enregistrement de lignes de réponse aléatoires, c'est à dire erronées.

La figure 1.6 illustre la chaîne de détection en TEP.

1.2.5 Atténuation, LORs diffusées et aléatoires

Il existe différentes sortes de LORs qui ne sont pas des lignes réelles et qui génèrent des artefacts dans les images reconstruites. Il est nécessaire de corriger ces effets. Le premier effet est celui de l'atténuation des tissus du patients. En effet, les photons doivent parcourir une

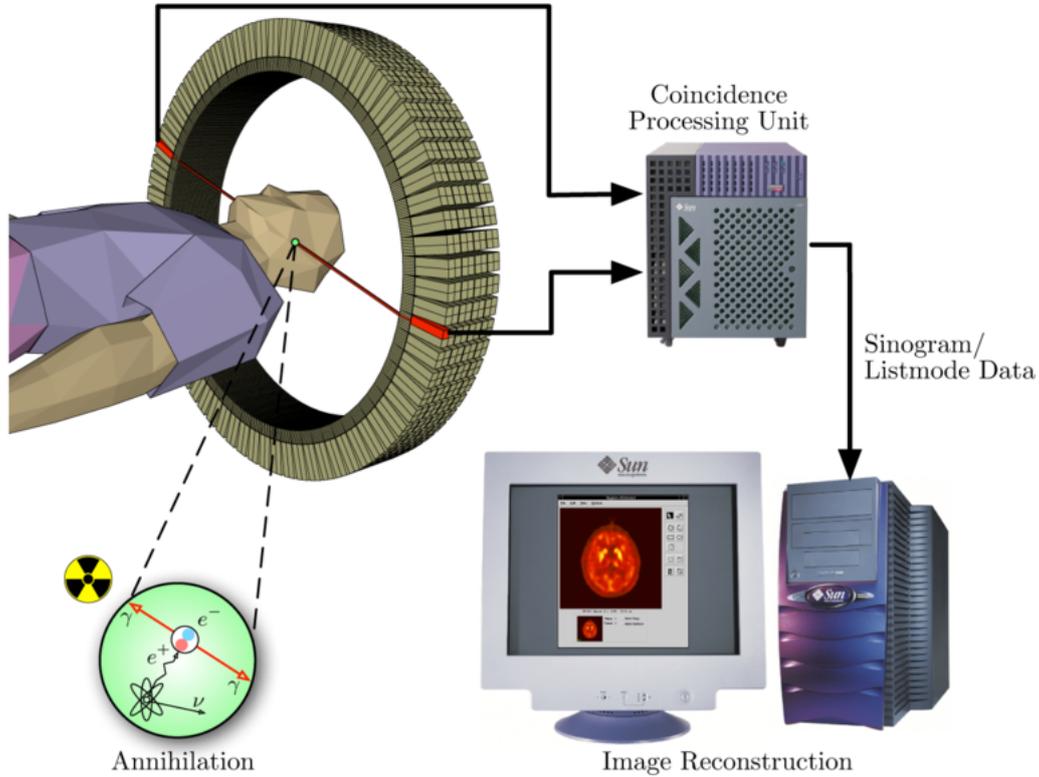


FIG. 1.6 – Illustration de la détection d'une LOR.

certaines distance entre leur lieu d'émission (l'annihilation entre le positon et un électron) et le détecteur, situé à plusieurs dizaines de centimètres de là. En chemin, le photon traverse les tissus du patient pour ensuite parcourir l'air séparant le patient de la couronne de détecteurs. Lors de cette traversée, le photon peut interagir avec la matière de plusieurs façons. Il peut être dévié, ralenti, absorbé. Si l'on ne tient pas compte des facteurs d'atténuation des tissus du patient (les poumons, par exemple, sont caractérisés par un facteur d'atténuation plus faible que le squelette, puisque les tissus mous sont moins denses que les os), les images obtenues sont incorrectes tant qualitativement que quantitativement. La figure 1.7 illustre ce phénomène avec une image de corps entier acquise sur TEP 18F-FDG, sans (colonne de gauche) et avec correction (à droite). La carte de correction (au centre) peut être obtenue soit grâce à une acquisition TEP par transmission (en bas), soit grâce à une acquisition scanner TDM (en haut) du patient, qui fournit l'information de densité des tissus. La correction d'atténuation a un impact significatif sur la sensibilité et la précision du diagnostic [77].

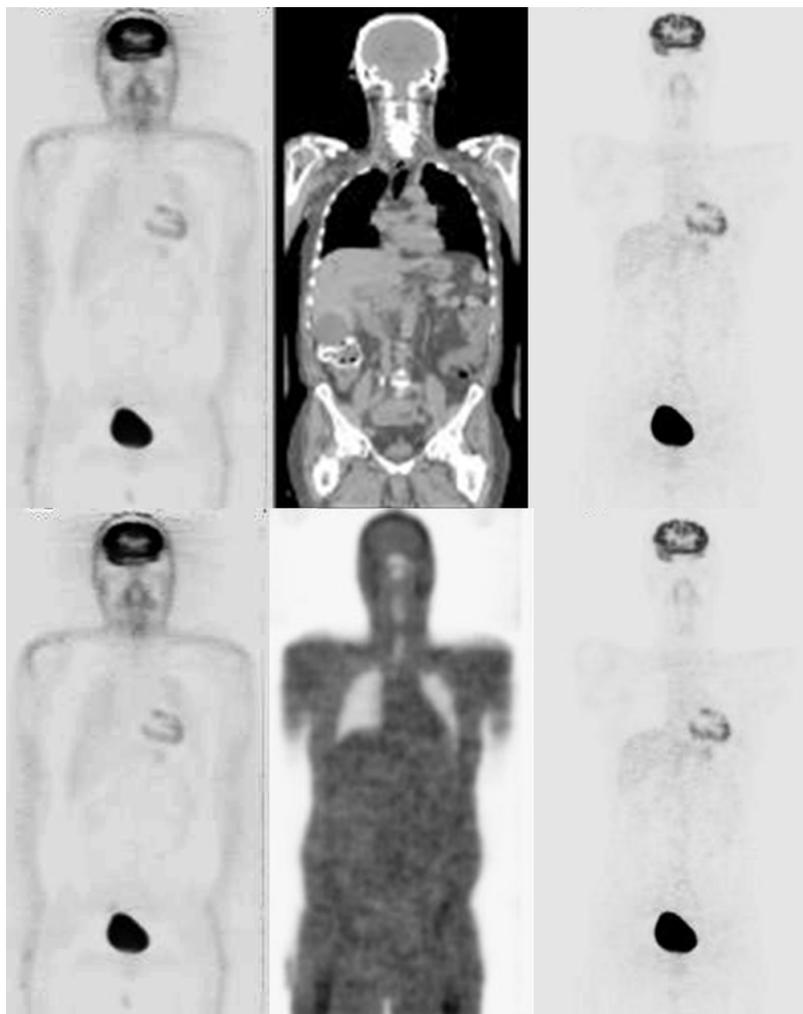


FIG. 1.7 – Images TEP corps entier, sans correction d’atténuation (colonne de gauche) et avec correction (colonne de droite), en utilisant une image TDM (en haut) ou une image de transmission (en bas).

Un autre effet qu’il faut corriger est la présence de LORs erronées. Il en existe deux types différents, illustrées dans la figure 1.8 : celles provenant de photons diffusés et celles provenant de détections aléatoires ou fortuites. Les photons diffusés (il s’agit de la diffusion Compton [32]) sont en général déviés de leur trajectoire et arrivent sur les détecteurs avec une énergie réduite. Bien qu’une fenêtre énergétique soit utilisée, elle ne suffit pas à éliminer totalement ces diffusés. Ceci peut avoir pour effet de créer des lignes de réponses virtuelles qui ne correspondent pas à une annihilation. Les coïncidences diffusées ne représentent qu’environ 10 % des coïncidences

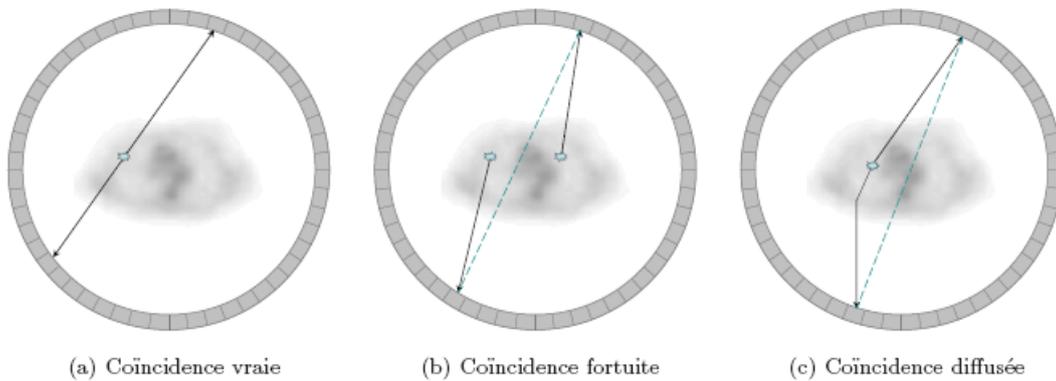


FIG. 1.8 – Illustration des LORs erronées : les fortuites (b) et les diffusées (c) par rapport à une LOR vraie (a).

totales en mode d'acquisition 2D, mais ce chiffre peut atteindre 40 % en mode 3D (voir section 1.5.1) et il est donc nécessaire de corriger cet effet. Pour ce faire, le premier type de techniques utilise l'information en énergie en combinant des données acquises dans au moins deux fenêtres en énergie. Les méthodes les plus avancées se fondent sur un calcul statistique de la distribution du diffusé pour un patient donné (simulations Monte Carlo) [6].

Les lignes de réponses dites fortuites sont créées à partir de deux détections correspondant à deux annihilations différentes, mais dont les photons ont été détectés au sein de la même fenêtre temporelle. Les coïncidences fortuites peuvent être directement mesurées dans une fenêtre temporelle décalée. Cette technique est la plus largement utilisée et présente l'avantage de mesurer la distribution spatiale des coïncidences aléatoires, mais elle augmente le bruit [20]. D'autres approches existent pour corriger les coïncidences aléatoires en soustrayant *a posteriori* de l'acquisition, une estimation de ces coïncidences aléatoires. Cette dernière peut être obtenue à partir d'un sinogramme décalé et lissé, par une acquisition de calibration ou estimée directement, chaque façon de procéder possédant différents avantages et inconvénients [20].

1.3 Imagerie multimodalités

Il est rapidement devenu évident que les informations de natures différentes, anatomiques et fonctionnelles, une fois combinées, permettent d'améliorer le diagnostic des patients. En effet,

l'information fonctionnelle permet de déterminer la nature exacte de structures anatomiques anormales visibles et de détecter des changements pathologiques avant les modifications des lésions anatomiques. D'un autre côté, l'apport en terme de localisation spatiale et de résolution de l'image anatomique permet de mieux positionner et interpréter le contenu de l'image fonctionnelle, dont la faible résolution, le flou résultant et le bruit gênent l'interprétation. Sur cette dernière, les structures anatomiques sont plus difficiles à localiser, voir absentes. L'utilisation de l'imagerie anatomique pour se repérer dans l'image fonctionnelle a été une première étape. Ensuite, grâce aux systèmes multimodalités permettant l'acquisition de l'image anatomique et fonctionnelle lors d'un même examen, sur le même statif, il est devenu possible d'obtenir des fusions et des recalages précis des deux images [155], ce qui permet notamment une bonne correction d'atténuation. Bien que la fusion des images et leur recalage offre encore des défis à relever, notamment à cause des mouvements respiratoires, les plateformes diagnostic disponibles permettent la fusion de l'image anatomique et de l'image fonctionnelle. Ces dernières sont performantes sous réserve de bonne corrélation entre les deux examens successifs. Il reste toutefois de nombreux défis à relever : la gestion des mouvements du patient et des effets respiratoires sur les images ou le manque de corrélation entre les deux modalités d'imagerie pour certaines zones d'intérêt, ce qui est souvent le cas notamment en ce qui concerne les tumeurs. Grâce à la multimodalités, on peut envisager d'utiliser l'information anatomique pour améliorer ou corriger l'image fonctionnelle, notamment dans le cadre de la correction des effets de volume partiel [15] [16] ou des acquisitions dynamiques et de la correction de mouvements respiratoires et cardiaques.

L'association d'un scanner TEMP avec un scanner TDM est plus ancien, mais l'association d'un TEP avec un TDM date de l'année 2000. Le journal *Time Magazine* a d'ailleurs élu le TEP/TDM *invention de l'année 2000* [80]. Dorénavant, les systèmes commercialisés sont systématiquement des machines TEP/TDM ou TEMP/TDM. L'aspect multimodalités est entré dans le cadre de la routine clinique. On parle déjà de TEP/IRM, même si ces derniers ne sont pas encore disponibles dans le commerce pour les examens corps entier (il existe des modèles dédiés pour le cerveau). L'immense intérêt offert par la fusion d'images anatomiques et fonctionnelles ne date certes pas de l'année 2000, mais dorénavant, cette fusion est devenue

plus rapide, plus facile et plus accessible.

La France a accumulé un retard important dans l'intégration de cette modalité au cours de ces dix dernières années par rapport à ses partenaires européens. Soixante appareils étaient en fonctionnement fin 2007. La carte sanitaire prévoit l'installation de 77 systèmes TEP/TDM sur l'ensemble du territoire français (<http://sfbmn.ifrance.com>), ce qui correspond à un système pour 800,000 habitants et une augmentation de 300% par rapport à l'année 2003 (25 systèmes étaient installés à cette date). À terme, chaque région devrait être équipée d'au moins un scanner TEP [14]. Il existe également une vingtaine de cyclotrons répartis de façon homogène sur le territoire français, ce qui permet un approvisionnement correct des centres d'examens en produits radiopharmaceutiques.

Dans la suite de ce chapitre, nous nous concentrerons essentiellement sur la physique de l'imagerie TEP, bien que plusieurs éléments essentiels, comme la correction d'atténuation par exemple, aient été significativement modifiés par l'introduction des scanners multimodalités.

1.4 L'examen 18F-TEP

Un examen TEP se déroule de la façon suivante : on injecte au patient le traceur correspondant au phénomène que l'on souhaite observer, typiquement en oncologie du 18F-FDG pour observer la consommation de sucre. Etant donné que les muscles consomment du sucre, on demande au patient de rester calme et immobile, et de ne pas parler jusqu'à l'examen. Environ une heure après l'injection, on procède à l'examen proprement dit, au cours duquel le patient passe d'abord une acquisition anatomique (TDM), qui ne prend qu'une vingtaine de secondes, puis le lit est décalé pour placer le patient dans le scanner TEP. L'acquisition TEP elle-même dure beaucoup plus longtemps : environ 5 minutes par "pas" de 15 à 18 cm, ce qui signifie environ 20 minutes pour un examen dit "corps entier", c'est à dire en réalité des cuisses jusqu'à la tête. Il est en effet nécessaire d'enregistrer suffisamment de LORs pour obtenir une image exploitable par le clinicien, c'est à dire possédant un contraste et un rapport signal sur bruit suffisants. Le fait que l'acquisition soit longue pose évidemment le problème des mouvements respiratoires, puisqu'on ne peut pas demander au patient de retenir sa respiration (voir 1.6.3).

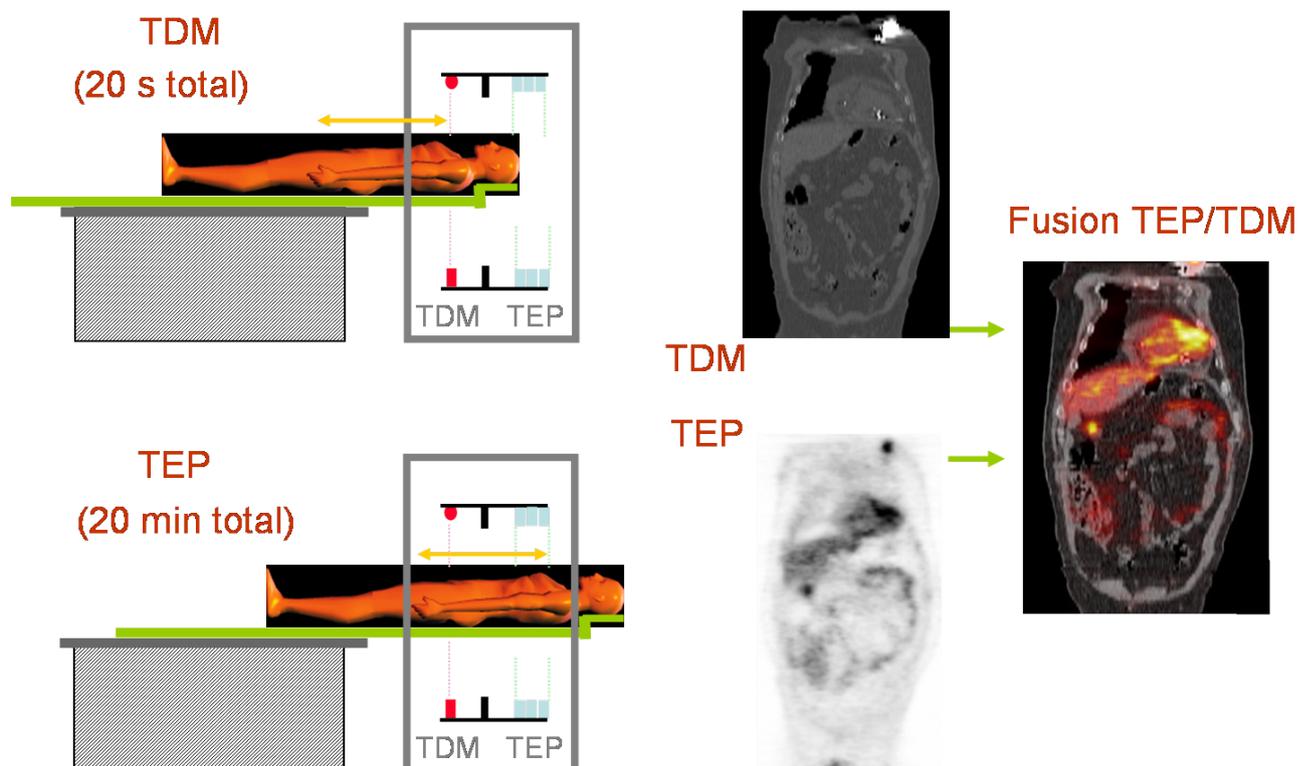


FIG. 1.9 – Illustration d’un examen TEP/TDM classique et les images obtenues, avec leur fusion.

L’acquisition anatomique TDM est actuellement utilisée en routine clinique pour la fusion des données et l’aide à la localisation de l’information fonctionnelle, ainsi que pour la correction d’atténuation (voir section 1.2.5). La figure 1.9 illustre l’examen TEP effectué en routine clinique ainsi que la fusion des données anatomiques et fonctionnelles.

1.5 De l’acquisition à l’image

1.5.1 Acquisition

Modes 2D et 3D

En TEP, il est possible de procéder à des acquisitions et reconstructions correspondantes en mode 2D ou 3D. En mode 2D, seules des tranches sont acquises, et un volume 3D complet du patient est ensuite reconstruit par empilement des tranches acquises indépendamment les unes des autres. On utilise en général des *septa*, sortes de ”collimateurs” de plomb qui dépassent vers

l'intérieur de l'anneau de détecteurs, pour empêcher les photons incidents d'arriver jusqu'aux détecteurs. Ceci permet d'assurer que seuls les photons arrivant dans le plan défini par l'anneau de détecteurs et les anneaux directement adjacents puissent former des coïncidences. En mode 2D, la machine ne détecte que les plans directs, ou obliques adjacents (reconvertis en plans directs), alors qu'en mode 3D, tous les plans aussi bien directs qu'obliques, peuvent être détectés (voir figure 1.10). Le développement d'algorithmes de reconstruction 3D est survenu

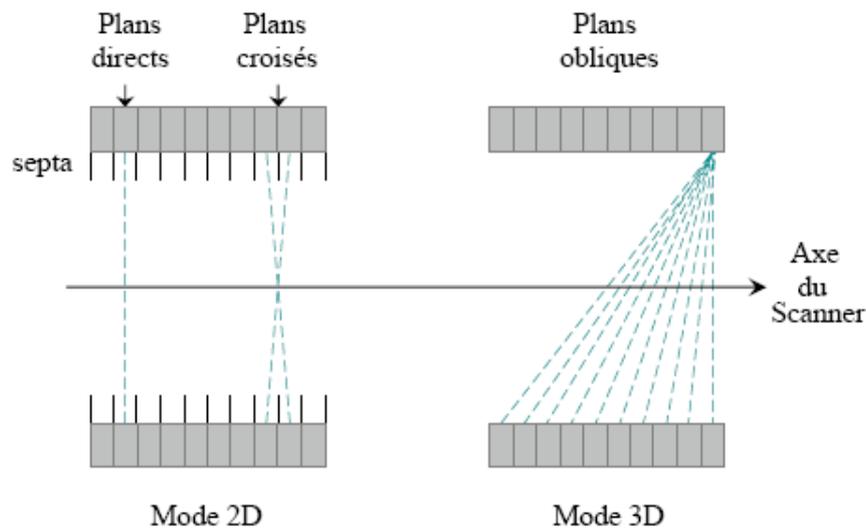


FIG. 1.10 – Comparaison entre les deux modes de fonctionnement, 2D et 3D.

à la suite de l'amélioration significative de la sensibilité géométrique de détection et de la généralisation de l'usage des tomographes en mode 3D [39]. Dans ce mode, on considère toutes les lignes de réponse possibles, y compris entre détecteurs n'appartenant pas aux mêmes anneaux ou anneaux adjacents de détecteurs. La reconstruction d'image qui s'ensuit prend en compte ces lignes de réponse, et on obtient une reconstruction réellement 3D du volume imagé. Cette approche permet parfois d'obtenir un rapport signal sur bruit plus élevé mais il est nécessaire de procéder à des corrections de coïncidences diffusées car le mode 3D s'accompagne d'une augmentation significative du nombre de détections de ces événements et donc d'une augmentation du bruit de fond (environ 40 % de coïncidences diffusées en mode 3D contre 10 % en mode 2D).

Stockage des données : sinogrammes et list mode

Dans le cas des sinogrammes ou des matrices de projection, les données brutes sont triées dans des histogrammes. Cette réorganisation des données acquises a pour but d'accélérer le traitement des données mais a pour conséquence la perte d'une partie de l'information due au sous-échantillonnage. La figure 1.11 illustre la formation d'un sinogramme à partir des données de projection. Même si dans certains cas, cette perte peut être faible [39], la

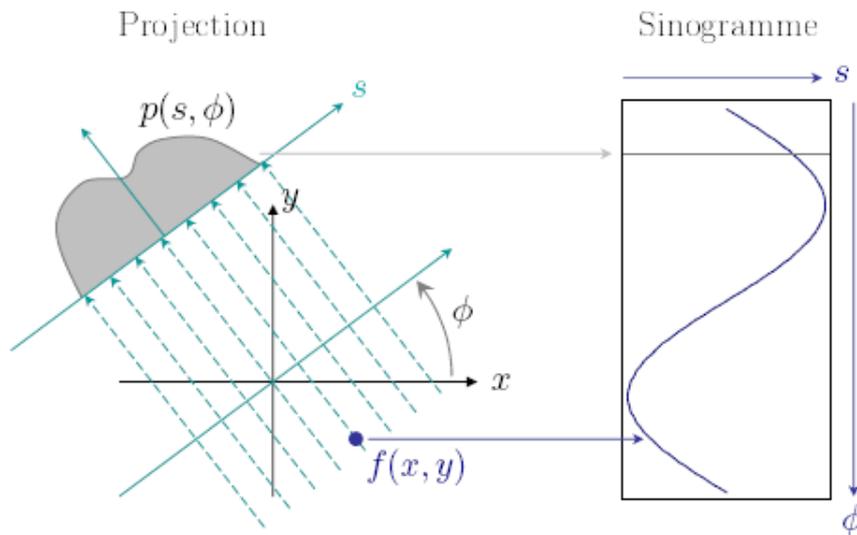


FIG. 1.11 – Une projection $p(s, \phi)$ représente l'intégrale selon toutes les LORs ayant une même direction ϕ . Les projections sont organisées en sinogramme de manière à ce que chaque projection dans une direction donnée ϕ corresponde à une ligne du sinogramme.

reconstruction d'image ne sera optimale que dans le cas où, sans compromis, chaque paire de photons détectés est replacé le plus précisément possible dans l'espace de l'acquisition. Ceci est rendu possible par l'acquisition des données sous forme *list mode*, c'est à dire une liste d'évènements (de détections), classés chronologiquement, avec pour chaque ligne les informations de coordonnées spatiales et temporelles ainsi que d'énergie. Ce format possède également des qualités intrinsèques intéressantes pour les traitements dynamiques 4D [39] étant donné le classement chronologique des évènements.

1.5.2 Reconstruction tomographique

Principe de la tomographie

La tomographie, terme issu de la racine grecque *tomê* pour coupe, donc représentation en coupes, est une technique qui consiste à reconstruire le volume d'un objet, le corps humain dans notre contexte, à partir d'une série de mesures déportées à l'extérieur de l'objet. Le résultat est une reconstruction de certaines propriétés de l'intérieur de l'objet, selon le type d'information que fournissent les capteurs, les lignes de réponse sur lesquelles se situent les désintégrations radioactives dans notre cas. La tomographie, d'un point de vue mathématique, se décompose en deux étapes. Tout d'abord elle nécessite un modèle direct, c'est-à-dire un modèle des phénomènes physiques mesurés. Le modèle inverse, ou reconstruction vient ensuite pour restituer le caractère tri-dimensionnel, en se fondant sur les résultats du modèle direct.

Rétro-projection filtrée

La rétroprojection filtrée (FBP pour Filtered BackProjection) est une approche purement analytique qui considère l'ensemble des lignes de réponses comme des projections, et qui fournit un résultat direct et rapide, mais de qualité faible par rapport aux reconstructions optimisées plus récentes dont nous discutons dans la section suivante (voir figure 1.12)

Algorithmes itératifs

Un algorithme de reconstruction itératif possède les caractéristiques suivantes [39] [131] : un modèle physique du processus d'acquisition comprenant une description des aspects déterministes (diffusés, réponse impulsionnelle des détecteurs, temps mort, etc.) et des aspects statistiques, un modèle des connaissances *a priori* sur la distribution spatio-temporelle du traceur, un critère qui définit la solution choisie, et un algorithme qui permet de s'en approcher. Le principal intérêt des algorithmes itératifs de reconstruction d'images TEP est donc la possibilité d'intégrer de nombreuses optimisations et informations *a priori* (une carte de la variation de la résolution spatiale au sein du champ de vue, la géométrie particulière au scanner, etc). Ces algorithmes sont basés sur des modèles algébriques ou probabilistes de type

bayésien et reposent souvent sur des étapes d'estimations / corrections qui sont assimilables à des rétroprojections / projections. Une image virtuelle est reconstruite à partir des données, un sinogramme virtuel est calculé à partir de cette image virtuelle, et est ensuite comparé au sinogramme réel. L'écart est utilisé pour modifier l'image reconstruite et ce cycle d'estimation/correction est itéré. Une des difficultés est de trouver un compromis entre convergence vers la solution et le biais et la variance résultants des itérations successives. Il est en effet bien reconnu que plus le nombre d'itérations est élevé, plus l'image est bruitée et il faut donc faire appel à des contraintes comme par exemple un post-filtrage par une gaussienne 3D ou un arrêt des itérations avant d'atteindre la convergence [39].

Il existe différents algorithmes, on peut citer OSEM (*Ordered Subsets Expectation Maximization*) [73] et OPL-EM (*One-Pass List-mode Expectation Maximization*) [133], qui travaille sur les données list-mode et qui est une évolution du ML-EM (*Maximum Likelihood Expectation Maximization*) [144], qui travaille lui sur les données brutes de projection. Citons également RAMLA 3D [21], développé comme une alternative à convergence plus rapide que celle de ML-EM, bien qu'il soit mathématiquement démontré qu'ils convergent tous les deux vers la solution ML. La figure 1.12 illustre les différences obtenues sur des images cliniques entre la rétroprojection filtrée et un algorithme itératif (ici OSEM).

Avec l'augmentation des puissances de calcul et la baisse des coûts des processeurs, les algorithmes itératifs dont le principal désavantage est le temps de calcul (complexité en N^4 environ, avec N de l'ordre du nombre d'échantillons de chaque variable [39]), sont devenus une réalité clinique essentielle [131]. En effet, toutes les machines commerciales utilisent dorénavant un algorithme itératif de reconstruction d'images et l'approche par rétroprojection filtrée n'est presque plus utilisée. C'est pourquoi nous limiterons l'évaluation des méthodes aux images reconstruites par des algorithmes itératifs et non par rétroprojection filtrée.

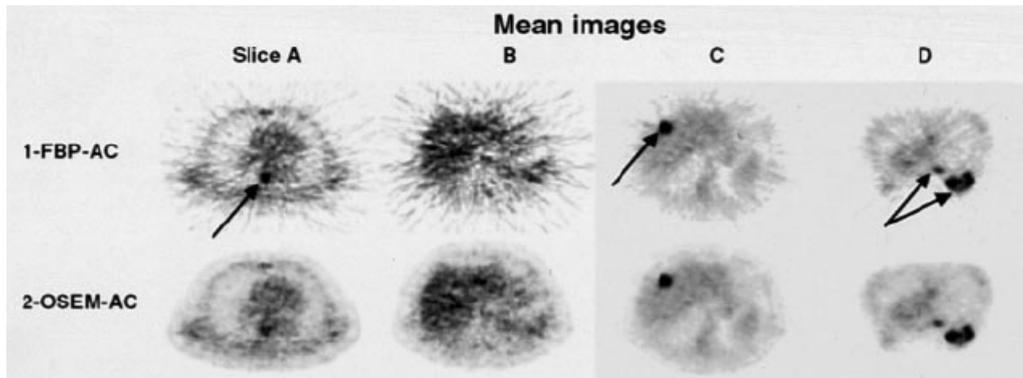


FIG. 1.12 – Images TEP cliniques reconstruites par rétroprojection filtrée (FBP) ou un algorithme itératif (OSEM), les deux avec correction d'atténuation (AC).

1.5.3 Temps de vol

Principe

L'idée du temps de vol (TOF pour *Time of Flight*) remonte aux premiers développements des scanners TEP (les années 60) et a même débouché sur des modèles commercialisés dans les années 80. Toutefois, ces modèles ont été abandonnés pour des questions de limites technologiques des composants électroniques utilisés. Depuis, l'amélioration en résolution temporelle des détecteurs, entre autres facteurs, a permis au principe du temps de vol de revenir sur le devant de la scène, des modèles commerciaux sont à nouveaux disponibles.

Le temps de vol consiste à utiliser l'information du décalage temporel existant entre deux détections de photons en coïncidence pour réduire l'incertitude quant à la position exacte de l'annihilation qui a généré l'émission des deux photons en question. En effet, dans le schéma de fonctionnement du scanner TEP classique, on ne connaît pas la position exacte, dans le champ de vue, de l'annihilation qui a généré l'émission des deux photons détectés en coïncidence, seulement la LOR reliant les deux détecteurs. Grâce à l'information de décalage temporel entre les deux détections, on peut estimer le temps mis par chaque photon pour arriver au détecteur, et ainsi définir une probabilité de position de l'annihilation sur la LOR, probabilité en général modélisée par une gaussienne. La figure 1.13 illustre le principe du temps de vol.

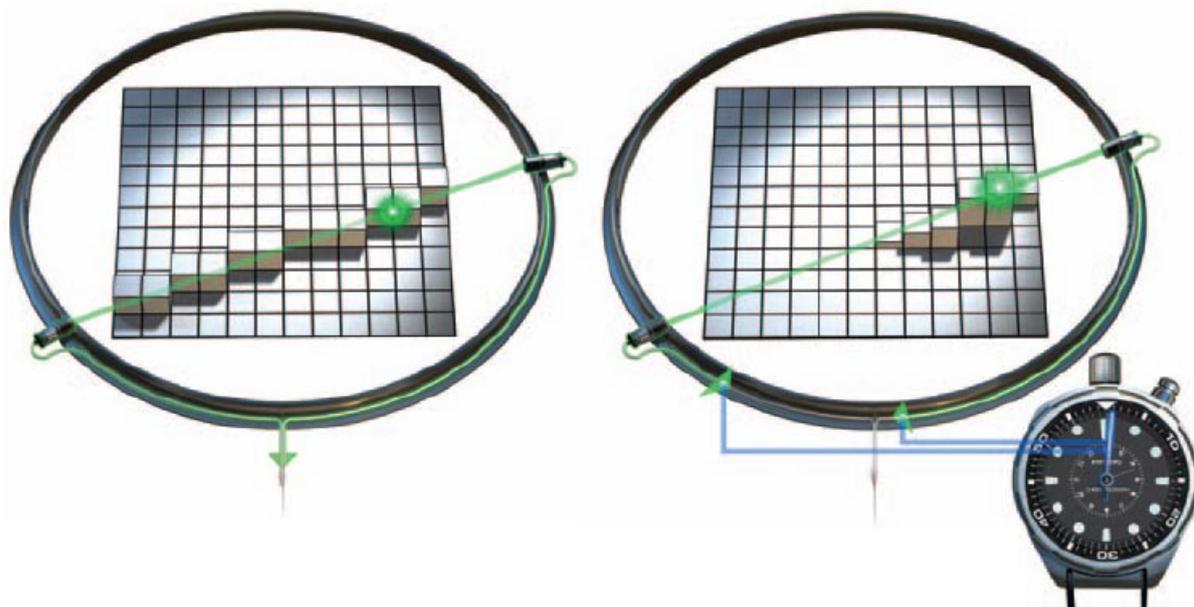


FIG. 1.13 – Illustration du principe de fonctionnement du temps de vol (image issue de la brochure Philips GEMINI TF).

Pourquoi le TOF revient-il sur le devant de la scène ?

La précision temporelle nécessaire pour que ce principe puisse fonctionner en pratique et apporter un bénéfice effectif à la qualité des images est d'environ 500 à 750 picosecondes (soit 10^{-12} s), ce qui était possible dans les années 80, mais seulement avec des matériaux de scintillateurs incapables d'atteindre les performances des systèmes TEP standards en terme de résolution spatiale ou de sensibilité [88]. Les systèmes disponibles à l'époque ont de plus été retirés du marché juste avant le début des années 90, c'est à dire au moment où les examens au ^{18}F -FDG pour les applications d'oncologie commençaient à s'imposer en clinique. De nouveaux scintillateurs disponibles depuis quelques années permettent désormais de considérer à nouveau la technologie du temps de vol pour améliorer les performances des scanners par rapport aux scanners TEP classiques, là où dans les années 80, l'implémentation du principe du temps de vol, associé aux détecteurs disponibles à l'époque, permettait seulement de rejoindre les performances des systèmes standards. De plus, les systèmes électroniques se doivent d'être non seulement très rapides mais aussi parfaitement stables, ce qui n'était pas

le cas à l'époque [88]. Enfin, les algorithmes de reconstructions itératifs doivent être modifiés pour incorporer l'information supplémentaire disponible, ce qui nécessite des développements spécifiques, mais également des puissances de calcul supérieures, qui n'étaient pas facilement disponibles il y a deux décennies, mais qui sont devenues facilement atteignables de nos jours. Philips propose un système TOF TEP/TDM commercial depuis juin 2006 (le GEMINI TF, qui utilise des scintillateurs au *Cerium doped Lutetium Yttrium Orthosilicate* (LYSO) et bénéficie d'une résolution temporelle de 600 ps), et les autres fabricants commencent également à proposer des systèmes TOF.

Bénéfices

La figure 1.14 illustre la différence de qualité entre un système classique et un système TOF, estimée ici pour la tâche précise de la quantification d'une région d'intérêt. Le gain en qualité d'image porte davantage sur l'amélioration du contraste et du rapport signal sur bruit que sur la résolution spatiale [88]. Ceci s'explique par la raison suivante : l'incertitude sur la position de l'annihilation peut être ramenée à quelques centimètres avec une résolution temporelle de l'ordre de 600 ps. Ceci reste largement plus grand que la résolution spatiale du scanner (environ 5 à 6 mm), mais bien plus petit que la largeur des patients, estimée entre 20 cm pour un patient maigre et 35 cm pour un gros patient [88]. C'est pourquoi la résolution spatiale n'est pas réellement améliorée, au contraire du rapport signal sur bruit qui est significativement augmenté, ce qui permet, au choix, d'obtenir des images moins bruitées et mieux contrastées pour des durées d'acquisition identiques, ou des images de qualité identique pour des durées réduites d'acquisition [88].

Le gain observé est donc d'autant plus grand, en regard des performances obtenues avec un système classique, que le patient est gros, en raison de l'atténuation supérieure, de la diminution des coups réels et de l'augmentation des diffusés. Or c'est justement pour ces patients que l'amélioration de l'image est la plus nécessaire.

Il est intéressant également de noter que la reconstruction semble converger plus vite et avec un meilleur contraste [88].

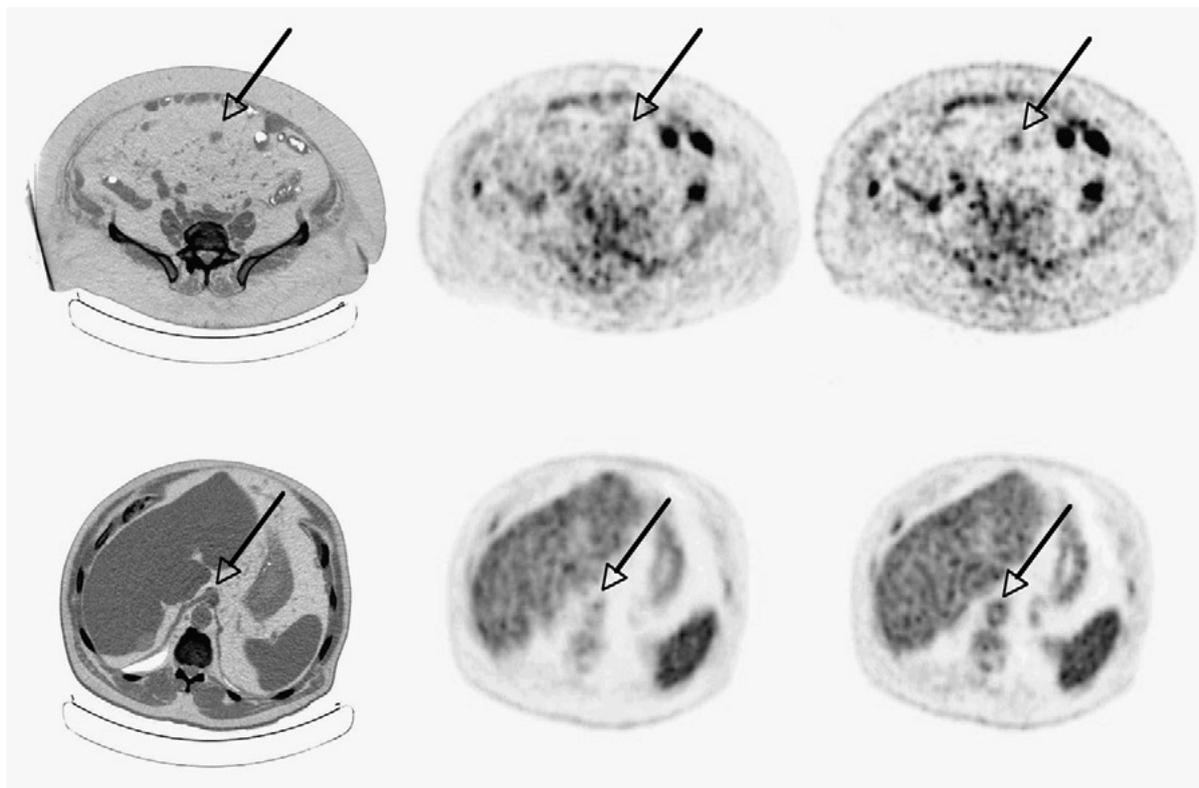


FIG. 1.14 – Comparaison de la qualité d'image entre un système TEP classique au centre, et un système TEP TOF à droite (image CT à gauche), pour des conditions d'acquisitions et de reconstruction identiques. Les images du haut illustrent un cas de cancer du colon bien mieux visible avec le TOF. En bas, illustration de l'amélioration sur la structure de l'aorte, mieux visible en TOF. Image issue de [88].

1.6 Principales limitations

Plusieurs facteurs limitent sérieusement l'utilisation quantitative et automatique des images TEP, notamment dans le cadre du diagnostic pour lequel il est nécessaire d'extraire une information quantitative précise sur un objet d'intérêt pour, par exemple, décider du statut bénin ou malin d'une tumeur. Est également concerné le suivi thérapeutique, pour lequel il faut pouvoir déterminer si l'activité au sein d'un objet d'intérêt a évolué, et si oui quantifier cette évolution. Enfin, la radiothérapie guidée par l'image souffre également de ces limitations car elles réduisent la capacité des utilisateurs à définir les volumes cibles fonctionnels.

1.6.1 Bruit

Le bruit dans les images TEP provient de deux sources distinctes. La première est directement liée à la nature statistique du phénomène observé, c'est à dire les désintégrations radioactives du marqueur associé à la molécule injectée au patient. Ce phénomène est modélisable par une loi de Poisson. La deuxième source de bruit est la reconstruction itérative de l'image. En effet, si le bruit des données brutes acquises par le scanner TEP suit une loi de Poisson, le bruit que l'on retrouve dans les images après reconstruction n'est plus "de Poisson", et dépend fortement de l'algorithme de reconstruction utilisé [129] [130]. Chaque reconstruction introduit un bruit et des artefacts spécifiques à son mode opératoire. Les algorithmes de reconstruction des images, bien qu'ils aient bénéficié de nombreuses améliorations au cours des années, produisent toujours des images imparfaites, aux propriétés variables, souvent bruitées et entachées d'artefacts liés à l'approche utilisée. La section 4.1.3 illustre les différences entre différents algorithmes de reconstruction.

Dans la mesure où notre approche consiste à segmenter les images reconstruites, la modélisation du bruit de Poisson ne sera pas utilisée dans nos algorithmes. Par contre, une étude de robustesse spécifique aux images reconstruites par différents algorithmes est nécessaire. Dans la suite de cette thèse, lorsque nous parlerons de bruit, il s'agira en général de la variation des valeurs des voxels dans une région donnée (variance), et une moyenne éventuellement biaisée, par rapport à une valeur réelle supposée. En particulier, dans le contexte des méthodes de segmentation statistique que nous utilisons pour développer les algorithmes de segmentation des images TEP, le bruit désigne la façon dont les valeurs des voxels se répartissent dans une classe donnée, caractérisée en général par un type de distribution statistique définie par des paramètres tels que la moyenne et la variance.

1.6.2 Résolution spatiale

La résolution spatiale mesure la capacité d'un imageur à différencier deux points voisins. Ainsi, un imageur affecté d'une résolution spatiale de un millimètre est capable de produire une image de deux objets séparés par un millimètre de distance. Avec une résolution spatiale

moins précise, les deux objets seront fusionnés en une seule tache indifférenciée. Cela ne signifie pas pour autant que des objets supérieurs à un millimètre de diamètre seront parfaitement résolus. Pour définir la résolution spatiale, on utilise généralement la notion de 'largeur à mi-hauteur' (ou FWHM pour *Full Width at Half Maximum*) de la réponse impulsionnelle (ou PSF pour *Point Spread Function*) de l'imageur. Dans le cas des scanners TEP, la PSF est tri-dimensionnelle. La notion de largeur à mi hauteur permet de visualiser la résolution spatiale et de considérer l'image obtenue comme une convolution par cette PSF de l'objet observé. La figure 1.15 illustre la notion de FWHM en 1D.

Le principal défaut de l'imagerie TEP est sa résolution spatiale relativement médiocre par

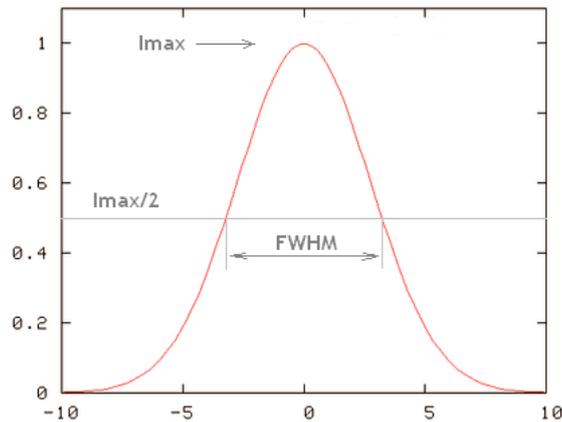


FIG. 1.15 – Illustration de la largeur à mi-hauteur ou FWHM.

rapport aux autres modalités d'imagerie médicale. En effet, là où l'imagerie anatomique par rayons X (TDM) ou par résonance magnétique (IRM) permettent d'obtenir des résolutions sub-millimétriques, il faut se contenter d'une résolution spatiale qui peine à descendre sous cinq millimètres pour un scanner TEP classique actuel. Une telle résolution spatiale a plusieurs conséquences gênantes, en particulier pour le domaine de l'oncologie.

Résolution spatiale intrinsèque

La résolution spatiale dite "intrinsèque" est celle découlant directement des propriétés physiques du scanner, la taille de son champ de vue, la taille, le nombre et le positionnement des détecteurs, etc. Il faut faire la différence entre la résolution spatiale du scanner et la taille des

voxels de l'image, qui sont deux éléments bien distincts. Il est en effet tout à fait possible de reconstruire des images TEP avec des voxels dont la dimension est fixée arbitrairement (par exemple 5 ou 2 mm de côté), mais cela n'améliore évidemment en rien la résolution spatiale intrinsèque du scanner. Cette dernière est fixée, bien qu'elle varie avec la position dans le champ de vue. Elle est la plus précise au centre du champ de vue, et se dégrade lorsque l'on s'en éloigne. Il est possible, grâce à une source radioactive ponctuelle que l'on déplace dans le champ de vue, de générer une carte de cette résolution spatiale en fonction de la position d'un objet dans le champ de vue. On peut alors utiliser cette carte dans la reconstruction de l'image, de façon à tenir compte de cette variation pour améliorer la qualité de l'image [53]. Malgré toutes les améliorations possible du scanner, il faut souligner que le phénomène physique détecté n'est lui même qu'une approximation du lieu de détection. En effet, avant son annihilation avec un électron, le positon parcourt une distance comprise entre 1 et 3 mm suivant son énergie d'émission. De plus, du fait que sa quantité de mouvement au moment de son annihilation n'est pas strictement nulle, il existe une variabilité de plus ou moins 0.5° de l'angle entre les deux photons émis. Ces deux éléments combinés font que la mesure est intrinsèquement entachée d'erreur.

Effets de volume partiel

La résolution spatiale a une conséquence que l'on nomme effet de volume partiel (EVP, ou PVE pour *Partial Volume Effect*) [147]. L'EVP désigne deux phénomènes distincts et particulièrement importants en imagerie d'émission. Le premier concerne l'effet de la faible résolution spatiale sur l'aspect visuel des frontières entre différentes régions de l'image. Les transitions entre les régions sont floues et étalées au lieu d'être nettes et brusques, et plus l'objet est petit, plus il est difficile à différencier du fond environnant, en particulier si le contraste est faible. Le deuxième phénomène est nommé spécifiquement *spillover* ou contamination croisée et concerne l'impact quantitatif des EVP. Il désigne le fait que deux régions d'intensités différentes se contaminent mutuellement. L'EVP est illustré dans la figure 1.16. On observe dans cette figure les deux effets : d'une part, le profil de la structure imagée est étalé, les frontières de l'objet sont rendues difficiles à discerner. D'autre part le signal de

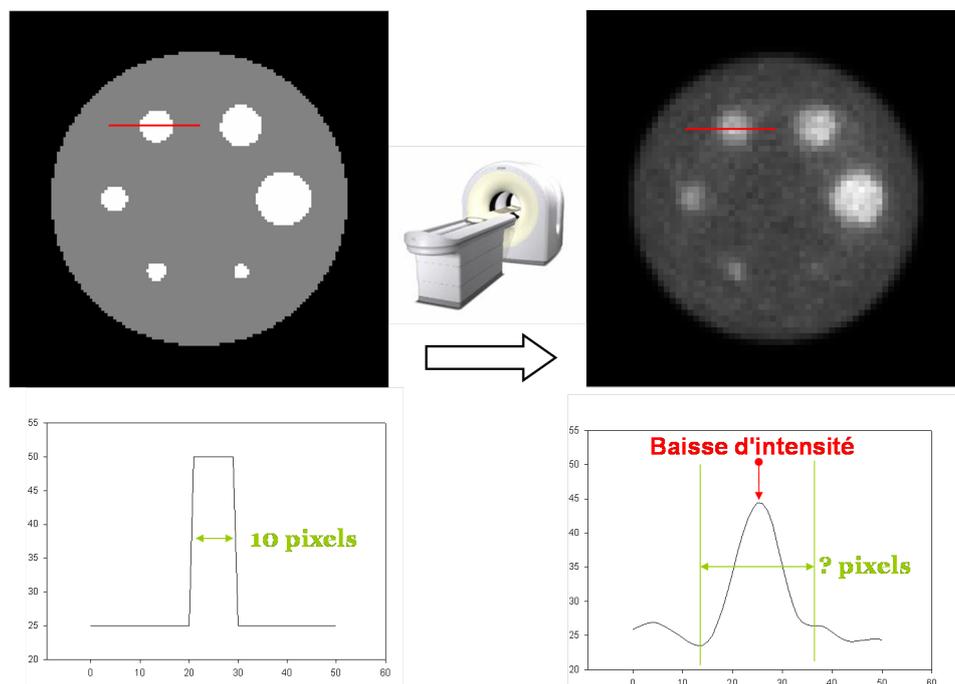


FIG. 1.16 – Illustration des effets de volume partiel. En haut à gauche, un objet idéal contenant des sphères. A droite, l'image acquise par un scanner TEP de cet objet (simulation). En bas, les profils correspondants aux barres rouges sur chaque image illustre la perte d'intensité et l'étalement du signal dus aux effets de volume partiel.

l'objet dans son ensemble, est réduit. Typiquement, lorsque l'objet est d'un diamètre inférieure à 2 cm (environ quatre fois la résolution spatiale), même le signal au centre de l'objet, et pas seulement sur ses bords, est diminué significativement par rapport à la réalité.

Il existe de nombreuses méthodes permettant de corriger ces effets, que l'on peut classer en deux catégories : les méthodes travaillant en post-reconstruction ou les méthodes s'appuyant sur la reconstruction elle-même. On trouve une revue récente de ces méthodes dans [137].

Echantillonnage spatial

En plus des EVP, il faut considérer le fait que les images TEP sont reconstruites sur une matrice (ou grille) de voxels dont le volume est défini arbitrairement, avec une influence non négligeable sur la qualité de représentation des objets dans l'image ainsi que de la qualité générale de l'image résultante. Plus les voxels sont petits, mieux les objets sont représentés

sur l'image, mais inversement, moins le nombre de LORs passant par ce voxel est grand, ce qui réduit la statistique par voxel et en conséquence diminue d'autant le rapport signal sur bruit. Ainsi, plus les voxels sont petits, meilleur est l'échantillonnage spatial d'un objet, mais la qualité de l'image s'en trouve dégradée. Il faut trouver un compromis entre le rapport entre la taille des voxels et l'objet imagé, et la qualité de l'image. En routine clinique, la taille des voxels est généralement située entre 2 et 5 mm.

1.6.3 Mouvements et respiration

Les mouvements et la respiration du patient génèrent des artefacts particulièrement sévères [162] qui peuvent entraîner des sous-estimations d'activité allant jusqu'à 150 % dans des tumeurs pulmonaires [114] et des sur-évaluations de volume de 100 % pour des lésions dont le diamètre est inférieur à 15 mm [168]. En effet, contrairement à une acquisition anatomique qui dure quelques secondes et pour laquelle il est envisageable de demander au patient de bloquer sa respiration pour éviter les artefacts, l'acquisition TEP dure une vingtaine de minutes en moyenne. Ceci est nécessaire pour obtenir une image de qualité suffisante et interprétable par l'oncologue : il faut enregistrer suffisamment de lignes de réponse. Cette durée explique que l'image reconstruite *a posteriori*, est une visualisation moyennée sur l'intégralité de l'acquisition, et donc de tous les cycles respiratoires du patient. Lorsque ce dernier respire, sa cage thoracique se soulève et redescend, ses poumons se gonflent et se vident, son diaphragme se déplace, etc. Tous les objets imagés dans la région du thorax sont donc rendus flous par ces mouvements, ainsi que les organes de l'abdomen avec une amplitude plus faible. La figure 1.17 illustre les artefacts résultant de la respiration du patient. C'est pourquoi de très nombreux travaux se sont penchés sur la prise en compte et la correction de ces mouvements respiratoires, pour produire des images de meilleure qualité tant d'un point de vue qualitatif que quantitatif, l'arrivée de la multimodalités TEP/TDM ayant permis des progrès dans ce domaine particulier [93] [94]. Une revue récente des méthodes de correction est disponible dans [164]. La deuxième partie de la figure 1.17 illustre l'efficacité de telles approches.



FIG. 1.17 – Illustration des effets de la respiration sur des images simulées, et de la performance des méthodes de correction. A gauche, une image sans respiration, au centre l'image incorporant les effets de respiration et à droite l'image corrigée par la correction élastique, au sein de la reconstruction [93]. D'un point de vue quantitatif, la récupération d'activité dans les tumeurs simulées est de l'ordre de 72-90%, l'information de position est corrigée à hauteur de 78-100% et l'étalement (FWHM) de 67-92%.

1.7 Contexte clinique, besoins et objectifs

La tomographie d'émission (TEP et TEMP) est utilisée dans plusieurs domaines de la médecine, en particulier la neurologie, la cardiologie et l'oncologie. La majorité des applications concerne l'oncologie. Il existe principalement trois traitements contre le cancer : la chirurgie, la chimiothérapie et la radiothérapie, souvent utilisés de concert suivant des protocoles précis, adaptés à chaque type de cancer. Les chiffres publiés début 2008 par l'institut de veille sanitaire français montrent que le nombre de cas de cancer était en augmentation, bien que la mortalité associée diminue. Les raisons principales de cet état de fait sont la combinaison de l'amélioration du dépistage et des comportements et environnements cancérogènes (tabac, alcool, habitudes alimentaires, expositions à des produits de plus en plus nombreux et variés, etc.). Ceci étant compensé par la précocité et l'amélioration de l'efficacité des traitements, qui entraînent de meilleures chances de survie.

Le dépistage et le diagnostic, le traitement et le suivi des patients en oncologie reposent aujourd'hui largement sur l'analyse préconisée d'examen TEP au 18F-FDG. Cet outil est maintenant considéré comme indispensable dans ce domaine [76] [84]. Notre travail trouve un intérêt dans plusieurs applications de la TEP en oncologie. Les sections qui suivent détaillent ces dernières.

1.7.1 Analyse semi-quantitative et SUV

L'exploitation des images TEP, comme nous l'avons vu dans la section précédente, est compliquée par l'aspect bruité, flou et la qualité variable des images en fonction des modes d'acquisition ou des paramètres de reconstruction d'image. De nombreux travaux ont cherché à standardiser l'extraction de paramètres quantitatifs des images, pour réduire la variabilité et la subjectivité liées à l'analyse visuelle et manuelle des images. La difficulté vient du fait que la fixation du FDG dans les tissus dépend de nombreux paramètres, notamment la quantité injectée au patient et la voie d'administration, la masse du patient, la perfusion du tissu, le temps écoulé entre l'injection et l'examen, etc. Une notion qui a été proposée est l'utilisation d'une valeur standardisée, nommée SUV pour *standard uptake value* [152], pour permettre la comparaison d'examens de multiples patients au sein d'études ou de différents examens d'un même patient. Le SUV est obtenu par le ratio entre la concentration de radioactivité du tissu (par exemple en kBq/ml) au temps T et la dose injectée (MBq) divisée par le poids du patient (Kg). Il a été proposé d'utiliser la surface du corps (*Body Surface Area* ou BSA) ou encore la masse maigre à la place du poids puisque la distribution du FDG est plus faible dans les tissus adipeux. Par exemple, pour les tumeurs pulmonaires, la valeur seuil d'un SUV de 2.5 a été proposée pour différencier une accumulation maligne d'une bénigne [7] [65] mais des études ont souligné les limites de la valeur pratique de ce seuil pour diagnostic [103] [43].

Malheureusement, la détermination du SUV elle-même est sujette à une grande variabilité, car les résultats peuvent varier de façon significative suivant la méthode employée pour calculer cette valeur, en fonction des normalisations utilisées [121]. De nombreuses études ont souligné les limites de l'utilisation du SUV comme mesure quantitative utile en diagnostic [66] [90] [72], en particulier dans le cadre d'étude multi-centriques [12]. Il a même été montré que dans certains cas, l'utilisation du SUV peut mener à un diagnostic erroné concernant l'évolution d'un patient au cours du traitement [57].

1.7.2 Utilisation diagnostique de la TEP en oncologie

Depuis les années 90 et jusqu'à aujourd'hui, l'utilisation majoritaire de l'imagerie TEP en oncologie est le diagnostic [141] [83] [79]. Les notions de sensibilité (Se) et spécificité (Sp) sont souvent utilisées dans ce contexte et sont définies à l'aide des faux négatifs (FN), vrais négatifs (VN), faux positifs (FP) et vrais positifs (VP) :

$$Se = \frac{VP}{VP + FN} \quad (1.2)$$

$$Sp = \frac{VN}{VN + FP} \quad (1.3)$$

L'idéal est d'avoir une parfaite sensibilité et spécificité, c'est à dire qu'on ne détecte aucun faux positif (spécificité de 100 %) et aucun faux négatif (sensibilité de 100 %). Une mesure de l'impact de l'examen TEP est la modification de la prise en charge du patient, c'est à dire le fait d'avoir éventuellement modifié le traitement ou d'avoir entraîné des examens supplémentaires. La force de l'imagerie TEP est de pouvoir visualiser des pathologies mises en évidence par le phénomène physiologique observé, dans le cas du ^{18}F -FDG la consommation de sucre des cellules, qu'il n'est pas forcément possible d'appréhender avec d'autres modalités d'imagerie ou d'examens non invasifs. Il s'agit d'une imagerie particulièrement sensible mais assez peu spécifique, en particulier, il est souvent difficile de faire la différence entre l'inflammation autour de la tumeur et la tumeur elle même. L'aspect diagnostique de la TEP reste largement qualitatif, avec une détection visuelle de la tumeur et une estimation grossière des paramètres de taille, position, activité, bien que les progrès réalisés avec les différentes méthodes de correction des images (atténuation, diffusées, aléatoires, etc.) permettent de mieux exploiter l'aspect intrinsèquement quantitatif de cette imagerie fonctionnelle. En effet, chaque élément (ou voxel) de l'image contient une information liée, malheureusement de façon non déterministe et non linéaire, à l'activité radioactive qu'il contient. La difficulté réside dans la définition de la région dont on va ensuite extraire cette information quantitative, ce qu'on appelle quantification. De nombreuses approches ont été proposées, mais aucune n'a réellement résolu les difficultés et le consensus actuel est qu'il n'existe pas de méthode standard pour le contourage des tumeurs sur les images TEP.

Le diagnostic en TEP peut concerner la différenciation entre cellules malignes et bénignes,

un bilan d'extension ou un diagnostic de récurrence. La différenciation entre cellules malignes et bénignes a lieu dans un stade précoce de prise en charge des patients. Au contraire, le bilan d'extension a lieu une fois le diagnostic cancéreux établi, pour orienter les choix thérapeutiques, en particulier pour mettre en évidence une extension de la maladie qui n'a pas été détectée au cours d'un examen précédent. Enfin, le diagnostic de récurrence est utilisé pour la détection des deux types de récurrences : celles sur le site de la tumeur primaire, ou les métastases à distance.

1.7.3 Suivi thérapeutique

Grâce à sa nature qualitative mais aussi quantitative, l'imagerie TEP est de plus en plus utilisée pour évaluer le pronostic [46] [160] et la réponse thérapeutique aux traitements [169] [136] [33]. Plusieurs examens au cours du temps peuvent ainsi être réalisés, notamment avant et après un traitement par chimiothérapie ou radiothérapie par exemple, de façon à évaluer l'évolution de la maladie. Dans certains cas, la nature qualitative de l'imagerie peut suffire, par exemple si une tumeur est clairement visible avant traitement, et a complètement disparu après traitement. Mais dans certains cas, la réponse du patient peut n'être que partielle, et plus subtile à quantifier. Une exploitation robuste et précise de la nature qualitative de l'imagerie TEP devient alors indispensable pour être capable de quantifier la réponse et de classer le patient dans une catégorie, comme par exemple répondant, répondant partiel, ou non répondant. Le problème est que la qualité limitée des images nuit encore à une utilisation généralisée de la TEP pour cette quantification. En effet, les méthodes disponibles pour quantifier la tumeur, c'est à dire obtenir les paramètres de taille/volume et d'activité, sont trop peu robustes et précises pour pouvoir quantifier des variations faibles entre deux images différentes. Une des solutions proposées est actuellement de se baser sur la valeur de SUV maximale pour évaluer la réponse thérapeutique.

1.7.4 Radiothérapie

La radiothérapie est un des trois traitements possibles du cancer, avec la chimiothérapie et la chirurgie. Les pathologies nécessitant une radiothérapie représentent une part significative des patients traités en oncologie (environ 60 à 70 %). Pour ce qui est de la France, d'après

le rapport de la commission d'orientation sur le cancer (janvier 2003), la radiothérapie est en effet une des trois principales modalités de traitement avec la chirurgie et la chimiothérapie, souvent en association avec ces dernières. 30 à 40 % des rémissions sont obtenues grâce (au moins en partie) à l'utilisation de la radiothérapie. Il existe en France presque 400 appareils répartis dans plus de 180 établissements (chiffres 2006) et depuis le début du plan cancer (2003-2006) le parc national de machines a augmenté de 12 %. La radiothérapie représentait 17 % des dépenses de l'assurance maladie pour les cancers en France pour l'année 2001 et s'est donc imposé comme un outil majeur pour le traitement de nombreux cancers. Aux Etats-Unis, la tendance est la même et la radiothérapie constitue un domaine dans lequel les nouvelles technologies se développent rapidement et pour lequel le développement d'applications et d'outils est important. Une étude de l'organisme américain IMV [47] portant sur les sites de radiothérapie aux Etats-Unis, a récemment montré une augmentation de 5 % de la fréquentation de ces services en 2007 par rapport à 2006. Si le nombre de patient traités est donc relativement stable, il est intéressant de noter ces autres chiffres, autrement plus significatifs : en 2004, 15 % environ des sites utilisaient des systèmes guidés par l'image (IGRT). Aujourd'hui, ce chiffre est passé à plus de 50 %, et 87 % des sites ont adopté un système de délivrance à intensité modulée (IMRT). Enfin, le budget de ces mêmes sites est également en augmentation significative : le pourcentage de sites disposant d'un budget supérieur à 1.5 million de dollars pour l'année 2008 a doublé, passant de 15 % en 2003 à plus de 30 %. Cette étude confirme également la faible utilisation en routine clinique de l'imagerie fonctionnelle dans la planification et l'exécution du traitement : si presque la totalité des sites (97 %) utilisent l'imagerie anatomique TDM, seuls 11 % utilisent l'IRM ou la TEP.

L'objectif de la radiothérapie est de délivrer une dose de rayonnements ionisants maximale dans les tissus tumoraux tout en réduisant au minimum l'irradiation résultante des tissus environnants [60]. Une organisation internationale (*International Commission on Radiation Unit*) a défini dans ce contexte différentes marges de sécurité concernant la définition des volumes cibles en radiothérapie basée sur l'imagerie anatomique [59] : le volume tumoral macroscopique (*Gross Tumor Volume*, GTV), le volume tumoral anatomoclinique (*Clinical Tumor Volume*, CTV) et le volume cible prévisionnel (*Planning Target Volume*, PTV). Ce dernier correspond

aux incertitudes liées aux mouvements, avec notamment le volume cible interne (*Internal Target Volume*, ITV) et la marge de mise en place (*set-up margin*). La définition du GTV est toutefois basée sur l'hypothèse fautive que la cible est homogène et nécessite une dose uniforme d'irradiation.

C'est pourquoi, la tendance actuelle dans les travaux de recherche concernant la radiothérapie externe, conformationnelle notamment, est de s'intéresser de plus en plus à l'information complémentaire apportée par la TEP pour définir les régions cibles du traitement par rayons, mais aussi pour décider du traitement (radiothérapie seule, ou combinée avec d'autres traitements, voire renoncer à la radiothérapie), et évaluer la réponse au traitement. Il s'agit du concept de volume cible biologique (*biological target volume*, BTV). Bien que l'information anatomique soit toujours la source principale servant à définir le traitement de radiothérapie [52], l'intuition selon laquelle l'utilisation complémentaire de l'information fonctionnelle a le potentiel d'améliorer la planification et les résultats du traitement de radiothérapie [3] [56] [19] a été confortée par les résultats apportés par de récentes études [1] [158]. La première confirme la stabilité des zones d'hétérogénéité et d'activité maximale au sein des tumeurs durant la radiothérapie, suggérant que la connaissance de ces zones peut améliorer le traitement. La deuxième apporte la preuve que délivrer des doses supérieures aux tumeurs est faisable et améliore la survie des patients, et que cette augmentation de dose est fortement dépendante de la dose délivrée aux tissus sains environnants. La récente conférence ASTRO'08 (*American Society for Therapeutic Radiology and Oncology*) a été notamment marquée par la présentation de nombreuses études cliniques soulignant l'impact significatif et favorable de l'intégration des données TEP dans la définition des cibles (en comparaison avec les données anatomiques seules) [92] [143] [68], par exemple en ce qui concerne la variabilité inter-radiothérapeutes pour la définition des volumes d'intérêt [107] [105] mais aussi plusieurs interventions soulignant la nécessité d'intégrer l'information fonctionnelle fournie par la TEP pour définir un "volume cible biologique" et ainsi optimiser la dosimétrie et réduire les doses délivrées aux organes à risques [4] [108] [64]. Ces mêmes intervenants ont toutefois tous en commun de poser l'interrogation suivante : comment utiliser l'information fonctionnelle ? Quelle valeur de SUV fixer pour définir le volume fonctionnel ?

Plusieurs études soulignaient également l'utilité et l'impact de la TEP dans le cadre du traitement par radiothérapie pour établir le traitement, suivre et évaluer la réponse thérapeutique pour éventuellement établir un nouveau planning (*re-planning*) en fonction de l'évolution du volume tumoral [55] [67] [149].

L'utilisation de l'information fonctionnelle pour définir les volumes tumoraux, qui serviront de base à la définition d'un plan de traitement et les volumes cibles associés, nécessite la segmentation précise et robuste des images fonctionnelles, or aucune méthode proposée jusqu'à présent n'est en mesure d'offrir une segmentation correcte des tumeurs réelles souvent hétérogènes. Notre travail permet en partie de combler ce manque et de répondre au besoin de la communauté avec une méthode automatique permettant d'exploiter l'imagerie TEP dans le cadre de la planification de traitement par radiothérapie. En effet, les systèmes commerciaux délivrant les rayons possèdent la capacité technique d'irradier les tumeurs suivant un contour défini, et ce avec une grande précision de l'ordre du millimètre. Il ont même la capacité de délivrer des doses non homogènes et adaptées aux différentes régions au sein même d'une tumeur. Ce qui manque pour y parvenir : le paramètre d'entrée, à savoir le contour, qui pour l'instant est toujours défini à la main sur l'image anatomique.

Cette nouvelle approche connue sous le nom de "*dose painting*" [101] peut être mise en oeuvre grâce à une alliance entre la précision des systèmes délivrant les doses et la segmentation précise des volumes à irradier prenant en compte l'information fonctionnelle, menant à une optimisation de la dosimétrie. Il est ainsi envisageable de délivrer des doses proportionnelles à l'activité de chaque région au sein même de la tumeur. Une segmentation des images TEP robuste et précise permettrait de pallier ce manque. La section 6.2.1 détaille l'utilisation de notre approche dans le contexte de la radiothérapie.

Notons que s'il est attendu que l'utilisation de l'information fonctionnelle dans le cadre de la radiothérapie permette une détermination plus précise et appropriée des volumes cibles afin d'augmenter l'efficacité des rayonnements et diminuer l'irradiation des tissus sains, de prendre en compte l'hétérogénéité biologique des tumeurs, et d'aider les études d'escalade de doses [61], aucune étude n'a pour l'instant démontré un bénéfice clinique pour les patients.

Chapitre 2

Etat de l'art de la segmentation en TEP

2.1 Introduction

Dans ce chapitre nous allons aborder la problématique de la segmentation d'images, apporter quelques définitions, et passer en revue les méthodes de segmentation ayant déjà fait l'objet de travaux pour les applications en imagerie d'émission. Nous parlerons également de certaines approches intéressantes n'ayant jusqu'à présent pas ou très peu été utilisées sur des images TEP.

2.1.1 Objectif

Cet état de l'art met en lumière les pistes de segmentation d'images TEP qui ont été explorées de manière exhaustives ou au contraire délaissées. Il s'agit de définir ce que doit offrir une technique automatique de segmentation d'image d'émission, voir si de telles méthodes existent, et si oui, si elles ont déjà fait l'objet d'application en TEP et TEMP.

Comme nous l'avons détaillé dans la section 1.6, les images d'émission sont à la fois bruitées et floues. Nous sommes donc à la recherche d'une approche permettant de prendre en compte ces deux aspects de façon cohérente et pertinente, sans multiples étapes de pré-traitement. En particulier, nous souhaitons éviter des étapes de débruitage et de déconvolution. Le débruitage

comme la déconvolution peuvent être réalisées à l'aide de nombreuses méthodes différentes (au moins aussi nombreuses que les méthodes de segmentation elles-mêmes), ce qui complique le choix et nécessite une étude complète en elle-même. En effet, le débruitage est souvent associé à une perte significative d'information, ce qui est particulièrement dommageable pour une imagerie intrinsèquement quantitative comme la tomographie d'émission. D'un autre côté, la déconvolution est un domaine également riche, mais les méthodes proposées impliquent souvent une diminution du rapport signal sur bruit et sont délicates à paramétrer, ne serait-ce parce que la résolution spatiale en TEP n'est pas la même à différentes positions dans le champ de vue.

2.1.2 Définitions

Segmentation et classification

En traitement d'images, la *segmentation* est usuellement définie comme l'opération consistant à partitionner une image en zones homogènes ou régions. Une région est alors définie comme un ensemble de pixels qui partagent une caractéristique commune comme l'intensité, la couleur, la texture, etc. La *classification* consiste à labelliser les régions formées par la segmentation. Bien que la segmentation puisse donner directement le résultat de la classification, suivant la complexité des données et la méthode de segmentation utilisée, il peut être nécessaire de recourir à des post-traitements supplémentaires ou à l'apport d'informations *a priori* pour obtenir une classification pertinente.

Dans le but de séparer un objet du fond dans une image, l'être humain se base sur des connaissances de haut niveau qui l'aident à détecter ce qui l'intéresse : le cerveau repère en général les formes dans leur ensemble, et réalise des opérations diverses (rotation, étirement, etc.), tout en cherchant dans la mémoire visuelle des formes similaires. Il s'agit d'une approche intégrée dans le sens où l'image est analysée dans son ensemble directement. Au contraire, la segmentation automatisée d'image se heurte généralement au problème que les algorithmes généralement mis en oeuvre travaillent au niveau de la brique élémentaire de l'image, le pixel (ou le voxel dans le cas de volumes 3D), et ne peuvent pas, comme l'humain, considérer

facilement l'image dans son intégralité directement et chercher dans une mémoire. La méthode qui se rapproche le plus du fonctionnement du cerveau humain est celle de la classification d'images utilisant des bases d'apprentissage et des approches fondées sur la fouille de données. Du point de vue des algorithmes et systèmes informatiques usuellement utilisés pour la segmentation automatique d'images, l'image est définie comme un ensemble de pixels organisés suivant une grille, chacun ayant une valeur, souvent un niveau de gris ou des informations de couleurs RVB. Dans le cas des images TEP, il s'agit d'une valeur réelle positive ou nulle, généralement codée sur 32 bits. On ne dispose généralement que de ces informations (ensemble des pixels, avec leur position dans la grille et leur valeur) pour réaliser la segmentation. Il s'agit donc de définir des critères de sélection des pixels en utilisant le mieux possible cette information limitée. La célèbre image de la figure 2.1 est fréquemment utilisée dans la



FIG. 2.1 – La fameuse photo de Léna, rendue populaire par la communauté scientifique de l'imagerie.

communauté scientifique et informatique de traitement d'images pour tester algorithmes de déconvolution, de débruitage, d'analyse spectrale ou encore de segmentation. Elle contient en effet un mélange intéressant de détails, de régions homogènes, d'ombrages et de textures [113]. Cette image permet de bien illustrer le problème général de la segmentation : en effet, même un enfant sans beaucoup d'expérience est capable de séparer la jeune femme, c'est à dire Léna et son chapeau, du fond de l'image, contenant le miroir et le reflet de Lena, et le mur. Pourtant, obtenir le même résultat par un traitement automatique est un défi, car les niveaux de gris sont bien souvent les mêmes au sein des différentes régions. Comment

faire alors la différence entre ces deux régions, en travaillant au niveau du pixel et non en considérant, comme un humain, l'image dans son ensemble? Dans le cadre de notre travail, les images sont, du moins d'un certain point de vue, moins complexes à segmenter que celle de Léna. La difficulté de segmentation des images acquises en tomographie d'émission, comme nous l'avons vu dans la section 1.6, ne réside pas dans le grand nombre de régions ou la complexité des objets à reconnaître, mais plutôt dans l'aspect bruité et flou d'une structure d'intérêt, une tumeur et son environnement immédiat par exemple, chacune de ces deux régions principales pouvant présenter des inhomogénéités importantes, susceptible de mettre en échec des approches simplistes de segmentation.

En général, les méthodes de segmentation des images sont regroupées en deux grandes familles : "l'approche contours" et "l'approche régions". On peut toutefois en rajouter d'autres, comme l'approche markovienne ou celle reposant sur une forme d'intelligence artificielle, avec ou sans apprentissage et fouille de données. Suivant le degré de complexité des objets et des images, on est en général amené à rechercher des régions "homogènes" pour résoudre la segmentation, et pour cela, il est nécessaire de définir ce que "homogène" signifie, et ce pour chaque application. On introduit alors des critères qui permettent de déterminer le regroupement de pixels, ou, suivant l'approche utilisée, de différencier des pixels adjacents.

L'approche contours

Cette première famille de méthodes s'intéresse aux variations des niveaux de gris dans l'image pour détecter des transitions entre différentes régions. Un contour est alors défini comme une frontière entre deux milieux. On utilise généralement des méthodes dérivatives pour détecter ces contours. De très nombreux filtres permettant de mettre en évidence les contours ont été proposés (Prewitt et Sobel, Laplacien, MDIF, Canny, Deriche, Perona Malik, etc.).

On peut également classer les modèles déformables dans cette catégorie. Ces méthodes utilisent des initialisations de modèles (courbes 2D et 3D, surface...) qui se déforment itérativement pour s'adapter à la forme de l'objet à séparer du fond. Suivant le bruit présent dans les images, les différents filtrages et lissages utilisés peuvent être plus ou moins efficaces, effacer en partie les contours recherchés, les rendre flous, ce qui complique leur segmentation ultérieure.

L'approche régions

Dans cette famille de méthodes on ne s'intéresse pas à la détection des contours des objets, mais plutôt à attribuer à chaque pixels ou voxels une appartenance à une région de l'image. On cherche donc à diviser l'image en régions distinctes suivant un critère donné, ou au contraire à fusionner les régions dans une image tout d'abord sur-segmentée. On peut citer l'algorithme très connu des K-moyennes (K-Means) ou du Fuzzy clustering (Fuzzy C-Means), ainsi que la croissance de régions. Ces algorithmes nécessitent parfois des graines d'initialisation, et souvent le nombre de régions recherchées *a priori* dans l'image.

Les approches bayésiennes et/ou markoviennes

Cette approche est probabiliste et statistique. Elle permet, dans les différents modèles qu'elle met en oeuvre, d'intégrer deux paramètres pour la décision de classification d'un pixel. Bien que son approche soit plus "région" que "frontière", elle mérite sa propre catégorie à cause de sa spécificité probabiliste et des méthodes d'estimation des paramètres qu'elle utilise. Elle fonde en effet sa décision, comme d'autres approches régions, sur l'intensité du pixel et la corrélation spatiale entre les pixels (deux pixels voisins et d'intensité proche ont plus de chance d'appartenir à la même classe), et également suivant la répartition statistique des pixels dans l'image, avec des modèles de bruit suivant des lois différentes (en général gaussienne), ce qui permet de segmenter des images fortement bruitées.

L'approche par intelligence artificielle

Cette approche est particulière dans la mesure où elle est rarement citée en segmentation d'images. Cette voie est encore expérimentale, et fait souvent appel à de longs pré-traitements d'apprentissage. On peut citer les méthodes de classification par réseaux neuronaux par exemple, les algorithmes génétiques ou les systèmes experts. Cette approche est plus souvent utilisée dans des systèmes de classifications d'images, de fouille de données ou de reconnaissance de formes, que de différenciation de zones dans une même image.

2.2 Approches manuelles

L'immense majorité des travaux s'intéressant spécifiquement à la détermination de volumes ou à la délimitation de tumeurs sur des jeux de données TEP utilisent des approches manuelles ou semi-automatiques basées sur des seuillages plus ou moins optimisés. Au lieu de tenter de faire faire le travail à une machine, on utilise un expert humain, en général à l'aide d'un logiciel de visualisation. Cet expert sélectionne la zone qu'il juge pertinente à l'aide de la souris ou du clavier, de façon manuelle ou semi-automatique (par exemple avec des régions d'intérêt de tailles et de formes arbitraires). Très majoritairement utilisée en routine clinique d'analyse diagnostique des images TEP, cette méthode est couramment utilisée par les médecins [91]. L'avantage significatif de la segmentation manuelle repose sur le fait que l'expert humain qui la réalise combine la capacité d'analyse intégrée du cerveau et ses connaissances de haut niveau d'ordre visuel, des informations *a priori* médicales (des connaissances générales en anatomie et en oncologie, mais aussi les informations contenues dans le dossier spécifique au patient dont il regarde l'image), et un entraînement et une expérience dans le déchiffrement de ce type spécifique d'images. Ceci lui permet de différencier ce qui relève de l'artefact ou du bruit de ce qui est d'intérêt comme une tumeur, ainsi que de faire facilement la différence entre une zone d'accumulation normale de glucose et une zone pathologique. Les inconvénients de cette approche sont toutefois rédhibitoires. Tout d'abord, elle est fortement subjective [91] et les résultats peuvent donc varier grandement d'un expert à l'autre (variabilité inter utilisateurs), ou pour un même expert d'une segmentation à l'autre (variabilité intra utilisateur) [157]. De plus dans certains cas, l'oeil humain peut être facilement trompé par l'aspect visuel d'une image, et le résultat peut donc dépendre de l'interprétation et de la visualisation des données (par exemple une image trop contrastée), là où un algorithme informatique travaillant sur les valeurs réelles des pixels ne sera pas influencé par cet aspect de visualisation de l'image. De plus, en TEP, les frontières des objets sont floues et difficiles à identifier, et même un humain peut dans certains cas avoir beaucoup de mal à séparer correctement un objet du fond. Enfin, la segmentation manuelle peut se révéler très laborieuse et longue, en particulier en trois dimensions, et si la tumeur est hétérogène et nécessite en plus la segmentation de régions

différentes en son sein. L'approche visuelle et manuelle d'un expert semble plus adaptée à la détection de tumeurs dans une image, qu'à la segmentation précise de l'objet d'intérêt ainsi détecté.

2.3 Approches semi-automatiques : seuillages

Le seuillage est sans doute la plus primaire et la plus simple des méthodes de segmentation. Elle consiste à regrouper les pixels d'une image suivant leur valeur en un nombre défini de classes. Aucune notion de corrélation spatiale n'est prise en compte à ce niveau. C'est à l'utilisateur de fixer un seuil (empirique, défini par une autre méthode, ou simplement arbitraire), puis on affecte une classe aux pixels dont la valeur est supérieure au seuil, et une autre classe aux pixels dont la valeur est inférieure à ce seuil. On peut définir ainsi plusieurs seuils si on désire diviser l'image en plus que deux classes. Le formalisme est donc très simple. Différents seuils peuvent être définis, par exemple la médiane (au lieu de la moyenne) des valeurs des voxels de l'image, en utilisant un histogramme (on définit alors souvent une ligne de partage des eaux en se basant sur la forme et la répartition du dit histogramme), etc.

L'avantage de cette technique est qu'elle est très simple et très rapide à mettre en oeuvre. Elle est de plus applicable directement aux images 3D. Malheureusement elle est très peu robuste face au bruit et à une inhomogénéité des valeurs dans l'image, ce qui est souvent le cas en tomographie d'émission. Elle n'utilise pas d'autre information que la valeur de chaque pixel, sans se soucier des relations spatiales pouvant exister entre eux. Le seuil doit en plus être défini par l'utilisateur *a priori* et il peut l'être de nombreuses façons, ce qui peut avoir un impact important sur le résultat. De nombreux travaux utilisant de simples seuillages ont été réalisés pour la segmentation de volumes en TEP mais ils font alors en général appel à des méthodes de croissance de régions (voir ci-après) pour incorporer une corrélation spatiale des voxels.

La combinaison de croissance de région et de seuillage permet d'incorporer une certaine corrélation spatiale des pixels dans la méthode de seuillage. L'idée est d'étendre le seuillage simple à une méthode de croissance de région, pour laquelle le critère de croissance utilisé

est celui de la valeur du pixel par rapport au seuil défini. Cette approche nécessite un ou plusieurs points d'initialisation (graine), une notion de connexité et un critère d'arrêt (ou de progression suivant le point de vue adopté). On commence par définir un seuil, puis une graine d'initialisation. On fait ensuite croître une région à partir de cette graine, en ajoutant les pixels voisins itérativement s'ils satisfont la condition, à savoir que leur valeur doit être supérieure ou égale au seuil pour être inclus dans la région. Comme précédemment, le seuil peut être défini de plusieurs façons. La croissance de la région peut elle aussi être définie de plusieurs façons suivant la connexité utilisée. Les deux connexités le plus souvent utilisées sont la 4- et la 8-connexité. La croissance de région s'étend directement au cas 3D (on utilise alors une 26-connexité et non une 8-connexité). Etant donné que la technique n'utilise que les valeurs des voxels et leur relation spatiale et qu'elle est totalement déterministe, elle est logiquement sensible au bruit, même si le semi automatisme (la graine est sélectionnée dans la lésion et le seuillage doit être fixé en fonction des informations *a priori*) peut en limiter les effets.

2.3.1 Seuillages fixes

Il s'agit d'une des approches les plus utilisées en TEP pour déterminer le volume, la majorité des études publiées sur la détermination des volumes en TEP faisant en effet appel à une approche fondée sur un seuillage. Toutefois, la façon de définir le seuil varie grandement d'une étude à l'autre. Certains travaux font état d'un seuil égal à un *standard uptake value* (SUV) de 2.5, d'autres utilisent un pourcentage (entre 40 et 75 %) de la valeur maximum au sein de la lésion. Ce voxel d'intensité maximal dans la lésion est en général choisi comme graine, mais malheureusement cette valeur est sujette à une forte variabilité à cause du bruit et en cas de forte hétérogénéité, se trouver séparé d'autres parties de la tumeur qui peuvent en conséquence ne pas être incorporées au volume segmenté. C'est pourquoi certaines études suggèrent l'utilisation d'une petite région d'intérêt au sein de la tumeur plutôt que la seule valeur maximale [37]. Malheureusement cela est difficile à rendre robuste dans le cas de tumeurs hétérogènes ou trop petites. Il est dorénavant accepté qu'utiliser un seuillage fixe est inapproprié [75]. La figure 2.2 illustre les variations obtenues sur une sphère homogène par un

seuillage fixe de 42 % du maximum dans la lésion.

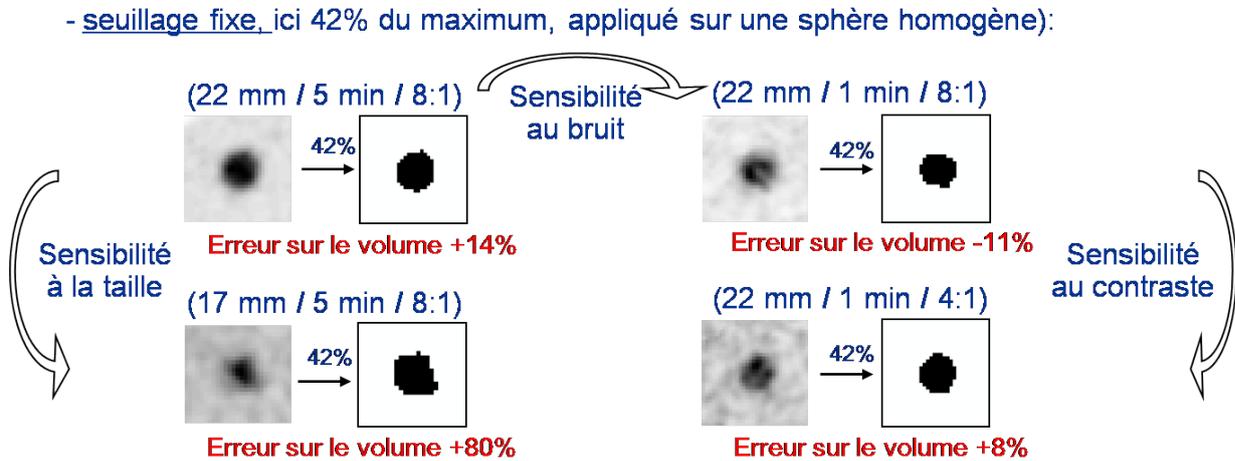


FIG. 2.2 – Illustration des résultats obtenus sur des sphères homogènes avec un seuillage fixe. La notation de chaque configuration est : taille (mm) / durée d’acquisition (min) / contraste sphère :fond. Ces résultats illustrent bien la faible robustesse de la méthode à des changements de taille, de contraste ou de bruit.

2.3.2 Seuillages adaptatifs

Les études les plus récentes font appel à des seuillages adaptatifs prenant en compte de façon variable la valeur du fond pour déterminer le seuil et dont les résultats sont nettement supérieurs aux seuillages fixes. L’approche la plus récente nommée RTL pour *Relative Threshold Level* [159] a généralisé les précédentes approches de seuillages adaptatifs basant leur décision sur le rapport entre l’activité de l’objet et celui du fond environnant [51] [38] [115] [37] [82]. Malheureusement, ces approches déterminent les seuils à utiliser en s’appuyant en règle générale sur des acquisitions de fantômes avec des sphères de taille variable et des ratios de sphères sur fond différents, pour en déduire la valeur de seuillage à appliquer pour un ratio et une taille de lésion donnée. Ces méthodes ont montré qu’elles étaient capables de déterminer des volumes avec une certaine précision pour des configurations de qualité bonne à moyenne et pour des volumes supérieurs à 4 ml, toutefois la plupart des études ne considèrent qu’un nombre de cas cliniques limités, ou des études sur fantômes avec des configurations

menant à des images de haute qualité, ou des volumes pour lesquels les EVP sont faibles. Deux approches utilisées en contexte cliniques sont présentées ici en détails :

Seuillage adaptatif proposé par Nestle

Le modèle proposé par Nestle dans [115], qui sera dorénavant noté $Tbckg$ dans cette thèse, propose le seuil suivant :

$$I_{seuil} = \alpha * I_{moyenne} + I_{fond} \quad (2.1)$$

I_{seuil} est la valeur absolue d'intensité en dessous de laquelle les voxels ne sont pas gardés dans le volume. $I_{moyenne}$ est la moyenne des voxels sélectionnés avec un seuillage fixé à 70% du maximum, et I_{fond} est la moyenne des voxels d'une région d'intérêt manuellement placée sur le fond, loin de l'objet à segmenter mais en restant dans le même organe (par exemple le poumon), ce qui peut parfois être compliqué par la position et la taille de la tumeur par rapport à l'organe de référence.

Cette méthode soulève plusieurs problèmes : tout d'abord, le placement manuel de la région d'intérêt sur le fond la rend utilisateur-dépendante, en particulier si le fond n'est pas homogène. Ensuite, l'utilisation d'un seuillage à 70% du max pour $I_{moyenne}$ peut donner des résultats aberrants avec certaines tumeurs hétérogènes, de diamètre inférieur à 2 cm (pour lesquelles les EVP sont très importants) ou dont le contraste avec le fond est faible. Enfin, le paramètre α doit être optimisé pour chaque modèle de scanner, chaque variation des paramètres d'acquisition ou de reconstruction, ce qui rend son utilisation compliquée, notamment dans le cadre d'études multi-centriques.

Seuillage adaptatif proposé par Daisne

Le modèle proposé par Daisne dans [37], qui sera dorénavant noté $TSBR$ dans cette thèse, propose le seuil suivant :

$$seuil = a + b * \frac{1}{SBR} \quad (2.2)$$

seuil est le pourcentage d'intensité maximale de la lésion définissant les voxels gardés dans le volume. *SBR* est obtenu par le rapport entre la moyenne des 8 voxels voisins de celui possédant l'intensité maximale au sein de l'objet, et la moyenne d'une région d'intérêt manuellement placée sur le fond, loin de l'objet à segmenter.

Cette méthode souffre des mêmes inconvénients que la précédente : tout d'abord, le placement manuel de la région d'intérêt sur le fond la rend utilisateur-dépendante, comme dans le cadre de *Tbckg*. Ensuite, la façon dont est calculé *SBR* peut donner des résultats aberrants avec certaines tumeurs hétérogènes, pour lesquelles la valeur moyenne associée à la tumeur serait réduite à la zone d'intensité maximum de la lésion, empêchant la prise en compte des régions de la tumeur d'intensité inférieure. Enfin, les paramètres *a* et *b* doivent être optimisés pour chaque modèle de scanner, chaque variation des paramètres d'acquisition ou de reconstruction, ce qui rend son utilisation également compliquée.

La plupart de ces techniques sont donc fortement dépendantes du système sur lequel les études sur fantômes ont été menées et leurs paramètres doivent être optimisés pour chaque configuration et protocole d'acquisition, du modèle de scanner utilisé à la durée d'acquisition, en passant par les paramètres de l'algorithme de reconstruction (nombre d'itérations, nombre de subsets, post-filtrage, etc.), ce qui en complique fortement l'usage généralisé dans un même centre d'examens, et *a fortiori* dans différents centres n'utilisant pas forcément les mêmes protocoles et équipements.

Seuillage adaptatif *relative threshold level*

L'approche RTL [159] ne nécessite pas l'optimisation de paramètres au préalable comme *Tbckg* [115] et *TSBR* [37], mais elle nécessite par contre la connaissance exacte de la résolution spatiale du scanner (et fait également l'hypothèse qu'elle est uniforme dans le champ de vue), ainsi que l'identification manuelle des tissus du fond et de l'objet d'intérêt. Elle ne résout donc qu'en partie les critiques formulées à l'égard des deux précédentes approches et fonctionne globalement sur le même principe.

Seuillage adaptatif raffiné par contour adaptatif

Une approche récente a suggéré un seuillage basé sur une croissance de région, suivie d'un raffinement de la segmentation à l'aide d'un contour adaptatif [99]. Elle a le mérite de présenter une façon originale de définir le seuillage initiale. Cette approche propose en effet de définir un premier volume grossier en utilisant une croissance de région et à utiliser la courbe de l'évolution du volume de l'objet en fonction du seuil, que l'on fait varier. Lorsque la courbe a une inflexion forte, on fixe le volume, et on raffine ensuite la segmentation à l'aide d'un contour adaptatif (*dual front active contour*) initialisé avec ce volume. La figure 2.3 illustre la définition initiale du volume et les résultats obtenus sur une tumeur clinique, qui a malheureusement été choisie homogène et assez facile à segmenter. Il est donc difficile d'évaluer la pertinence de la méthode pour segmenter des tumeurs plus grande et plus complexes. Il est toutefois probable qu'étant donné que l'initialisation du contour adaptatif est très importante d'après les auteurs, si cette initialisation est effectuée à l'aide du volume initial qui est déterminé par un seuillage, alors dans le cas de tumeurs complexes en forme et / ou en distribution d'activité, cette méthode risque d'échouer.

Mise en évidence des limites des seuillages adaptatifs

Globalement, ces approches fournissent donc de bons résultats sur les objets approximativement sphériques et à la distribution d'activité homogène, ce qui est normal puisqu'elles sont optimisées sur ce type de données (en utilisant des images de fantôme contenant des sphères homogènes sur fond homogène). Elles sont malheureusement prises en défaut lors de l'application sur des objets plus complexes, dont l'activité est hétérogène, comme illustré dans la figure 2.4.

De plus, la plupart des études ayant permis la mise au point de ces seuillages adaptatifs prennent en compte les variations de taille d'objet et de contraste, mais négligent souvent la variabilité due au bruit (c'est à dire des durées d'acquisition variables). Une étude récente [118] a d'ailleurs montré que les seuils "idéaux" peuvent varier de façon significative entre différents centres d'examen utilisant les mêmes modèles de scanner, ne serait-ce qu'à cause de proto-

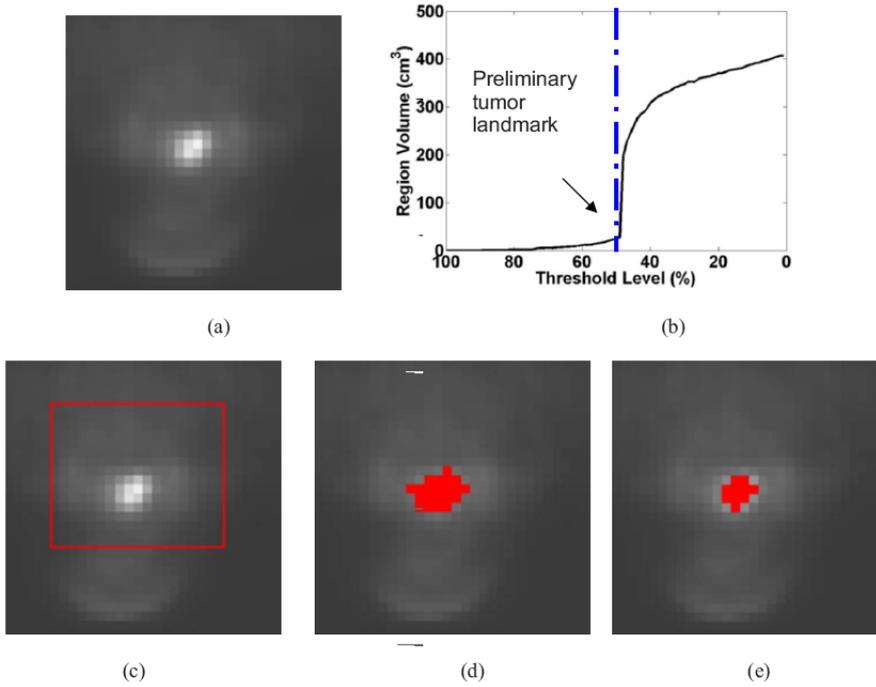


FIG. 7. The segmentation result on a clinical PET head-and-neck tumor image. (a) One slice of the original 3D PET image, (b) plot of the grown region volume (in voxel) as a function of the threshold level T , (c) the manual drawn ROI, (d) the initial tumor volume defined at a threshold level of 49%, just before the sharp volume increase, (e) the final result after running the dual-front active contour model.

FIG. 2.3 – Figure issue de [99]. La tumeur utilisée pour tester l’approche sur un cas clinique a été choisie suffisamment grande et surtout avec une activité totalement homogène.

coles d’acquisition et de reconstruction différents, ce qui confirme les critiques portées à la faible robustesse des méthodes basées sur des seuillages. Enfin, si la validation des techniques proposées dans ces travaux est effectuée de façon sérieuse sur des fantômes, elles sont souvent uniquement appliqués à un nombre réduit de cas cliniques (ou de simulations réalistes) pour explorer leur comportement dans de telles conditions, cas qui sont souvent choisis pour leur simplicités, avec des tumeurs de taille modeste et à la distribution d’activité souvent très homogène. C’est par exemple le cas dans [99], où une simple tumeur O.R.L. est choisie pour illustrer le fonctionnement de la méthode (voir figure 2.3), et l’approche n’est donc pas du tout testé sur des tumeurs plus complexes à segmenter, notamment en terme d’hétérogénéité de leurs distributions d’activité.

Une étude récente [161] a comparé les résultats de plusieurs seuillages proposés dans diverses publications sur quatre patients avec des tumeurs pulmonaires. L’intérêt de cette étude est de disposer de données d’histologie, et donc de pouvoir confronter les résultats en termes de

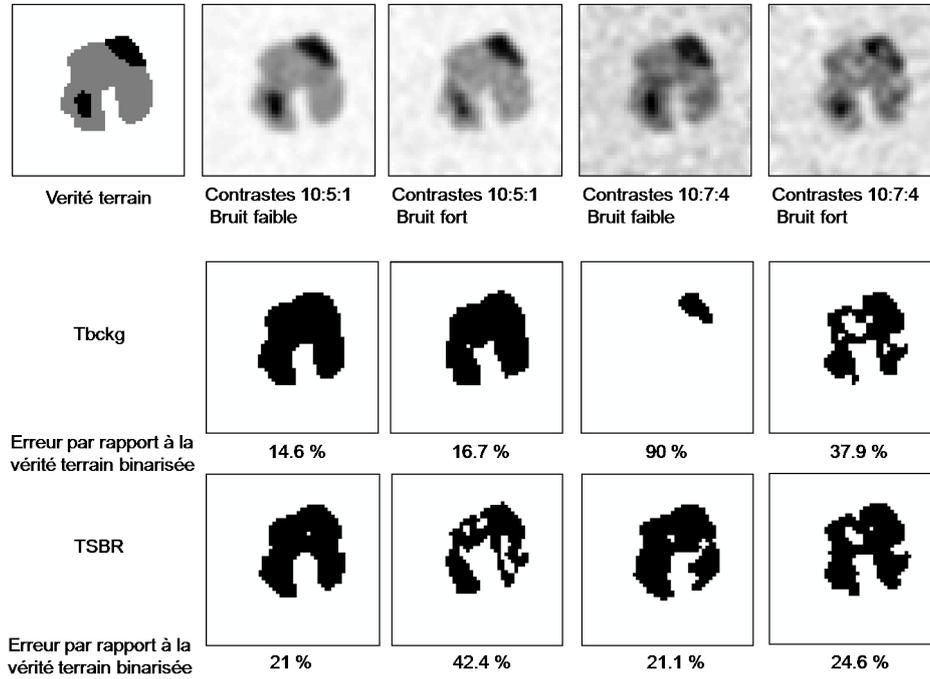


FIG. 2.4 – Illustration des résultats obtenus sur un objet hétérogène avec les deux seuillages adaptatifs décrits précédemment. Ces résultats montrent qu’une approche binaire, même adaptative par rapport au fond, est insuffisante pour rendre compte correctement de la complexité de la tumeur à segmenter, et que les résultats peuvent varier de façon très significative suivant les configurations de bruit et de contrastes.

volume, obtenus par les méthodes, à ceux mesurés en histologie. Malheureusement les tumeurs sont là aussi très homogènes, et ne présentent donc pas beaucoup de difficultés à une méthode de segmentation, bien que l’étude montre des différences de performance significatives entre les méthodes comparées.

2.4 Approches automatiques

2.4.1 Clustering

Il s’agit ici de regrouper les pixels de l’image en les classifiant par rapport à un critère donné. On peut citer les très connus K-Moyennes (K-means) [106], et sa version floue, le Fuzzy C-

Means (FCM) [49] [8]. Dans le cas du K-Means, l'appartenance à chacune des classes est déterminée par la moyenne des caractéristiques des éléments, ce qui permet de calculer une nouvelle moyenne, et ceci est répété jusqu'à stabilisation. Dans le cas du FCM, on attribue à chaque pixel un degré d'appartenance à un centroïde d'une classe suivant son intensité et la distance qui en découle par rapport à l'intensité du centroïde de la classe. Le FCM nous intéresse plus spécifiquement à cause de la notion de flou.

L'initialisation nécessite le nombre de clusters recherchés dans l'image. Ensuite, les centroïdes de chaque classe sont déterminés, par exemple en utilisant une sélection itérative divisant l'image par seuillage, ou alors au hasard (comme pour le K-Means classique) dans l'algorithme standard du FCM. Ensuite, on effectue pour chaque pixel un calcul de distance par rapport au centroïde de leur classe. Dans le cas de la segmentation d'image, cette distance est calculée non pas sur la position du pixel, mais sur son intensité par rapport à celles des pixels centroïdes. Ainsi, chaque pixel se voit attribuer une appartenance (entre 0 et 1) pour chacun des centroïdes, et on attribue à chaque pixel la classe dont la distance avec son centroïde est la plus faible. Pour ce faire, on fait appel à une méthode itérative dont le but est de minimiser la fonction objectif suivante :

$$J_f = \sum_{j=1}^c \sum_{i=1}^p [u_j(x_i)]^b \|x_i - m_j\| \quad (2.3)$$

Avec c le nombre de clusters, p le nombre de pixels, m_j le centroïde du cluster j , et b un paramètre de flou (1 signifiant pas de flou et 2 signifiant un flou maximum, ce paramètre étant généralement fixé à 2) et avec la condition suivante :

$$\sum_{j=1}^c u_j(x_i) = 1 \quad (2.4)$$

Il s'agit donc de calculer les valeurs de m_j et de $u_j(x_i)$ pour tout i qui minimisent J_f . Ceci est réalisé à chaque itération grâce aux formules de ré-estimation suivantes :

$$m_j = \frac{\sum_{i=1}^p [u_j(x_i)]^b x_i}{\sum_{i=1}^p [u_j(x_i)]^b} \quad (2.5)$$

$$u_j(x_i) = \frac{\frac{1}{\|x_i - m_j\|^2} \frac{1}{b-1}}{\sum_{j=1}^c \frac{1}{\|x_i - m_j\|^2} \frac{1}{b-1}} \quad (2.6)$$

FCM a été utilisé pour la segmentation de tumeurs cérébrales dans [171] de façon itérative pour résoudre le problème des sur-segmentations de régions dues au bruit. Les résultats qualitatifs présentés dans cette étude ne permettent pas de réellement apprécier les performances de l'algorithme concernant la détermination des volumes tumoraux. Ce dernier, non modifié, a aussi été appliqué en TEP, en particulier dans [170] pour développer une correction d'atténuation d'images corps entier fondée sur les images TEP segmentées, une toute autre application que celle de la détermination précise de volumes d'intérêt. Ces travaux ne présentent pas de résultats concrets de détermination de volume ou de quantification, mais souvent uniquement des appréciations qualitatives.

Un inconvénient de la méthode est qu'elle ne prend pas en compte la relation spatiale entre les voxels. Il est toutefois possible de modifier la fonction objectif J_f pour incorporer une distance spatiale, de façon à prendre en compte la corrélation spatiale entre voxels, par exemple dans [63] qui utilise un parcours de Peano. La nature du bruit n'est pas prise en compte et on peut craindre une relative variabilité des résultats face à des images plus ou moins bruitées. Toutefois la notion d'appartenance floue aux centroïdes permet une certaine modélisation du flou existant au niveau des frontières des objets d'intérêt.

Une autre modélisation faisant appel à la logique floue a été utilisée par Huglo dans sa thèse [74], et évaluée sur un ensemble de données significatif, sur fantômes et sur dix patients, dont les tumeurs ont été étudiées en histologie pour connaître leur véritable volume. La méthode proposée a montré des résultats pertinents et supérieurs à ceux des seuillages, avec notamment une dispersion plus faible en fonction des variations du bruit de fond et des contrastes, renforçant notre intuition que l'utilisation d'une modélisation floue est la voie à suivre dans le cadre de la segmentation des images d'émission.

2.4.2 Watersheds

Le principe est de faire une analogie entre l'image et ses niveaux de gris, et un terrain géographique et ses altitudes. On transforme le gradient d'une image en surface topologique et on remplit alors progressivement ce terrain avec de l'eau. La ligne au niveau de laquelle l'eau passe d'une région à l'autre permet d'en déterminer la frontière. Il existe deux algorithmes : "top-down" et "bottom-up". Ce dernier a tendance à réduire la précision de l'algorithme et "top-down" est celui généralement utilisé. L'analogie avec l'eau sur une surface topologique réside dans l'utilisation d'une analyse des points faibles le long des frontières des régions, détectées par une descente de gradient. L'image est d'abord filtrée par un filtre lissant préservant les contours pour obtenir une "edge map" ou carte des contours. Un premier travail utilise un filtrage préalable, suivi d'un watershed itératif pour segmenter des images de transmissions de courtes acquisitions et très bruitées [135]. Le principal inconvénient de cette méthode est qu'elle génère en général une sur-segmentation, et qu'elle nécessite donc une étape supplémentaire de post-traitement pour fusionner les régions suivant un critère à définir. L'inconvénient de cette sur-segmentation peut se transformer en intérêt dans la mesure où la hiérarchie de régions ainsi produite offre une certaine flexibilité, et les régions obtenues peuvent être fusionnées ou isolées, mais cela nécessite toutefois l'intervention d'un utilisateur, ou la mise au point d'un critère défini suivant des informations *a priori*. La sur-segmentation peut être résolue en fixant une profondeur minimale par défaut, mais cette profondeur est aussi fixée arbitrairement par l'utilisateur. Le fait que cette méthode doit être appliquée sur des données filtrées au préalable, de préférence avec un filtre lissant préservant les contours induit un résultat dépendant fortement du filtrage utilisé. Les performances dépendent en grande partie des nombreux paramètres à ajuster (pré- et post-traitement(s), filtrage par rapport au bruit, fusion de régions, etc.). Dans [156], une approche fondée sur les watersheds mais modifiée car les résultats en appliquant directement la segmentation watersheds étaient insuffisants, a été évaluée précisément en utilisant des données de fantôme TEP. Cette étude rapporte des performances similaires à celles d'un seuillage adaptatif tout en réduisant la dépendance à l'opérateur et les calibrations nécessaires. Par contre, la méthode est décrite comme souffrant

d'une variabilité dans la détermination des volumes et les résultats sur les sphères inférieures à 4 mL ne sont pas présentés, ce qui laisse penser que la méthode ne parvient pas à les segmenter avec précision.

2.4.3 Filtrages et détection de contours

Une approche contour, fondée sur le *Gradient Vector Flow* (GVF) a été utilisée pour segmenter des images TEP de foie avec succès dans [102]. Mais le foie est un objet largement plus grand que les tumeurs que l'on se propose de traiter dans notre travail, et les effets de volume partiel sont sans doute négligeables par rapport à la taille de l'objet à segmenter. Toutefois le contraste est relativement faible, et la méthode réussit tout de même à produire un contour satisfaisant du foie. La méthode paraît mal adaptée à la segmentation d'objets de petite taille pour lesquels les effets de volume partiel sont nettement plus significatifs, ou pour des images plus bruitées que celles considérées dans cette étude.

Une approche également fondée sur l'utilisation d'un gradient a été utilisée pour une étude dont la validation est complète et sérieuse, sur des images de fantôme ainsi qu'avec une base de données de tumeurs réelles avec résultats d'histologie [58]. Toutefois, pour fonctionner correctement sur les images TEP, la méthode nécessite l'application de deux étapes de pré traitement qui sont sujettes à caution : un débruitage (pour supprimer le bruit) et une déconvolution (pour supprimer le flou). Nous ne souhaitons pas avoir recours à de tels traitements dans notre méthodologie, pour réduire la variabilité due au choix des méthodes de débruitage et de déconvolution, qui sont au moins aussi nombreuses que les approches de segmentation elles-mêmes, sans parler du problème de perte éventuelle d'informations pertinentes lors du débruitage, ou l'ajout de bruit lors de la déconvolution, deux problèmes que les auteurs de cette étude eux mêmes soulèvent dans la discussion.

2.4.4 Modèles déformables

L'idée repose sur l'utilisation d'une forme (une courbe, une surface, un volume) *a priori* et initialisée suffisamment proche de la forme désirée, puis de la faire évoluer itérativement vers le résultat le plus proche possible de l'objet recherché dans l'image. Cette évolution se fait

en général grâce à des contraintes *a priori* (sur la forme) combinées à des contraintes basées sur le contenu de l'image. Les modèles déformables forment une famille de plusieurs méthodes plus ou moins complexes. On peut citer le modèle des contours actifs « snakes » ou les fronts de propagation « Level Set », une méthode plus évoluée.

Un « snake » est une courbe qui se déplace sous l'influence de forces internes et externes, et nécessite des informations *a priori* sur la forme de l'objet à segmenter. Les forces internes dépendent des caractéristiques du contour, tandis que les forces externes dépendent de l'image. L'influence de ces forces est étudiée de façon à atteindre un équilibre en minimisant une position d'énergie. Il existe aussi des forces externes introduites par l'utilisateur pour guider le « snake » vers la structure d'intérêt. Dans le cadre de la segmentation d'images, le potentiel de l'image est exprimée comme le négatif de la norme du gradient de l'image, car l'objectif est que le « snake » soit attiré vers les contours. Pour plus de détails sur le formalisme mathématique et les différentes implémentations des « snakes », voir [142]. Les « snakes » ont plusieurs inconvénients : d'abord, les résultats sont très sensibles à l'initialisation. Ceci peut toutefois être réglé par l'introduction d'une nouvelle force de pression, mais le « snake » devient alors instable. Un autre problème est l'impossibilité du « snake » à changer de topologie. Enfin, les divers coefficients de courbure, d'élasticité, et les poids appliqués aux forces externes doivent être choisis de manière empirique, suivant le type d'image à segmenter et les informations *a priori* de l'objet à segmenter. Cela nécessite de très bien connaître le modèle et d'avoir une bonne expérience dans ce domaine.

Un modèle plus évolué est le front de propagation "Level Set". Ce modèle ne nécessite pas une connaissance *a priori* de la topologie de l'objet d'intérêt. Il s'agit d'un modèle de propagation d'une interface fermée. La vitesse de propagation dépend de plusieurs paramètres, qui forment un modèle qui doit être adéquat et à choisir. Le front croît dans l'image en s'adaptant et en se collant aux parois de l'objet. Le principe du "Level Set" est de considérer un front de propagation comme "une courbe lisse, continue et fermée, définie dans le plan et engendrant une famille de courbes obtenues par le mouvement de la courbe suivant sa normale à une vitesse de propagation, qui est fonction de sa courbure" [142]. Cette courbe peut être généralisée à une interface 2D ou 3D. On définit alors l'interface à l'aide d'une fonction qui représente toutes les

courbes de niveau de l'image. Elle est donc fonction de la distance entre tout point de l'image et l'interface. Pour une définition mathématique et une implémentation détaillée, voir [142]. Une étude s'est penchée sur l'utilisation de modèles déformable pour déterminer un volume en TEP [167] et présente des résultats concernant la précision avec laquelle le volume est calculé. Il fait appel à une courbe qui évolue grâce à une fonction de vitesse (dont la définition n'est pas donnée par le compte rendu). Une étude plus récente a utilisé une méthode fondée sur les "Level Set" pour une approche de la segmentation dédiée à la radiothérapie et a présenté de bonnes performances sur fantôme et des résultats préliminaires sur données cliniques. Un intérêt de l'approche est de prendre en compte simultanément les données TDM et TEP pour influencer l'évolution du "Level Set" [50]. D'autres travaux utilisent des modèles déformables en TEP, mais sont essentiellement utilisés pour déterminer des volumes 3D d'objets tels que le cerveau [153], et non pas des tumeurs millimétriques.

Qu'il s'agisse des "snakes" ou des "Level Set", un prétraitement de filtrage lissant et préservant les contours est nécessaire. L'automatisme de la méthode semble un peu limité dans la mesure où les différents éléments du modèle, surtout en ce qui concerne le terme de vitesse peuvent être définis de plusieurs façons. L'implémentation est donc délicate pour éviter des temps de calcul trop longs.

2.4.5 Inférence bayésienne et segmentation statistique d'image

Ce champ de méthodes est vaste et est fondé sur la formule de Bayes et l'utilisation de modèles probabilistes et statistiques pour classifier les éléments (les voxels par exemple) par rapport à des classes dont on estime les paramètres telles que la moyenne et la variance. On utilise en général un modèle *a priori* combiné à une estimation des paramètres de l'observation, pour contraindre la solution, souvent sous forme de modèles de Markov cachés tels que des champs, des chaînes ou des arbres. Les champs sont le plus souvent utilisés dans le contexte de la segmentation d'images grâce à leur modélisation intuitive de la corrélation spatiale, ainsi que les chaînes grâce au parcours d'Hilbert-Peano car ces dernières accélèrent grandement les temps de calcul par rapport aux champs. Les arbres trouvent des applications dans la segmentation multidimensionnelle car leur structure s'y prête naturellement. Plus de détails

sur l'inférence bayésienne et les modèles de Markov sont donnés dans la section 3.1.1 du chapitre 3.

Ce champ de méthodes a peu été considéré dans le cadre de l'imagerie d'émission, en particulier pour la segmentation de volumes d'intérêt tumoraux. Les rares études existantes utilisent des champs de Markov durs sans modélisation floue ce qui est peu pertinent dans le cadre de l'imagerie d'émission. Une première étude [29] concerne la segmentation d'images paramétriques de cerveau. Une autre étude [41] s'est intéressée à la segmentation de tumeurs dans des images corps entier mais les résultats semblent trop approximatifs pour finement délimiter les volumes tumoraux, comme illustré dans la figure 2.5. Les résultats sont trop grossiers pour la détermination exacte des volumes, bien que pour faire ressortir une accumulation pathologique, cela pourrait peut être suffire. Il est important de noter que le texte de l'article semble suggérer qu'au moins une partie de la segmentation a été initialisée ou ajustée à la main pour un résultat satisfaisant, ce qui était également le cas dans l'étude sur la segmentation d'images paramétriques du cerveau [29].

Une étude plus récente a été proposée, qui combine champs de Markov avec décompositions en ondelettes [111]. Malheureusement la validation de l'approche est réalisée avec des volumes homogènes dont les contrastes sont largement supérieurs à ceux observés en routine clinique. En effet, bien que les tailles des lésions utilisées soient pertinentes (entre 1.4 et 3.1 mL), le rapport d'activité entre les poumons et les lésions est dans cette étude de 40 pour 1 (20 :1 entre les lésions et les autres tissus mous), ce qui limite grandement la validité des résultats présentés en ce qui concerne leur robustesse par rapport aux problèmes de contraste et d'effets de volume partiel, ou même de bruit. De plus, l'approche ne propose que l'utilisation de champ de Markov durs, c'est à dire sans modélisation floue, et est particulièrement longue (30 minutes, pourtant implémenté en langage C et tournant sur une machine actuelle).

2.5 Conclusion

La conclusion de cet état de l'art de la segmentation d'images en TEP peut se résumer en deux points : premièrement, la plupart des travaux étudiant spécifiquement la détermination

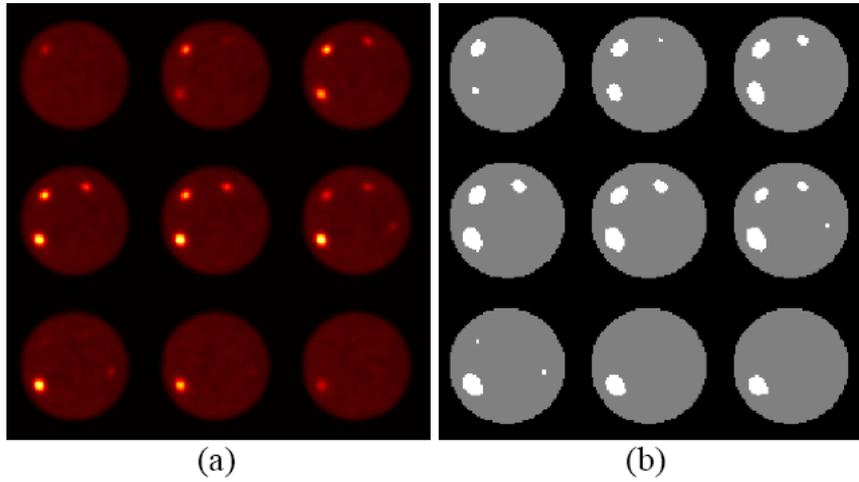


Fig. 2. (a) Transaxial slices of the hot-lesion phantom reconstructed with the OSEM. (b) After final segmentation.

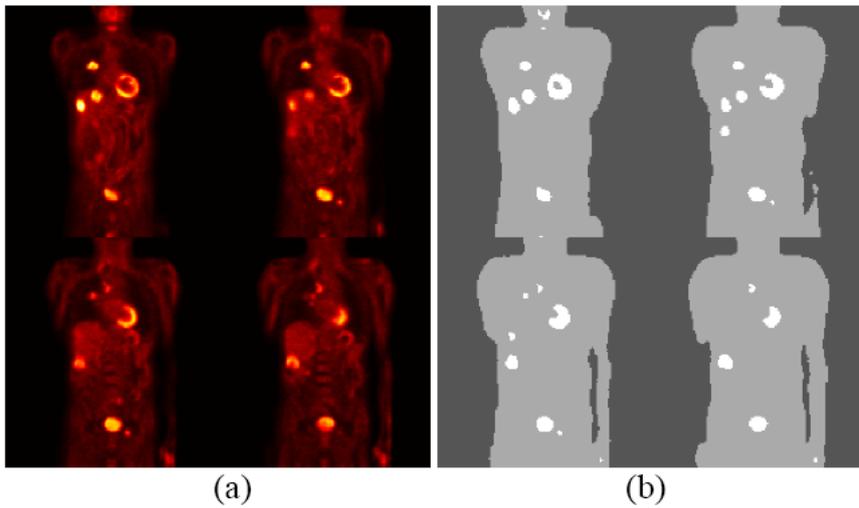


Fig. 3. (a) Coronal slices from a whole-body image. (b) Final segmented image shows the lesions, bladder and the heart separated from the remaining body tissue.

FIG. 2.5 – Figure issue de [41]. La segmentation est faite sur l'image corps entier et la segmentation des zones marquées est clairement trop grossière. Cela est particulièrement flagrant sur les images de fantôme.

des volumes en imagerie TEP utilisent des méthodes simples semi-automatiques à base de seuillages, fixes ou plus récemment adaptatifs. Ces dernières ne seront jamais capables de gérer correctement la variabilité du bruit ou du contraste que l'on constate dans les images

TEP. De plus, même validée sur fantômes avec sphères homogènes, rien ne peut assurer que ces méthodes pourront gérer des objets non sphériques et hétérogènes, comme le sont la plupart des tumeurs. Enfin, ces méthodes sont souvent système- et utilisateur- dépendants. Deuxièmement, une minorité de travaux ont exploré l'utilisation de méthodes de segmentation automatiques plus avancées mais le plus souvent pour des applications différentes (détection de lésion, filtrage, etc.) et non pour une détermination précise d'un volume d'intérêt. D'autre part, le champ des méthodes de segmentation statistique, pourtant prometteur et obtenant de bons résultats sur d'autres types d'imagerie, n'a été que très peu exploré pour la TEP. C'est ce que nous nous proposons d'examiner. Les chapitres suivants vont maintenant détailler les approches que nous avons considérées et adaptées à notre problématique.

Chapitre 3

Méthodologies développées

3.1 Approche choisie : combinaison du modèle statistique et de la mesure floue

3.1.1 Inférence bayésienne et segmentation statistique

Nous allons détailler ici le contexte dans lequel les méthodes de segmentation développées dans cette thèse se situent : celui de la segmentation statistique d'image reposant sur une inférence bayésienne. Dans ce cadre, les éléments de l'image (pixels ou voxels) sont considérés comme des réalisations de variables aléatoires. Pour les applications de segmentation statistique, on définit en général deux variables aléatoires, X et Y . Ces dernières représentent respectivement un ensemble d'étiquettes discrètes associées à chaque élément de l'image, c'est à dire la classe que l'on souhaite affecter à un voxel donné, et un ensemble de valeurs réelles correspondant aux observations liées à chaque élément, c'est à dire les valeurs de chaque voxel. X prend donc ses valeurs dans un ensemble C de valeurs discrètes (en général le nombre de classes recherchées dans l'image) et Y prend ses valeurs dans \mathfrak{R} .

L'objectif de la segmentation statistique d'image est de restituer le champ caché des étiquettes X à partir d'une probabilité *a priori* sur ce dernier et du champ des observations associées Y . On définit en général une loi sur le couple (X, Y) ce qui permet ensuite d'obtenir la loi de

X conditionnellement à Y grâce à la loi de Bayes :

$$P(X, Y) = P(Y|X)P(X) \tag{3.1}$$

Dans cette expression, $P(X)$ désigne la loi *a priori* sur les étiquettes, qui est généralement un modèle spatial, tandis que $P(Y|X)$ désigne la loi des observations conditionnellement à celle des étiquettes. Cette loi des observations s'appelle également le modèle de bruit et fait le plus souvent appel à l'utilisation de lois gaussiennes, mais cela dépend du domaine d'application. Il est possible d'obtenir la loi recherchée des étiquettes conditionnellement aux observations $P(X|Y)$:

$$P(X|Y) = \frac{P(X, Y)}{P(Y)} = \frac{P(Y|X)P(X)}{P(Y)} \tag{3.2}$$

Modèles

Il existe de très nombreux modèles pour définir ces lois [127]. Les plus classiques sont les champs et les chaînes de Markov, qui ont depuis été étendues aux arbres, ainsi qu'aux modèles couples et triplets, qui généralisent les modèles de base tout en permettant de modéliser des lois et des données plus complexes telles que les aspects non stationnaires, les bruits corrélés, les bruits à mémoire longue, etc.

Il existe trois algorithmes majeurs d'estimation des paramètres qui ont été développés pour les modèles cités ci-dessus. L'algorithme le plus ancien est l'EM (pour Expectation Maximization) [42]. Il est entièrement déterministe. Il existe une version stochastique d'EM, le SEM (pour Stochastic Expectation Maximization) [28]. Enfin, un algorithme hybride, dont une partie est stochastique et l'autre déterministe, l'ICE (pour Iterative Conditional Estimation) a également été proposé [124] [125].

L'implémentation pratique de chaque algorithme dépend du modèle sur lequel il est appliqué, et doit donc être détaillé pour chaque modèle et application. Nous présentons dans la suite quelques généralités sur les algorithmes d'estimation, et pour plus de détails, voir les sections correspondantes des sections 3.2 et 3.3.

Algorithme EM

L'algorithme EM repose sur l'idée de définir une suite de paramètres θ^n telle que la suite des probabilités $P_{\theta^n}(y)$ est croissante. On se donne un paramètre initial θ^0 , et θ^{q+1} est obtenu à partir de θ^q et y par :

$$\theta^{q+1} = \operatorname{argmax}_{\theta} E_{\theta^q} [\operatorname{Log}(P_{\theta}(X, Y) | Y = y)] \quad (3.3)$$

EM fonctionne pour un grand nombre de classes ainsi que dans le cas multidimensionnel. Il fournit de plus de bons résultats en général, si le nombre de classes est connu. Les limites de cet algorithme telles qu'identifiées dans [28] sont les suivantes. Il est déterministe et fondé sur le maximum de vraisemblance, ce qui signifie que le résultat après convergence dépend de l'initialisation des paramètres, et comme la fonction de vraisemblance n'est pas bornée, il se peut que l'algorithme reste bloqué dans un minimum local. De plus, il converge lentement, en particulier si les valeurs d'initialisation sont éloignées de la solution. Toutefois, il existe une norme sur l'espace des paramètres pour laquelle la convergence peut se faire correctement si l'initialisation est suffisamment proche de la solution [28].

Algorithme SEM

Comme l'indique son nom, le SEM [28] est une version stochastique d'EM [42]. Contrairement à l'EM, l'estimation des paramètres se fait de façon empirique par comptage sur une (ou une moyenne de plusieurs) réalisation(s) aléatoire(s) obtenue(s) par tirage(s) suivant la loi *a posteriori*. D'après [28], les estimations sont précises, et les perturbations introduites à chaque itération par les tirages aléatoires permettent de s'affranchir en partie du conditionnement des valeurs des paramètres à l'initialisation, et d'éviter la convergence vers un maximum local instable de la vraisemblance. De plus, la convergence est plus rapide qu'avec EM, car les tirages aléatoires l'empêche de stationner trop longtemps loin de la solution.

Algorithme ICE

L'algorithme ICE [125] a été proposé comme une version hybride du SEM [28] et de l'EM [42]. Dans cet algorithme, l'estimation des paramètres du modèle *a priori* se fait comme dans le

cas du EM, de façon déterministe, avec les mêmes formules de ré-estimation. A contrario, l'estimation des paramètres du bruit se fait comme dans le SEM, de façon stochastique sur une réalisation *a posteriori*.

Le déroulement de ICE est le suivant : on se donne un paramètre initial θ^0 , et θ^{q+1} est obtenu à partir de θ^q et y par :

$$\theta^{q+1} = E_{\theta^q}[\theta(X, Y) | Y = y] \quad (3.4)$$

Lorsque cette espérance est calculable, et par :

$$\theta^{q+1} = \frac{\theta(x^1, y) + \dots + \theta(x^s, y)}{s} \quad (3.5)$$

avec $x^1 \dots x^s$ simulés selon $p_{\theta^q}(x|y)$ lorsqu'elle ne l'est pas.

3.1.2 Travail sur l'image

Une remarque importante concernant l'approche envisagée dans ce travail : il s'agit de segmentation d'images, et nous travaillons donc sur les images après reconstruction, et non sur les données brutes (données list-mode ou sinogrames). Une autre approche pourrait consister à travailler sur la reconstruction d'image pour y intégrer des contraintes et des méthodes destinées à améliorer la résolution, mais une étape de segmentation d'image est toujours nécessaire, en particulier pour les applications que nous visons à automatiser, telles que la détermination des régions d'intérêt pour la radiothérapie (notions de "*dose sculpting*" ou "*dose painting*"). Cette segmentation peut être rendue plus facile par une reconstruction améliorée, qui réduit le bruit ou améliore la résolution spatiale mais ce n'est pas l'objet de ce travail. Il est important de comprendre que tout ce que nous proposons dépend fortement des étapes situées en amont, à savoir l'acquisition des données et leur reconstruction. Nous cherchons à obtenir une approche à la fois précise et robuste, qui puisse fonctionner de façon satisfaisante quel que soit l'algorithme de reconstruction choisi, ou le scanner utilisé pour procéder à l'acquisition des données. Toutefois, nous restreignons notre champ de validation à des reconstructions validées et utilisées en routine clinique. Nous n'envisageons par exemple pas de valider notre approche sur des images reconstruites par rétro-projection filtrée car cette méthode a maintenant été écartée de la routine clinique au profit des méthodes itératives, largement majoritaires dans

les machines commercialisées. Malgré tout, au sein de ces algorithmes itératifs, suivant les paramètres choisis pour un algorithme, ou entre deux algorithmes différents, il existe des différences significatives sur les images reconstruites, à la fois en terme d'aspect général, de bruit ou de contraste.

3.1.3 Segmentation et détectabilité

La détectabilité des tumeurs est un domaine qui fait l'objet de recherches spécifiques [154] [62]. La détectabilité d'une tumeur est en particulier affectée par la taille et le contraste de la tumeur. En effet, un petit nodule, de l'ordre du centimètre, sera facilement repérable si son activité est très élevée, donc avec un fort contraste par rapport aux tissus sains environnants. Dans le cas contraire, à cause des effets de volume partiel et du bruit, l'objet sera difficile à voir dans l'image. De même, une tumeur de grande taille mais affectée d'une fixation de faible intensité, peut être difficilement repérable à cause d'un contraste trop faible.

Dans notre contexte, nous ne cherchons pas à automatiser ou améliorer la détection des tumeurs, c'est à dire l'opération qui consiste à repérer, dans une image TEP corps entier, une anomalie ou un objet d'intérêt spécifique. Nous faisons l'hypothèse que l'objet d'intérêt à segmenter a déjà été repéré et isolé dans une boîte de dimension raisonnable, et c'est dans cette sélection que nous cherchons à délimiter avec la plus grande précision possible l'objet s'y trouvant. Cette boîte doit être suffisamment grande pour inclure l'objet et une partie suffisante de son environnement immédiat, mais suffisamment petite pour permettre un traitement rapide et éviter l'inclusion de structures à haute intensité non pathologique, comme le coeur par exemple, pour éviter d'avoir à utiliser des post-traitements, la nécessité d'informations *a priori* ou l'intervention de l'utilisateur. Notons que dans le cas des chaînes de Markov, cette boîte doit avoir des dimensions particulières : elle doit être cubique et chaque arête doit avoir une dimension en puissance de 2, de façon à pouvoir simplement transformer le volume tri dimensionnel représenté par la boîte en chaîne uni dimensionnelle avec le parcours d'Hilbert Peano (voir figure 3.1).

Cette étape de détection et d'isolement dans une boîte de sélection de l'objet à segmenter pourrait être automatisée.

3.1.4 Délimitation du volume et quantification

La quantification, c'est à dire l'extraction d'un paramètre quantitatif pertinent sur le plan clinique à partir d'une image, nécessite logiquement de déterminer une région à partir de laquelle calculer la valeur associée à la région, en général la moyenne des voxels contenus dans cette région. Toutefois il est important de souligner que la délimitation correcte du volume ne permet pas directement une quantification précise. En effet, sans correction d'EVP au préalable, le volume exact (ou suffisamment précis) de l'objet d'intérêt à quantifier contient les voxels des bords de l'objet, dont la valeur réelle a été réduite significativement par le "spillover". Ceci est d'autant plus important que l'objet est petit par rapport à la résolution spatiale. Pour des objets dont la taille est inférieure à deux centimètres, cet effet est majeur. Déterminer le volume précisément et la quantification exacte de ce volume sont donc deux opérations bien distinctes, car utiliser le volume précis pour en extraire une valeur moyenne d'activité mènera à une sous-estimation importante de l'activité réelle de l'objet correspondant (dans le cas d'un objet d'activité supérieure à son environnement sain, ce qui est souvent le cas pour les tumeurs). Cette différence significative entre détermination du volume et quantification a été mise en évidence dans [70]. Toutefois une piste de travail pour une détermination et quantification simultanée d'un objet d'intérêt en combinant la segmentation et la correction de volume partiel, a également été explorée, et est détaillée dans les perspectives, dans le chapitre 6 (voir section 6.2.2).

3.1.5 Prise en compte du bruit et du flou

Les méthodes de segmentation statistique offrent une panoplie de modèles et d'algorithmes permettant de modéliser les caractéristiques des images TEP. Le caractère à la fois bruité et flou des images peut en effet être pris en compte par la combinaison de l'aspect statistique (pour le bruit) et flou (pour la faible résolution spatiale) des approches bayésiennes présentées dans le chapitre précédent.

Nous avons décidé d'explorer la voie de la segmentation statistique combinée à la modélisation flou. La difficulté réside dans le choix des modèles et des paramètres des modèles pour obtenir

de bons résultats. Il existe en effet de nombreuses modélisations et des implémentations différentes. Nous avons choisi de travailler avec cette famille de méthodes car les algorithmes développés dans ce contexte ont montré de bons résultats dans de nombreux domaines d'application, de la reconnaissance de la parole [36] à l'imagerie radar et sonar, industrielle, astronomique ou médicale [127]. Dans ce domaine particulier, auquel ce rattache notre domaine d'application, ces méthodes de segmentation statistique ont essentiellement été appliquées dans le cadre de l'imagerie IRM, en particulier cérébrale [44]. Il existe très peu d'applications en imagerie fonctionnelle [29] [111], et aucune utilisant des approches floues, ce qui justifie notre intérêt de pousser plus loin l'exploration de cette famille de méthodes pour la tâche spécifique de segmentation d'images d'émission.

3.1.6 Idée générale et historique

L'idée générale derrière l'introduction d'une mesure floue dans le contexte bayésien a été introduite précédemment dans les travaux de Caillol, Pieczynski et Salzenstein [25] [139]. Son implémentation est fondée sur l'utilisation simultanée de mesures de Dirac pour modéliser les classes homogènes, ou "dures", et d'une mesure de Lebesgue pour modéliser la transition floue. Le modèle standard ne prévoit en effet que des classes dures bien différenciées, et ne permet pas que certains éléments puissent contenir un mélange de plusieurs classes. L'introduction d'une telle mesure floue permet de considérer, en plus d'une observation bruitée pour chaque élément de la vérité cachée, que chaque élément puisse contenir plusieurs classes à la fois : la partie statistique du modèle permet de considérer que pour un voxel clairement identifié à une classe, son observation soit bruitée. La mesure floue permet de plus de considérer que certains voxels puissent appartenir à deux classes, c'est à dire qu'elle modélise l'imprécision inhérente aux données cachées. Cette approche de la représentation du flou se distingue clairement de celle considérée dans le cadre du Fuzzy C-Means (voir l'état de l'art). Il est à noter que la modélisation à base de champs aléatoires cachés flous sans l'usage de mesures de Dirac, est plus ancienne [89].

3.1.7 Implémentation pratique

La mesure utilisée dans le cadre statistique standard peut être notée de la façon suivante :

$$\nu = \delta_1 + \delta_2 + \dots + \delta_C \quad (3.6)$$

Avec C le nombre de classes dures et δ_i la mesure de Dirac associée à la classe i . L'introduction de la mesure floue entraîne les considérations suivantes : premièrement, on considère que les éléments x_t prennent leurs valeurs dans l'intervalle $[0, 1]$ et non plus dans l'ensemble de classes discrètes $\Omega = 1, 2, \dots, C$. Deuxièmement, on définit une nouvelle mesure sur cet intervalle $[0, 1]$:

$$\nu = \delta_0 + \zeta + \delta_1 \quad (3.7)$$

Avec δ_0 et δ_1 les mesures de Dirac sur les classes dures 0 et 1 respectivement, et ζ la mesure de Lebesgue sur l'intervalle $]0, 1[$.

Il est nécessaire, pour implémenter cette approche, de discrétiser l'intervalle $]0, 1[$ en un certain nombre de niveaux de flou F_i auxquels on associe une valeur ε_i . Comme nous le verrons dans le chapitre suivant, cette discrétisation nécessite une mise en oeuvre et une optimisation spécifiques à l'application.

3.1.8 Limites et extensions possibles

On entrevoit ici une première limitation concernant ce modèle : il ne permet de modéliser que deux classes dures et une transition floue. Il faudra relaxer cette hypothèse si l'on veut appliquer cette approche à des tumeurs hétérogènes contenant clairement trois classes : fond, tumeur peu active (ou tissus nécrosés) et tumeur très active, par exemple.

3.2 Approche globale : chaînes de Markov

La première méthodologie de segmentation automatique développée dans le cadre de ce travail de thèse s'appuie sur une approche globale de la segmentation de l'image. En effet, le modèle *a priori* utilisé est celui des chaînes de Markov. Ce modèle a été choisi car nettement plus rapide que celui des champs. Les champs de Markov nécessitent des temps de calcul environ

vingt fois plus long. Il a également été montré que les chaînes sont capables de performances similaires dans de nombreuses situations [127]. Les chaînes peuvent également être étendues au cas flou, ce qui est indispensable pour segmenter correctement les images TEP.

3.2.1 Approche globale : hypothèse de Markov

La structure de la chaîne de Markov permet une prise en compte de la corrélation spatiale des voxels au sein de l'image et une certaine robustesse par rapport au bruit et aux changements de région. Dans plusieurs applications, les chaînes de Markov ont montré une nette supériorité face à des approches bayésiennes locales aveugles ou contextuelles [127]. L'hypothèse markovienne suppose que la structure des données cachées, c'est à dire le modèle *a priori*, suit une loi de chaîne de Markov. Ceci est rarement le cas en réalité, mais les résultats prometteurs de cette approche dans de nombreux domaines sont encourageants.

3.2.2 Modèle *a priori* : chaîne de Markov

Une chaîne de Markov est définie par la relation suivante :

$$P(x_t|x_1\dots x_{t-1}) = P(x_t|x_{t-1}) \quad (3.8)$$

Cette équation peut être traduite par : la probabilité d'obtenir un état donné à l'élément x_t connaissant tous les éléments précédents de la chaîne se résume uniquement à la connaissance de l'élément précédent x_{t-1} . On a ainsi une probabilité calculable directement, sans calcul itératif comme dans le cas des champs de Markov, et qui propage l'information tout le long de la chaîne, donnant une assurance de prendre une décision au niveau de l'élément x_t en tenant compte de toute l'information disponible. C'est pour cela que ce modèle peut être considéré comme une approche globale de la segmentation de l'image. Une chaîne de Markov est donc entièrement définie par la probabilité initiale $P(x_1)$ et les probabilités de transitions $P(x_{t+1}|x_t)$, c'est à dire la probabilité de rester dans un même état ou de changer d'état en passant d'un élément de la chaîne au suivant.

L'image à segmenter étant dans notre cas particulier un volume 3D de voxels, il est nécessaire de trouver une transformation qui permette d'associer les éléments d'une chaîne, par définition

uni-dimensionnelle, aux voxels du volume 3D, tout en préservant au mieux leur corrélation spatiale. Pour cela, il est possible d'utiliser une extension du parcours fractal d'Hilbert-Peano [146], déjà utilisé dans le cas 2D pour sa capacité à garantir au mieux la corrélation spatiale entre les pixels, au cas 3D [87]. Ce parcours assure que tous les éléments voisins dans la chaîne le sont également dans l'image. L'inverse n'étant malheureusement pas vrai (deux voxels voisins dans l'image peuvent se retrouver très éloignés dans la chaîne), ceci peut entraîner quelques difficultés, notamment pour les petits objets d'intérêt définis par un petit nombre de voxels dans les images TEP, ce que nous verrons en détails dans les résultats. La figure 3.1 illustre le parcours d'Hilbert-Peano en 2D et en 3D.

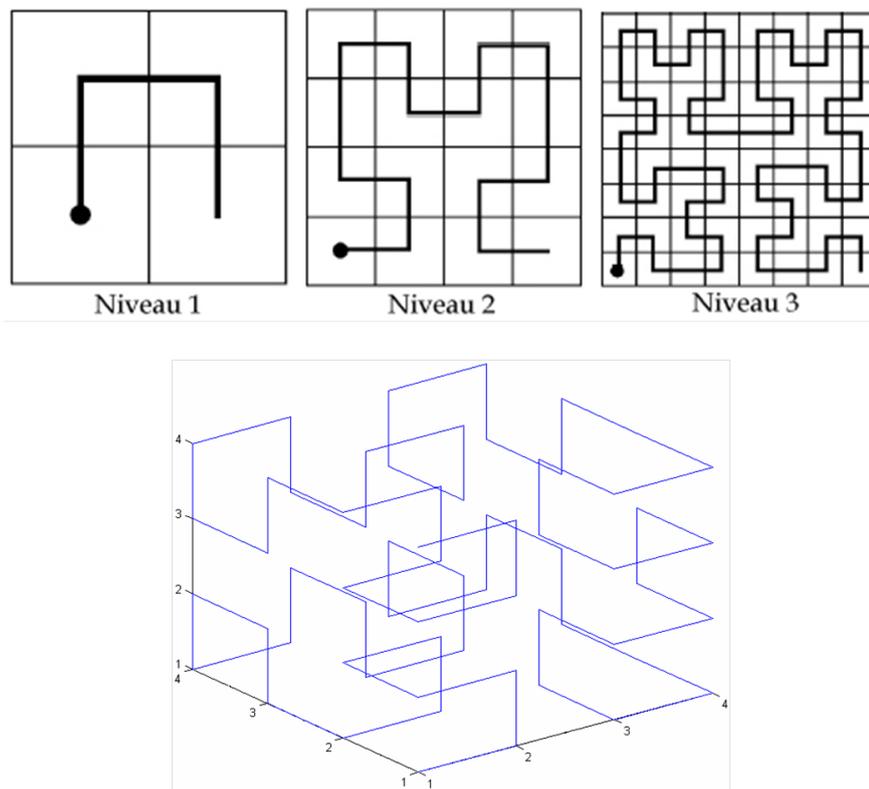


FIG. 3.1 – Le parcours fractal d'Hilbert-Peano en 2D et à plusieurs niveaux (en haut) et en 3D pour un volume 4x4x4 (en bas), utilisé pour transformer les images 2D et 3D en chaîne unidimensionnelle pour les traitement par chaîne de Markov.

3.2.3 Chaîne de Markov floue

L'extension au flou des chaînes de Markov a été proposé sous différentes formes [5] [110] [27]. Le modèle que nous utilisons a d'abord été développé et appliqué dans le cadre d'images astronomiques [31] [138]. C'est la première fois qu'un tel modèle flou de segmentation statistique d'image est utilisée dans le cadre de l'imagerie d'émission pour la segmentation d'objets d'intérêt, la seule occurrence de l'utilisation de modèles bayésiens et markoviens pour la segmentation d'images TEP étant limitée au cas des champs durs [29] [111]. Cette approche sera notée FHMC (pour *Fuzzy Hidden Markov Chains*) dans la suite de cette thèse.

La chaîne de Markov floue est définie comme un ensemble $X = (x_t)_{1 < t < T}$ prenant ses valeurs dans un ensemble continu (contrairement au cas "dur"), c'est à dire $x_t \in [0, 1]$. Pour définir les probabilités initiales et de transitions, il faut définir les densités initiales $P(x_1)$ et les densités de transitions $P(x_{t+1}|x_t)$. En faisant l'hypothèse d'une chaîne stationnaire, toutes les densités *a priori* peuvent être déduites d'une densité jointe $P(x_t, x_{t+1})$, définie sur la paire $(x_t, x_{t+1}) \in [0, 1]^2$. En accord avec la mesure floue $\nu \otimes \nu$ (voir section 3.7), la condition de normalisation est la suivante :

$$\int_0^1 \int_0^1 P(a, b) d(\nu \otimes \nu)(a, b) = 1 \quad (3.9)$$

Le modèle proposé est :

$$P(x_t = 0, x_{t+1} = 0) = p_{00}$$

$$P(x_t = 1, x_{t+1} = 1) = p_{11}$$

$$P(x_t = 0, x_{t+1} = 1) = p_{01}$$

$$P(x_t = 1, x_{t+1} = 0) = p_{10}$$

$$P(x_t \in]0, 1[, x_{t+1} \in \{0, 1\}) = P(x_t \in \{0, 1\}, x_{t+1} \in]0, 1]) = P(x_t \in \{0, 1\}, x_{t+1} \in \{0, 1\})$$

$$= 1 - |x_t - x_{t+1}|^r \text{ avec } r \in \mathfrak{R}$$

avec les conditions suivantes : $p_{01} = p_{10}$

$$\begin{aligned} &\text{et } p_{00} + p_{11} + p_{01} + p_{10} \\ &+ 2 \int_{]0,1[} P(0, a) da + 2 \int_{]0,1[} P(1, a) da + \int_{]0,1[} \int_{]0,1[} P(a, b) dadb = 1 \end{aligned} \quad (3.10)$$

Dans ce modèle, r contrôle l'homogénéité de l'image, c'est à dire que lorsque r augmente, la probabilité d'avoir deux éléments voisins similaires augmente. Dans notre implémentation, r a été fixé à 1. Ceci définit entièrement la densité jointe, qui est calculée en pratique à l'aide d'une discrétisation de l'intervalle $[0, 1]$ en un certain nombre de niveaux de flou F_i auxquels sont associés des valeurs ε_i . Dans notre implémentation, deux niveaux de flou F_1 et F_2 sont utilisés avec les valeurs associées $\varepsilon_1 = \frac{1}{3}$ et $\varepsilon_2 = \frac{2}{3}$. Enfin, la distribution initiale $P(x_1)$ correspond à la distribution marginale :

$$P(x_t) = \int_0^1 P(x_t, a) d\nu(a) = P(x_t, 0) + P(x_t, 1) + \int_{]0,1[} P(x_t, a) da \quad (3.11)$$

3.2.4 Lois des observations

Les lois des observations font appel à deux distributions pour les classes dures 0 et 1, dont les densités f_0 et f_1 sont définies par leurs moyennes et variances respectives (μ_0, σ_0^2) et (μ_1, σ_1^2) . Les densités des distributions des niveaux de flous F_i sont alors définies à partir de f_0 et f_1 comme suit :

$$\mu_{F_i} = \mu_0(1 - \varepsilon_i) + \varepsilon_i \mu_1 \quad (3.12)$$

$$\sigma_{F_i}^2 = \sigma_0^2(1 - \varepsilon_i) + \varepsilon_i \sigma_1^2 \quad (3.13)$$

avec ε_i la valeur associée à chaque niveau de flou F_i

3.2.5 Estimation des paramètres

Pour estimer les distributions initiales et de transitions, ainsi que les moyennes et variances des lois d'observation, nous utilisons l'algorithme SEM avec une réalisation *a posteriori*. Ceci signifie qu'une réalisation de la chaîne floue X est réalisée suivant sa loi *a posteriori* $P_{X_t}^y$. Pour cela nous devons rappeler d'abord les procédures de calcul *forward* et *backward* [45] sur la chaîne de Markov. Les densités forward et backward sont définies ainsi :

$$fwd_t(x_t) = P(x_t, y_1, \dots, y_t) \quad (3.14)$$

$$bwd_t(x_t) = \frac{P(y_{t+1}, \dots, y_T | x_t)}{P(y_{t+1}, \dots, y_T | y_1, \dots, y_t)} \quad (3.15)$$

Les procédures de calcul récursives sont alors analogues à celles du cas "dur" :

$$fwd_t(x_t) = P(y_t|x_t) \int_0^1 fwd_{t-1}(a)P(x_t|a)d\nu(a) \quad (3.16)$$

$$bwd_t(x_t) = P(y_t|x_t) \int_0^1 bwd_{t+1}(a)P(y_{t+1}|a)P(a|x_t)d\nu(a) \quad (3.17)$$

Enfin, la simulation de X est réalisée suivant sa densité de transition *a posteriori* $P^{y_{t+1}}(x_{t+1}|x_t)$ et sa densité initiale $P_{X_t}^Y(x_t)$:

$$P_{X_t}^Y(x_t) = fwd_t(x_t)bwd_t(x_t) \quad (3.18)$$

$$P^{y_{t+1}}(x_{t+1}|x_t) = \frac{P(x_{t+1}|x_t)P(y_{t+1}|x_{t+1})bwd_{t+1}(x_{t+1})}{\int_0^1 P(a|x_t)P(y_{t+1}|a)bwd_{t+1}(a)d\nu(a)} \quad (3.19)$$

pour tout $(x_t, x_{t+1}) \in [0, 1]^2$

Appelons $R = (r_1, \dots, r_t)$ cette réalisation *a posteriori*. Les valeurs des paramètres du modèles sont ré-estimées sur une partition de R définie comme suit :

$$Q_0 = \{t|r_t = 0\}$$

$$Q_1 = \{t|r_t = 1\}$$

Les probabilités initiales sont estimées par :

$$P(x_1 = c) = \delta(r_1, c) \quad (3.20)$$

Les probabilités de transitions sont estimées par :

$$P(x_{t+1} = d|x_t = c) = \frac{\sum_{t=2}^T \delta(r_{t-1}, c)\delta(r_t, d)}{\sum_{t=2}^T \delta(r_{t-1}, c)} \quad (3.21)$$

Les moyennes et variances des deux classes dures sont estimées par la méthode des moments (et celles des niveaux de flou sont déterminés par 3.14) :

$$\mu_c = \frac{\sum_{t=1}^T y_t \delta(r_t, c)}{\sum_{t=1}^T \delta(r_t, c)} \quad (3.22)$$

$$\sigma_c^2 = \frac{\sum_{t=1}^T (y_t - \mu_c)^2 \delta(r_t, c)}{\sum_{t=1}^T \delta(r_t, c)} \quad (3.23)$$

pour $c \in 0, 1$

3.2.6 Segmentation

Il est possible d'adapter le critère du MPM (*Maximum a Posteriori Mode*) [109] au contexte flou. Etant donnée une observation Y , la décision bayésienne déterminant $S(X_t)$ inclut la minimisation de l'espérance conditionnelle [45] en chaque site t pour obtenir le X_t optimal :

$$s(y_t)^{optimal} = \operatorname{argmin}_{X_t=s(x_t)} E[L_t(X_t, S(X_t)) | Y = y] \quad (3.24)$$

La fonction de perte $L_t(x_t, S(x_t)) = |x_t - S(x_t)|$ modélise la perte associée à l'attribution d'une classe $S(x_t)$ erronée à la place de la vraie classe x_t , et pour une variable X stationnaire, l'erreur est approximée par :

$$E[L_t(X_t, S(X_t))] = \frac{1}{T} \sum_{t=1}^{t=T} L_t(x_t, S(x_t)) \quad (3.25)$$

Enfin, le calcul de l'espérance conditionnelle E est possible grâce à la connaissance de la distribution *a posteriori* $P_{X_t}^y$ (voir section précédente) :

$$E[L_t(X_t, S(X_t)) | Y = y] = P_{X_t}^y(0)L_t(0, S(X_t)) + P_{X_t}^y(1)L_t(1, S(X_t)) + \int_{]0,1[} P_{X_t}^y(a)L_t(a, S(X_t)) da \quad (3.26)$$

La segmentation consiste alors à affecter au pixel x_t la classe ou le niveau de flou qui minimise la fonction ci-dessus.

3.3 Approche locale adaptative

3.3.1 Approche locale sans hypothèse markovienne

Cette approche est plus simple et plus ancienne que l'approche utilisant une modélisation markovienne de la loi *a priori*. Il s'agit d'une approche dite "aveugle", c'est à dire travaillant "pixel par pixel" avec une prise en compte inexistante du voisinage spatial, bien qu'il existe plusieurs extensions permettant justement d'intégrer le voisinage, soit dans la loi *a priori* elle-même (modélisation dite "contextuelle"), soit dans l'estimation des paramètres (modélisation dite "adaptative") [123]. Ces modélisations ont été étendues au cas flou dans [26]. Nous avons décidé de nous intéresser à l'utilisation de cette modélisation pour résoudre le problème particulier de la segmentation des objets de petite dimension dont le diamètre est inférieur à

2 cm [71], qui posent problème à la méthode de segmentation par chaînes de Markov à cause de l'utilisation du parcours d'Hilbert-Peano (voir la section 4.3) [70].

3.3.2 Estimation adaptative

L'estimation adaptative des paramètres du modèle *a priori* (les probabilités *a priori* d'appartenir à une classe donnée), pour chaque voxel, est un des aspects les plus importants de cette modélisation. En effet, l'approche dite "aveugle" ou "pixel par pixel" est très faible car elle ne permet pas de prendre en compte les relations spatiales entre les voxels, contrairement à l'approche par chaîne de Markov, qui intègre cette corrélation spatiale au sein de sa modélisation. Il existe une modélisation dite "contextuelle" [26], qui, contrairement à l'approche dite "aveugle", permet d'intégrer au modèle un ou plusieurs voisins de chaque voxel.

Modélisation contextuelle

Cette modélisation est mal adaptée à notre application, car plus on intègre de voisins plus le nombre de paramètres à estimer augmente [26], ce qui mène en général à l'inclusion de seulement un ou deux voisins dans le modèle. Or dans notre cas nous souhaitons intégrer toute l'information spatiale du voisinage d'un voxel donné pour sa classification. Idéalement, on souhaite prendre en compte au moins les 26 premiers voisins en 3D d'un voxel donné pour sa classification, voir plus si cela se révèle insuffisant.

Estimation adaptative des paramètres *a priori*

Cette approche prend en compte le voisinage spatial au sein de l'étape d'estimation des paramètres [123]. Il s'agit de déterminer les probabilités *a priori* de chaque voxel au lieu d'une seule probabilité pour chaque classe, en fonction de la distribution *a posteriori* des voisins du voxel. La taille de la fenêtre d'estimation adaptative, un cube dans notre cas, doit être optimisée en fonction de la taille de l'image, de la taille des objets imagés et de l'échantillonnage spatial (les dimensions des voxels). Le cube d'estimation doit être suffisamment petit pour mener à de bonnes propriétés locales mais suffisamment grand pour que l'information du voisinage soit incorporée de façon pertinente dans l'estimation. Une telle optimisation a été

réalisée pour notre application (voir section 4.3.1) et a mené à une taille optimale de 3x3x3 pour des objets de taille inférieure à 4 cm de diamètre et pour des tailles de voxels de 2 ou de 4 mm de côté. Cette approche sera notée FLAB (pour Fuzzy Locally Adaptive Bayesian) dans la suite de cette thèse.

3.3.3 Loi *a priori*

L'algorithme qui en résulte repose donc sur l'utilisation d'une loi *a priori* suivant une densité h sur $[0, 1]$ pour la mesure ν définie en 3.7 :

$$\begin{aligned} h(0) &= P[x_t = 0] = p_{0,t} \\ h(1) &= P[x_t = 1] = p_{1,t} \\ h(\varepsilon) &= P[x_t = \varepsilon] = 1 - p_{0,t} - p_{1,t} \text{ pour } \varepsilon \in]0, 1[\\ &\text{avec la condition de normalisation suivante : } h(0) + h(1) + \int_0^1 h(\varepsilon) = 1 \quad (3.27) \end{aligned}$$

Les probabilités *a priori* $p_{c,t}$ sont indicées par t car elles sont différentes pour chaque voxel, contrairement au cas "aveugle".

3.3.4 Loi des observations

La loi des observations fait appel à deux distributions pour les classes 0 et 1, dont les densités f_0 et f_1 sont définies par leurs moyennes et variances respectives (μ_0, σ_0^2) et (μ_1, σ_1^2) . Les densités des distributions des niveaux de flous F_i sont alors définies à partir de f_0 et f_1 comme suit :

$$\mu_{F_i} = \mu_0(1 - \varepsilon_i) + \varepsilon_i\mu_1 \quad (3.28)$$

$$\sigma_{F_i}^2 = \sigma_0^2(1 - \varepsilon_i) + \varepsilon_i\sigma_1^2 \quad (3.29)$$

avec ε_i la valeur associée à chaque niveau de flou F_i

Comme première approximation, ces lois ont été modélisées par des gaussiennes, mais l'utilisation du système de Pearson permet de considérer d'autres distributions (voir section 3.5.1)

3.3.5 Estimation des paramètres

Comme dans le cas des chaînes de Markov floues, l'estimation des paramètres du modèle (probabilités *a priori* de chaque voxel et moyennes et variances des classes) se fait avec l'algorithme SEM, à l'aide d'une réalisation *a posteriori*. La loi *a posteriori* permettant de générer cette réalisation est donnée par la probabilité de chaque classe c connaissant l'observation y_t :

$$P(x_t = c|y_t) = \frac{p_{t,c}f(y_t|c)}{p_{t,0}f(y_t|0) + p_{t,1}f(y_t|1) + (1 - p_{t,0} - p_{t,1}) \int_0^1 p_{t,a}f(y_t|a)da} \quad (3.30)$$

Appelons $R = (r_1, \dots, r_t)$ la réalisation *a posteriori* obtenue en utilisant $P(x_t \in]0, 1[|y_t) = 1 - P(x_t = 0|y_t) - P(x_t = 1|y_t)$ et en procédant au tirage aléatoire parmi 0, 1 ou F (pour les pixels flous). Les valeurs des paramètres du modèles sont ré-estimées sur une partition de R définie comme suit :

$$Q_0 = \{t|r_t = 0\}$$

$$Q_1 = \{t|r_t = 1\}$$

Les probabilités *a priori* de chaque voxel t sont alors estimées en utilisant le cube d'estimation C_t centré sur le voxel t :

$$p_{c,t} = \frac{1}{\text{Card}(C_t)} \sum_{j \in C_t} \delta(r_j, c) \text{ pour } c \in \{0, 1\} \quad (3.31)$$

Les moyennes et variances des deux classes dures sont estimées par la méthode des moments (et celles des niveaux de flou sont déterminés par 3.14) :

$$\mu_c = \frac{\sum_{t \in Q_c} y_t}{\text{Card}(Q_c)} \quad (3.32)$$

$$\sigma_c^2 = \frac{\sum_{t \in Q_c} (y_t - \mu_c)^2}{\text{Card}(Q_c)} \quad (3.33)$$

$$\text{pour } c \in \{0, 1\}$$

3.3.6 Segmentation

Pour procéder à la segmentation nous utilisons le *Maximum Posterior Likelihood* tel que suggéré dans [26]. Il s'agit de calculer la densité *a posteriori* en chaque voxel telle que définie

par l'équation 3.30 et d'affecter à chaque voxel la classe ou le niveau de flou qui maximise cette densité. Pour ce faire, la procédure est la suivante :

Calculer pour chaque voxel la probabilité *a posteriori* $P(x_t = c|y_t)$

Si $P(x_t = c|y_t)$ est maximale avec $c = 1$ ou $c = 0$, affecter la classe 1 ou 0

Sinon, choisir le niveau de flou F_i qui maximise $P(x_t = \varepsilon_i|y_t)$

3.4 Approche locale étendue à trois classes

3.4.1 Problématique

Bien que l'approche locale adaptative décrite dans la section précédente soit bien adaptée à la segmentation de structures homogènes, puisque ce type de configurations ne nécessite que deux classes dures et une transition floue (il s'agit uniquement d'extraire l'objet du fond), elle est limitée dans son application à des tumeurs réelles, ces dernières présentant souvent des distributions d'activité fortement hétérogènes en leur sein. Un exemple de tumeur fortement hétérogène est illustré dans la figure 3.2. En effet, la méthode à deux classes dures ne peut

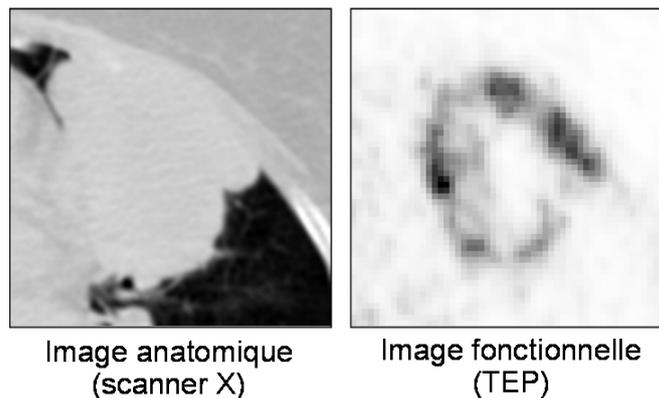


FIG. 3.2 – Illustration d'une tumeur pulmonaire dont la distribution d'activité est fortement hétérogène. Noter que l'imagerie anatomique TDM ne permet pas d'apprécier les variations au sein de la tumeur, qui ne sont mises en évidence que sur l'imagerie fonctionnelle.

correctement segmenter de telles tumeurs, lorsque le contraste entre les régions les plus actives

et les plus faibles de la tumeur est supérieur au contraste entre le fond et les régions faibles de la tumeur. En effet, lorsque le contraste entre les régions au sein de la tumeur est faible en comparaison avec le contraste séparant la tumeur du fond, la méthode à deux classes est en général capable d'extraire correctement la tumeur dans son ensemble. Par contre, en cas d'hétérogénéité forte, les contrastes font que les niveaux de flou se trouvent attribués par erreur à des zones correspondant en réalité à une troisième classe dure. La figure 3.3 illustre cette difficulté. Dans les deux premières situations, la méthode à deux classes dures et une transition

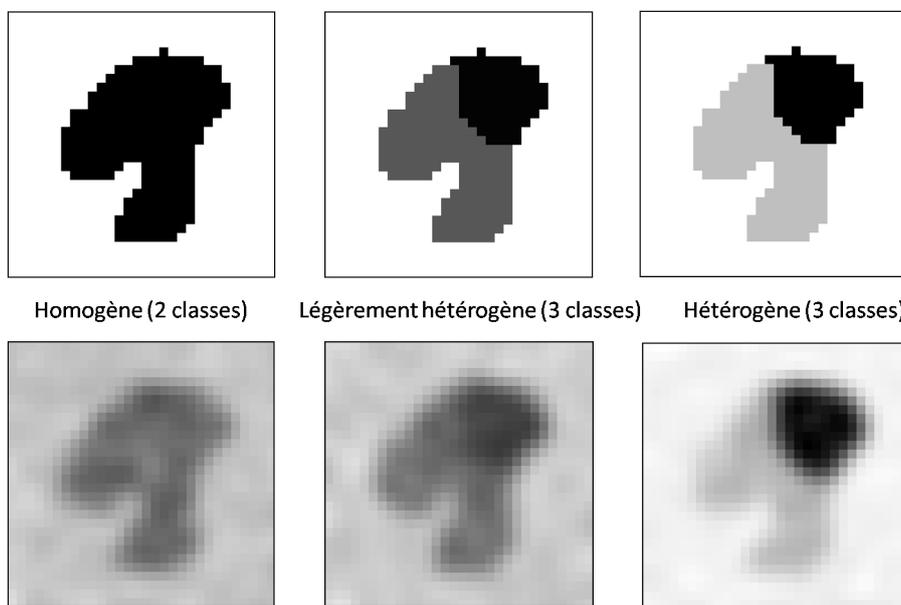


FIG. 3.3 – Différents cas de figure (en haut, la vérité terrain, en bas l'image TEP simulée correspondante) : tumeur homogène (à gauche), légèrement hétérogène (au centre) et fortement hétérogène (à droite).

floue peut suffire à extraire la tumeur du fond, mais il est impossible de définir directement deux régions distinctes au sein de la tumeur dans le deuxième cas, à moins d'utiliser un post-traitement. Par exemple, on peut ré-appliquer la même approche de segmentation à deux classes dures à la sélection restreinte à la segmentation de la tumeur obtenue précédemment sur l'image entière, ce qui est illustré dans la figure 3.4. Toutefois, dans le troisième cas, il est nécessaire de mettre en place une méthodologie à trois classes, car l'utilisation de la méthode à deux classes associera, suivant les contrastes, la zone faible de la tumeur au fond ou aux

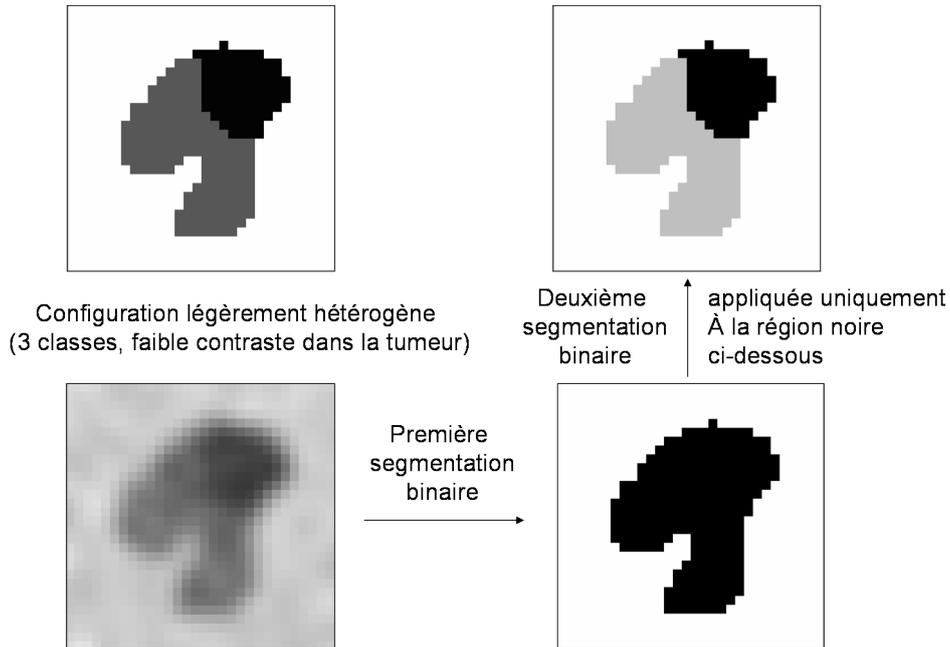


FIG. 3.4 – Exemple de segmentation en deux phases, appliquée à une tumeur faiblement homogène pour différencier deux régions distinctes en son sein.

niveaux de flou de la transition. De plus, une méthode à trois classes a l'avantage de pouvoir donner directement le résultat souhaité dans le deuxième cas de figure, plutôt que de devoir ré-appliquer la segmentation une deuxième fois. Il s'agit donc de considérer directement trois classes : une associée au fond (les tissus sains), et deux classes pour la tumeur, une pour les zones d'activité élevée, et une pour les zones d'activité réduite. Notons qu'il est également possible de considérer une approche à trois classes dans le contexte d'une tumeur homogène sur un fond fortement hétérogène. Les classes seraient alors affectées différemment (une classe pour la tumeur et deux pour le fond). Notons également que pour certaines tumeurs très complexes, on peut imaginer que trois classes soit insuffisant pour rendre compte de la complexité et de la réalité de la distribution d'activité dans la tumeur, toutefois cela devrait être assez rare. D'après notre expérience, les tumeurs que nous avons observées dans nos données cliniques peuvent le plus souvent être réduites à trois classes, même en cas d'hétérogénéité très forte. Le cas de multiples tumeurs possédant chacune leur activité propre pourraient nécessiter plus que trois classes si on souhaitait les segmenter toutes d'un coup, mais dans la mesure où nous

faisons l'hypothèse que chaque tumeur est isolée et segmentée séparément, ce n'est pas une limite.

3.4.2 Extension à trois classes dures

Le passage de deux à trois classes dures entraîne la nécessité de considérer plusieurs transitions floues différentes. En effet, dans le cas de deux classes dures, une seule transition floue suffit à modéliser les "mixels" pouvant exister dans l'image. Dans le cas de la présence de trois classes dures, il peut exister jusqu'à quatre zones différentes de flou. Trois zones correspondant à un mélange de deux classes dures parmi les trois existantes, et une quatrième correspondant à un mélange des trois classes dures à la fois, avec la possibilité que chaque classe dure ait une participation différente à la zone en question. Toutefois, cette quatrième configuration est plutôt rare dans le cas des images que l'on a à traiter et ne devrait concerner qu'un nombre très réduit de voxels dans un volume 3D typique contenant une tumeur hétérogène. C'est pourquoi nous limiterons notre modélisation aux trois premières configurations. La figure 3.5 illustre les différentes zones considérées.

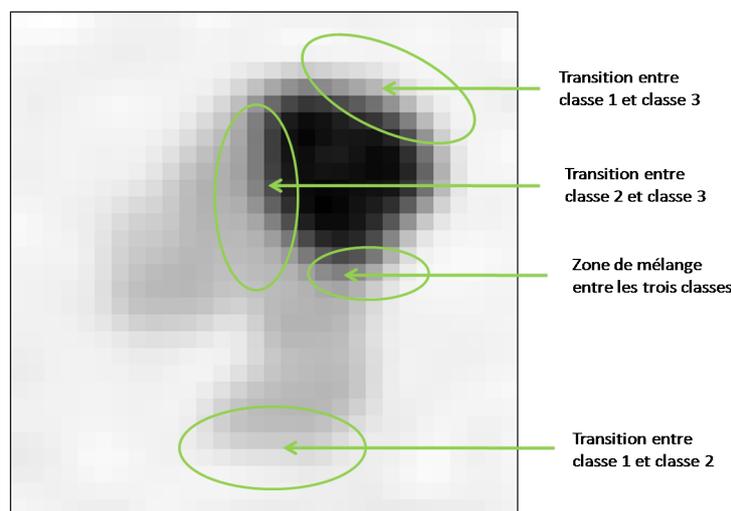


FIG. 3.5 – Illustration des différentes zones envisageables dans le cadre d'une tumeur fortement hétérogène.

3.4.3 Modélisation des transitions floues

Nous limitons notre modélisation à trois classes dures et trois transitions floues. L'hypothèse qui est posée ici est que chaque voxel de l'image peut être classé soit dans une des trois classes dures, soit dans une des trois transitions floues entre chaque paire de classes dures. On peut représenter la modélisation utilisée par le triangle de la figure 3.6. Un voxel peut être classifié comme étant sur un des trois sommets du triangle (classes dures), ou quelque part le long d'un des trois segments (transitions floues), mais pas sur la surface au centre du triangle (quatrième configuration évoquée précédemment). A chacun des sommets du triangle, on associe une mesure de Dirac, et à chaque segment, une mesure continue de Lebesgue. La mesure utilisée devient donc :

$$v = \delta_1 + \delta_2 + \delta_3 + \zeta_{12} + \zeta_{23} + \zeta_{13} \quad (3.34)$$

Chaque transition floue est donc différente, et doit être découpée en niveaux de flou (2 dans

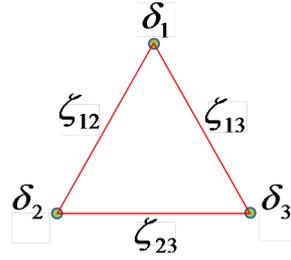


FIG. 3.6 – Représentation de la modélisation adoptée pour la segmentation à trois classes dures et trois transitions floues.

notre cas), comme dans le cas binaire. L'algorithme qui en résulte est donc l'utilisation de la loi *a priori* suivant une densité h sur $[0, 1]$ pour la mesure ν définie en 3.34. On définit tout d'abord les probabilités *a priori* en 1, 2 et 3 avec deux variables différentes correspondant aux deux segments liés à la mesure de Dirac située sur les sommets du triangle :

$$\begin{aligned} p_{1,t} &= p_{112,t} + p_{113,t} \\ p_{2,t} &= p_{212,t} + p_{223,t} \\ p_{3,t} &= p_{313,t} + p_{323,t} \end{aligned}$$

avec : $p_{c,t}$ la probabilité du voxel t d'appartenir à la classe c (3.35)

Les probabilités *a priori* $p_{c,t}$ sont indicées par t car elles sont différentes pour chaque voxel, contrairement au cas "aveugle".

On en déduit la densité h suivante :

$$\begin{aligned} h(\varepsilon) = & p_{112,t}\delta(\varepsilon) + p_{113,t}\delta(\varepsilon) + p_{212,t}\delta(\varepsilon - 1) \\ & + p_{223,t}\delta(\varepsilon) + p_{313,t}\delta(\varepsilon - 1) + p_{323,t}\delta(\varepsilon - 1) + h_{12}(\varepsilon) + h_{23}(\varepsilon) + h_{13}(\varepsilon) \end{aligned} \quad (3.36)$$

avec $h_{AB}(\varepsilon)$ la densité associée à la transition floue entre les classes A et B

En intégrant sur $[0, 1]$ on obtient :

$$\int_0^1 h(\varepsilon)d\varepsilon = p_{1,t} + p_{2,t} + p_{3,t} + \int_0^1 (h_{12}(\varepsilon) + h_{23}(\varepsilon) + h_{13}(\varepsilon))d\varepsilon = 1 \quad (3.37)$$

3.4.4 Loi des observations

La loi des observations fait appel à trois distributions pour les classes 1, 2 et 3, dont les densités f_1 , f_2 et f_3 sont définies par leurs moyennes et variances respectives (μ_1, σ_1^2) , (μ_2, σ_2^2) et (μ_3, σ_3^2) . Les densités des distributions des niveaux de flous i de la transition floue entre les classes dures A et B (F_{AB_i}) sont alors définies à partir de f_A et f_B comme suit :

$$\mu_{F_{AB_i}} = \mu_A(1 - \varepsilon_{AB_i}) + \varepsilon_{AB_i}\mu_B \quad (3.38)$$

$$\sigma_{F_{AB_i}}^2 = \sigma_A^2(1 - \varepsilon_{AB_i}) + \varepsilon_{AB_i}\sigma_B^2 \quad (3.39)$$

avec ε_{AB_i} la valeur associée à chaque niveau de flou F_{AB_i}

Comme première approximation, ces lois ont été modélisées par des gaussiennes, mais l'utilisation du système de Pearson permet de considérer d'autres distributions (voir section 3.5.1)

3.4.5 Estimation des paramètres

Comme dans le cas des chaînes de Markov floues, l'estimation des paramètres du modèle (probabilités *a priori* de chaque voxel et moyennes et variances des classes) se fait avec l'algorithme SEM, à l'aide d'une réalisation *a posteriori*. La loi *a posteriori* permettant de générer cette réalisation est donnée par la probabilité de chaque classe c connaissant

l'observation y_t :

$$P(x_t = c|y_t) = \frac{p_{t,c}f(y_t|c)}{\sum_{c=1}^3 p_{t,c}f(y_t|c) + p_{fuzzy}} \quad (3.40)$$

Avec p_{fuzzy} :

$$p_{fuzzy} = \frac{1}{2}p_{t,(12)} \sum_i F_{(12)_i} + \frac{1}{2}p_{t,(23)} \sum_i F_{(23)_i} + \frac{1}{2}p_{t,(13)} \sum_i F_{(13)_i} \quad (3.41)$$

Où i est le nombre de niveaux de flou et $p_{t,(AB)}$ est la probabilité a priori du voxel t d'appartenir à la transition floue entre les classes A et B , et c prend 6 valeurs possibles (les trois classes dures et les trois transitions floues).

Appelons $R = (r_1, \dots, r_t)$ la réalisation *a posteriori* obtenue par tirages aléatoires pondérés par la distribution *a posteriori* ci-dessus. Les valeurs des paramètres du modèles sont ré-estimées pour les classes dures sur une partition de R définie comme suit :

$$\begin{aligned} Q_1 &= \{t|r_t = 1\} \\ Q_2 &= \{t|r_t = 2\} \\ Q_3 &= \{t|r_t = 3\} \\ Q_{F_{12}} &= \{t|r_t = F_{12}\} \\ Q_{F_{23}} &= \{t|r_t = F_{23}\} \\ Q_{F_{13}} &= \{t|r_t = F_{13}\} \end{aligned} \quad (3.42)$$

Les probabilités *a priori* de chaque voxel t sont alors estimées en utilisant le cube d'estimation C_t centré sur le voxel t :

$$p_{c,t} = \frac{1}{Card(C_t)} \sum_{j \in C_t} \delta(r_j, c) \quad (3.43)$$

Les moyennes et variances des classes dures sont estimées par la méthode des moments (et celles des niveaux de flou sont déterminés par 3.38) :

$$\mu_c = \frac{\sum_{t \in Q_c} y_t}{Card(Q_c)} \quad (3.44)$$

$$\sigma_c^2 = \frac{\sum_{t \in Q_c} (y_t - \mu_c)^2}{Card(Q_c)} \quad (3.45)$$

pour $c \in \{1, 2, 3\}$

3.4.6 Segmentation

Pour procéder à la segmentation nous utilisons le *Maximum Posterior Likelihood* tel que suggéré dans [26]. Il s'agit de calculer la densité *a posteriori* en chaque voxel telle que définie par 3.40 et d'affecter à chaque voxel la classe ou le niveau de flou qui maximise cette densité. Pour ce faire, la procédure est la suivante :

Calculer pour chaque voxel la probabilité *a posteriori* $P(x_t = c|y_t)$

Si $P(x_t = c|y_t)$ est maximale avec $c = 1, 2$ ou 3 , alors affecter la classe c

Sinon, choisir le niveau de flou F_{AB_i} qui maximise $P(x_t = \varepsilon_{AB_i}|y_t)$

3.5 Modélisation de la loi des observations

L'approche choisie pour développer les méthodes de segmentation adaptées aux images d'émission se base une modélisation statistique et probabiliste. Ceci permet en pratique de modéliser le bruit dans les images comme une loi statistique donnée, dont les paramètres (moyenne, variance, paramètre de forme, etc.) doivent être estimés. Dans de nombreuses applications, des lois gaussiennes sont considérées car elles permettent de modéliser un certain nombre de phénomènes physiques d'acquisition des images de façon approximative mais satisfaisante. Dans certains cas, par exemple en imagerie radar, d'autres lois telles que les distributions gamma ou K-distributions sont utilisées car elles modélisent mieux les répartitions des valeurs des pixels dans l'image entraînant une amélioration significative des résultats de segmentation [54].

Pour l'application à des images TEP, nous avons d'abord envisagé comme première approximation d'utiliser la simple hypothèse gaussienne en ce qui concerne la forme des distributions considérées pour modéliser la loi des observations conditionnellement aux étiquettes ($P(Y|X)$). Bien que cette hypothèse soit utilisée avec succès dans différents contextes d'imagerie, il est possible qu'elle ne soit pas suffisamment réaliste pour notre application. Pour tenter de mettre à l'épreuve cette hypothèse de distribution gaussienne pour les classes dans les images TEP,

nous allons examiner l'utilisation du système de Pearson pour affiner les distributions utilisées.

3.5.1 Système de Pearson

Intérêt pour notre application

Dans le cas où l'hypothèse de bruit gaussien se révèle une simplification grossière, l'utilisation du système de Pearson [85] pourrait permettre d'utiliser des distributions différentes, potentiellement mieux adaptées, et menant donc à des résultats de segmentation plus fins, comme cela a été le cas dans d'autres applications. Le fait que chaque scanner TEP et chaque algorithme de reconstruction d'image présentent leurs propres particularités en terme de structure de bruit, de post-filtering et d'artefacts, il est également possible que l'hypothèse de bruit gaussien convienne pour l'un mais soit insuffisante pour l'autre. L'étude et l'utilisation du système de Pearson dans notre application permet d'une part d'apporter une meilleure compréhension des distributions des valeurs des voxels dans les images reconstruites avec certains algorithmes itératifs, et d'autre part, d'améliorer potentiellement les résultats de segmentation de nos méthodes qui permettent d'intégrer des formes de distributions différentes de la simple loi gaussienne. En effet, le système de Pearson a la particularité de pouvoir s'intégrer directement dans les méthodes d'estimation de paramètres habituellement utilisées en segmentation statistique telles que EM, SEM ou ICE [40]. Il a donc le potentiel de rendre l'approche FLAB plus robuste et universelle par rapport à l'utilisation de différents scanners et différents algorithmes de reconstruction.

Présentation du système de Pearson

Le système de Pearson contient huit distributions différentes qui peuvent s'exprimer sous la forme suivante : soit une distribution de densité f sur R . Elle appartient au système de Pearson si elle vérifie l'équation suivante :

$$\frac{1}{f(y)} \frac{df(y)}{dy} = -\frac{y + a}{c_0 + c_1 y + c_2 y^2} \quad (3.46)$$

Les variations des coefficients c_0, c_1, c_2 et a déterminent les différentes formes de distributions, aussi bien que les paramètres définissant la distribution. Ces paramètres, déduits directement des données observées y , permettent de déterminer quelle distribution est statistiquement la plus proche de celle des données, grâce au graphe de Pearson (figure 3.7). Les huit familles

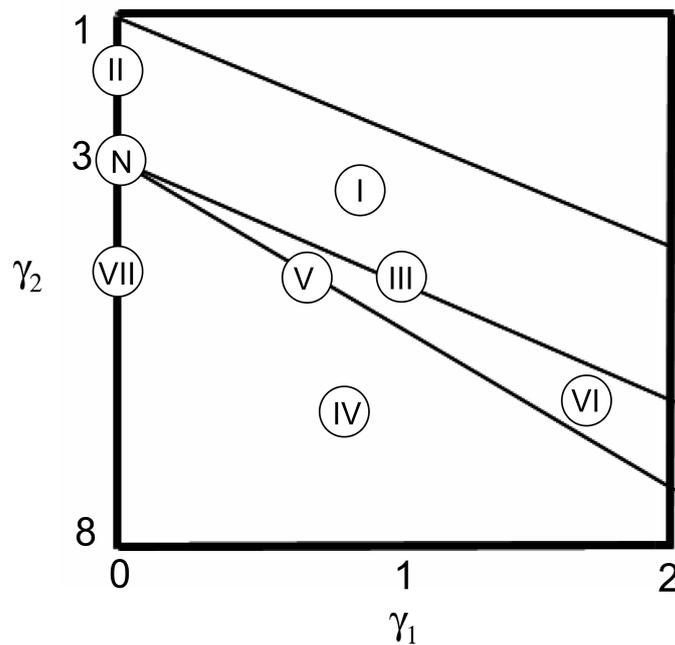


FIG. 3.7 – Le graphe de Pearson permet de localiser une distribution en fonction des paramètres γ_1 et γ_2 .

de lois contenues dans le système de Pearson sont les suivantes : bêta du premier type, type II, gamma, type IV, gamma inverse, bêta II, type VII, et gaussienne. Le graphe de Pearson résume sous forme graphique les règles permettant de classer une distribution parmi ces huit familles. En effet, les paramètres γ_1 et γ_2 qui graduent les axes des abscisses et des ordonnées du graphe sont définis ainsi :

$$\gamma_1 = \frac{\mu_3^2}{\mu_2^3} \tag{3.47}$$

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} \tag{3.48}$$

Où μ_n désigne le moment statistique d'ordre n , définit par :

$$\mu_1 = E[Y] \quad (3.49)$$

$$\mu_n = E[(Y - E(Y))^n] \text{ pour } n > 1 \quad (3.50)$$

Y désigne une variable aléatoire dont la distribution appartient au système de Pearson. Les paramètres γ_1 et γ_2 sont liées aux caractéristiques statistiques des distributions. En effet, $\sqrt{\gamma_1}$ donne une mesure de l'asymétrie de la distribution ("skewness") quant à γ_2 , il s'agit de la mesure de l'aplatissement ("kurtosis").

Le lien entre les coefficients c_0, c_1, c_2 et a et les moments μ_n sont obtenus grâce aux formules suivantes :

$$a = \frac{(\gamma_2 + 3)\sqrt{\gamma_1\mu_2}}{(10\gamma_2 - 12\gamma_1 - 18)} - \mu_1 \quad (3.51)$$

$$c_0 = \frac{\mu_2(4\gamma_2 - 3\gamma_1) - \mu_1(\gamma_2 + 3)\sqrt{\gamma_1\mu_2} + \mu_1^2(2\gamma_2 - 3\gamma_1 - 6)}{(10\gamma_2 - 12\gamma_1 - 18)} \quad (3.52)$$

$$c_1 = \frac{(\gamma_2 + 3)\sqrt{\gamma_1\mu_2} - 2\mu_1(2\gamma_2 - 3\gamma_1 - 6)}{(10\gamma_2 - 12\gamma_1 - 18)} \quad (3.53)$$

$$c_2 = \frac{(2\gamma_2 - 3\gamma_1 - 6)}{(10\gamma_2 - 12\gamma_1 - 18)} \quad (3.54)$$

De plus les paramètres γ_1 et γ_2 permettent d'obtenir un coefficient λ dont la valeur permet de classer la distribution de Y dans une des huit familles du système de Pearson. Ce coefficient λ est donné par :

$$\lambda = \frac{\gamma_1(\gamma_2 + 3)^2}{4(4\gamma_2 - 3\gamma_1)(2\gamma_2 - 3\gamma_1)(2\gamma_2 - 3\gamma_1 - 6)} \quad (3.55)$$

Les règles permettant de classifier la distribution F_Y dans l'une des huit familles en fonction de λ , γ_1 et γ_2 sont les suivantes :

$$\begin{aligned}
 \lambda < 0 &\Leftrightarrow F_Y \in F_1 \\
 \gamma_1 = 0 \text{ et } \gamma_2 < 3 &\Leftrightarrow F_Y \in F_2 \\
 2\gamma_2 - 3\gamma_1 - 6 = 0 &\Leftrightarrow F_Y \in F_3 \\
 0 < \lambda < 1 &\Leftrightarrow F_Y \in F_4 \\
 \lambda = 1 &\Leftrightarrow F_Y \in F_5 \\
 \lambda > 1 &\Leftrightarrow F_Y \in F_6 \\
 \gamma_1 = 0 \text{ et } \gamma_2 > 3 &\Leftrightarrow F_Y \in F_7 \\
 \gamma_1 = 0 \text{ et } \gamma_2 = 3 &\Leftrightarrow F_Y \in F_8
 \end{aligned} \tag{3.56}$$

L'intégration du système de Pearson se fait directement au sein des algorithmes de segmentation statistique. Il s'agit d'ajouter une étape de détection de la loi permettant de sélectionner quel type de loi utiliser, et quels paramètres calculer, et ce à chaque itération de l'estimation des paramètres. Le type de loi ainsi que les paramètres définissant cette loi donnée une fois sa famille sélectionnée, peuvent d'ailleurs évoluer au fur et à mesure des itérations de l'estimation. Pour chaque classe, le type de distribution et les paramètres qui la définissent sont estimés grâce au calcul des paramètres utilisant les équations décrites précédemment. Les moments d'ordre n de 1 à 4 de chaque classe c sont estimés grâce aux moments empiriques de la façon suivante :

$$\mu_{1,c} = \frac{\sum_{t \in Q_c} y_t}{\text{Card}(Q_c)} \tag{3.57}$$

$$\mu_{n,c} = \frac{\sum_{t \in Q_c} (y_t - \mu_{1,c})^n}{\text{Card}(Q_c)} \text{ pour } n > 1 \tag{3.58}$$

Où Q_c représente une partition des données observées y_t (t de 1 à T , le nombre de voxels).

Chapitre 4

Validation et optimisation des méthodes

Les méthodes présentées dans les chapitres précédents ont été appliquées sur différents ensembles de données pour optimiser les différents paramètres et valider la robustesse de l'approche. En effet, nous voulons faire en sorte que l'approche sélectionnée puisse être appliquée avec des performances acceptables sur une image TEP, quel que soit le modèle de scanner ou l'algorithme de reconstruction utilisés. Nous commençons par présenter les données, puis les résultats obtenus.

4.1 Données de validation et d'évaluation

4.1.1 Fantôme

Dans ce travail nous utilisons un fantôme contenant des sphères de taille variable, ce qui permet d'évaluer les performances des méthodes en ce qui concerne la segmentation d'objets sphériques homogènes. Ce fantôme est le NEMA IEC [86]. Il s'agit d'un cylindre de plexiglas d'environ 20 cm de hauteur et de 20 cm de diamètre. Il contient six sphères de diamètres 37, 28, 22, 17, 13 et 10 mm. Il est possible de remplir le fantôme ainsi que les sphères avec une solution d'eau et de fluor 18 (ou autre), et de placer des concentrations différentes dans les sphères et le fond, de façon à créer le contraste désiré. Bien qu'il ne s'agisse là que d'objets

parfaitement sphériques avec des distributions d'activité homogènes tant dans le fond que dans les sphères, il s'agit d'une première approximation pour évaluer les performances des méthodes de segmentation des images TEP. La figure 4.1 illustre le fantôme en question et les images obtenues par acquisition TEP/TDM.

L'inconvénient des acquisitions de fantôme est de ne pas maîtriser parfaitement certains paramètres, comme l'homogénéité de l'activité au sein du fantôme. En effet, pour obtenir une distribution homogène, il est nécessaire de remuer l'eau contenue dans le fantôme en le remuant car on ne peut pas y introduire de quoi remuer le contenu. Ceci pour tenter de répartir le produit radioactif, qui est sous forme liquide également, le plus uniformément possible. Mais il est impossible d'atteindre une homogénéité parfaite, et il peut aussi exister des impuretés voir des particules de matière étrangère ou de la poussière à l'intérieur du fantôme, ce qui peut mener à des artefacts ou des inhomogénéités dans le fantôme. On est également fortement dépendant de la structure physique réelle du matériel, qu'on ne peut pas modifier à volonté. Le principal intérêt des acquisitions réelles est de pouvoir relativement facilement obtenir des images acquises avec différents modèles de scanners et algorithmes de reconstruction, pour étudier la robustesse des méthodes considérées.

4.1.2 Protocole d'acquisition commun

Les acquisitions du fantôme ont toutes été réalisées suivant un protocole similaire pour assurer la meilleure homogénéité possible des paramètres utilisés pour générer les images des différents modèles de scanner, et pouvoir étudier rigoureusement la robustesse des méthodes. Ceci a été réalisé en utilisant les caractéristiques et paramètres normalement considérés en routine clinique pour les équipements utilisés. Deux contrastes différents entre les sphères et le fond (environ 4-5 pour 1 et 8-10 pour 1) ont été considérés. Chaque acquisition a consisté en un scan anatomique TDM suivi d'une acquisition TEP de 5 minutes en mode 3D et enregistrement au format listmode de façon à pouvoir reconstruire plusieurs acquisitions de 1 ou 2 minutes à partir de l'acquisition réalisée. Les images furent ensuite reconstruites avec les paramètres et les algorithmes utilisés en routine clinique dans les centres où furent réalisées les acquisitions, à savoir l'algorithme RAMLA 3D pour le Philips GEMINI, OSEM pour le GE et le Siemens,

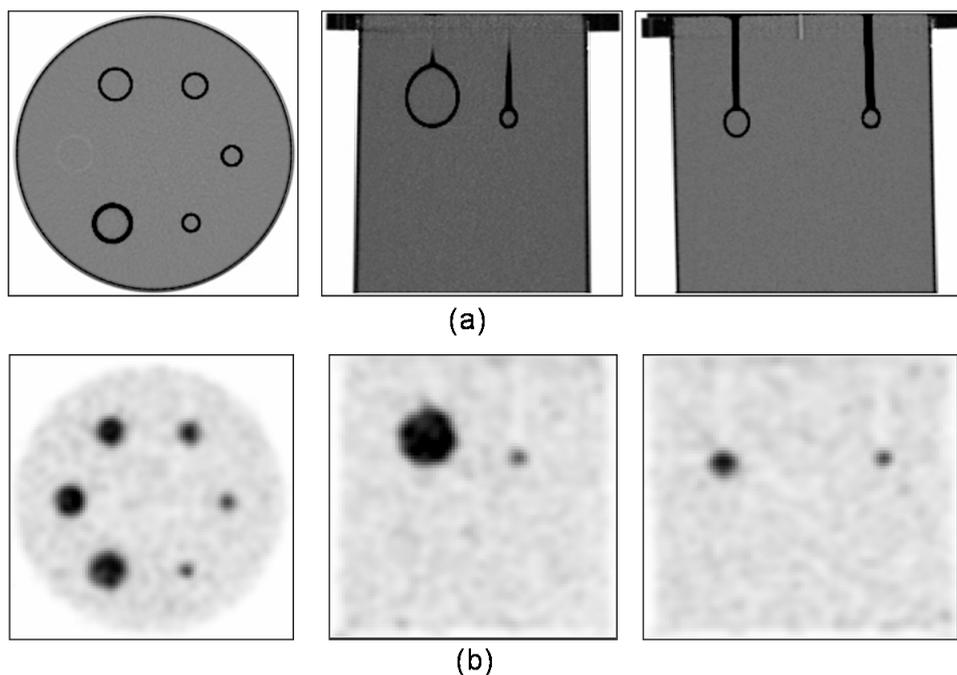


FIG. 4.1 – Illustration du fantôme IEC. (a) Acquisition TDM et (b) TEP (vues axiales, sagittales et coronales) du fantôme réalisées sur le TEP/TDM Philips Gemini.

et TF ML-EM pour le Philips GEMINI TF. Deux tailles de voxels furent à chaque fois considérées, à savoir $2 \times 2 \times 2$ mm et $4 \times 4 \times 4$ ou $5 \times 5 \times 5$ mm avec éventuellement des acquisitions spécifiques à chaque taille de voxels. Plus de détails sur chaque acquisition spécifique aux différents modèles de scanners sont détaillés plus loin.

4.1.3 Acquisitions du fantôme IEC obtenues avec différents scanners

Modèles de scanners considérés et paramètres généraux

Nous résumons ici les paramètres considérés pour chaque scanner, présentés dans la figure 4.2. Les paramètres utilisés sont : deux contrastes (de 8 à 10 pour 1 et 4 à 5 pour 1), trois durées d'acquisitions (1, 2 et 5 minutes) pour étudier la robustesse au bruit statistique, et deux tailles de voxels différentes (2 et 4 ou 5 mm de côté) pour étudier l'impact de l'échantillonnage spatial et du bruit résultant.

Scanner	Contrastes	Dimensions des voxels (mm)	Durées d'acquisition (min)	Reconstruction
Philips Gemini	4:1 8:1	2 x 2 x 2 4 x 4 x 4	1, 2, 5	RAMLA 3D
Philips TOF	4:1 10:1	2 x 2 x 2 4 x 4 x 4	1, 2, 5	TOF MLEM
Siemens Biograph	4:1 8:1	5.33 x 5.33 x 2 2 x 2 x 2	1, 2, 5	OSEM
GE Discovery	5:1 9:1	2.15 x 2.15 x 4.25	1, 2, 5	OSEM, FBP

FIG. 4.2 – Paramètres considérés pour chaque ensemble de données acquis avec différents types de scanner TEP/TDM.

Philips Gemini

La résolution spatiale du Philips Gemini (FWHM) est de 4.9 mm au centre du champ de vue [104]. Deux contrastes différents entre les sphères et le fond (4 et 8 pour 1) ont été obtenus en introduisant 7.4 kBq/mL dans le fond et 29.6 ou 59.2 kBq/mL dans les sphères. Chaque acquisition a consisté en un scan anatomique TDM haute résolution suivi d'une acquisition TEP de 5 minutes en mode 3D et enregistrement au format listmode de façon à pouvoir reconstruire plusieurs acquisitions de 1 ou 2 minutes à partir de l'acquisition réalisée. Des acquisitions séparées ont été réalisées pour chaque contraste et chaque taille de voxels. En effet les acquisitions destinées à être reconstruites avec une grille de voxels de taille 4x4x4 mm ont été réalisées suivant un protocole corps entier et un champ de vue complet. Les acquisitions destinées à être reconstruites avec une grille de voxels de 2x2x2 mm ont été réalisées avec un protocole d'acquisition du cerveau et un champ de vue réduit. Les images furent reconstruites avec l'algorithme itératif RAMLA 3D avec des paramètres optimisés auparavant [165], à savoir 2 itérations, un facteur de relaxation de 0.05 et un post filtrage gaussien 3D de FWHM de 5 mm. La sphère de 28 mm ayant été remplacée par une sphère de plastique dont le volume exact est inconnu, elle n'a pas été utilisée pour l'analyse des données et la comparaison des

performances des méthodes. De plus, le tube capillaire de remplissage de la sphère de 37 mm ayant été cassé et raccourci, cette sphère ne se trouve pas dans le même plan que les autres sphères, c'est la raison pour laquelle elle semble plus petite que ce qu'elle devrait dans les coupes axiales présentées dans cette thèse. Ceci n'empêche toutefois pas son utilisation dans l'analyse des performances car chaque sphère est isolée et traitée séparément.

Philips Gemini TF

La résolution spatiale du Philips Gemini TF (FWHM) est de 4.8 mm au centre du champ de vue et la résolution temporelle du système de coïncidences de 585 ps [150]. Deux contrastes différents entre les sphères et le fond (4 et 10 pour 1) ont été obtenus en introduisant 7 kBq/mL dans le fond et 28 ou 70 kBq/mL dans les sphères. Chaque acquisition a consisté en un scan anatomique TDM haute résolution suivi d'une acquisition TEP de 5 minutes en mode 3D et enregistrement au format listmode de façon à pouvoir reconstruire plusieurs acquisitions de 1 ou 2 minutes à partir de l'acquisition réalisée. Comme pour les acquisitions du Philips Gemini, deux types d'acquisitions différents ont été considérés : un champ de vue complet dans le cas des images en 4x4x4 mm, et un champ de vue réduit dans le cas des images en 2x2x2 mm. Ces images ont été reconstruites avec l'algorithme itératif ML-EM TF et les paramètres cliniques standards.

GE Discovery LS

Deux contrastes différents entre les sphères et le fond (5 et 9 pour 1) ont été obtenus en introduisant 7 kBq/mL dans le fond et 35 ou 63 kBq/mL dans les sphères. Chaque acquisition a consisté en un scan anatomique TDM haute résolution suivi d'une acquisition TEP de 5 minutes en mode 3D et enregistrement au format listmode de façon à pouvoir reconstruire plusieurs acquisitions de 1 ou 2 minutes à partir de l'acquisition réalisée. Les images ont ensuite été reconstruites avec des voxels de 2x2x2 mm avec l'algorithme OSEM en utilisant 2 itérations et 28 *subsets*. Noter que la sphère de 28 mm était manquante dans ce fantôme.

Siemens Biograph

Deux contrastes différents entre les sphères et le fond (4 et 8 pour 1) ont été obtenus en introduisant 7 kBq/mL dans le fond et 28 ou 56 kBq/mL dans les sphères. Chaque acquisition a consisté en un scan anatomique TDM haute résolution suivi d'une acquisition TEP de 5 minutes en mode 3D et enregistrement au format listmode de façon à pouvoir reconstruire plusieurs acquisitions de 1 ou 2 minutes à partir de l'acquisition réalisée. Les images ont été reconstruites avec l'algorithme itératif *Fourier rebinning* (FORE) OSEM [148] et les paramètres cliniques standards (4 itérations, 8 subsets) et un post-filtrage gaussien 3D de FHMW 5 mm. Des grilles de voxels de 2.03x2.03x2 mm et 5.33x5.33x2 mm ont été utilisées pour les reconstructions.

Images

Nous avons tenté d'obtenir les données les plus homogènes possibles entre les différents scanners, malgré l'absence d'échantillonnage avec des voxels de 4 ou 5 mm de côté pour les données du système GE Discovery LS. Notons également que les données Siemens n'ont pas été reconstruites, comme pour le Philips avec des voxels de 4 mm, mais 5.3 mm, ce qui est une différence relativement significative. Dans la mesure où les données des systèmes autres que Philips proviennent de centres différents par l'intermédiaire de collaborations à distance, nous n'avons eu que peu de contrôle sur ces acquisitions. Toutefois, les données utilisées sont tout de même suffisamment homogènes pour pouvoir être comparées. Les figures 4.3, 4.4, 4.5 et 4.6 illustrent visuellement les différences entre les différentes acquisitions pour chacun des systèmes. Comme on peut le constater dans ces figures, la représentation des objets, ici des sphères homogènes de diamètre variable, peut varier considérablement lorsque le système ou les paramètres d'acquisition et de reconstruction changent. En effet, le changement d'un seul paramètre peut avoir une répercussion importante sur l'aspect général de l'image, et plus particulièrement la représentation des objets d'intérêt, ici des sphères parfaitement homogènes, qui une fois imagées et reconstruites avec des voxels relativement grands par rapport à leur diamètres respectifs, ne sont plus ni vraiment des sphères, ni vraiment homogènes.

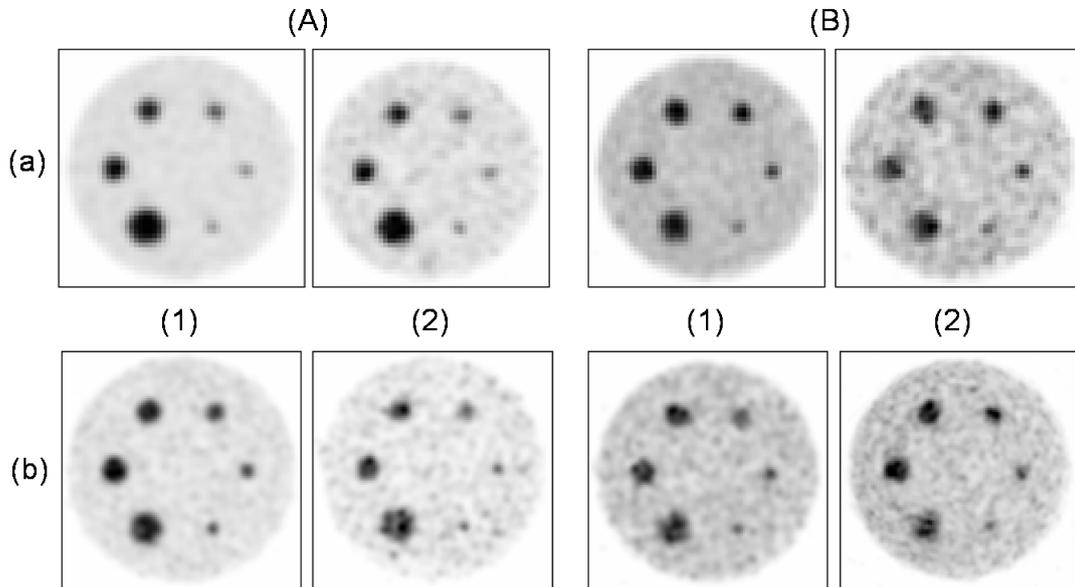


FIG. 4.3 – Images obtenues sur le Philips Gemini. Contrastes (A) 8 :1 et (B) 4 :1. Voxels de (a) 4x4 mm et (b) 2x2 mm. Acquisitions de (1) 5 min et (2) 1 min.

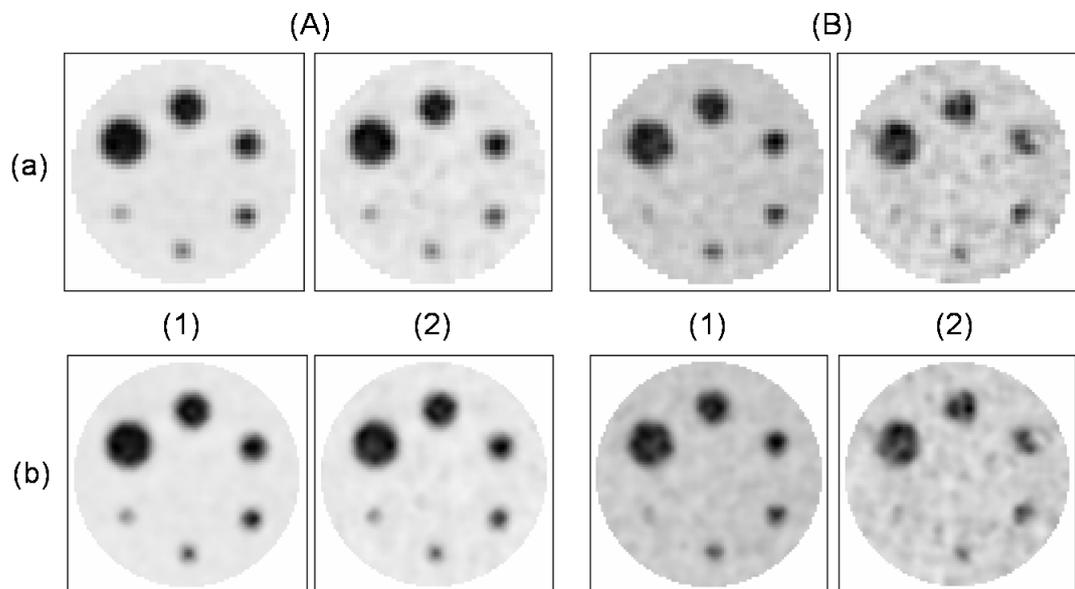


FIG. 4.4 – Images obtenues sur le Philips Gemini TF. Contrastes (A) 10 :1 et (B) 4 :1. Voxels de (a) 4x4 mm et (b) 2x2 mm. Acquisitions de (1) 5 min et (2) 1 min..

Si on considère le changement de contraste de 8 à 4, on peut parfois rendre la détection des plus petites sphères difficiles. On rend également nettement plus délicate la séparation de l'objet de

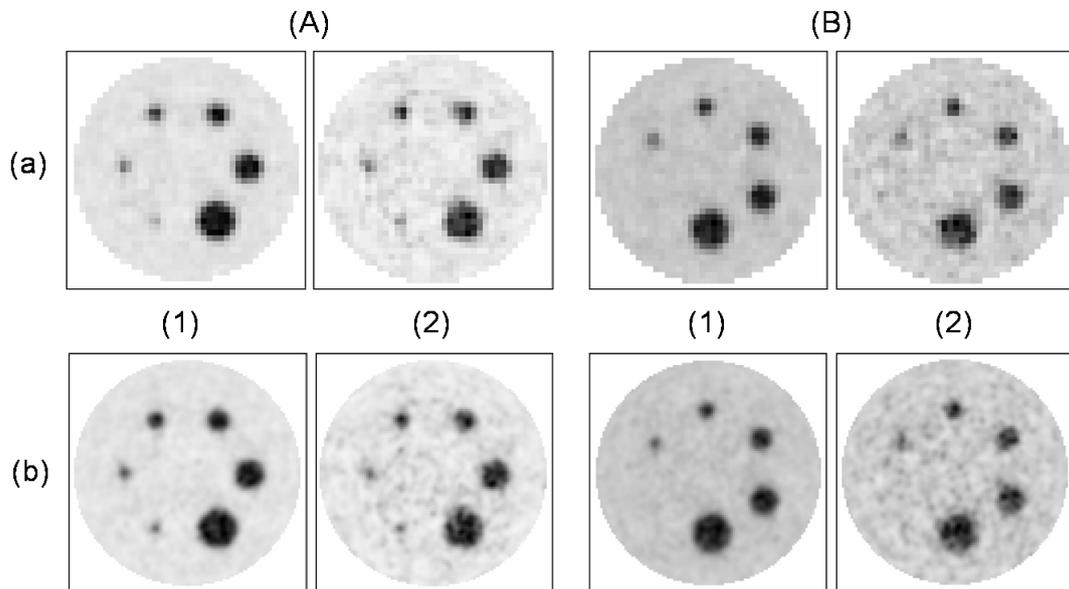


FIG. 4.5 – Images obtenues sur le Siemens Biograph. Contrastes (A) 8 :1 et (B) 4 :1. Voxels de (a) 5.3x5.3 mm et (b) 2x2 mm. Acquisitions de (1) 5 min et (2) 1 min.

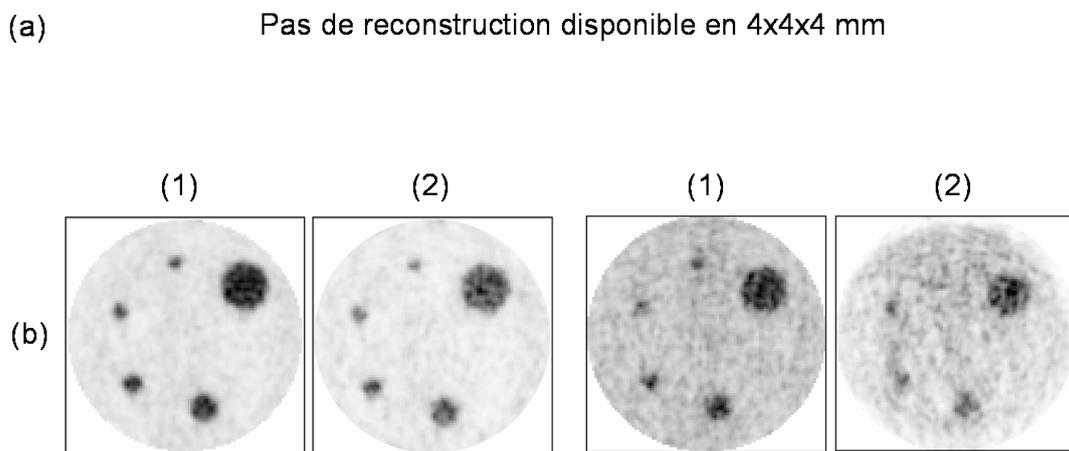


FIG. 4.6 – Images obtenues sur le GE Discovery LS. Contrastes (A) 9 :1 et (B) 5 :1. Voxels de (b) 2x2 mm. Acquisitions de (1) 5 min et (2) 1 min.

son fond environnant. En particulier, les seuillages fixes habituels échouent totalement dans cette configuration (voir les résultats présentés plus loin dans ce chapitre).

En changeant la taille des voxels, on obtient les changements visuels les plus importants, car on

modifie non seulement l'échantillonnage spatial des objets d'intérêt, mais on divise également les statistiques, ce qui augmente très significativement le bruit. On obtient donc des régions qui apparaissent bien plus hétérogènes qu'elles ne devraient, car rappelons que ces études de fantômes ne considère que des sphères homogènes sur un fond homogène également. On a donc un compromis entre la représentation des objets, et la fidélité qualitative et quantitative des distributions d'activité dans chaque région.

Quant à la durée d'acquisition, elle a un impact direct sur la qualité globale de l'image, avec des textures de bruit qui changent radicalement d'une configuration à l'autre, même sans considérer des tailles de voxels différentes.

Les images reconstruites avec l'algorithme RAMLA 3D dans le cas du Philips Gemini présentent un aspect plus lissé que les images reconstruites avec OSEM (pour les données GE et Siemens), qui présentent une structure de bruit plus "poivre et sel". On peut également constater qu'entre les données Philips reconstruites avec des voxels de 4 mm de côté, et les données Siemens reconstruites avec des voxels de 5.3 mm de côté, la définition des sphères est clairement à l'avantage des données Philips, illustrant l'importance de l'impact de la taille des voxels utilisée, malgré une résolution spatiale relativement similaire pour les deux appareils (environ 5 mm). La sphère de 10 mm notamment, reste visible sur toutes les configurations dans les images du Philips, mais est quasiment indétectable dans le cas des images Siemens sur les configurations de contraste à 4 :1.

Enfin, les images issues du scanner temps de vol présentent des images de qualité légèrement améliorées grâce aux performances concernant le contraste et le rapport signal sur bruit, liées à la technologie employée (voir section 1.5.3).

4.1.4 Images synthétiques d'objets hétérogènes

Nous avons également étudié le comportement de la méthode FLAB sur un objet synthétique (non fondé sur une tumeur réelle), avec trois classes, présentant une région d'intensité forte isolée dans la région d'intensité moyenne, et une autre région de forte intensité dans la région d'intensité moyenne, mais en contact avec le fond. Elles sont présentées dans la figure 4.7. Ainsi, les trois types de transitions floues sont présentes. Deux configurations de contrastes

et deux intensités différentes de bruit ont été utilisées pour générer quatre images de qualité variable. De plus, le même objet binarisé a été généré avec une résolution spatiale variable, de 6 à 12 mm (du TEP au TEMP) pour étudier la robustesse de la méthode au flou. Toutes les images ont été reconstruites avec un seul échantillonnage spatial de voxels de 2x2x2 mm.

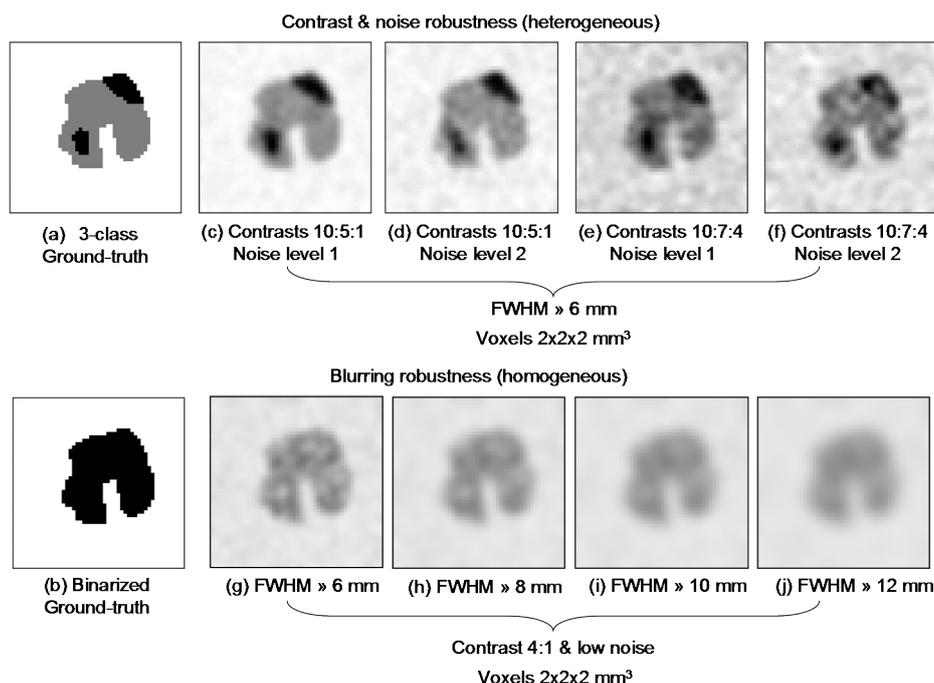


FIG. 4.7 – Objets synthétiques générés avec différentes conditions pour étudier la robustesse de l’algorithme FLAB.

4.2 Analyse et mesure de performance

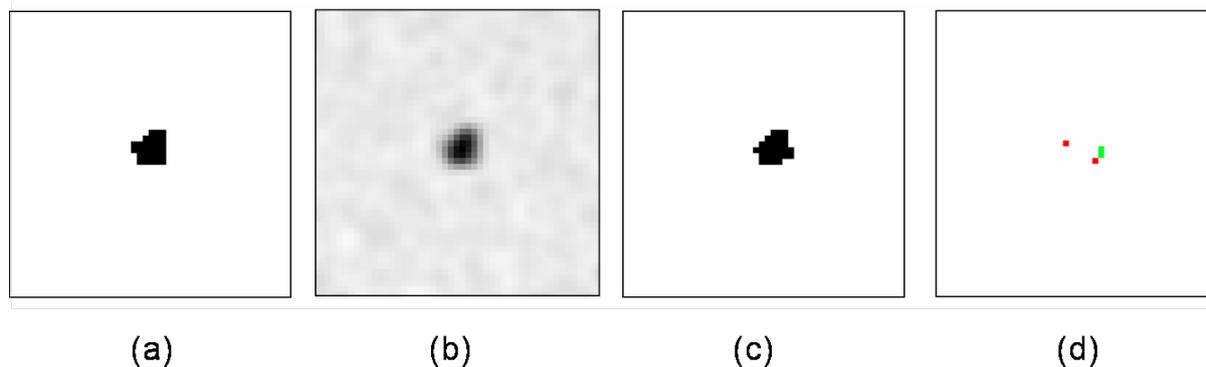
Rappelons que, dans la mesure où l’on se place dans un contexte de détermination de volume d’un objet particulier, il ne s’agit pas ici de détecter l’objet dans une image corps entier (voir 3.1.3). Nous partons du principe que les méthodes développées ici ont pour objectif d’être utilisées une fois qu’une lésion a été repérée par le clinicien dans un scan corps entier puis isolée. Les objets d’intérêt sont donc systématiquement isolés dans une boîte de sélection tri-dimensionnelle, et les méthodes de segmentation sont appliquées uniquement sur cette boîte. Par exemple, dans le cas des images de fantôme contenant plusieurs sphères de diamètre

variable, chaque sphère est isolée dans une boîte et traitée séparément des autres : on ne segmente pas le volume entier du fantôme.

Pour ce qui est de l'évaluation des performances des méthodes, lorsqu'une vérité terrain voxel à voxel est disponible, par exemple dans le cas de la plupart des simulations, on utilise une mesure d'erreur de classification (EC), à savoir le nombre de voxels mal classés par rapport au nombre de voxels définissant l'objet d'intérêt à segmenter (et non par rapport au nombre total de voxels de l'image) :

$$EC = \frac{VMC}{VO} \quad (4.1)$$

Où VMC est le nombre de voxels mal classés, et VO est le nombre de voxels définissant l'objet d'intérêt. Ceci donne en effet une mesure relative à la taille de l'objet, ce qui permet de mieux estimer les performances des méthodes car se tromper sur un voxel a moins d'impact sur une grande tumeur que sur une petite. Noter que dans le cas d'objets simulés hétérogènes, on peut calculer une EC par rapport à l'ensemble de l'objet (après binarisation de la vérité terrain et de la carte de segmentation), et pour chaque classe de l'image séparément, pour avoir une estimation de la capacité de l'algorithme à délimiter les régions au sein de l'objet en plus de son contour externe. La figure 4.8 illustre comment nous évaluons les performances d'une segmentation d'un objet d'intérêt et son intérêt par rapport à d'autres mesures plus simples. Cet exemple, illustré ici en deux dimensions mais le problème est identique en 3D, met en évidence les problèmes liés aux mesures de performances basées uniquement sur le volume ou l'erreur de classification par rapport à l'image entière. Dans le cas de l'erreur sur le volume, comme il y a autant de voxels du fond classés en tant qu'objet et inversement, l'erreur sur le volume final est nulle car les deux types d'erreurs de classifications s'annulent. Or il y a bien une erreur de segmentation. C'est évidemment un cas particulier, mais cela illustre bien les limites de cette mesure de performance, qui est par contre tout à fait valide dans les cas où le volume est entièrement sous-évalué (tous les voxels mal classés sont des voxels de l'objet classés en fond) ou inversement. Mais si les erreurs sont des deux types à la fois, cela fausse l'impression que l'on peut avoir de la performance d'une méthode. D'un autre côté, considérer le nombre de voxels mal classés par rapport au nombre total de voxels dans l'image (comme cela est souvent fait dans les travaux courants en segmentation statistique



$$\begin{aligned}
 EV &= (2+(-2)) / 29 = 0 \% \\
 EC_{\text{globale}} &= (2+2) / (50 \times 50) = 0.16 \% \\
 EC &= (2+2) / 29 = 8.89 \%
 \end{aligned}$$

FIG. 4.8 – Illustration des différences de résultats suivant la mesure utilisée pour évaluer les performances des méthodes de segmentation : (a) vérité terrain ; (b) image TEP simulée ; (c) segmentation ; (d) différences entre segmentation et vérité terrain. En rouge, les voxels de l'objet classés comme fond (entraînent une sous-évaluation du volume), en vert, ceux du fond classés comme objet (entraînent une sur-évaluation du volume). EC est l'erreur de classification telle que définie dans ce travail. EC globale est la même erreur mais calculée par rapport à l'ensemble de l'image et non l'objet. EV est l'erreur sur le volume.

d'image) est peu pertinent, notamment dans les cas où l'objet est de petite taille par rapport à la boîte de sélection. Ici, la boîte en question a été prise grande par rapport à l'objet (50 par 50 voxels), elle pourrait être plus petite mais cela a été choisi de façon à mettre bien en évidence le problème. En effet, ici l'erreur obtenue est seulement de 0.16 % ce qui paraît très peu. Pourtant 4 voxels sur les 29 que compte l'objet d'intérêt à segmenter ont été mal classés, ce qui indique tout de même une erreur de presque 9 % si on se réfère à la mesure que nous avons choisi. Noter que le désavantage de cette mesure est qu'elle ne donne pas d'indication quant à la sur- ou sous-évaluation du volume, à moins de séparer les deux types d'erreurs. Lorsqu'une telle vérité terrain n'est pas disponible, par exemple pour certaines acquisitions réelles, on utilise l'erreur commise par rapport au volume réel connu ou estimé (dans le cas des tumeurs réelles) de l'objet d'intérêt. De plus, dans le cas des tumeurs réelles de patients pour lesquels des résultats d'histologie sont disponibles, on compare le diamètre maximal mesuré

sur la carte de segmentation avec celui obtenu par la mesure réalisée en histologie.

Les graphiques présentés dans ce chapitre sont généralement présentés sous forme de barres d'erreurs pour différentes configurations de contrastes et de tailles d'objets en abscisse. A chaque barre d'erreur est associée une barre noire de variabilité, qui illustre les résultats obtenus dans différentes qualités statistiques d'image, et donc différents niveaux de bruit, et sont à interpréter comme une mesure de la robustesse au bruit d'une méthode : plus la barre de variation noire est petite, plus la méthode est robuste par rapport au bruit. La figure 4.9 illustre comment lire les graphiques. En règle générale, l'erreur figurant en ordonnée est limitée

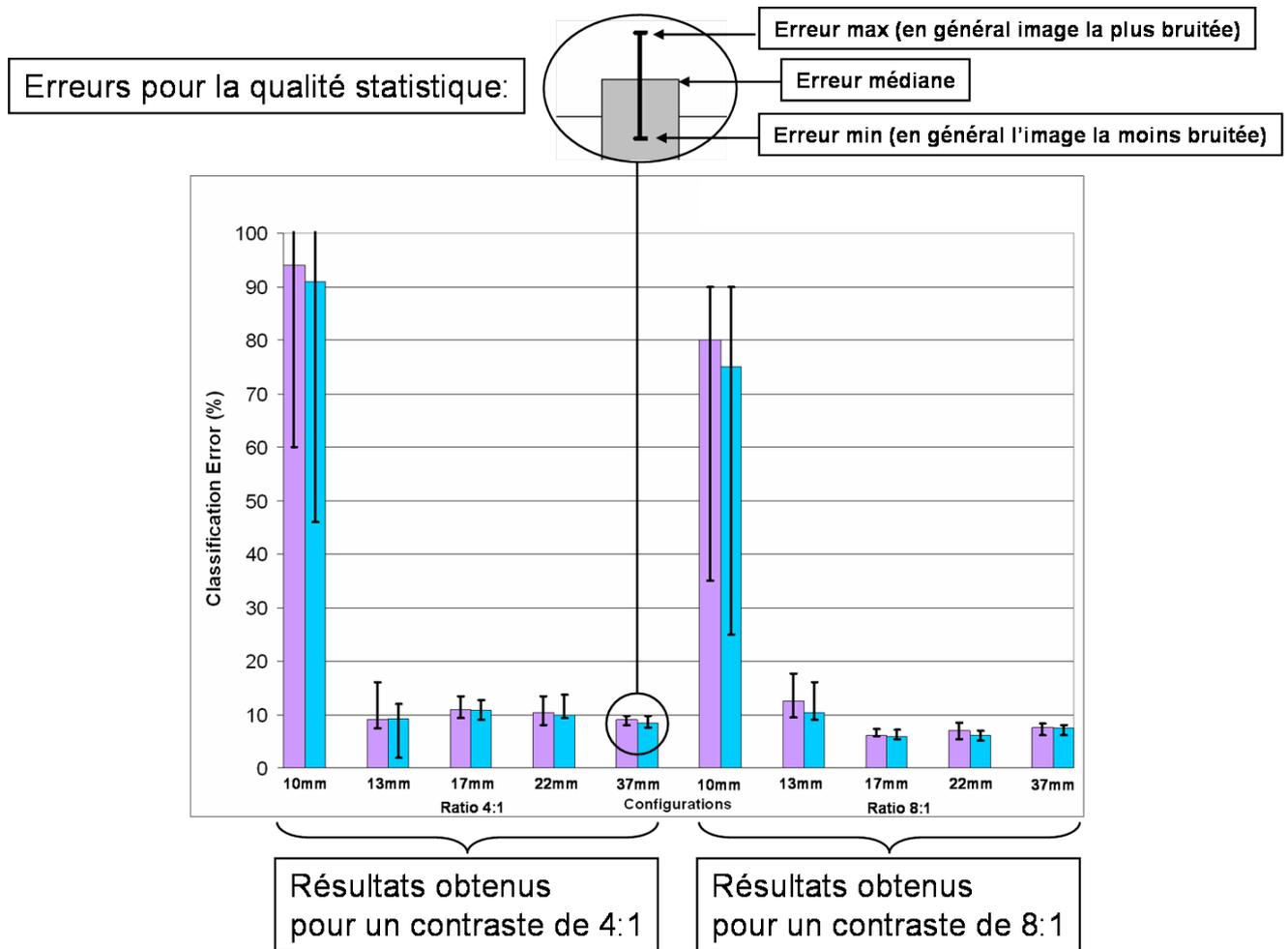


FIG. 4.9 – Comment lire les graphiques d'erreurs présentés dans la suite.

à 100% bien que certains résultats dépassent largement cette valeur. En effet, des échecs de

segmentation peuvent parfois mener à des valeurs de plusieurs centaines de %, en particulier pour de petits objets d'intérêt. Nous n'affichons pas de telles valeurs sur les graphes, car déjà à partir de plus de quelques dizaines de %, la segmentation peut être considérée comme un échec.

4.3 Optimisation sur sphères homogènes

Les résultats sur fantôme permettent de juger la capacité des méthodes de segmentation testées à extraire un objet homogène d'un fond homogène. Ceci n'est qu'un cas idéal sur des objets sphériques, toutefois cela permet d'avoir une bonne idée de la robustesse des méthodes par rapport au contraste et au bruit, car il est relativement facile d'étudier des configurations différentes en terme de temps d'acquisition (et donc de bruit), et de contraste. De plus, l'étude sur fantôme permet de tester les limites des méthodes en terme de taille d'objet par rapport à la résolution spatiale du scanner, ainsi que l'influence de paramètres tels que la taille des voxels de l'image reconstruite. En effet, un échantillonnage spatial utilisant des voxels plus petits permet de mieux représenter les objets de petite taille, de l'ordre du centimètre, mais l'image est plus bruitée. Passer par exemple de 4 mm de côté à 2 mm de côté, divise le volume d'un voxel par 8 (8 mm^3 au lieu de 64), ce qui revient approximativement à diviser le nombre de lignes de réponse passant par ce voxel du même ordre de grandeur. Pour un même temps d'acquisition, la statistique de l'image étant ainsi fortement diminuée, le bruit est significativement plus important. On peut constater la différence de qualité entre les images dans les figures 4.3, 4.4, 4.6 et 4.5. Des voxels plus grands permettent d'obtenir des contrastes et un rapport signal sur bruit plus élevés, mais les objets de petite taille sont représentés par très peu de voxels ce qui complique leur segmentation. En effet, une sphère de 10 mm de diamètre est très mal échantillonnée par une grille de voxels de 4 mm de côté. La figure 4.10 illustre les différents résultats obtenus pour une sphère de 10 mm avec différents contrastes et tailles de voxels dans l'image reconstruite. Un autre intérêt des images acquises du fantôme est de servir à l'optimisation d'une méthode de segmentation. En effet, l'approche par chaînes de Markov floues (FHMC, voir section 3.2) et l'approche locale adaptative (FLAB, voir section

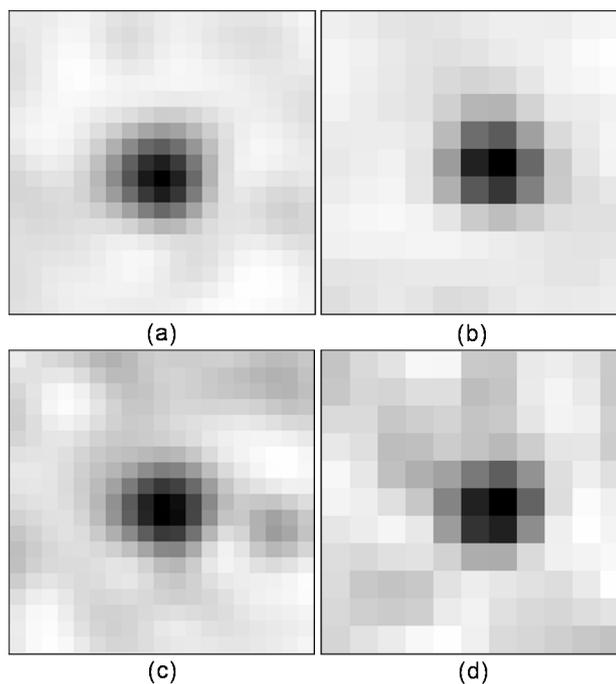


FIG. 4.10 – Différentes représentations d’une sphère de 10 mm sur les acquisitions réelles du fantôme IEC (5 minutes d’acquisition, ce qui correspond au bruit le plus faible considéré dans cette étude) sur Philips Gemini et reconstruites par Ramla 3D. (a) contraste de 8 :1, voxels de 2x2x2 mm ; (b) contraste de 8 :1, voxels de 4x4x4 mm ; (c) contraste de 4 :1, voxels de 2x2x2 mm ; (d) contraste de 4 :1, voxels de 4x4x4 mm.

3.3) utilisent des paramètres qu’il faut optimiser pour une application donnée. Pour déterminer la meilleure configuration de ces paramètres, l’erreur moyenne sur l’ensemble des sphères et pour l’ensemble des paramètres d’acquisition (contraste, bruit, taille de voxels) a été considérée pour chaque combinaison de ces paramètres. La combinaison de paramètres minimisant cette erreur fut ensuite choisie comme implémentation pratique des algorithmes.

Les sections qui suivent contiennent des résultats typiques obtenus et publiés dans différents articles, lors de l’optimisation des méthodes.

4.3.1 Optimisation du nombre de niveaux de flou et valeurs associées

La figure 4.11 illustre les résultats obtenus en faisant varier le nombre de niveaux de flou utilisés dans la segmentation et pour déterminer le volume définitif de l’objet d’intérêt, dans

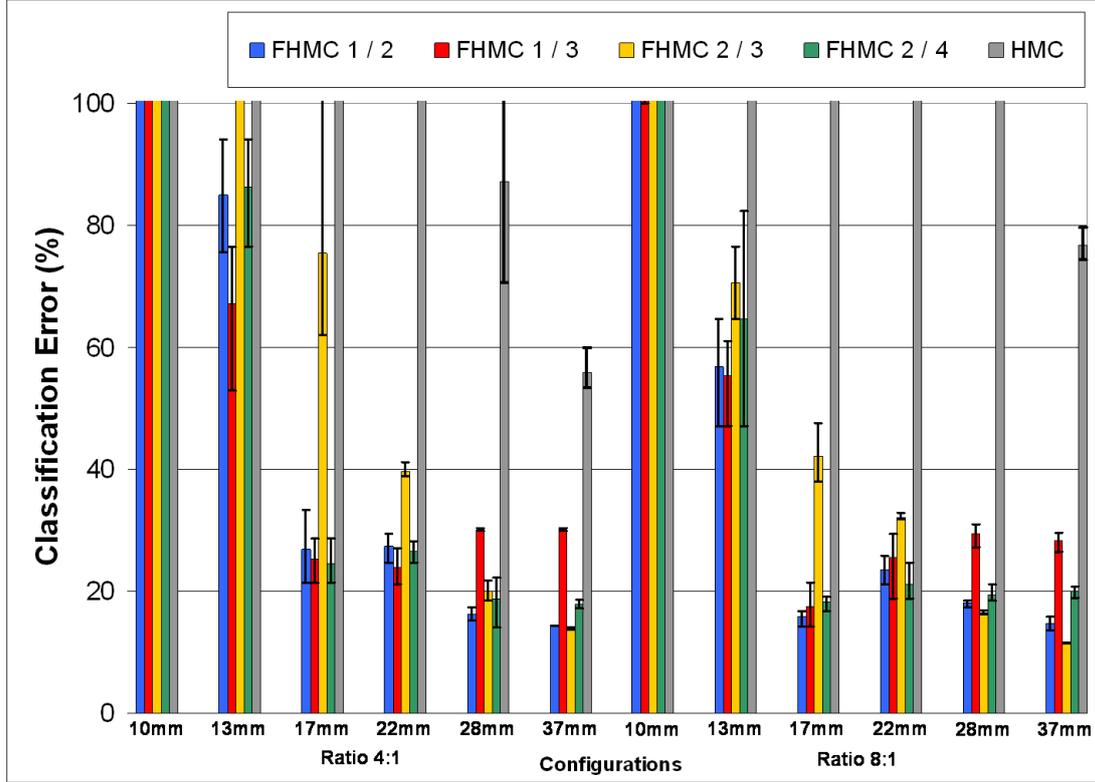


FIG. 4.11 – Erreurs de classification obtenus par les chaînes de Markov floues (FHMC) sur les images du fantôme en considérant différentes configurations de niveaux de flou. Pour comparaison, les résultats obtenus par une approche par chaînes dures (HMC), c'est à dire avec seulement deux classes dures sans modélisation floue, sont également présentés.

le cas des chaînes de Markov (FHMC). En effet, les deux approches FLAB et FHMC partagent une modélisation floue similaire, c'est à dire un espace de valeurs associées aux voxels flous qui est divisé en un certain nombre de niveaux de flous F_i , auxquels sont associées des valeurs ε_i (voir section 3.1.7). Lorsqu'on utilise une telle approche pour segmenter un volume d'intérêt sur une image TEP, outre la question du nombre de classes (pour l'instant, comme nous étudions les résultats sur fantôme, et donc sur des sphères homogènes, ce nombre est fixé à deux), il est nécessaire de définir combien de niveaux de flou nous utilisons, quelle valeur donner à ces niveaux de flou, et enfin comment utiliser la carte de segmentation obtenue pour définir le volume fonctionnel segmenté. Nous avons donc étudié l'influence du nombre de niveaux de flous, les valeurs qui leur sont associées, et comment utiliser ces niveaux de

fou pour définir une segmentation définitive. L'approche la plus intuitive consiste à considérer deux niveaux de fou, un représentant les voxels de l'objet dont les valeurs ont été réduites par la présence adjacente du fond, et l'autre associé aux voxels du fond dont les valeurs ont été rehaussées par la présence de l'objet (l'effet de *spillover* associé aux effets de volume partiel). Pour déterminer le volume final de l'objet, on peut alors associer les voxels classés en classe dure 1 et le deuxième niveau de fou pour représenter le volume fonctionnel, tandis que les voxels classés comme classe dure 0 et le premier niveau de fou sont eux considérés comme les voxels du fond. Les résultats présentés dans la figure 4.11 démontrent que cette approche est effectivement la plus performante si l'on considère l'ensemble des sphères. En effet, certaines autres configurations donnent de meilleurs résultats si on considère une sphère de façon isolée. Mais comme l'algorithme doit être automatique, le nombre de niveaux de fou ne doit pas être variable en fonction de l'objet à segmenter, et c'est donc la configuration donnant l'erreur minimale en considérant l'ensemble des sphères qui a été retenue. Dans cette figure, la notation FHMC x / y signifie que y niveaux de fou sont utilisés au total dans la segmentation, et que x niveaux sont conservés pour calculer le volume de l'objet. HMC désigne les chaînes de Markov sans modélisation floue (deux classes dures uniquement). On peut voir dans cette figure qu'utiliser plus que deux niveaux de fou, par exemple trois voire quatre, n'apporte pas de précision supplémentaire sur la définition du volume, notamment à cause du nombre réduit de voxels. On peut noter par exemple, que la configuration 1 / 3 donne de bons résultats sur les plus petites sphères que 1 / 2, mais par contre génère une sous-évaluation importante dans le cas des grandes sphères, avec des erreurs de classification nettement supérieures à la configuration 1 / 2. Il se pose de plus la question de la valeur associée à un niveau de fou. On peut en effet définir la discrétisation de l'intervalle $[0, 1]$, et donc les valeurs des ε_i de plusieurs façons. Par exemple $\{0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}, 1\}$ où $N - 1$ est le nombre de niveaux de fou, ce qui mène à $\varepsilon_i = \frac{i}{N}$, et donc dans le cas de deux niveaux de fou, $\varepsilon_1 = \frac{1}{3}$ et $\varepsilon_2 = \frac{2}{3}$. Toutefois, une autre façon de discrétiser l'intervalle est proposée dans [27] notamment, où les valeurs 0.25 et 0.75 (la médiane du sous intervalle considéré) sont suggérées pour deux niveaux. Nous avons donc testé les deux implémentations, et dans le cas de deux niveaux de fou, utiliser $\varepsilon_1 = \frac{1}{3}$ et $\varepsilon_2 = \frac{2}{3}$ donne de meilleurs résultats que $\varepsilon_1 = \frac{1}{4}$ et $\varepsilon_2 = \frac{3}{4}$ (voir figure 4.12 pour l'illustration

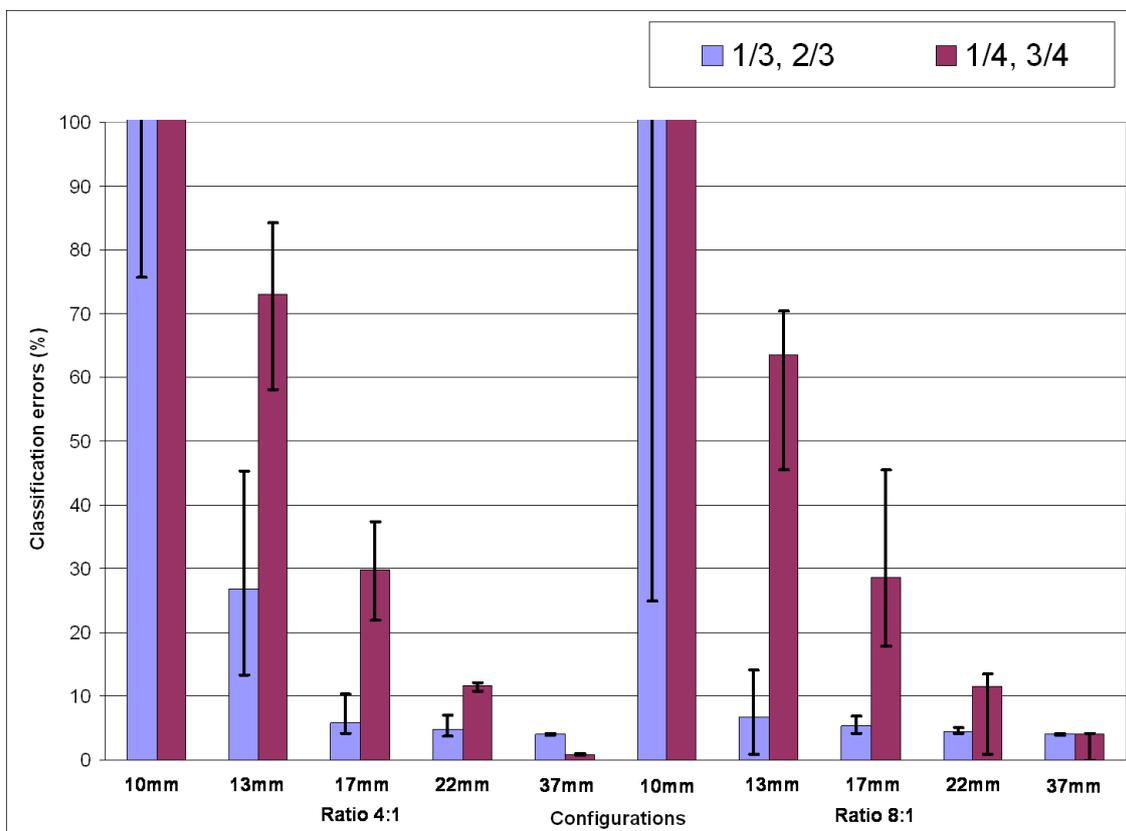


FIG. 4.12 – Erreurs de classification obtenues par l’approche locale adaptative (FLAB) sur les images du fantôme en considérant une implémentation avec 2 niveaux de flou et différentes valeurs pour chaque niveau de flou.

dans le cas de FLAB). C’est donc cette implémentation qui a été choisie pour FHMC et FLAB.

4.3.2 Optimisation de la loi des observations

Le type de loi utilisée pour modéliser la probabilité d’observation, autrement dit la loi du bruit, peut changer les résultats de segmentation. Nous avons testé l’hypothèse simple gaussienne ainsi que le modèle de Pearson dans le cas de FLAB pour étudier les éventuelles améliorations. Plus de résultats seront présentés dans la section dédiée à l’étude du bruit dans les images TEP en considérant la variabilité des images obtenues avec différents scanners et algorithmes de reconstruction et les protocoles d’acquisitions associés (4.6).

La figure 4.13 illustre la comparaison entre l'utilisation de lois gaussiennes ou de lois bêta

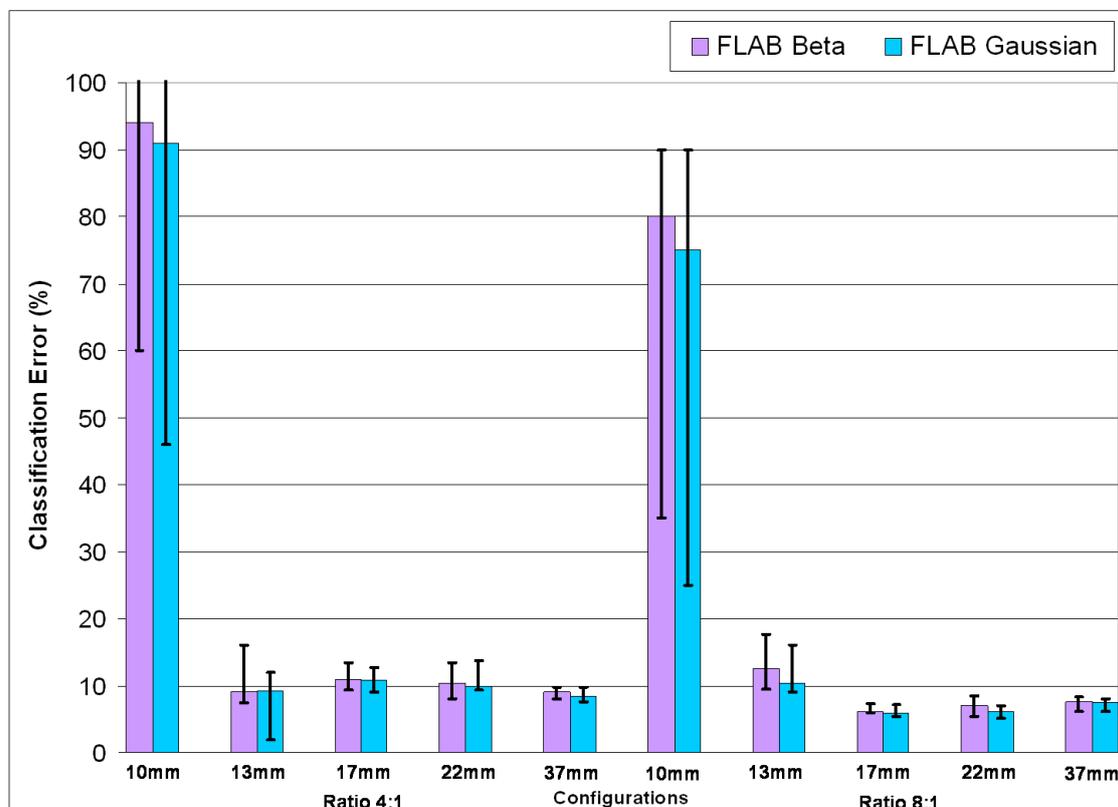


FIG. 4.13 – Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini (avec voxels de $2 \times 2 \times 2$ mm) en considérant des lois gaussiennes ou bêta I dans l'implémentation de FLAB.

I pour les distributions des voxels du fond et de l'objet d'intérêt, dans le cas de l'approche locale (FLAB) et sur des données acquises du fantôme. Le modèle de Pearson a été utilisé pour déterminer les types de lois associées à chaque classe des images. Dans le cas des images de fantôme, on peut facilement déterminer les classes et obtenir les distributions des voxels associés à chaque classe, ce qui permet de calculer les paramètres d'intérêt du système de Pearson γ_1 et γ_2 , puis de repérer les distributions associées sur le graphe de Pearson (voir figure 3.7). Cette étude a montré que bien que le système de Pearson détermine des distributions bêta I (pour le fond et l'objet, bien qu'avec des paramètres légèrement différents), ces dernières sont très proches de la distribution gaussienne sur le graphe de Pearson. De

fait, en implémentant l'algorithme avec des lois bêta I à la place des lois gaussiennes dans l'approche locale FLAB, peu de différence (de l'ordre de 5 à 10% de variation) fut constatée sur les résultats de segmentation des sphères du fantôme. Comme l'indique la figure 4.13, les résultats étaient globalement même un peu moins bons notamment sur les petites sphères. La loi gaussienne a donc été conservée dans l'implémentation pratique de l'algorithme.

4.3.3 Optimisation de la taille du cube d'estimation dans le cas de FLAB

Dans le cas de l'approche locale, la taille du cube d'estimation peut avoir une influence significative sur l'estimation des paramètres. Nous avons testé plusieurs tailles de cube, trois ou cinq voxels de côté.

Les figures 4.14 et 4.15 illustrent les différences de performances de l'approche locale, en

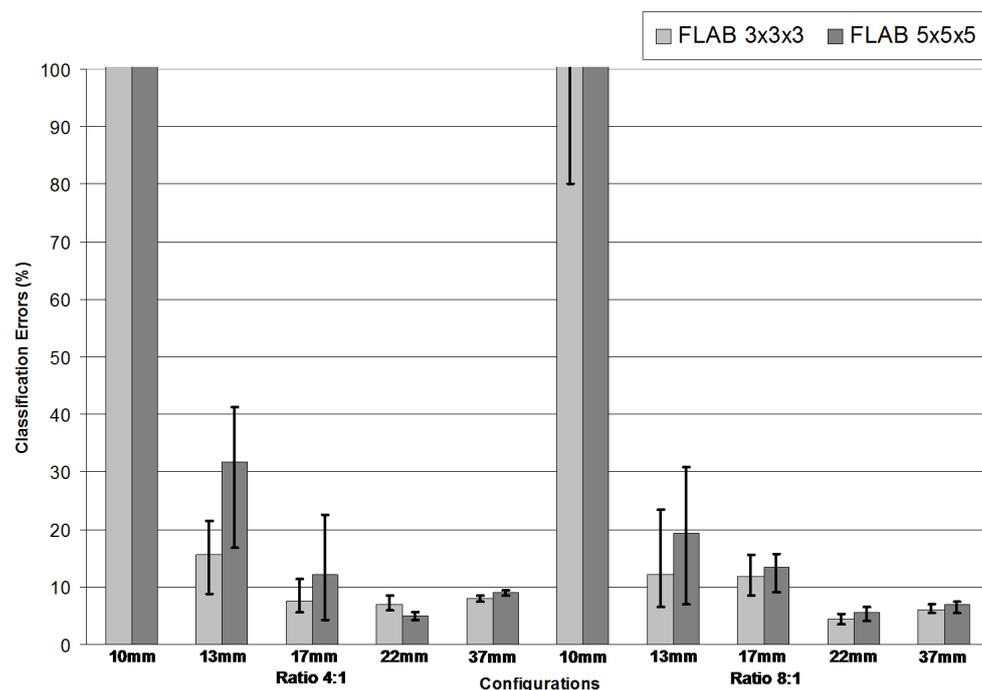


FIG. 4.14 – Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini en considérant un cube d'estimation de taille 3x3x3 ou 5x5x5 voxels dans l'implémentation de FLAB. Images avec voxels de 4x4x4 mm.

considérant un cube d'estimation adaptatif des probabilités *a priori* de deux tailles différentes

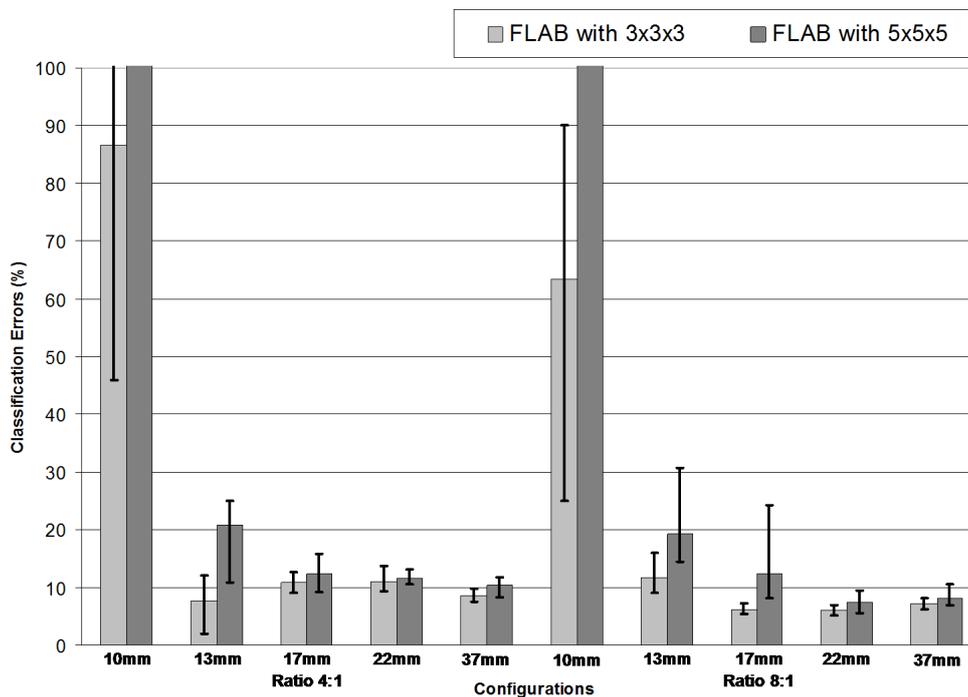


FIG. 4.15 – Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini en considérant un cube d'estimation de taille 3x3x3 ou 5x5x5 voxels dans l'implémentation de FLAB. Images avec voxels de 2x2x2 mm.

(trois ou cinq voxels de côté respectivement), sur les données d'acquisitions réelles. Considérant que la taille du cube doit être déterminée en fonction de la taille des objets d'intérêt (entre 10 mm et 50 mm environ pour la plupart des tumeurs) et de la taille des voxels généralement utilisée en routine clinique (entre 2 et 5 mm de côté), et qu'il doit être suffisamment petit pour pouvoir suivre des contours dans le cas d'objets non sphériques, ainsi que détecter correctement des petites régions au sein d'objets plus grands, nous avons testé deux tailles de cube d'estimation : trois ou cinq voxels de côté. Les résultats résumés dans les figures 4.14 et 4.15 montrent que dans tous les cas, c'est globalement le cube de taille 3x3x3 qui est optimisé. Les différences sont notables sur les petites sphères, pour lesquelles un cube de cinq voxels est visiblement trop grand, tandis que sur les grandes sphères, l'utilisation d'un cube de trois voxels semble suffisant pour donner globalement les mêmes résultats qu'avec un cube de cinq voxels. De plus, dans la mesure où les temps de calcul nécessaires à l'algorithme sont linéairement dépendants du nombre de voxels à explorer dans le voisinage, utiliser un cube de

trois voxels (27 voisins) est environ vingt fois plus rapide qu'un cube de cinq (524 voisins).

4.3.4 Importance de l'estimation adaptative dans FLAB

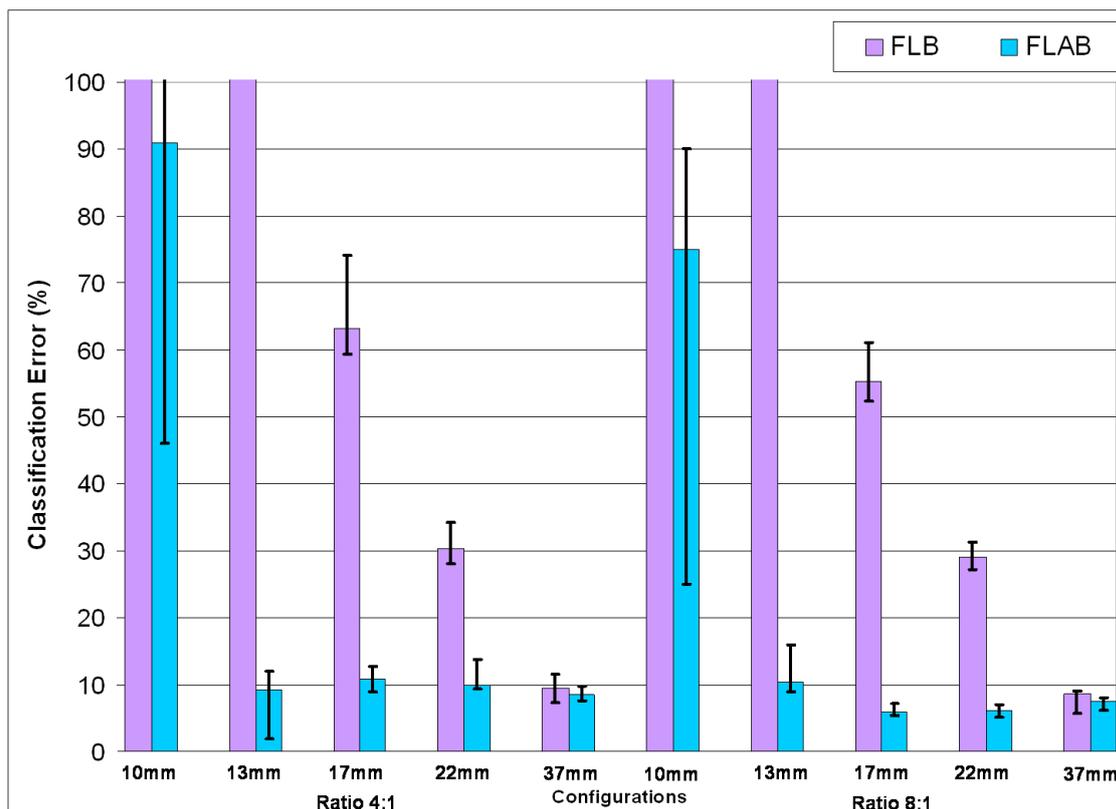


FIG. 4.16 – Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini (avec voxels de 2x2x2 mm) en considérant l'approche adaptative (FLAB) avec un cube d'estimation de taille 3x3x3 et l'approche "aveugle" sans estimation adaptative (FLB).

Pour illustrer l'importance de l'estimation adaptative des probabilités *a priori*, nous illustrons dans la figure 4.16 la différence de performance entre une approche "aveugle", c'est à dire sans estimation adaptative (FLB), et l'approche adaptative (FLAB) utilisée avec un cube de 3x3x3 sur le même jeu de données. Les résultats sont particulièrement éloquents, avec des erreurs nettement plus importantes avec l'approche "aveugle". Cette dernière parvient à extraire la plus grosse sphère avec une performance similaire (bien que légèrement moins bonne, elle

reste inférieure à 10%) à celle de FLAB, mais sa robustesse par rapport au bruit est moins bonne, comme l'indiquent les barres de variabilité. Pour les sphères de diamètre inférieur, les performances de FLB se dégradent très vite, avec une EC de l'ordre de 30% pour la sphère de 22 mm, et déjà supérieure à 50% pour celle de 17 mm, là où FLAB restent aux alentours des 10% voir en dessous jusqu'à la sphère de 13 mm.

4.3.5 Reproductibilité de FLAB

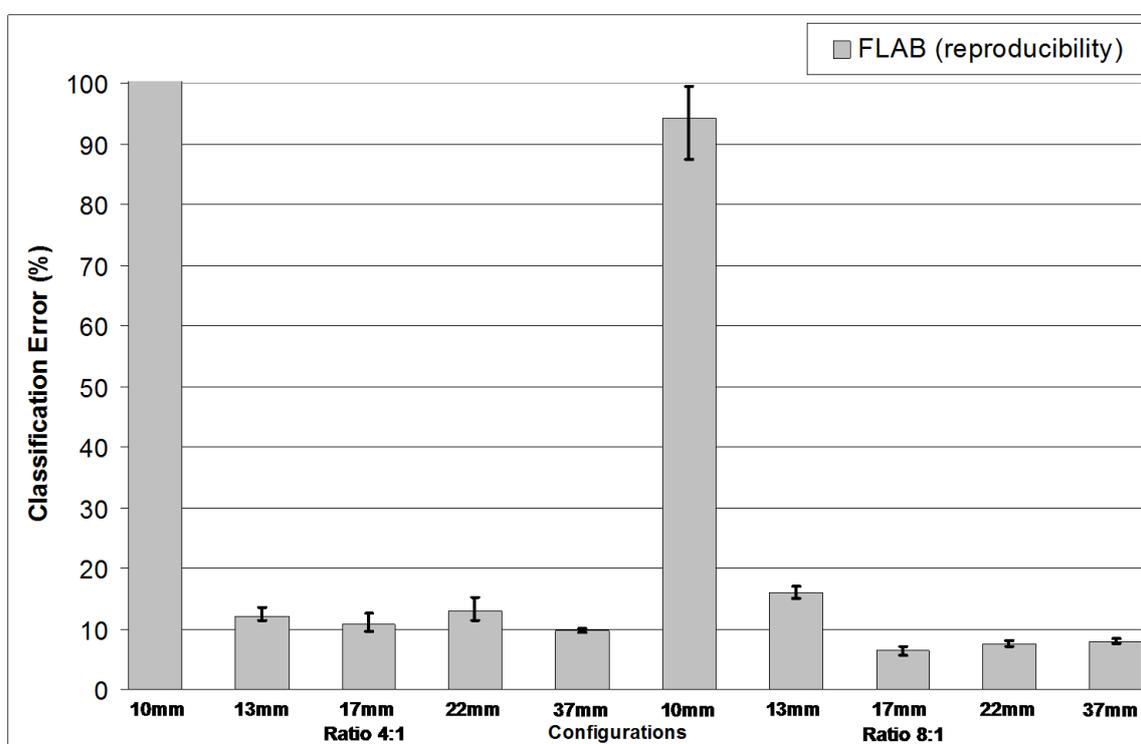


FIG. 4.17 – Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini (avec voxels de 2x2x2 mm) en considérant l'approche adaptative (FLAB) optimisée, sur cinq acquisitions indépendantes de 1 min. Contrairement aux autres graphiques, les barres noires de variation indiquent ici l'écart type des erreurs obtenues sur les cinq réalisations indépendantes de une minute chacune.

la figure 4.17 illustre non pas la robustesse mais la reproductibilité de FLAB, en appliquant l'approche sur différentes réalisations d'une même acquisition, c'est à dire cinq images différentes de même qualité statistique (une minute d'acquisition pour chaque image) mais

correspondant à des acquisitions indépendantes. Dans cette figure, contrairement aux autres, les barres de variabilité correspondent à l'écart type des erreurs sur l'ensemble de ces réalisations de même qualité statistique, et non à la robustesse par rapport au bruit (qui est la représentation adoptée pour les autres figures, où les barres de variabilité illustre la variabilité des erreurs en fonction des durées d'acquisition, une, deux ou cinq minutes). La reproductibilité de FLAB a donc été étudiée en considérant cinq acquisitions indépendantes de 1 minute chacune. L'erreur moyenne et l'écart type associé sont donnés pour chaque configuration de sphère et de contraste dans la figure 4.17. Il s'agit donc de vérifier qu'en considérant différentes réalisations d'une même vérité, la variabilité statistique des voxels due à la reconstruction d'image et à la nature même des acquisitions (les cinq acquisitions ne faisant qu'une minute chacune, elles sont assez bruitées) ne handicape pas la méthode. Ici, le faible écart type (de l'ordre de 2 à 4 % sur les sphères dont le diamètre est supérieur à 10 mm) démontre clairement la reproductibilité de la segmentation par FLAB.

4.4 Erreurs de classification et de volume sur les acquisitions de fantôme

4.4.1 Resultats

Les résultats sont ici exprimés en erreur de classification par rapport à la vérité terrain voxel à voxel générée grâce aux acquisitions TDM correspondantes et recalées sur les images TEP. Ils sont également condensés en un seul tableau comparant toutes les méthodes considérées. Les mêmes résultats sont présentés en erreur par rapport au vrai volume de la sphère, et méthode par méthode dans l'annexe (section 7.4).

Les résultats obtenus sur les images réelles sont présentés dans les figures 4.18 et 4.19 pour les tailles de voxels 4 et 2 mm respectivement. On compare ici les performances des approches d'un seuillage fixe à 42 % (T42), le Fuzzy C-Means (FCM) avec deux clusters, et les deux approches statistiques floues, l'approche par chaînes (FHMC) et l'approche locale adaptative (FLAB). Ces deux dernières étant ici utilisées avec les paramètres précédemment optimisés, à savoir deux niveaux de flou avec les valeurs associées $1/3$ et $2/3$, et dans le cas de FLAB, un

cube d'estimation de 3x3x3 voxels. Noter que les erreurs concernant la sphère de 28 mm sont manquantes (voir section 4.1.3) car celle-ci était exclue de l'analyse (voir section 4.1.3). En

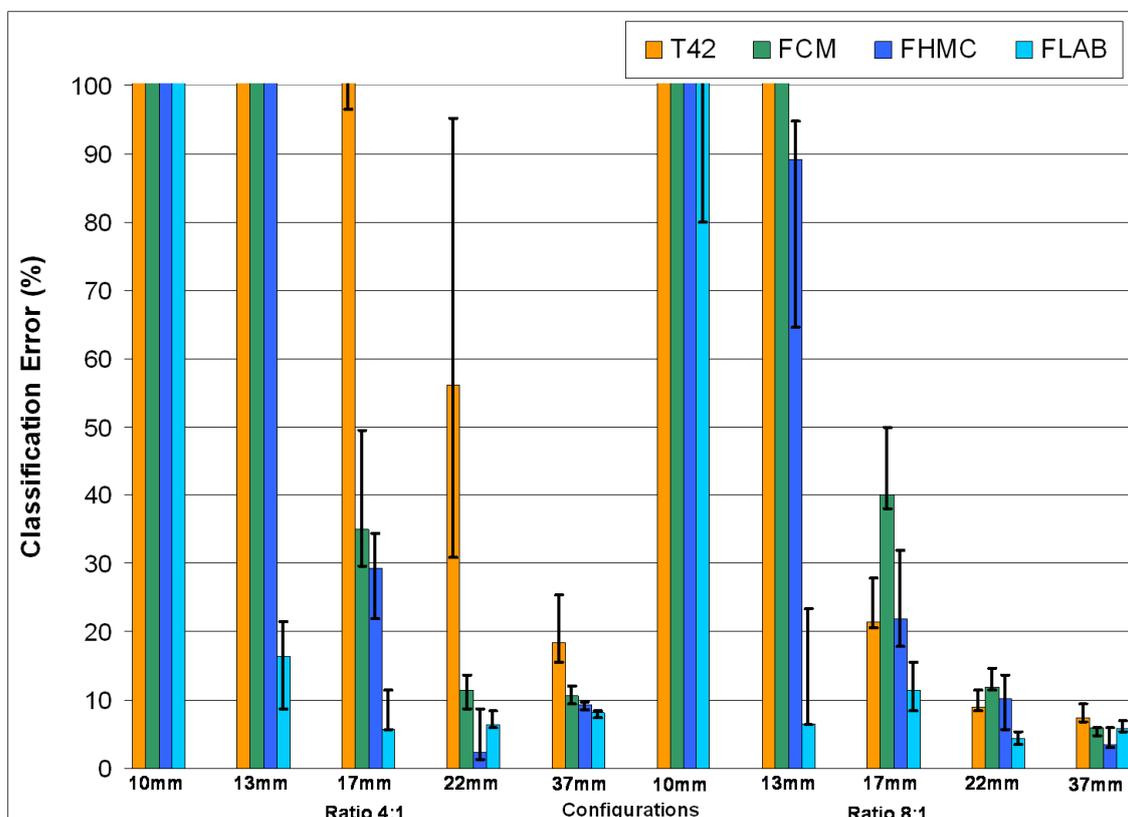


FIG. 4.18 – Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.

observant les figures 4.18 et 4.19, on observe que la qualité statistique des images reconstruites avec des voxels de 2 mm a été suffisamment conservée par la reconstruction RAMLA 3D pour permettre non seulement une segmentation correcte, mais également une amélioration des résultats par rapport à la reconstruction avec des voxels de 4 mm. En effet, les erreurs en valeur absolue sont plus faibles (autour de 10% au lieu de 20), mais elles sont aussi globalement meilleures avec des voxels plus petits pour FLAB. Ceci démontre que la robustesse au bruit de FLAB et FHMC leur permet de bénéficier de l'amélioration de l'échantillonnage spatial pour mieux délimiter les objets qui sont définis par plus de voxels, tout en résistant à l'augmentation du bruit résultant et donc d'éviter de mal classer des voxels du fond dont la valeur pourrait

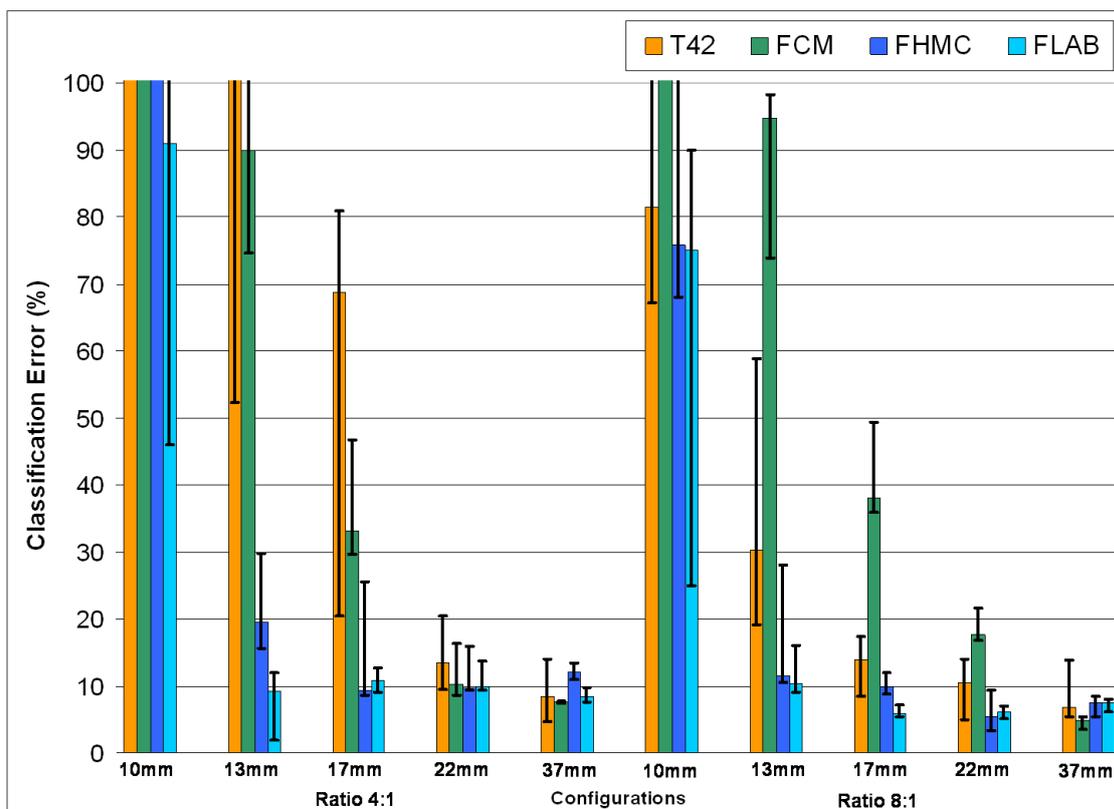


FIG. 4.19 – Erreurs de classification obtenues sur les sphères du fantôme IEC imagé par le Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.

être proche de celle de la sphère à cause du bruit. Sur cet ensemble de données, FLAB fait preuve d'une plus grande efficacité que FHMC sur les sphères de 13 et 17 mm, en particulier avec des voxels de 4 mm, configuration dans laquelle FHMC génère des erreurs supérieures à 20 et 60 % pour 17 et 13 mm respectivement, tandis que FLAB permet de rester dans tous les cas sous la barre de 20 % dans le cas des voxels de 4 mm, et sous 10 % avec des voxels de 2 mm. La sphère de 10 mm reste toujours problématique, particulièrement avec un faible contraste. L'erreur minimale est atteinte par FLAB avec environ 50 % d'erreur avec le contraste de 4 pour 1 et 25 % dans le cas d'un contraste de 8 pour 1, avec des voxels de 2 mm. Avec des voxels plus grands, aucune méthode ne parvient à extraire la plus petite sphère du fond. Les performances de FCM et T42 s'améliorent également de façon significative quand la taille des voxels diminue, mais leur variabilité par rapport au bruit reste importante, comme le

montrent les barres de variabilité. De plus, si elles offrent des performances similaires à FHMC et FLAB sur les sphères supérieures à 2 cm, leurs résultats sont totalement aberrants sur les plus petites sphères, avec d'énormes variabilité suivant la durée d'acquisition. Par exemple, l'erreur de T42 varie de 20 à 80 % entre 5 et 1 min d'acquisition sur la sphère de 17 mm et un contraste de 4 pour 1 et des voxels de 2 mm.

4.4.2 Conclusions de l'étude sur fantôme

L'étude du comportement des méthodes développées sur les différentes images de fantôme contenant des objets sphériques a permis de tirer plusieurs conclusions importantes. Tout d'abord, les méthodes statistiques floues (FLAB et FHMC) ont globalement des performances nettement supérieures à un simple clustering type Fuzzy C-Means ou à une approche de seuillage, ce qui était espéré. Elles démontrent non seulement une précision correcte sur l'ensemble des sphères exceptée la plus petite, qui reste un défi, mais également une robustesse supérieure vis à vis des variations des paramètres d'acquisitions (bruit et contraste notamment) avec des performances acceptables dans la plupart des configurations étudiées, là où le seuillage ou le FCM peuvent présenter d'immenses variations entre deux contrastes ou deux niveaux de bruit différents. L'influence de la taille des voxels est également particulièrement importante. Il est important de noter que dans le cas de FLAB, il vaut mieux utiliser des petits voxels avec un protocole d'acquisition adapté pour bénéficier du meilleur échantillonnage spatial, quitte à augmenter le bruit, pour lequel la méthode est suffisamment robuste.

Il convient donc de noter que même un cas idéal comme celui du fantôme peut mettre en évidence de grandes différences de performances entre différentes approches de segmentation, contrairement à ce que l'on pourrait penser *a priori* étant donné la relative simplicité de la configuration étudiée. En effet, les objets à segmenter sont ici des sphères parfaites et dont l'activité est homogène, et on est donc loin des configurations observées dans le cas de tumeurs réelles. Il n'y a ici aucune hétérogénéité de forme ou d'activité, et on pourrait penser que n'importe quelle approche de segmentation pourrait suffire, or ce n'est clairement pas le cas.

Troisièmement, l'approche locale adaptative (FLAB) présente un comportement plus stable

que celle fondée sur les chaînes de Markov (FHMC) sur l'ensemble des sphères. Ses performances sont en effet meilleures, en particulier sur les petites sphères. Ceci est important à souligner, car ce n'est pas forcément la conclusion que l'on aurait pu tirer *a priori* d'une comparaison entre ces deux modélisations, la deuxième étant en général considérée comme plus puissante que la première en segmentation d'image. Ici, il semble que la faiblesse de l'approche par chaîne de Markov soit sa modélisation du voisinage spatial via le parcours d'Hilbert Peano, trop approximative pour l'application à des petits objets peu contrastés et une image dans l'ensemble très bruitée. Notons que l'utilisation de champs de Markov à la place de la modélisation par chaîne aurait peut être également pu résoudre ce problème spécifique de corrélation spatiale. Toutefois, dans la mesure où l'approche locale adaptative a donné des résultats encourageants, c'est sur cette approche que nous avons décidé de bâtir notre méthode de segmentation des tumeurs réelles. En effet, de meilleures performances sur des petites sphères suggèrent que FLAB sera également plus performant que FHMC pour la délimitation précise des petites régions au sein d'une tumeur hétérogène, ou pour définir un contour précis d'une tumeur dont la forme serait complexe et dont les structures du contour peuvent être petites par rapport à la résolution spatiale du scanner. Notons que cette observation de supériorité de l'approche locale adaptative par rapport aux chaînes de Markov est en accord avec l'observation faite dans [140], notamment concernant l'homogénéité et la corrélation du bruit. Les images traitées ici sont clairement des régions homogènes sur fond homogènes. Non pas au sens de l'homogénéité de l'activité radioactive (bien que cela soit aussi le cas), mais au sens de l'homogénéité et de la corrélation du bruit telles que définies dans [140]. Ainsi, des images de tumeurs réelles, même hétérogènes (au sens clinique) et avec des formes non sphériques, restent des zones connexes et des régions de taille relativement importantes dans l'image, ce qui correspond à des situations correspondant à un facteur d'homogénéité élevé et une corrélation du bruit élevée également. Nos résultats sont donc en accord avec les conclusions de [140].

4.5 Validation sur objets non sphériques et non homogènes

Les résultats présentés ici concernent les performances de segmentation sur les objets synthétiques présentés dans la figure 4.7 (voir section 4.1.4). FLAB est utilisé ici avec l'implémentation optimisée telle que définie auparavant, c'est à dire utilisant des lois gaussiennes, un cube d'estimation de 3x3x3 voxels, et deux niveaux de flou avec les valeurs $\varepsilon_1 = \frac{1}{3}$ et $\varepsilon_2 = \frac{2}{3}$. Les résultats de différentes approches concurrentes sont présentés pour comparaison. Les résultats concernant la robustesse au bruit et au contraste sur l'objet à trois classes sont présentés dans la figure 4.20 et ceux concernant la résolution spatiale sur l'objet binarisé dans la figure 4.21. Les résultats sont nettement à l'avantage de FLAB, en particulier pour les configurations les plus difficiles. Les seuillages, qu'ils soient fixes ou adaptatifs, montrent clairement leurs limites, en particulier dans leur capacité à s'adapter à différentes configurations de contrastes. FCM permet, comme FLAB, de définir directement une segmentation à trois classes, mais se montre bien moins robuste sur les configurations difficiles. Ici, nous avons considéré l'objet initiale comme une vérité terrain à trois classes, avec différentes configurations d'observations plus ou moins complexes à segmenter. Pour évaluer les performances des algorithmes binaires (les seuillages), nous avons calculé les erreurs de classifications par rapport à la vérité terrain binarisée, c'est à dire sans différencier les zones noires et grises, sinon les erreurs calculées ne seraient pas vraiment pertinentes. Malgré cela les résultats démontrent les limites des seuillages, même adaptatifs, pour gérer les objets hétérogènes. Il est intéressant de noter les différences importantes entre les deux types de seuillages adaptatifs. Bien qu'ils fonctionnent à peu près sur le même principe, le fait qu'ils définissent la valeur de la tumeur par une région d'intérêt avec un seuillage initiale à 70 % (Tbckg) ou avec les 8 voxels entourant le voxel d'intensité maximale entraînent des différences énormes, comme par exemple dans le cas (e) où Tbckg ne conserve qu'une toute petite région de la tumeur, résultant en une erreur de 90 %, alors que TSBR fonctionne mieux (avec tout de même une erreur de plus de 21 %). Les erreurs des deux seuillages adaptatifs s'échelonnent entre 15 et 40 %. Le seuillage fixe à 42 % a un comportement relativement stable avec des erreurs entre 20 et 30 %. Quant au seuillage à 50 % il fonctionne mieux sur les configurations les plus difficiles (e) et (f) avec des

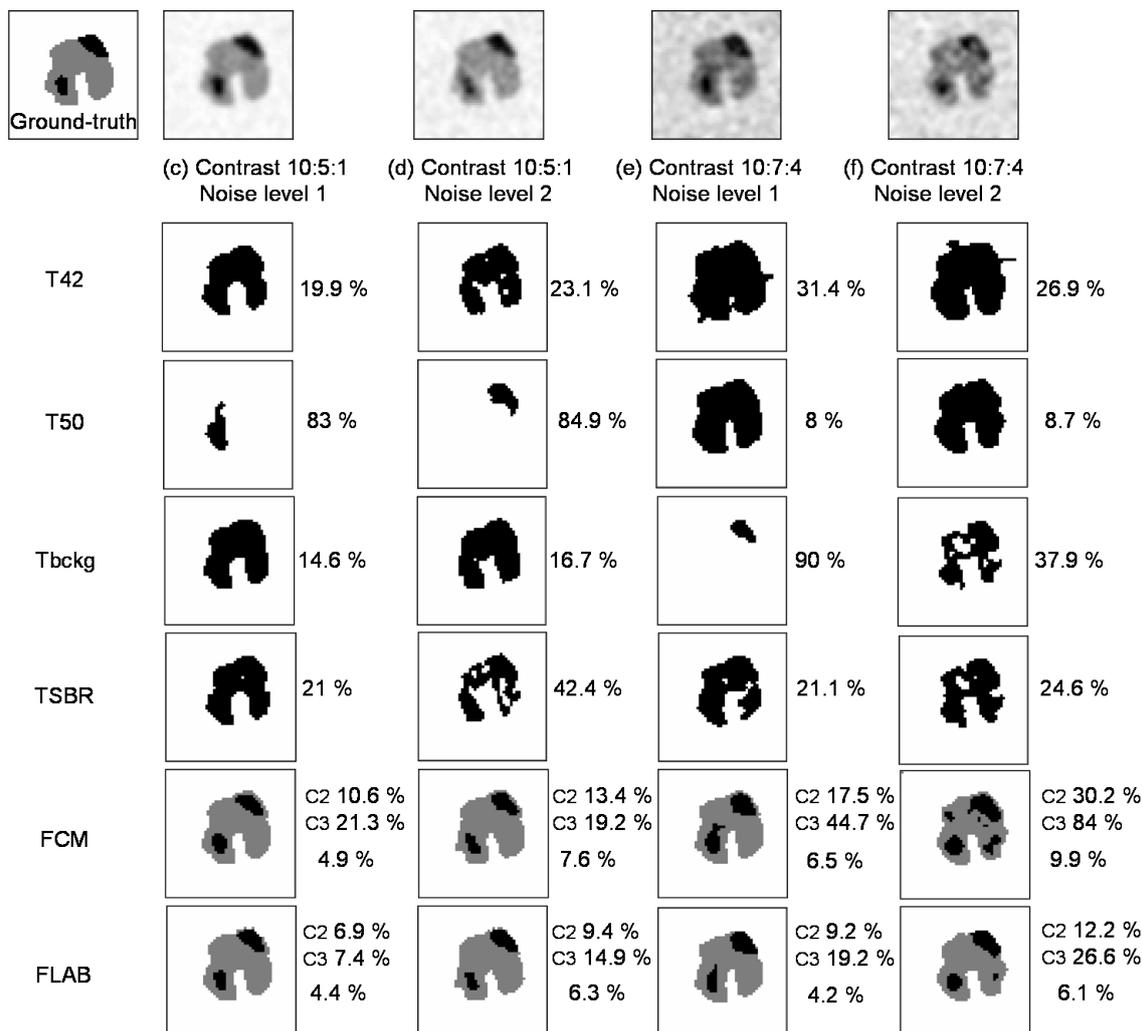


FIG. 4.20 – Résultats de FLAB et FCM obtenus sur l'objet synthétique à trois classes, avec deux configurations de contrastes et deux niveaux de bruit différents (c à f).

erreurs de 8 et 9 % respectivement, car sur les deux configurations les plus faciles (c) et (d), le contraste à l'intérieur de la tumeur est trop important et seules des petites régions de la tumeur sont conservées, menant à des erreurs supérieures à 80 %.

FCM et FLAB ont des performances supérieures, ne serait-ce que par leur capacité à segmenter en trois classes les objets d'intérêt. Les erreurs présentées sont donc calculées différemment que dans le cas des seuillages binaires. Les erreurs nommées C2 et C3 correspondent aux erreurs de classification pour les régions grises et noires respectivement. L'erreur indiquée en dessous est l'erreur de classification calculée par rapport à l'ensemble de la tumeur. FLAB

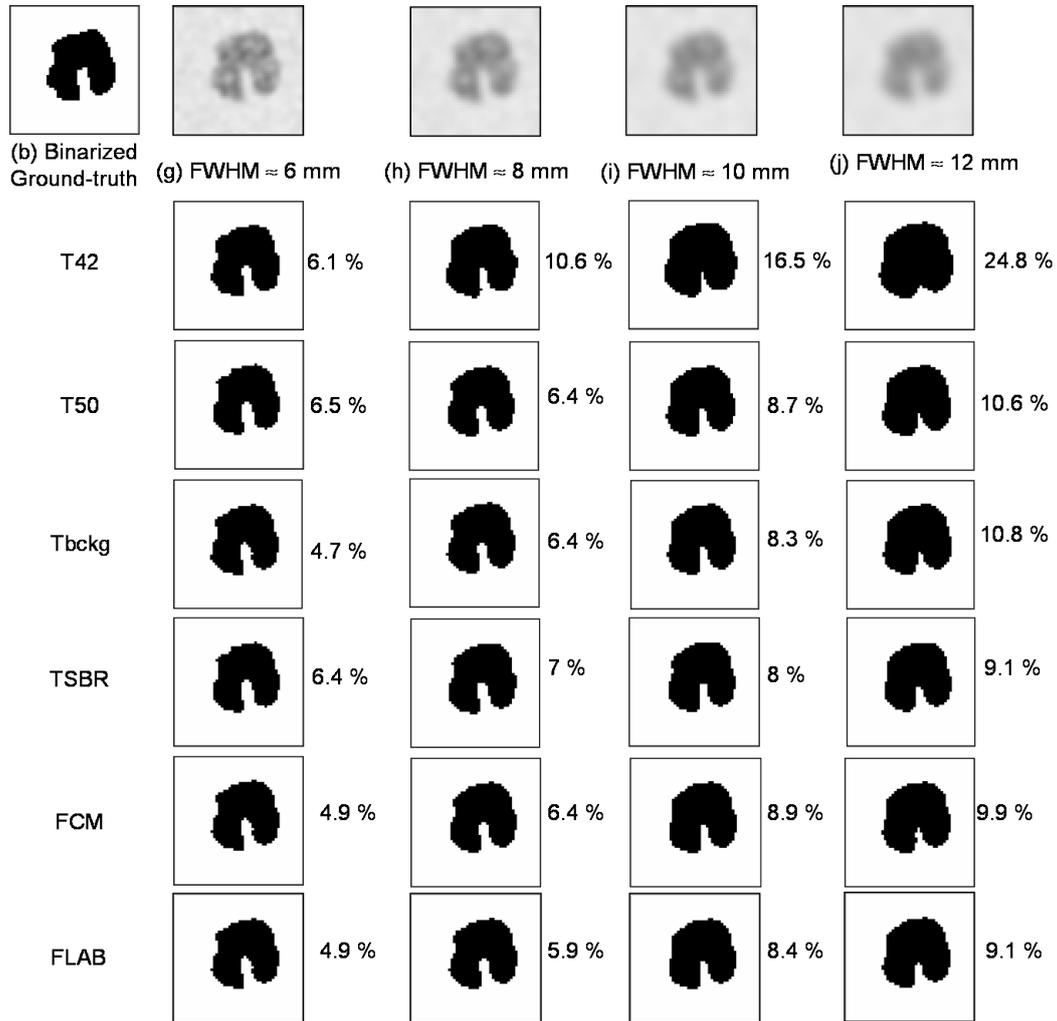


FIG. 4.21 – Résultats de FLAB et FCM obtenus sur l'objet synthétique binarisé, convolué par des PSF de largeur 4 (g) à 10 (j) mm.

obtient des performances supérieures en terme de précision de la segmentation, mais aussi en terme de robustesse, comme la stabilité des erreurs commises le démontre. On peut noter une supériorité flagrante de FLAB pour ce qui est de la segmentation des régions noires, en particulier pour les situations difficiles à faibles contrastes (e) et (f), avec 45 et 84 % d'erreur pour FCM contre 20 et 26 % pour FLAB. FLAB est également capable de mieux restituer la vérité terrain à la transition entre les classes 1 et 3 (voxels du fond blancs et voxels noirs de la tumeur), alors que FCM introduit des voxels de la seconde classe (gris) systématiquement entre les deux.

La robustesse par rapport à la dégradation de la résolution spatiale, du moins sur un objet binaire, est similaire pour toutes les méthodes, même si les meilleurs résultats sont tout de même obtenus par FLAB. Toutefois, les différences sont ténues. A part T42, les méthodes semi-automatiques ou automatiques montrent qu'elles peuvent être utilisées en TEMP, même si leurs performances sont diminuées par un flou plus important. Globalement, l'erreur comise est doublée pour les seuillages adaptatifs, FCM et FLAB (de 4.9 % à 9.1 % pour FLAB, 4.9 à 9.9 pour FCM, de 4.7 à 10.8 pour Tbckg), mais l'erreur reste tolérable. T42 est ici beaucoup moins robuste à cause du contraste de l'image, passant de 6 à 25 %.

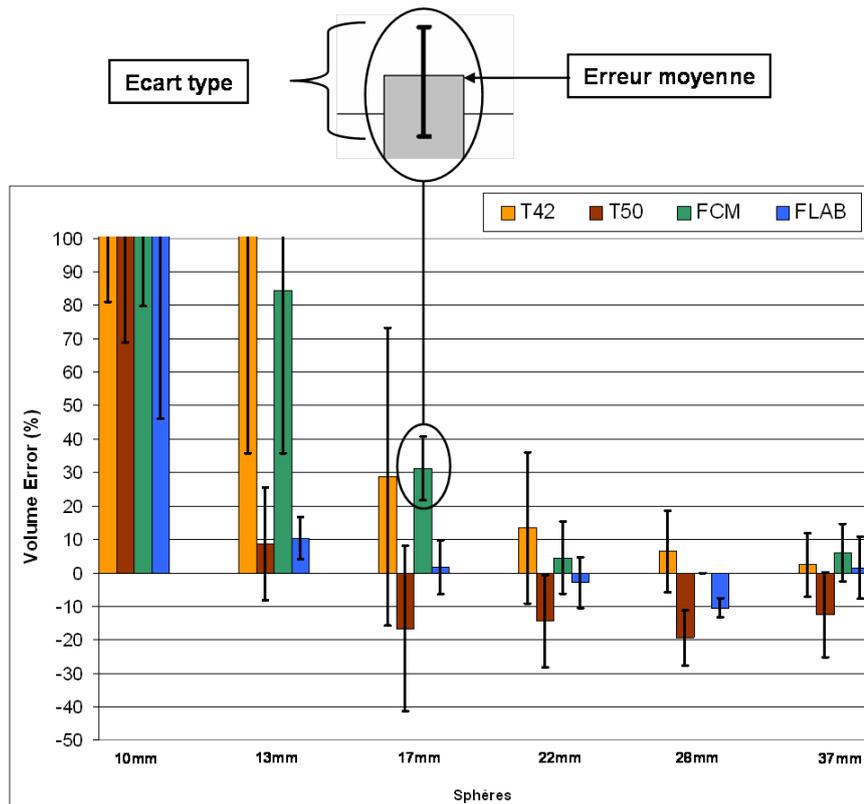
4.6 Résultats de l'étude sur la robustesse

L'utilisation du système de Pearson dans le cadre de FLAB a permis de comparer les performances de l'algorithme avec une simple loi gaussienne ou en utilisant les variantes offertes par le système de Pearson. Sur les images de fantôme IEC réelles, le système de Pearson a détecté des lois bêta I proches de la loi gaussienne. Les résultats de segmentation n'ayant pas été améliorés dans ce cas, l'utilisation de loi gaussienne a été conservée (voir figure 4.13).

Nous avons également analysé les différentes acquisitions de fantômes réalisées avec d'autres scanners et reconstruites avec différents algorithmes de reconstruction, pour étudier la variabilité des résultats obtenus par FLAB comparée à celle des autres approches. Le but étant d'estimer si FLAB peut être utilisé sur des images TEP issues de différents centres, et donc acquises avec différents modèles de scanner et reconstruites avec différents algorithmes itératifs. Il s'agit donc de vérifier que FLAB est suffisamment robuste pour être appliqué à des images issues de différents scanners, et donc que l'approche est suffisamment universelle pour ne pas nécessiter une optimisation spécifique de ses paramètres pour chaque modèle de scanner, protocole d'acquisition ou algorithme de reconstruction.

Les résultats sont ici exprimés en erreur de volume par rapport au vrai volume de la sphère, ainsi qu'en erreur de classification calculée par rapport à une vérité terrain voxel à voxel obtenue par recalage des images TEP avec les acquisitions haute résolution TDM. Contrairement

aux résultats présentés dans les sections précédentes, les erreurs sont ici calculées pour l'ensemble des données pour une sphère donnée. Pour chaque sphère, on a donc une erreur moyenne et un écart type, calculés sur l'ensemble des configurations (tous les contrastes, durées d'acquisitions, scanners, algorithmes...). Il faut donc lire la figure 4.23 comme indiqué dans la figure 4.22 ci-dessous : On compare ici les performances de FLAB à celles obtenues avec deux



Pour chaque sphère,
erreur moyenne et écart type calculés pour l'ensemble des données

FIG. 4.22 – Comment lire les résultats de l'étude de robustesse sur les différents scanners considérés.

seuillages fixes (42 et 50 %) et Fuzzy C-Means. Aucune approche de seuillage adaptatif n'a été considéré ici car ces méthodes sont justement conçues pour être utilisées après avoir été optimisées au préalable sur des acquisitions de fantôme contenant des sphères de diamètre variable (voir section 2.3.2). Les résultats sont nettement en faveur de FLAB, avec un avantage pour FLAB concernant les plus petites sphères. L'erreur moyenne de classification en considérant

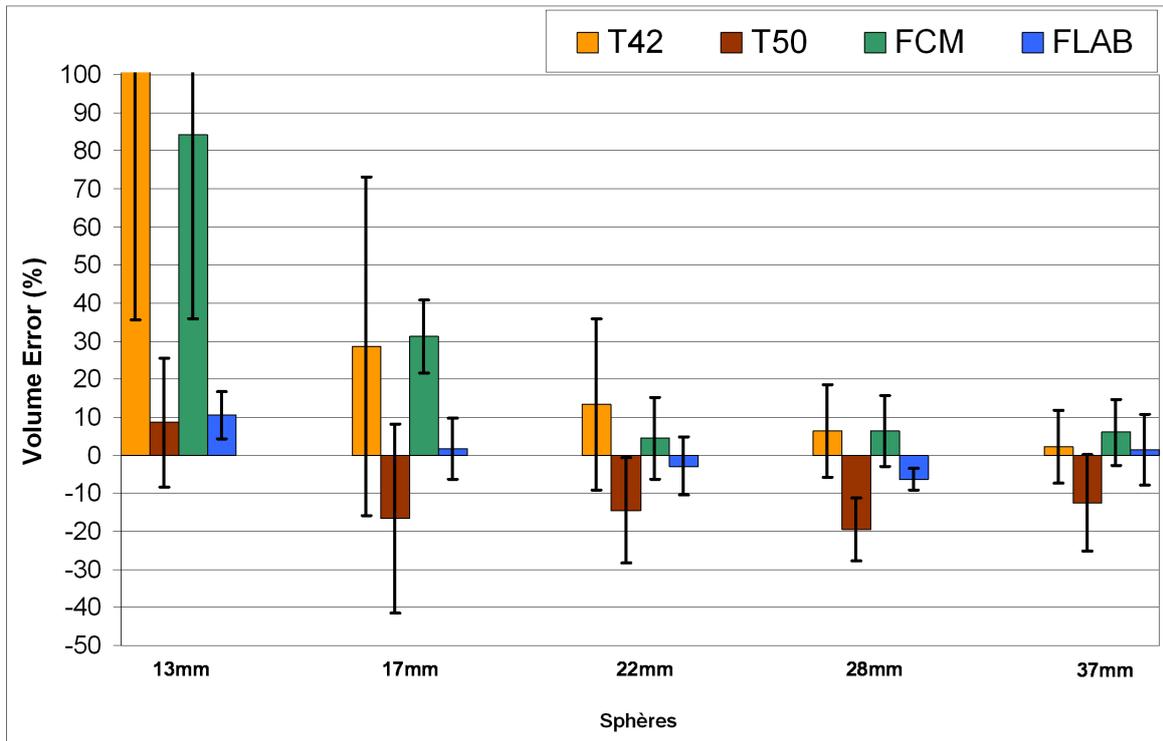


FIG. 4.23 – Erreur de volume moyenne et écart type pour chaque méthode considérée, sur l'ensemble des données disponibles pour l'étude de robustesse.

les sphères dont le diamètre est supérieure à 13 mm (la sphère de 10 mm étant exclue pour éviter un biais trop important) pour le seuillage fixe à 42 % est supérieure à 37 % avec un écart type de presque 30 %. L'erreur de classification moyenne associée au seuillage à 50 % est nettement plus faible avec 19 %, ainsi qu'affecté d'une variabilité plus faible (17 %). Ceci s'explique facilement par le fait que le seuillage à 50 % est plus restrictif, et a donc tendance à systématiquement sous-estimer le volume des sphères, et à nettement moins sur-évaluer les volumes des plus petites sphères, ce qui fait baisser son erreur moyenne en considérant l'ensemble des sphères. FCM obtient une erreur moyenne de presque 30 %, mais une variabilité moindre de 15 %. Les meilleurs résultats sont obtenus par FLAB avec une erreur moyenne inférieure à 9 % et une variabilité inférieure à 5 %. Ces résultats démontrent sans appel la robustesse de FLAB largement supérieure à celle des seuillages fixes et du Fuzzy C-Means. Là où ces méthodes font preuve non seulement d'erreurs importantes dans la segmentation des

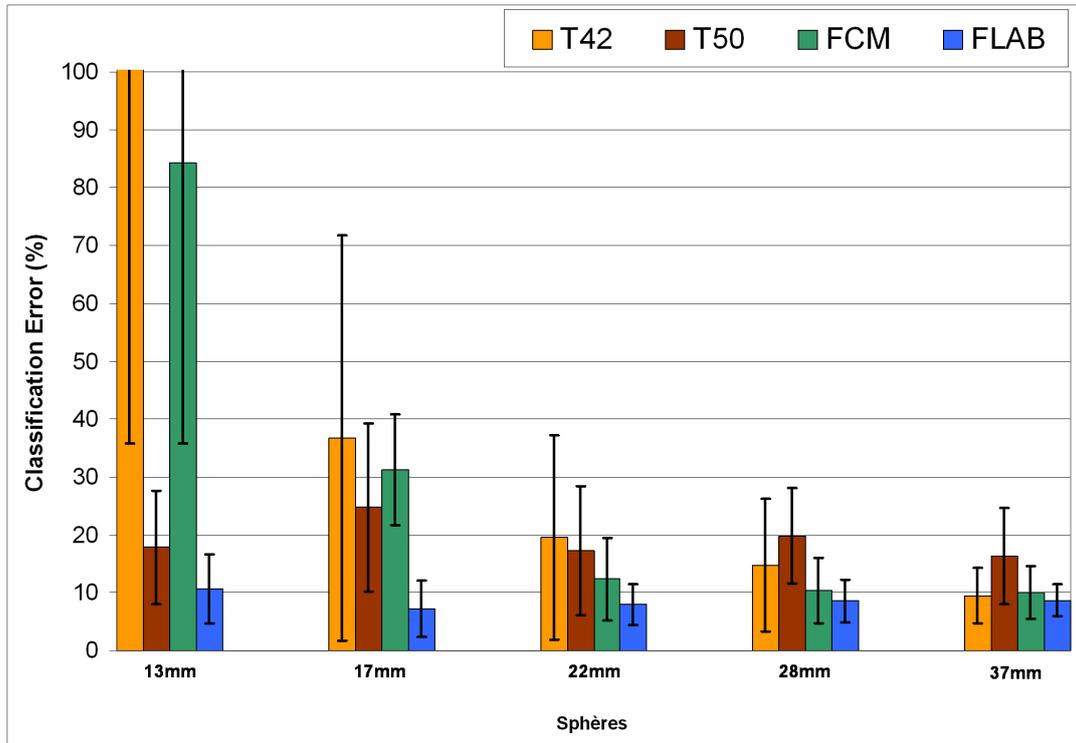


FIG. 4.24 – Erreur de classification moyenne et écart type pour chaque méthode considérée, sur l'ensemble des données disponibles pour l'étude de robustesse.

sphères, surtout les plus petites, mais également d'une très grande variabilité en fonction du scanner et de l'algorithme utilisé, ou des paramètres d'acquisition, FLAB obtient une erreur acceptable ($< 10\%$ en moyenne) et d'une variabilité très faible (seulement 5%).

Chapitre 5

Performances des méthodes dans les applications cliniques

5.1 Images de tumeurs réalistes

Une étape cruciale de validation des approches développées concerne la segmentation de tumeurs réelles. L'évaluation des résultats obtenus est toutefois nettement plus complexe car la vérité terrain est plus difficile à obtenir et à estimer.

5.1.1 Intérêt, avantages et inconvénients

Les images cliniques présentent l'immense avantage de constituer une base de données réaliste. C'est ce que nous visons comme objectif à long-terme applicatif des méthodes développées et testées dans ce travail. L'inconvénient qu'elles présentent est à la hauteur de leur avantage : on ne connaît pas la vérité terrain facilement. En effet, la seule solution pour obtenir une relative vérité terrain des objets d'intérêt au sein du patient imagé par le scanner TEP est une procédure chirurgicale au cours de laquelle la tumeur est retirée puis examinée lors d'une étude histologique. La tumeur est figée dans de la paraffine, puis découpée en tranches fines pour être analysée. Ceci ne peut toutefois être envisagé que sur un nombre limité de patients, et est également sujet à plusieurs approximations, notamment le changement éventuel de forme ou de taille de la tumeur (généralement estimé à 10% environ) une fois retirée de son

”environnement naturel”. C’est la raison pour laquelle les tumeurs choisies sont généralement de type O.R.L. plutôt que pulmonaires.

Sans histologie, même l’image anatomique TDM ou IRM ne permet pas d’obtenir cette vérité terrain, puisque l’information n’est pas de même nature. Malgré la résolution fine que l’on peut obtenir avec ce type d’imagerie, on ne peut déduire une vérité terrain appropriée pour évaluer la segmentation sur l’image fonctionnelle du fait de la différence de nature entre les deux modalités, complémentaires mais non superposables voxel à voxel. En effet, la vérité terrain dont nous avons besoin pour évaluer la segmentation des images d’émission est celle concernant le fonctionnement métabolique observé dans le cadre de l’acquisition TEP, l’accumulation de glucose par exemple, et non le contour anatomique correspondant à la densité des tissus. Enfin, notons que la vérité terrain déterminée par histologie n’est pas non plus une correspondance voxel à voxel parfaite avec l’image à segmenter, mais uniquement des informations de dimensions, comme le diamètre maximal ou le volume total. Seules les simulations permettent d’obtenir une vérité terrain avec une correspondance parfaite voxel à voxel entre l’image à segmenter et l’image de vérité terrain.

5.1.2 Images cliniques

Nous avons travaillé sur ensembles de données. Le premier ensemble d’images correspond aux examens de patients suivant un traitement par radiothérapie dans le centre du Royal Hospital de Belfast, en Irlande. La planification de traitement de radiothérapie étant une des applications visées par les applications de ce travail de thèse, il est intéressant de considérer ces patients dont les tumeurs sont *a priori* destinées à subir un tel traitement, ce qui nécessite une segmentation des objets d’intérêt destinés à recevoir les doses. Les tumeurs concernées dans cet ensemble de données sont essentiellement des tumeurs pulmonaires. Les acquisitions de ces patients ont été réalisées sur un système TEP/TDM GE Discovery LS, et reconstruites avec l’algorithme OSEM.

Un autre ensemble d’images est issu du centre de médecine nucléaire du CHU Morvan, à Brest. Ces tumeurs sont essentiellement ORL ou pulmonaires, ainsi que abdominales. Les acquisitions de ces patients ont été réalisées sur un système TEP/TDM Philips Gemini, et

reconstruites avec l'algorithme RAMLA.

Un troisième ensemble d'images acquises au service de l'hôpital de Maastricht (MAASTRO) aux Pays-bas, correspond à des examens de patients ayant par la suite été opérés pour retirer la tumeur, et dont les tumeurs ont fait l'objet d'une étude histologique. Cette dernière consiste à extraire la tumeur du patient, à la figer rapidement pour éviter que sa forme et son volume ne change trop, puis la découper en tranches de façon à étudier les tissus pour identifier les tissus tumoraux, et définir des propriétés de mesure du volume tumoral, tels que le diamètre maximal de la tumeur. Ceci nous donne la possibilité de comparer le diamètre maximal déterminé automatiquement sur la segmentation de l'image à celui obtenu par l'histologie. Il convient de souligner la rareté de telles données et le peu d'études publiées jusqu'à présent utilisant des informations issues d'histologie pour comparer les performances de méthodes de segmentation d'images TEP [157].

5.1.3 Images simulées de tumeurs

GATE

Le logiciel GATE (*Geant 4 Application for Tomography Emission*) est une plate-forme de simulation TEP et TEMP développée depuis 2001 dans le cadre d'une collaboration internationale regroupant 23 laboratoires et quelques 1200 utilisateurs. Cette plate-forme fondée sur la librairie Geant 4 est un outil puissant, à la fois flexible (capable de simuler une large gamme de systèmes d'acquisitions), précis (qui reproduit correctement les propriétés des images) et maintenue par un réseau de développeurs. GATE a été configuré pour simuler les acquisitions de plusieurs imageurs TEP et TEMP. Le LaTIM (INSERM U650) en particulier, a validé les performances de simulation (résolution spatiale, taux de comptage..) pour la géométrie du scanner Philips Gemini [96]. GATE permet de simuler de façon très réaliste les interactions au niveau particulaire des photons émis par la matière radioactive avec l'environnement, lui aussi simulé. Il est ainsi possible de simuler une acquisition à partir d'une carte d'activité et d'une carte d'atténuation, le détail de toutes les désintégrations radioactives, leurs annihilations avec la matière environnante, leur transport dans la matière, jusqu'à leur détection

par les détecteurs du scanner, et ce entièrement *in silico*. Les données générées à l'issue de la simulation le sont au format habituel (projections ou listmode), qu'il est possible ensuite de reconstruire comme s'il s'agissait de données réelles avec les algorithmes constructeurs.

Pour évaluer la capacité des algorithmes à segmenter des tumeurs dont la forme n'est pas nécessairement sphérique, et la distribution d'activité peut être hétérogène, des objets simulés ont été générés de la même façon que les simulations de fantômes, de façon à simuler des tumeurs réalistes dont la vérité terrain est disponible, pour pouvoir évaluer rigoureusement les performances de segmentation des méthodes testées. Ceci est particulièrement nécessaire pour l'évaluation de l'approche à trois classes dures et trois transitions floues. Cette dernière permet de segmenter avec précision le contour extérieur d'une tumeur et également de délimiter à l'intérieur même de la tumeur des régions d'activité différentes. Ceci permet de rendre compte au sein de la carte de segmentation, de l'hétérogénéité de la distribution d'activité au sein même de la tumeur. Les tumeurs simulées servant à la validation de cette approche doivent donc présenter ces hétérogénéités. L'avantage de travailler sur les tumeurs simulées est en particulier de pouvoir faire varier l'intensité du bruit et les configurations de contrastes, pour étudier la robustesse de la méthode face à des images de qualité variable.

Nous avons ainsi généré une base de données de tumeurs simulées réalistes en prenant comme base les images de patients. Pour générer une vérité terrain qui est par la suite utilisée pour générer les tumeurs simulées, et par rapport à laquelle les erreurs sont calculées, les tumeurs réelles des patients sont segmentées à la main, coupe par coupe. Précisons ici que la précision et la variabilité de cette segmentation manuelle n'a pas d'importance car il s'agit uniquement de construire une base de données de tumeurs simulées présentant globalement les mêmes caractéristiques en terme de variabilité de taille, forme, et distribution d'activité. Les méthodes sont ensuite évaluées sur les simulations par rapport à la vérité terrain ainsi générée, et non par rapport aux images réelles.

Procédure de génération des images simulées de tumeurs

La figure 5.1 illustre le processus de simulation des images de tumeurs réalistes.

La première étape consiste à constituer une base de données d'images cliniques de patients

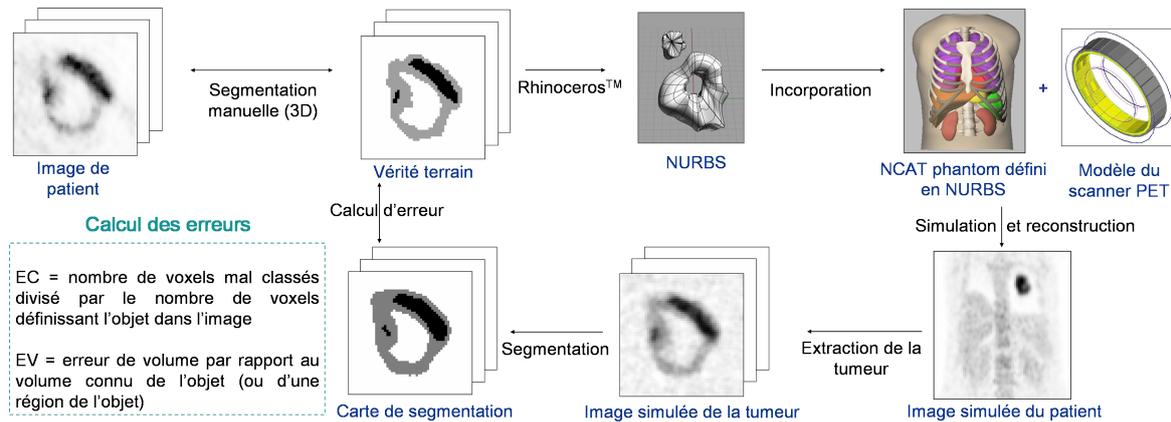


FIG. 5.1 – Procédure suivie pour obtenir des images simulées réalistes de tumeurs, et évaluer la qualité de la segmentation.

dont les tumeurs possèdent des caractéristiques de formes, de taille, de positionnement et d'activité souhaitées. L'étape suivante consiste à générer des modèles de tumeurs à partir des images de patients. Pour ce faire, on commence par segmenter à la main, coupe par coupe, les images cliniques pour établir ce qui sera la vérité terrain de la tumeur. Puis cette segmentation manuelle, sous forme de carte voxelisée, est transformée en NURBS (Non Uniform B-Splines) à l'aide du logiciel Rhinocerostm (société CADLINK), de façon à pouvoir être combinée à un fantôme réaliste du corps humain pour les simulations GATE, à savoir le fantôme NCAT [122]. Ce dernier simule les principaux organes et le squelette ainsi que les mouvements respiratoires, permettant de générer des images très réalistes, une fois combiné avec un modèle du scanner TEP et le moteur physique de simulation de particules (Geant 4) de GATE. Noter que dans le cadre de ce travail, les mouvements respiratoires n'ont pas été simulés. La dernière étape consiste à intégrer les tumeurs sous forme de NURBS dans le fantôme, et à procéder aux simulations. L'approche considérée consiste à simuler un corps entier NCAT avec GATE (ce qui demande de très longs temps de calculs) et les volumes tumoraux séparément, puis de procéder au mélange des lignes de réponse simulées, pour ensuite procéder à la reconstruction

de l'image. Ceci ne permet pas de modifier la morphologie du patient pour qu'elle corresponde au cas clinique utilisé pour récupérer la tumeur, mais cela n'a que peu d'importance dans la mesure où nous ne nous intéressons ici qu'à la segmentation d'une tumeur donnée sur une région bien définie. Toutefois cela peut changer les aspects de diffusés et d'atténuation en fonction de la morphologie du patient.

Une fois les simulations réalisées, il ne reste plus qu'à reconstruire ensuite les images à l'aide des algorithmes itératifs habituels. La segmentation obtenue sur ces images proches des images cliniques, mais pour lesquelles on dispose d'une vérité terrain voxel à voxel est ensuite comparée à cette dernière pour calculer des erreurs précises de volume et de classification (voir section 4.2).

5.2 Précision de segmentation sur tumeurs simulées et réelles pour la radiothérapie

5.2.1 Tumeurs simulées

Dans le cas des tumeurs simulées, il est possible de calculer des erreurs de classification précises par rapport à la vérité terrain voxelisée utilisée dans la simulation (voir section 4.2). Les résultats de précision de segmentation sont donc présentés dans la figure 5.5 sous forme d'erreurs de classification (erreur moyenne, erreur minimale et erreur maximale) pour l'ensemble des vingt tumeurs simulées, pour chacune des méthodes considérées, à savoir un seuillage adaptatif (Tbckg), un seuillage fixe à 42 % (T42), le Fuzzy C-Means (FCM) et notre algorithme FLAB. La figure 5.6 présente les mêmes résultats mais avec l'erreur moyenne et l'écart type.

Dans le cas des seuillages binaires, les erreurs ont été calculées par rapport à une vérité terrain binarisée dans le cas des tumeurs hétérogènes, et la vérité terrain originale dans le cas des tumeurs homogènes. Ceci a été fait, comme dans le cas de l'étude sur objets synthétiques (voir section 4.5) pour permettre une comparaison avec les approches FCM et FLAB, qui elles peuvent offrir une segmentation à trois classes dans le cas des tumeurs fortement hétérogènes.

Dans le cas de FCM et FLAB, les erreurs ont été calculées soit par rapport à la vérité terrain binarisée pour être comparées aux seuillages binaires (figure 5.5), mais également par rapport à la vérité terrain à trois classes. Ces derniers résultats sont présentés dans la figure 5.7 et permettent d’apprécier la précision de FCM et FLAB pour la segmentation des différentes classes des tumeurs hétérogènes.

Exemples représentatifs de résultats de segmentation

Les trois figures 5.2, 5.3 et 5.4 illustrent des résultats représentatifs des performances obtenues par les méthodes sur différents types de tumeurs. Précisons que si les images présentées ici sont des coupes uniques, tout a été fait en 3D, et les erreurs sont calculées sur volume total de la tumeur. Le premier cas illustré dans la figure 5.2 est celui d’une petite tumeur hépatique (dans

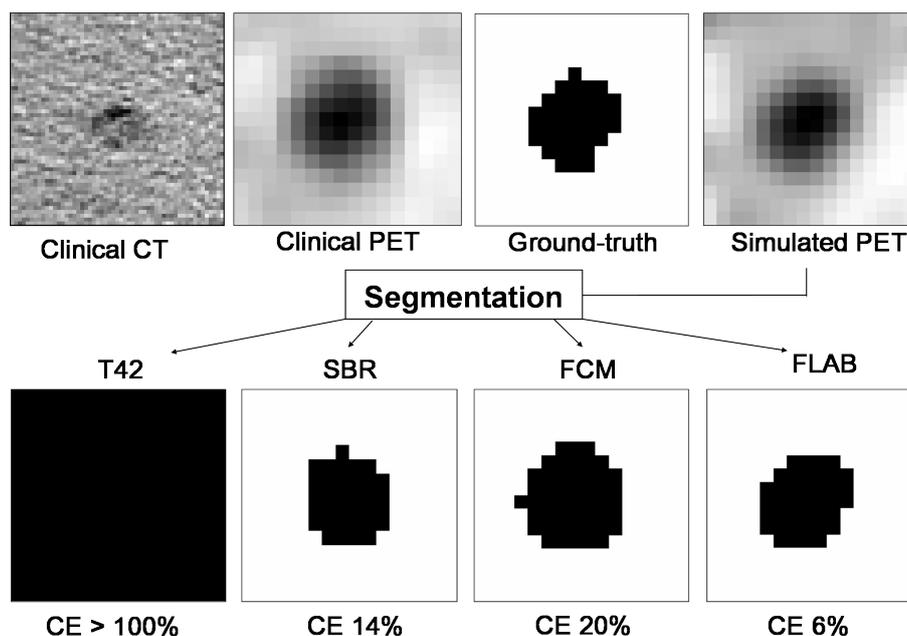


FIG. 5.2 – Illustration des performances des méthodes considérées sur une petite tumeur homogène au fond légèrement hétérogène.

le foie) (environ 2 cm de diamètre) au contraste faible avec le fond, qui est assez hétérogène. On constate l’échec complet du seuillage à 42 % (T42), à cause du contraste trop faible (la croissance de région emplit toute la boîte). Si FCM et le seuillage adaptatif obtiennent des erreurs de classification acceptables (20 et 14 % respectivement), la délimitation de la tumeur

est nettement plus convaincante avec FLAB, qui obtient une erreur de classification inférieure (6 %). Noter que dans le cas du seuillage adaptatif, étant donné l'hétérogénéité du fond, deux résultats ont été générés, le premier avec une région d'intérêt placée dans la partie claire du fond, et l'autre avec une région placée dans la partie plus sombre. Le meilleur résultat est celui présenté ici, l'autre ayant généré une erreur de 19 %. Le deuxième cas illustré dans la figure 5.3

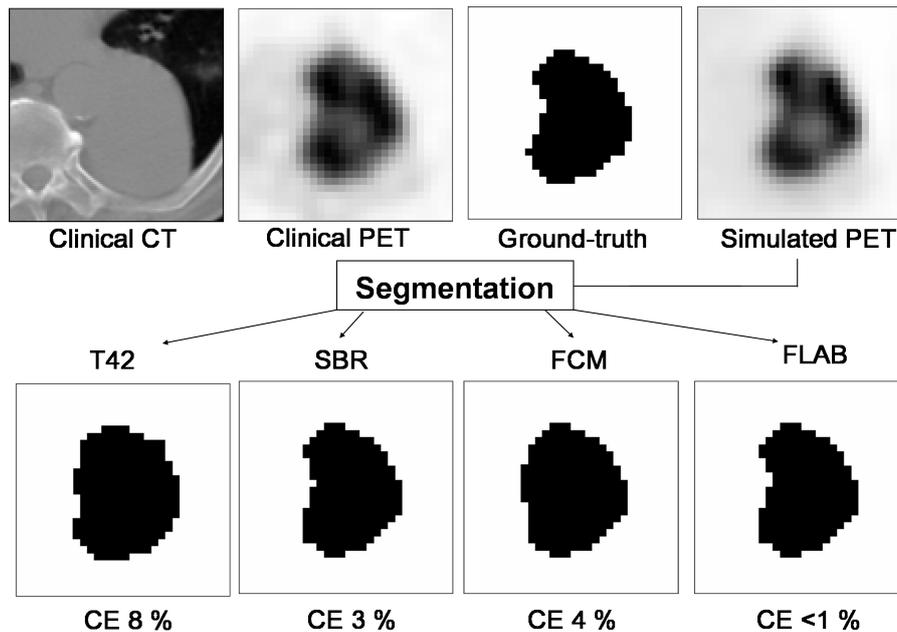


FIG. 5.3 – Illustration des performances des méthodes considérées sur une tumeur pulmonaire légèrement hétérogène et de taille moyenne (environ 4,5 cm).

est celui d'une tumeur de taille moyenne (environ 4,5 cm de diamètre) au contraste important entre la tumeur et le fond, et dont l'activité intérieure est légèrement hétérogène. On constate que si FLAB obtient toujours la meilleure segmentation (moins de 1 % d'erreur), les différences entre les approches sont faibles : 8 % pour T42 et 3 et 4 % respectivement pour SBR et FCM. Il s'agit ici d'un cas simple à segmenter malgré la présence d'hétérogénéité, car le contraste entre les différentes régions de la tumeur est insignifiant par rapport au contraste global entre la tumeur dans son ensemble et le fond. Le troisième cas illustré dans la figure 5.4 est celui d'une très grande tumeur pulmonaire (environ 6 cm de diamètre) au contraste important avec le fond et dont l'activité est très variable. Etant donné la taille de la tumeur, son coeur est nécrosé. Sur les pourtours, l'hétérogénéité est très importante, avec un contraste significatif

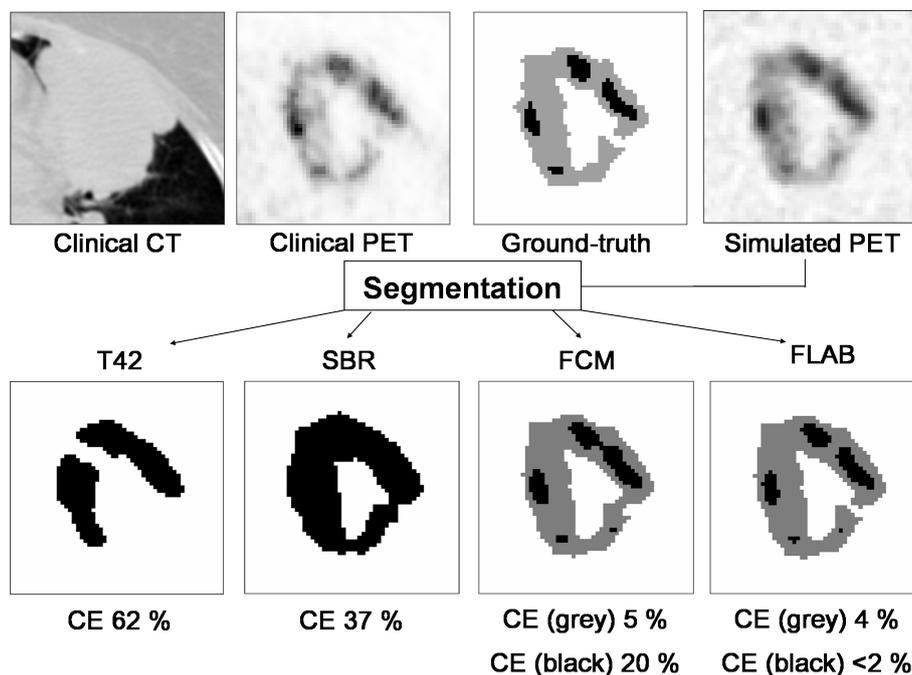


FIG. 5.4 – Illustration des performances des méthodes considérées sur une tumeur fortement hétérogène et de grande taille (environ 6 cm).

entre les régions d'activité forte au sein de la tumeur et le reste du volume tumoral, ce qui explique l'échec flagrant des méthodes binaires. On constate l'échec complet du seuillage à 42 % (T42), qui rate complètement certaines parties de la tumeur (62 % de sous évaluation du volume). Le seuillage adaptatif permet de récupérer l'ensemble de la tumeur, mais surestime fortement son volume total (+37 %) et surtout ne permet pas de définir un contour précis ni de la tumeur dans son ensemble, ni des régions d'intensité supérieure au sein de la tumeur elle-même. Si FCM et FLAB offrent toutes deux cette segmentation à trois classes, FCM manque de précision sur la classe 3, comme les erreurs de classification le montrent (20 % contre 2 % pour FLAB).

Erreurs sur l'ensemble des tumeurs simulées

Les résultats obtenus par les méthodes considérées sur l'ensemble des vingt tumeurs simulées sont à l'avantage de FLAB, qui obtient moins de 10 % d'erreur moyenne de classification, avec un minimum à 6 % et un maximum à 22 %. FCM et le seuillage adaptatif obtiennent des

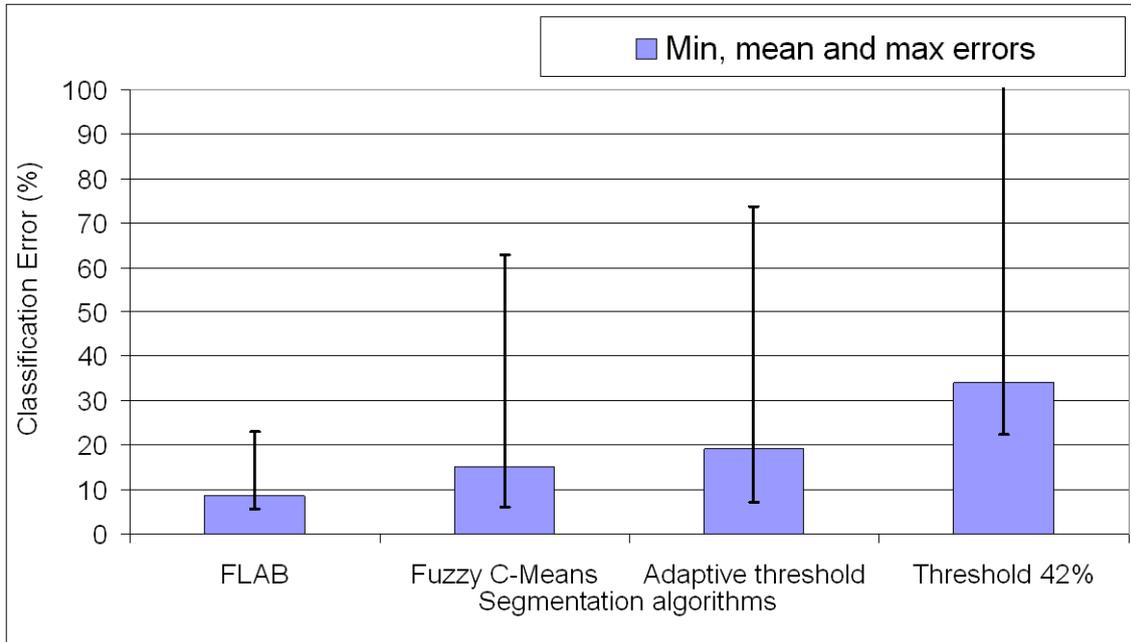


FIG. 5.5 – Erreurs de classifications obtenues par quatre méthodes de segmentation (FLAB, FCM, TSBR et T42) sur l'ensemble des 20 tumeurs simulées. Les barres bleues représentent l'erreur moyenne sur les 20 tumeurs, et la barre de variation indique l'erreur minimale et maximale obtenue.

résultats moins bons tant en erreur moyenne (environ le double pour le seuillage adaptatif, 20 % et un peu plus de 15 % pour FCM) qu'en terme de variabilité puisque FCM et le seuillage adaptatif obtiennent jusqu'à 60 et 70 % d'erreur respectivement sur certaines tumeurs, bien que sur certaines autres tumeurs, leurs performances descendent également à moins de 10 %. Le seuillage fixe, comme attendu, obtient les moins bonnes performances sur cet ensemble de tumeurs simulées, avec une erreur moyenne de 30 % environ, et une variabilité encore plus grande, avec un minimum à 20 % environ et un maximum supérieur à 100 %.

Les mêmes résultats sont présentés avec l'erreur moyenne et l'écart type calculés sur l'ensemble des 20 tumeurs dans la figure 5.6. L'écart type obtenu avec FLAB (8 %) est inférieur à celui obtenu avec les autres méthodes (15 % pour FCM et TSBR, 20 % pour T42). Les résultats détaillés classe par classe pour FLAB et FCM (figure 5.7) calculés pour les tumeurs hétérogènes à deux classes révèlent la même supériorité de FLAB par rapport à FCM que les performances constatées sur l'objet synthétique hétérogène utilisé dans la validation de FLAB (voir sections

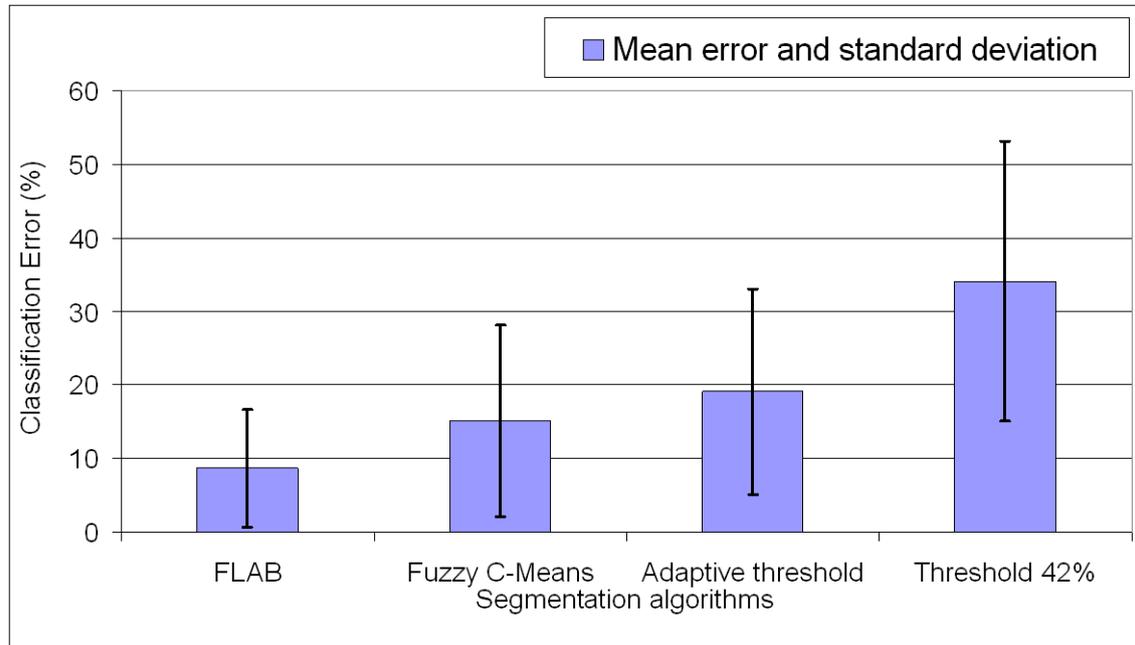


FIG. 5.6 – Erreurs de classifications obtenues par quatre méthodes de segmentation (FLAB, FCM, TSBR et T42) sur l'ensemble des 20 tumeurs simulées. Les barres bleues représentent l'erreur moyenne sur les 20 tumeurs, et la barre de variation indique l'écart type.

4.1.4 et 4.5). FLAB obtient une erreur de classification moyenne légèrement supérieure à 10 % pour les deux classes, avec des minimums et maximums de 5 et 23 % pour la classe 2, et 1 et 29 % pour la classe 3. FCM de son côté obtient une erreur moyenne de 17 et 24 % pour les classes 2 et 3 respectivement, tandis que les minimums et maximums atteints sont nettement plus élevés avec de 6 à 35 % pour la classe 2 et de 8 à 58 % pour la classe 3. Etant donné la faible capacité de FCM à segmenter les petites structures, constatée sur les résultats de performance les petites sphères des acquisitions de fantôme, il est normal que les performances de FCM sur les petites régions d'intensité plus élevées au sein de la tumeur soient moindre que la segmentation de la tumeur dans son ensemble, pour laquelle ses performances sont plus proches de FLAB. Ceci est d'ailleurs illustré visuellement sur l'exemple de la grande tumeur hétérogène présentée en figure 5.4.

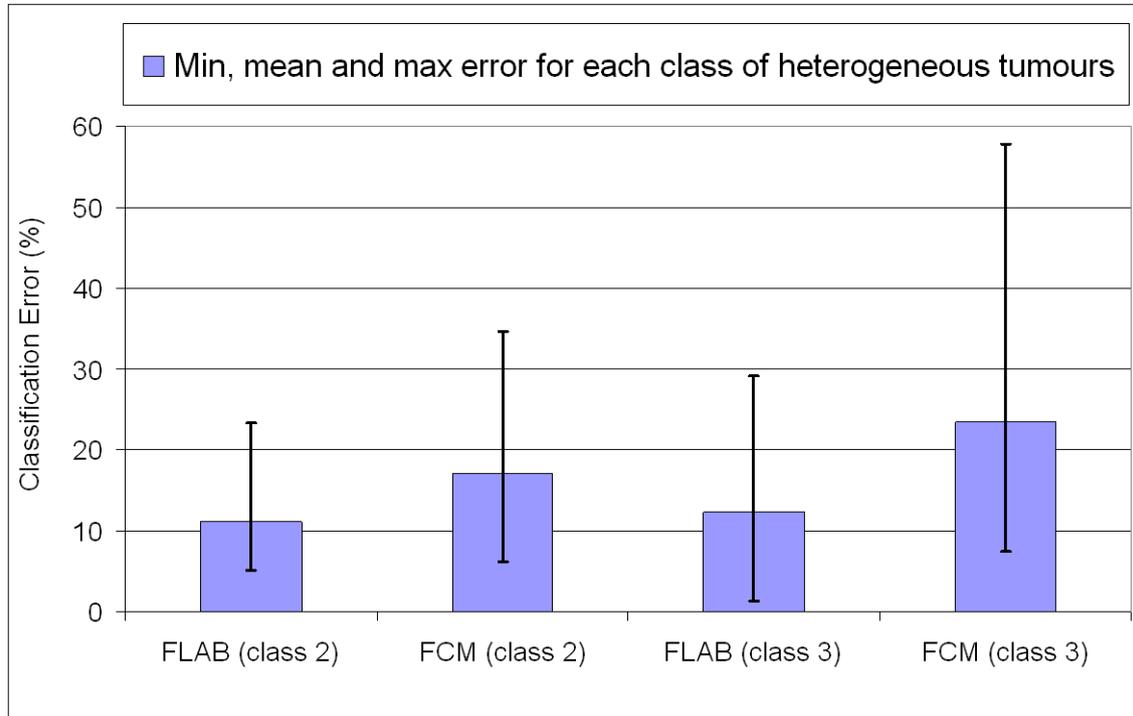


FIG. 5.7 – Erreurs de classifications obtenues par FLAB et FCM sur les tumeurs fortement hétérogènes pour les classes 2 et 3 définissant les tumeurs. Les barres bleues représentent l’erreur moyenne, et la barre de variation indique les erreurs minimales et maximales atteintes.

5.2.2 Tumeur réelles avec résultats d’histologie

Ces résultats sont cette fois obtenus en segmentant les images de tumeurs réelles (et non simulées), et en comparant le diamètre maximal mesuré sur la segmentation avec celui mesuré dans l’étude histologique. La figure 5.8 illustre visuellement les segmentations obtenues sur deux tumeurs de cet ensemble de données. La figure 5.9 donne l’erreur moyenne obtenue par chaque méthode par rapport au diamètre maximal de la tumeur, sur l’ensemble des 18 patients. L’écart type associé est la valeur représentative des différences entre chaque méthode. En effet l’erreur moyenne est relativement faible pour toutes les méthodes (moins de 5 %) car la proportion de sous-évaluation et de sur-évaluations se compensent. Les écart types obtenus par chaque méthode montrent l’avantage de FLAB (environ 10 %) par rapport aux autres approches dont la variabilité est largement supérieure (environ 25 % pour le Fuzzy C-Means et le seuillage adaptatif, et plus de 30 % pour le seuillage fixe).

5.2. PRÉCISION DE SEGMENTATION SUR TUMEURS SIMULÉES ET RÉELLES
POUR LA RADIOTHÉRAPIE

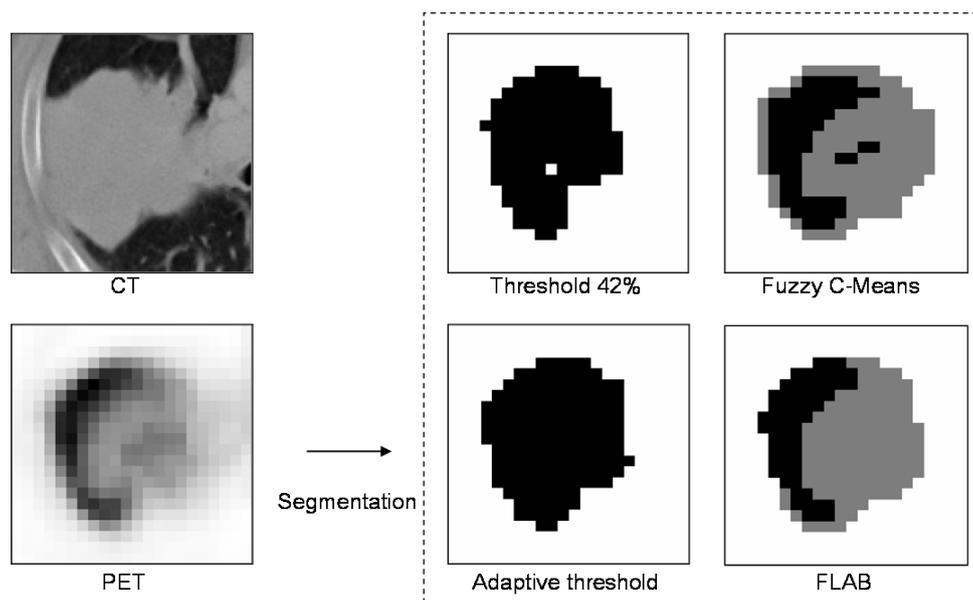


FIG. 5.8 – Illustration des segmentations obtenues sur une tumeur réelle des données de l'étude histologique.

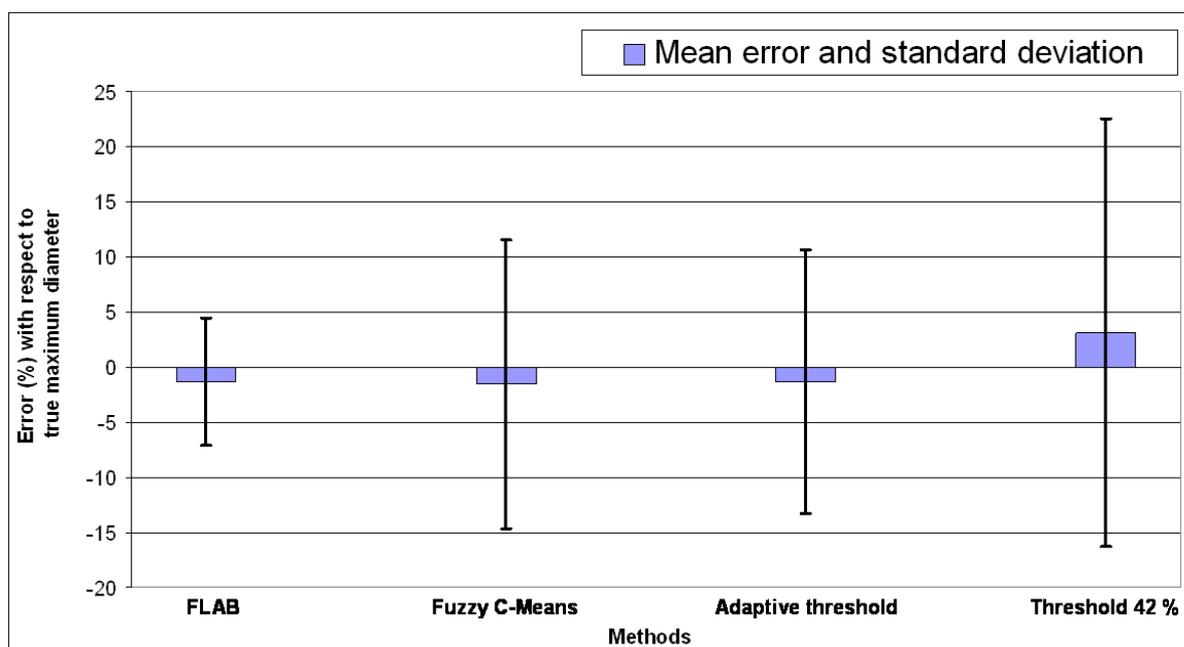


FIG. 5.9 – Erreur moyenne et écart type obtenu par chaque méthode par rapport au diamètre maximal de la tumeur mesuré par l'étude histologique.

5.3 Suivi du volume tumoral au cours du temps pour le suivi thérapeutique : résultats sur tumeurs simulées

5.3.1 Cas considérés

Une deuxième étude sur tumeurs simulées a consisté à évaluer la capacité des algorithmes à suivre l'évolution d'une tumeur dans le temps. Il s'agit de tester si une méthode de segmentation donnée est capable de correctement définir le volume d'une tumeur à différents instants dans le temps, par exemple au cours d'un traitement. Les évolutions de la tumeur que nous avons testé correspondent à des situations diverses, avec des combinaisons variées d'évolutions de volume et d'activité. Nous avons ainsi généré huit cas différents, chacun composés de quatre étapes, présentés dans la figure 5.10. Ces cas ont pour but de représenter des cas vraisemblables d'évolution de tumeurs, tout en présentant une variété d'évolution, tant en terme de forme et de volume que d'intensité. Ils ont été jugés pertinents par deux experts médecins nucléaires.

Cas 1

Le cas numéro 1 correspond à une tumeur dont le volume augmente légèrement à chaque étape, et dont la distribution d'activité change également de façon significative, tant au niveau de sa répartition que de son intensité générale. Le contraste initial est de 8 :4 :1 (stable à l'étape 2), puis augmente à 10 :7 :1 (étape 3) pour finir à 12 :1 (étape 4). L'hétérogénéité évolue également de façon importante au cours des étapes, pour finir complètement homogène.

Cas 2

L'évolution de la tumeur 2 débute avec la même tumeur que le cas numéro 1 mais évolue différemment, avec une diminution très significative du volume et de l'activité. A l'étape 2, la tumeur reste hétérogène mais son volume a été divisé par 2, bien que l'activité de la zone restante soit globalement inchangée (contraste 8 :4 :1). A l'étape 3, le volume a été à nouveau divisé par 2 et l'activité est dorénavant homogène avec un contraste de 6 :1. Enfin, à la dernière

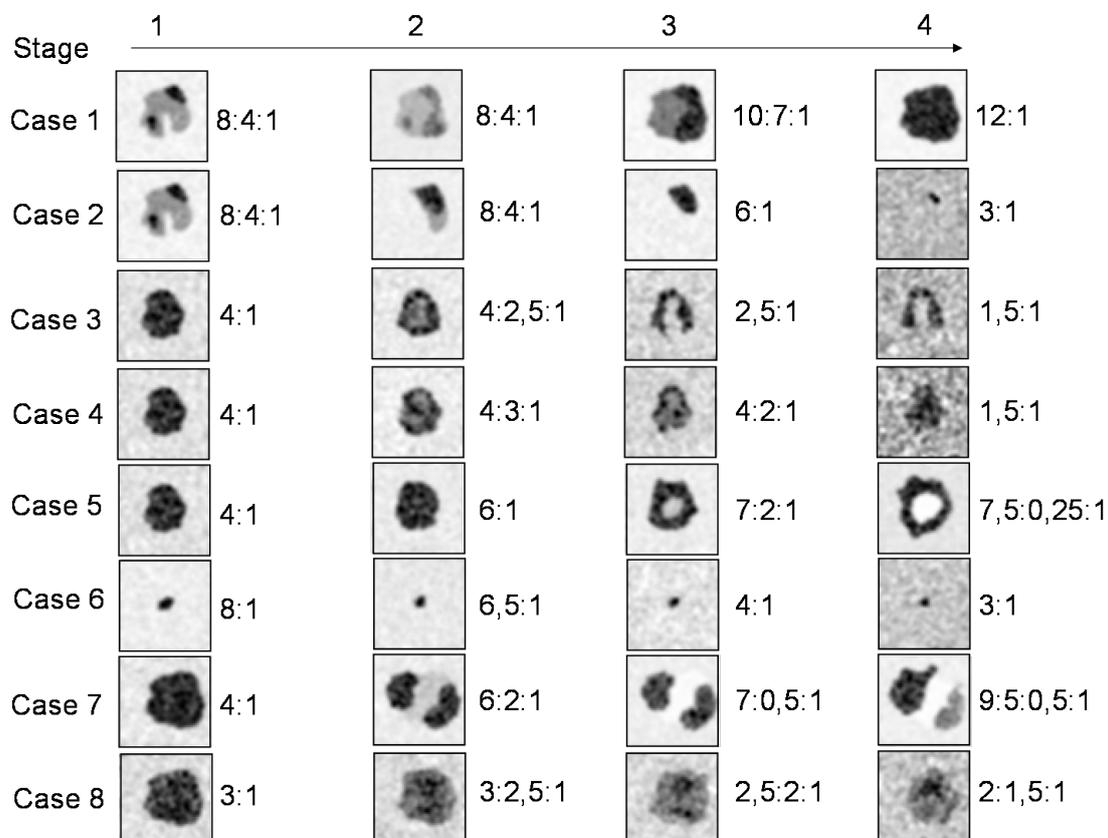


FIG. 5.10 – Les huit cas considérés pour l'évaluation de l'utilisation des méthodes pour le suivi thérapeutique.

étape, il ne reste plus qu'une petite tumeur ne représentant qu'une petite fraction du volume initial, et dont l'intensité a fortement diminué (contraste 3 :1).

Cas 3

Le cas numéro 3 présente une évolution commençant avec une tumeur homogène au contraste de 4 :1, puis évolue vers une tumeur présentant une diminution d'activité en son centre (contraste 1 :2,5 :4) résultant en une faible hétérogénéité. A l'étape 3, le centre est devenu inactif, réduisant fortement le volume de la tumeur (divisé par 2 environ), et le contraste global est tombé à 2,5 :1. A la dernière étape, le volume tumoral a légèrement diminué et son activité est devenue très faible, avec un contraste de 1,5 :1 seulement.

Cas 4

Le cas numéro 4 est plus simple. Il commence avec la même tumeur que le cas 3, mais le volume reste relativement stable (il diminue progressivement mais seulement de 20 % environ). Son activité commence homogène (contraste 4 :1), et devient légèrement hétérogène (4 :3 :1 puis (3 :2 :1), pour finir avec un volume homogène de très faible contraste (1.5 :1).

Cas 5

La tumeur 5 est la même à l'étape 1 que dans les cas 3 et 4 (homogène, contraste 4 :1). Elle évolue avec une forte augmentation du volume et de l'activité (volume + 11% et contraste 6 :1 à l'étape 2, +25 % à l'étape 3), et une nécrose se forme au centre (contraste 7 :2 :1). A l'étape 4, la tumeur présente un volume encore plus important et une intensité fortement augmentée sur les bords, mais le coeur nécrosé présente une intensité inférieure aux tissus sains environnants (contraste 7,5 :1 :0,25).

Cas 6

Dans le cas 6, on considère un petit objet dont le volume évolue mais reste faible (-50 % entre la première étape et la dernière). Son activité diminue également fortement, passant d'un contraste de 8 :1 à 6,5 :1 puis 4 :1, et enfin 3 :1.

Cas 7

Le cas 7 est le plus complexe considéré, et se situe aux limites du réalisme en ce qui concerne l'évolution naturelle des tumeurs. Ici on considère une tumeur volumineuse et homogène (contraste 4 :1), qui à partir de l'étape 2 se scinde en deux parties plus active (contraste 6 :1), avec une zone au centre dont l'activité diminue (contraste 2 :1) et devient nécrosée à l'étape 3 et 4, avec une intensité inférieure aux tissus sains environnants (contraste 1 :2). Dans le même temps, les deux parties plus actives, séparées à l'étape 2, évolue différemment. D'abord elles augmentent toutes deux d'intensité (contraste 7 :1) sans changer de volume à l'étape 3. Enfin à l'étape 4, la première partie augmente de volume (+ 50 %) et d'intensité (contraste 9 :1), tandis que la deuxième partie diminue de volume (-20 %) et d'intensité (contraste 5 :1).

Cas 8

Le dernier cas considéré considère une tumeur volumineuse évoluant de façon peu significative en terme de volume (- 15 %) mais dont l'activité est divisé par 2 entre l'étape 1 et l'étape 4, et une légère hétérogénéité qui se développe dans l'étape 2 et 3 mais avec des contrastes faibles de l'ordre de 2 :1.

5.3.2 Résultats pour chaque cas

Les figures 5.11 à 5.18 présentent les résultats du suivi de volume de chacun des huit cas considérés. Les figures présentent l'évolution du volume (en ordonnée) en fonction des quatre étapes simulées (en abscisse). Deux seuillages à 42 et 50 % du maximum respectivement (T42 et T50), ainsi qu'un seuillage adaptatif (Tbckg) sont comparés à FLAB. FLAB 2 indique l'utilisation de la version à deux classes dures uniquement, pour mettre en valeur ses limites dans le cas de forte hétérogénéité, et FLAB 2-3 indique l'utilisation appropriée de la méthode à trois classes dures lorsque cela est nécessaire (forte hétérogénéité).

Cas 1

Comme on peut le constater dans la figure 5.11, toutes les méthodes binaires ont tendance à sous-estimer le volume réel de la tumeur, notamment dans les deux premières étapes, à cause de l'hétérogénéité et des contrastes mis en jeu dans cette configuration, comme on avait pu le constater dans la section précédente (voir 4.5). Seule l'utilisation de la méthode à trois classes permet de suivre correctement l'évolution de la tumeur dans le temps, comme le montre l'accord entre la courbe grise (vérité) et bleue claire (FLAB utilisant 3 classes quand nécessaire). Le seuillage adaptatif se montre performant, avec une courbe située entre FLAB 2 et FLAB 2-3.

Cas 2

Dans le cas 2, on retrouve le même problème que dans le cas 1, à savoir une sous-évaluation nette du vrai volume de la part des méthodes binaires pour les deux premières étapes. Il est

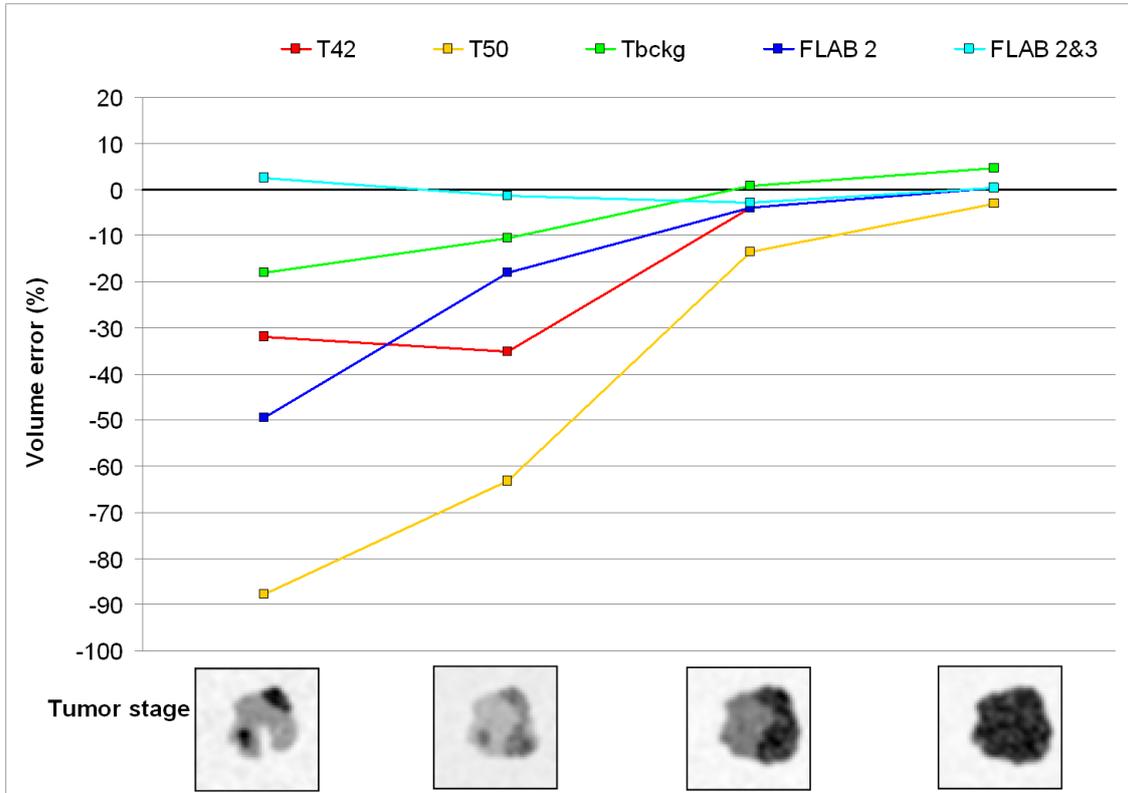


FIG. 5.11 – Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 1.

intéressant de constater que si les méthodes retrouvent bien le bon volume dans les étapes 3 et 4, T42 sur-estime tout de même significativement le volume dans la dernière étape, ce qui est normal car c'est un petit objet faiblement contrasté. A nouveau, l'utilisation appropriée de la méthode à trois classes permet de parfaitement estimer l'évolution du volume, là où l'utilisation seule de FLAB binaire est en échec dans les deux premières étapes. Comme dans le cas précédent, l'utilisation du seuillage adaptatif permet d'obtenir de meilleurs résultats qu'avec FLAB 2, mais de moins bons qu'avec FLAB 2-3 (à part sur le dernier stage).

Cas 3

Le cas 3 ne nécessite à aucune étape l'utilisation d'une approche à trois classes. Ici, l'hétérogénéité est suffisamment faible pour que FLAB2 suffise à extraire la tumeur dans son ensemble du fond. Toutes les méthodes permettent de bien suivre l'évolution du volume,

5.3. SUIVI DU VOLUME TUMORAL AU COURS DU TEMPS POUR LE SUIVI THÉRAPEUTIQUE : RÉSULTATS SUR TUMEURS SIMULÉES

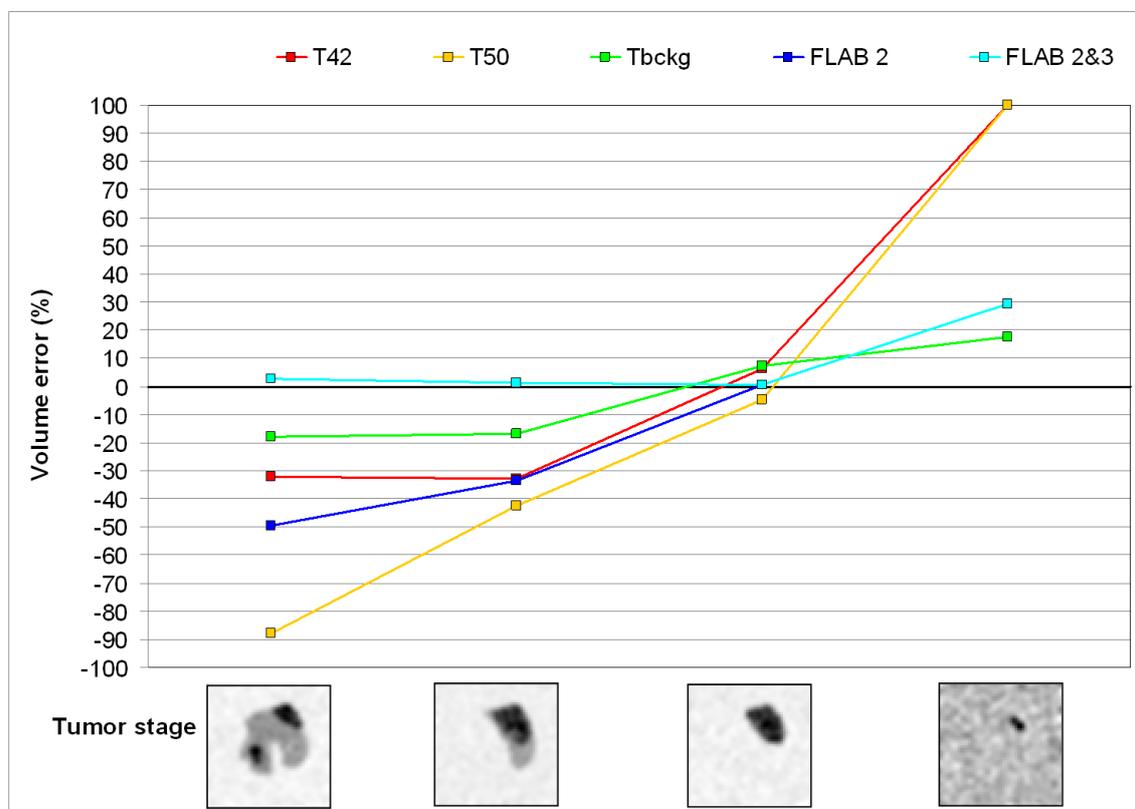


FIG. 5.12 – Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 2.

sauf à la dernière étape, où le contraste est devenu trop faible et pour laquelle les méthodes par seuillage sur-estiment le volume, en particulier T42, qui échoue totalement à extraire l'objet du fond, là où T50 réussit mieux, mais avec une sur-évaluation du volume suffisante pour fausser complètement l'analyse de l'évolution de la tumeur. Le seuillage adaptatif fonctionne très bien sur les trois premiers stages (moins de 10 % d'erreur), mais sous-estime fortement le volume (presque 100 %) à la dernière étape en raison du faible contraste associé à une hétérogénéité rendue plus importante par le bruit associé au contraste faible. En effet, la région utilisée pour calculer le rapport entre l'activité de la tumeur et celle du fond est déterminée par un seuillage initial de 70 % et seule une petite région autour du voxel d'intensité maximale est alors conservée, le rapport entre le fond et l'objet étant sur estimée, le seuillage adaptatif ne conserve que les régions d'intensité forte au sein de la tumeur, ratant les autres régions d'intensité plus faible.

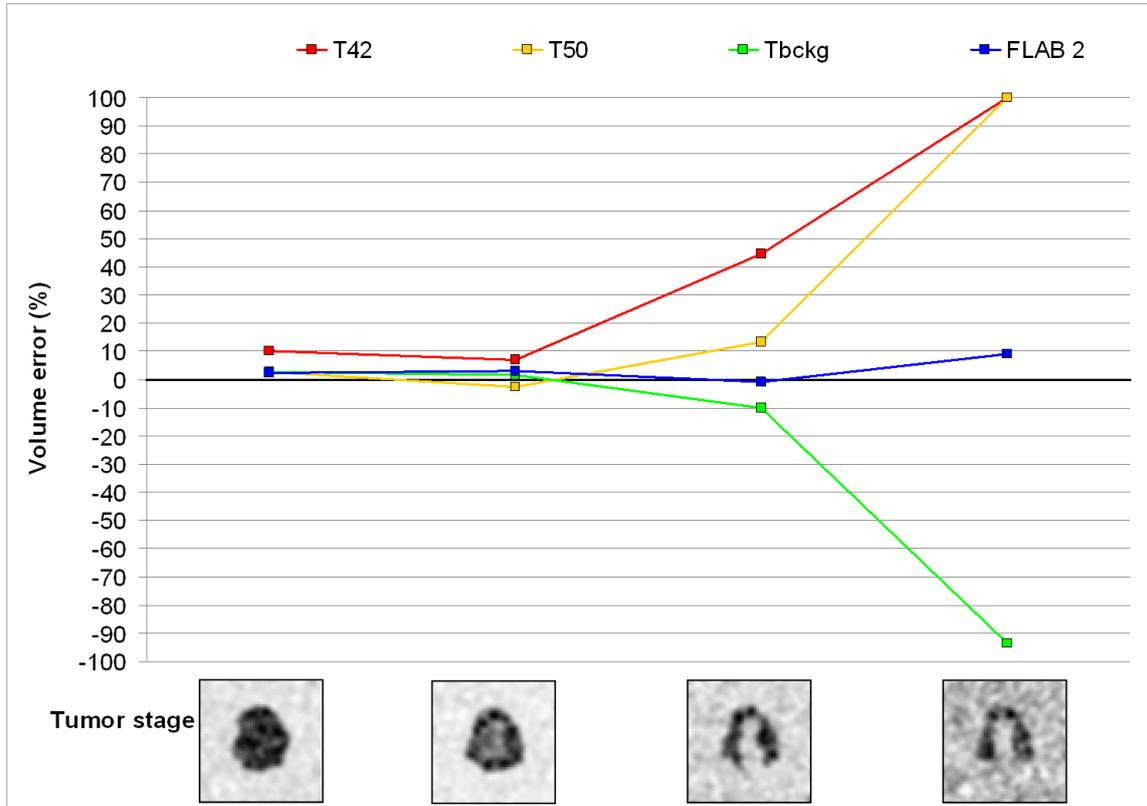


FIG. 5.13 – Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 3.

Cas 4

On observe exactement le même comportement dans le cas 4 que dans le cas 3, à l'exception près que T50 échoue avec la même intensité que T42. Le seuillage adaptatif quant à lui, fonctionne bien sur les deux premières étapes, mais sous-estime fortement le volume dans les deux derniers (-20 et -60 % respectivement), pour la même raison que dans le cas 3 : le contraste devenant faible et le bruit important, l'hétérogénéité entraîne une sélection faussée de la région utilisée pour calculer le rapport entre l'objet et le fond, ce qui conduit à la définition d'un seuil inapproprié à la sélection de l'ensemble de la tumeur.

Cas 5

Le cas 5 est plus complexe à étudier. T42 et T50 se révèlent les moins performants pour suivre l'évolution de la tumeur. En effet, T42 sur-estime tout d'abord le volume pour ensuite le sous-

5.3. SUIVI DU VOLUME TUMORAL AU COURS DU TEMPS POUR LE SUIVI THÉRAPEUTIQUE : RÉSULTATS SUR TUMEURS SIMULÉES

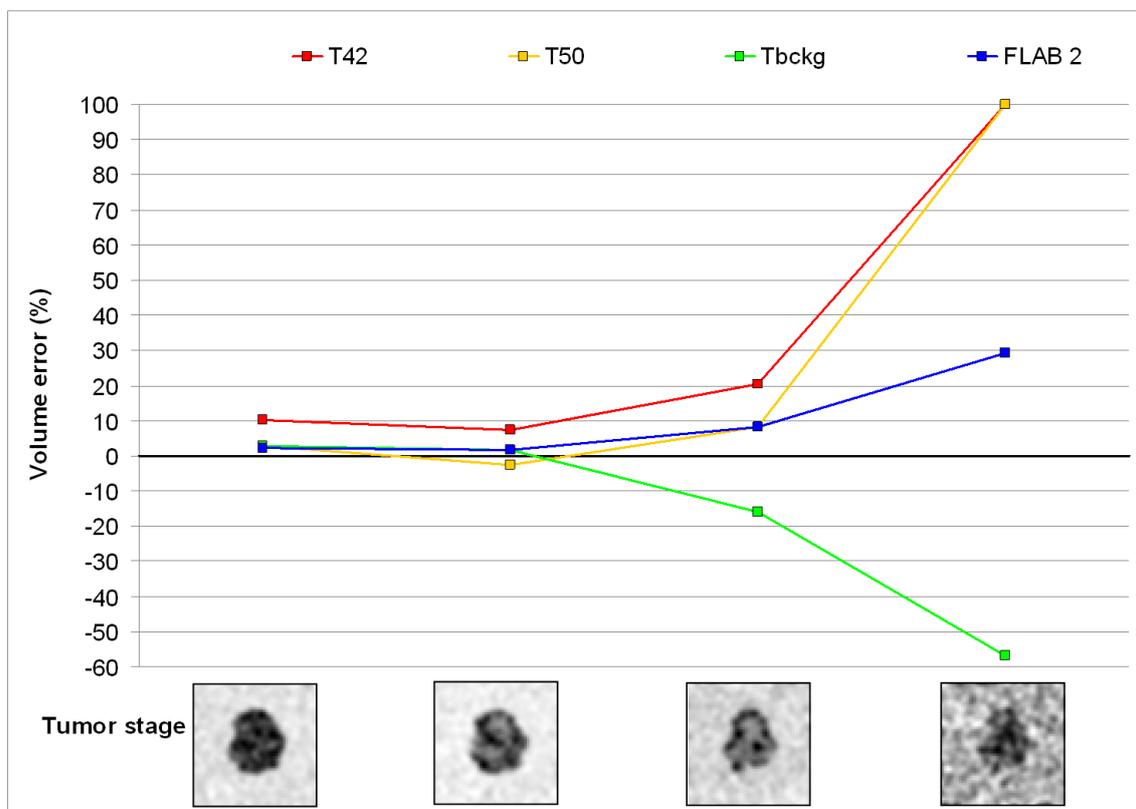


FIG. 5.14 – Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 4.

estimer, lorsque le coeur nécrosé apparaît. Enfin, à la dernière étape il sur-estime légèrement le volume. T50 de son côté permet de correctement suivre le volume dans les deux premiers stages, mais sous-estime le volume à l'étape 3, et continue à le sous-estimer à l'étape 4. Le seuillage adaptatif a à peu près le même comportement que FLAB : les deux méthodes sont en accord avec le vrai volume tout au long de l'évolution de la tumeur, sauf à l'étape 3, pour laquelle ils excluent le coeur nécrosé en cours de formation (mais qui devrait tout de même être considéré comme faisant partie de la tumeur étant donné que son intensité est supérieure aux tissus sains), d'où une sous-estimation du volume. Ceci vient du fait que le contraste entre le coeur en cours de nécrose et le fond est faible tandis qu'il est élevé entre la zone nécrotique et le reste de la tumeur. L'utilisation de la méthode à trois classes permet de suivre correctement le volume, y compris à cette étape. Noter qu'à la dernière étape, étant donné que le coeur est nécrosé et d'intensité inférieure aux tissus sains environnants, mais que le contraste global

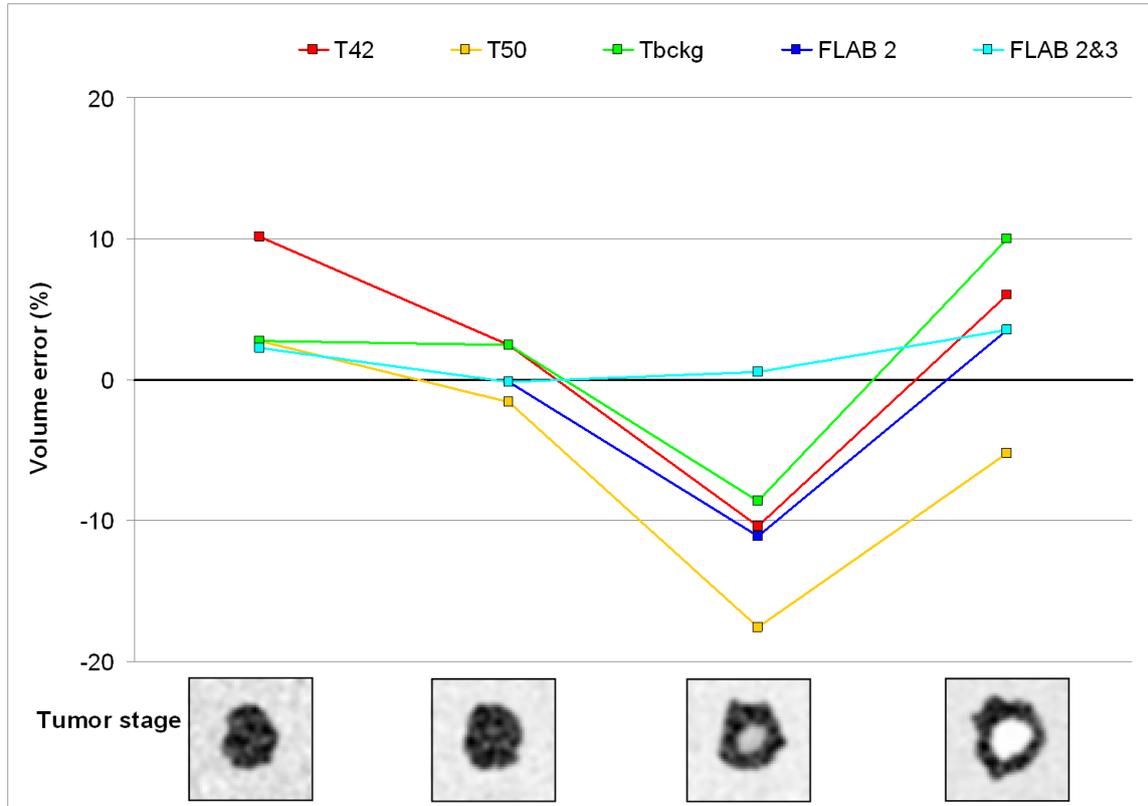


FIG. 5.15 – Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 5.

du reste de la tumeur par rapport au fond est élevé, la méthode binaire suffit à segmenter correctement la tumeur (puisque cette fois, le coeur nécrosé n'est plus à considérer comme de la tumeur). Il est important toutefois de souligner que toutes les méthodes restent entre -10 et +10 % d'erreur (à part T50 qui atteint -20 %), ce qui est raisonnable. Toutefois la mauvaise prise en compte du coeur nécrosé à l'étape 3 pourrait fausser l'évaluation de la réponse thérapeutique.

Cas 6

Toutes les méthodes permettent de suivre l'évolution de la petite tumeur du cas 6 jusqu'à l'étape 2, mais seul FLAB est capable de correctement estimer (malgré une légère sur-estimation) les deux dernières étapes. Les deux seuillages indiquent de façon erronée une augmentation forte du volume à l'étape 3 et 4, alors que celui-ci diminue. Quant au seuillage

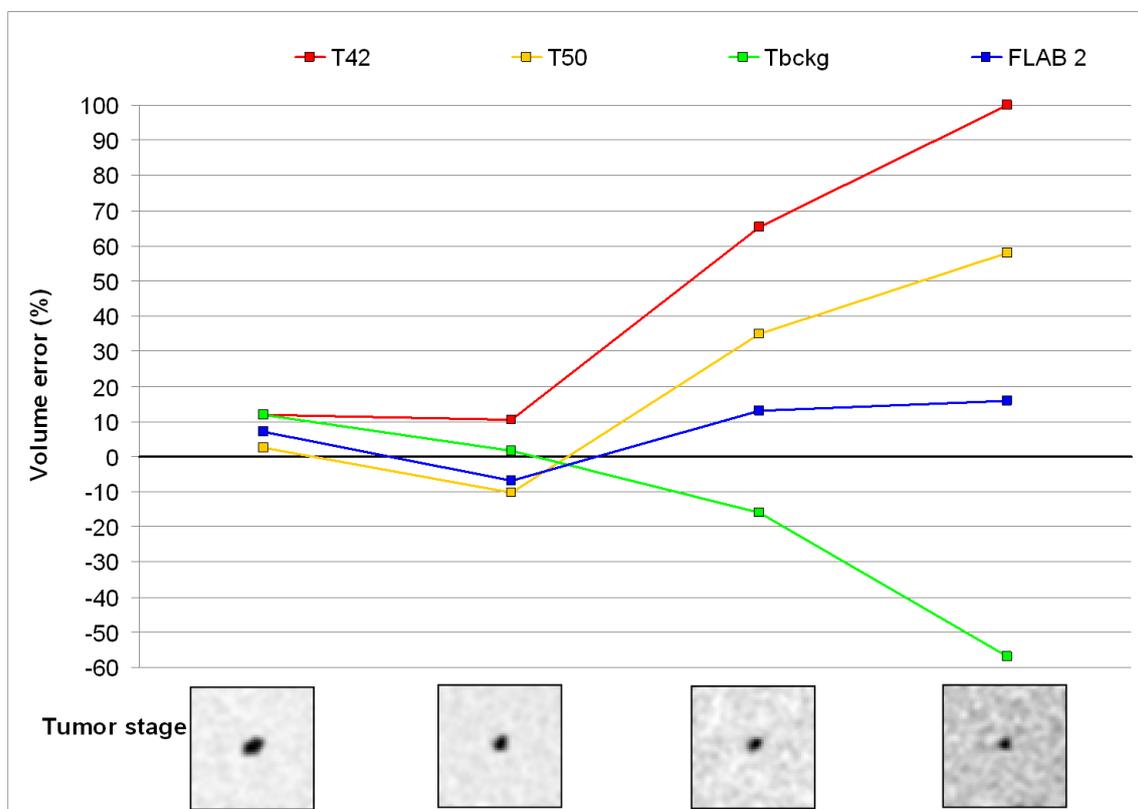


FIG. 5.16 – Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 6.

adaptatif il indique une trop forte diminution du volume, avec une sous-évaluation de 20 et 60 % aux étapes 3 et 4 respectivement, car le contraste bien que faible, est suffisant pour que la région d'intérêt initiale définie par un seuillage à 70 % corresponde en fait déjà quasiment au volume qu'il faudrait conserver, ce qui a pour effet de mener à un seuil trop faible pour agrandir ce volume initial. FLAB sous-estime cette diminution, mais permet tout de même l'estimation la plus proche de la réalité.

Cas 7

Comme constaté dans les cas hétérogènes considérés auparavant, seule l'utilisation de FLAB avec trois classes permet de suivre correctement l'évolution de la tumeur tout au long des 4 étapes. Les seuillages, y compris l'approche adaptative, estiment correctement la configuration de départ, mais sous-estiment à chaque étape suivante de façon significative le volume tumoral,

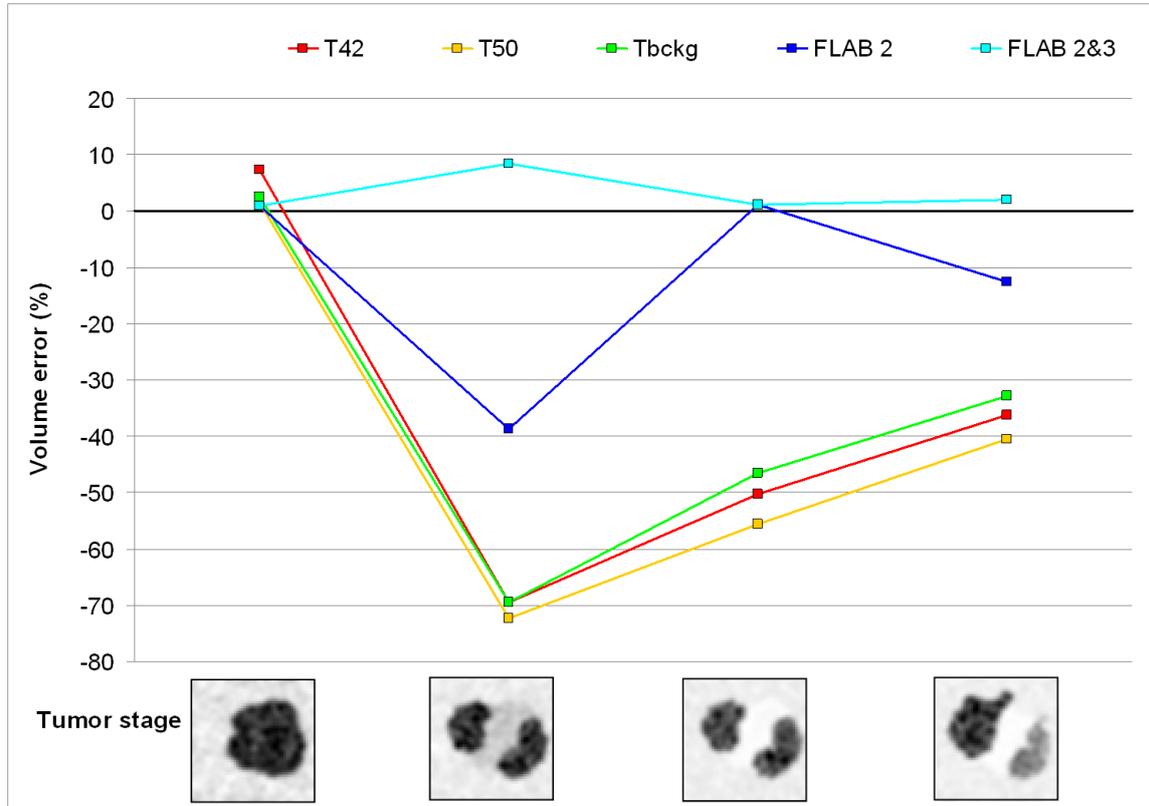


FIG. 5.17 – Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 7.

en particulier à l'étape 2. FLAB dans sa version binaire sous-estime également fortement l'étape 2 et légèrement l'étape 4. Une observation intéressante sur ce cas, est que les trois approches par seuillage ont le même comportement, à savoir qu'ils ne conservent que la partie gauche de la tumeur à l'étape 2, et que la partie droite à l'étape 3. Cela s'explique par le fait que le voxel d'intensité maximale (utilisé comme graine d'initialisation de la croissance de région) change de zone tumorale entre l'étape 2 et l'étape 3 et que la connexité entre les deux régions passe par une région centrale d'intensité beaucoup plus faible (et la croissance de région ne s'y propage donc pas). Il faudrait dans ce cas utiliser plusieurs graines d'initialisation pour résoudre cette limite des seuillages.

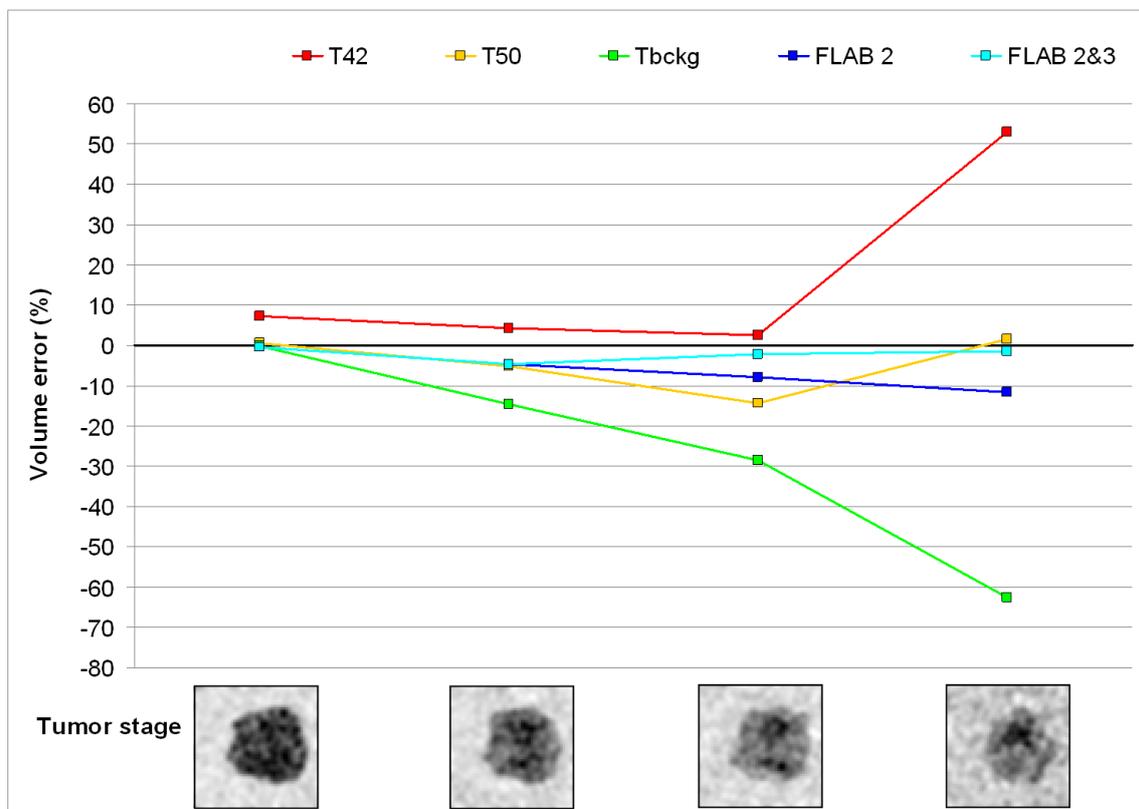


FIG. 5.18 – Résultats de l'évaluation de FLAB pour le suivi thérapeutique sur le cas 8.

Cas 8

Les résultats obtenus dans ce dernier cas montre que même une faible hétérogénéité peut mettre en évidence des différences entre les méthodes, si les contrastes considérés sont faibles. En effet, ici la méthode FLAB à trois classes permet d'affiner clairement le suivi de l'évolution du volume tumoral, à l'étape 3 et 4, pour lesquelles la méthode binaire sous-estime le vrai volume. T42 sur-estime fortement le volume, en particulier à la dernière étape où le contraste est le plus faible. T50 permet de correctement suivre le volume aux deux premières étapes et à la dernière, mais le sous-estime fortement à l'étape 3, du fait de l'hétérogénéité (seules les zones d'intensité élevées au sein de la tumeur sont conservées dans le volume). Le seuillage adaptatif estime correctement le volume à l'étape initiale, mais le sous-estime systématiquement dans les trois étapes suivantes (de -20 à -60 %) pour la même raison que dans les cas précédents : la sélection initiale avec un seuillage à 70 % ne conserve que les régions d'intensité maximale

au sein de la tumeur, et le contraste objet / fond calculé résultant est inapproprié pour que la croissance de région se propage dans le reste de la tumeur.

Performances sur l'ensemble des huit cas

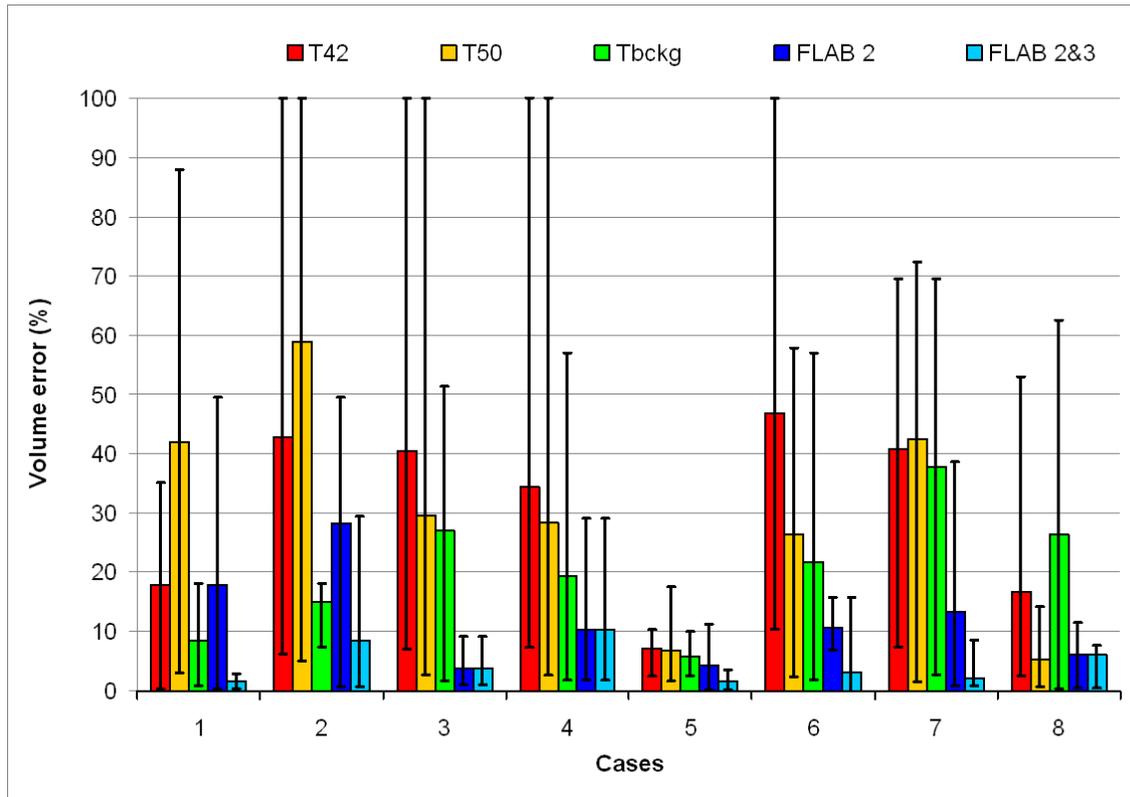


FIG. 5.19 – Erreurs moyennes et écart types calculés sur chacun des 8 cas, pour les méthodes considérées.

Comme on peut le constater dans la figure 5.19 qui présente les erreurs des méthodes sur chaque cas (moyenne et écart type calculés sur les quatre étapes pour chaque cas), Globalement, les résultats sont en faveur de FLAB, mais il apparaît toutefois que la version à deux classes dures est parfois mise en échec par certaines configurations, comme en témoignent les résultats de FLAB 2 dans les cas 1, 2, 5, 7 et 8. Au contraire, lorsque l'utilisateur peut utiliser également la version à trois classes dures de FLAB, les configurations complexes et hétérogènes sont bien prises en compte, ce qui permet un suivi quasi parfait de la plupart des configurations dans les huit cas considérés, et met en évidence la nécessité d'utiliser la

méthode à trois classes lorsque cela est nécessaire. La sélection du nombre de classes (2 ou 3) dans le cadre de FLAB est pour l'instant déterminée par l'utilisateur, mais il est envisageable de l'automatiser grâce à un algorithme de détection du nombre de classes dans l'image, tel que proposé dans [100], une approche basée sur l'utilisation du critère AIC (*Akaike's Information Criterion*) [2] et l'algorithme ML-EM, pour déterminer automatiquement le nombre de classes présentes dans une image, entre 1 et K_{max} le nombre maximal de classes.

Chapitre 6

Conclusion et perspectives

6.1 Méthode développée

6.1.1 Mérites et performances

Nous avons développé et validé une approche automatique, robuste et précise de segmentation en trois dimensions des volumes fonctionnels en imagerie d'émission (TEP ou TEMP). Cette approche est fondée sur la combinaison de la modélisation statistique dans le cadre de l'inférence bayésienne avec une mesure floue, pour prendre en compte la problématique spécifique aux images d'émission du bruit et du flou dû aux effets de volume partiel. Une attention particulière a été portée à l'étude de la robustesse de la méthode pour permettre son utilisation dans les différents centres cliniques utilisant des scanners ou des algorithmes de reconstruction d'images différents. En effet, contrairement à la majorité des travaux proposant une approche pour déterminer les volumes en TEP, nous avons en effet étudié les performances de notre approche sur une large gamme de paramètres d'acquisitions, de types de scanners ou de reconstructions, d'images simulées et réelles, avec des configurations plus ou moins difficiles de contraste et de bruit (un paramètre souvent négligé dans les évaluations des méthodes proposées) ainsi qu'exploré l'application à des images plus complexes et plus réalistes de tumeurs, dont certaines fortement hétérogènes. De plus, nous avons exploré et proposé différentes utilisations possibles de la méthode pour des applications spécifiques en oncologie, notamment la planification de radiothérapie guidée par l'image fonctionnelle, le diagnostic et

le suivi thérapeutique, pour lesquelles les utilisateurs cliniciens sont demandeurs d'outils de segmentation précis et robustes. Les résultats de validation de l'approche proposée démontrent sa nette supériorité par rapport à la majorité des approches proposées jusqu'à présent dans la communauté pour déterminer les volumes dans les images d'émission, notamment en ce qui concerne les tumeurs réalistes, souvent complexes à segmenter, par leur forme et / ou leur distribution d'activité, et qui tiennent en échec les méthodes basées sur seuillages généralement proposées, notamment dans les stations des industriels. L'utilisation de FLAB dans le cadre du diagnostic et le suivi thérapeutique nécessite la prise en compte des effets de volume partiels sur la quantification, et éventuellement la combinaison de l'approche proposée avec une méthode de correction de ces effets. Le domaine de la radiothérapie guidée par l'image pourrait bénéficier de l'utilisation directe de l'algorithme pour définir des plans de traitements précis à l'aide de l'image fonctionnelle (voir section 6.2.1).

6.1.2 Limites

L'approche proposée nécessite l'intervention de l'utilisateur pour la détection de la tumeur et la sélection de la région d'intérêt, introduisant une dépendance à l'utilisateur. Toutefois, dans le cas de la radiothérapie, les volumes tumoraux à traiter sont en général de grande taille et sont donc assez facilement repérable. Il est envisageable d'utiliser une étape préliminaire de détection automatique des tumeurs dans l'image corps entier [154]. Il restera toutefois le problème des zones à fixations élevées mais non pathologiques, à proximité de la tumeur, comme par exemple le coeur qui peut se retrouver en partie dans une sélection cubique d'une tumeur pulmonaire. Il est possible de permettre à l'utilisateur de définir la sélection autour de la tumeur non pas comme un cube, mais comme une zone tracée à main levée, au prix d'une plus grande variabilité, mais permettant ainsi d'exclure facilement les régions à fixation élevée, mais qui ne doivent pas être incluses dans le volume tumorale.

De plus, le choix de l'utilisation de la version binaire (à deux classes et une transition floue) ou celle à trois classes et trois transitions floues n'est pas encore automatisé, et il revient à l'utilisateur de décider quelle version appliquer à sa sélection. Dans la plupart des cas, l'estimation de l'hétérogénéité de la tumeur est facilement faite visuellement, mais certains

cas de tumeurs présentant une hétérogénéité "moyenne" peuvent être plus complexes. Il est toutefois envisageable de contourner cette limitation, soit en proposant à l'utilisateur les deux résultats de segmentation, en lui laissant la responsabilité de sélectionner la segmentation qui lui semble la plus appropriée, soit automatiser le choix de la méthode grâce à une approche d'estimation du nombre de classes présentes dans la zone sélectionnée au préalable comme celle présentée dans [100], une approche basée sur l'utilisation du critère AIC (*Akaike's Information Criterion*) [2] et l'algorithme ML-EM, pour déterminer automatiquement le nombre de classes présentes dans une image, entre 1 et K_{max} le nombre maximal de classes.

Enfin, il est important de garder à l'esprit que la détermination du volume, même la plus précise possible, ne débouche pas sur une quantification exacte de l'objet d'intérêt. En effet, extraire la moyenne des voxels de la zone délimitée par l'algorithme FLAB résulte en général en une sous-estimation (dans le cas d'un objet d'intérêt dont l'activité est supérieure aux tissus environnants, et une sur-estimation dans le cas contraire) de l'activité réelle inversement proportionnelle à la taille de l'objet, mais toujours significative, même dans le cas d'objets dont les dimensions sont de plusieurs centimètres. Pour obtenir une quantification précise d'un objet d'intérêt délimité, même avec précision, il reste nécessaire d'appliquer une phase de correction quantitative, telle que celle que nous proposons dans la section 6.2.2.

6.2 Perspectives et applications

L'approche développée a montré des performances prometteuses pour la segmentation précise, robuste et automatique de volumes d'intérêt sélectionnés sur les images tri-dimensionnelles TEP. Cette segmentation peut se révéler très utile dans des applications précises de l'imagerie TEP : la radiothérapie guidée par l'image, la détermination du volume tumoral et sa quantification pour le diagnostic et le suivi thérapeutique, une fois combinée avec une correction d'effets de volume partiel.

6.2.1 Application : radiothérapie guidée par l'image (IGRT)

La radiothérapie est aujourd'hui largement guidée par l'image (IGRT pour *Image Guided Radiotherapy*), que ce soit pour la définition du traitement (planning) avec notamment la modulation des doses, la gestion de la délivrance des rayonnements (avec le contrôle en temps réel nécessitant des modélisations 4D, notamment pour gérer les mouvements du patients (surtout respiratoires), et l'évaluation de la réponse thérapeutique.

Principe

Cette forme de traitement a pour objectif de délivrer une dose de rayons X, électrons, ions carbonés, ou protons dans les tissus tumoraux. C'est un domaine qui a fait d'importants progrès : les outils utilisés pour délivrer les doses sont en particulier devenus plus fiables et plus précis [13]. Notons qu'il existe différents systèmes tels que la 3D-CRT (*Tri-Dimensional Conformational Radiotherapy* ou Radiothérapie conformationnelle tri-dimensionnelle) qui permet la délivrance de rayonnements non modulés sur différents plans en trois dimensions, l'IMRT (*Intensity-Modulated Radiotherapy* ou Radiothérapie à intensité modulée), une version améliorée de la 3D-CRT qui permet de délivrer des doses adaptées à la forme de la cible et permettant d'épargner les organes à risques et tissus sains environnants, et le très récent VMAT (*Volumetric intensity Modulated Arc Therapy* ou thérapie par arc à intensité modulée volumique), qui permet d'atteindre le même résultat mais avec une seule rotation (jusqu'à 360°) au cours de laquelle l'intensité peut être modulée, ce qui permet des temps de traitement réduits (environ moitié moins longs) et une dosimétrie optimisée, avec notamment une diminution des doses délivrées aux organes à risques, d'après une étude récente [119] qui a comparé ces différentes techniques dans le cadre du traitement de la prostate. On parle également de ART (*Adaptive Radiotherapy* ou radiothérapie adaptative) pour les approches permettant d'adapter la délivrance du traitement en fonction de la respiration et des mouvements de la cible observée en temps réel et le *re-planning*, c'est à dire l'adaptation des séances successives en fonction de la réponse au traitement et des changements fonctionnels et anatomiques de la tumeur, que ce soit en terme de taille et de forme, de positionnement ou des aspects de

biologie.

La radiothérapie est toujours en pleine évolution technologique, avec pour objectif principal l'amélioration des modalités de délivrance de la dose et une meilleure définition des volumes cibles. Les retombées espérées en termes d'amélioration du traitement dépendent essentiellement, lors des différentes étapes d'une séance de radiothérapie, de l'intégration optimale d'informations pertinentes issues de plusieurs modalités d'imagerie. Les systèmes délivrant les rayons sont capables de le faire avec une précision de l'ordre du millimètre, en suivant des contourages très précis dans l'espace. Ils possèdent la capacité technique et physique d'irradier l'intérieur d'une tumeur différemment suivant les contours définis en amont par les utilisateurs. Des études lourdes et longues sont encore nécessaires pour prouver un bénéfice pour les patients en terme de survie de l'inclusion de ces nouvelles méthodologies et techniques de délivrance dans la pratique clinique.

Pratique clinique actuelle

Le problème est que ce contour est, pour l'instant, toujours largement défini sur l'image anatomique, ce qui néglige l'apport de l'information fonctionnelle, pourtant jugée critique pour cette application [13]. L'utilisation de l'imagerie fonctionnelle en radiothérapie représente une activité en pleine croissance et dont l'intérêt est devenu considérable depuis le développement et l'introduction massive des systèmes d'imagerie multimodalités TEP/TDM en routine clinique, considérés à présent comme l'état de l'art de l'imagerie médicale. Toutefois la communauté des radiothérapeutes est encore minoritaire à utiliser l'imagerie TEP ou IRM dans le cadre de la planification ou de la délivrance du traitement (environ 10 %) [47].

Dans le cadre de l'IMRT, les systèmes utilisés sont des accélérateurs linéaires (LINAC) capables de pivoter autour du patient et les rayons ionisants générés par les LINAC sont donc délivrés au patient par focalisation, mais il y a toujours une partie des rayons qui traversent les tissus du patient et y déposent une certaine énergie. Le LINAC est équipé d'un collimateur multi-lames qui permet d'adapter la forme du rayon au contour défini en amont, de façon à irradier la tumeur le plus précisément possible et limiter les doses délivrées aux alentours. La figure 6.1 illustre le fonctionnement de la radiothérapie conformationnelle.

Noter qu'il existe depuis quelques années une technique de thérapie par protons qui permet

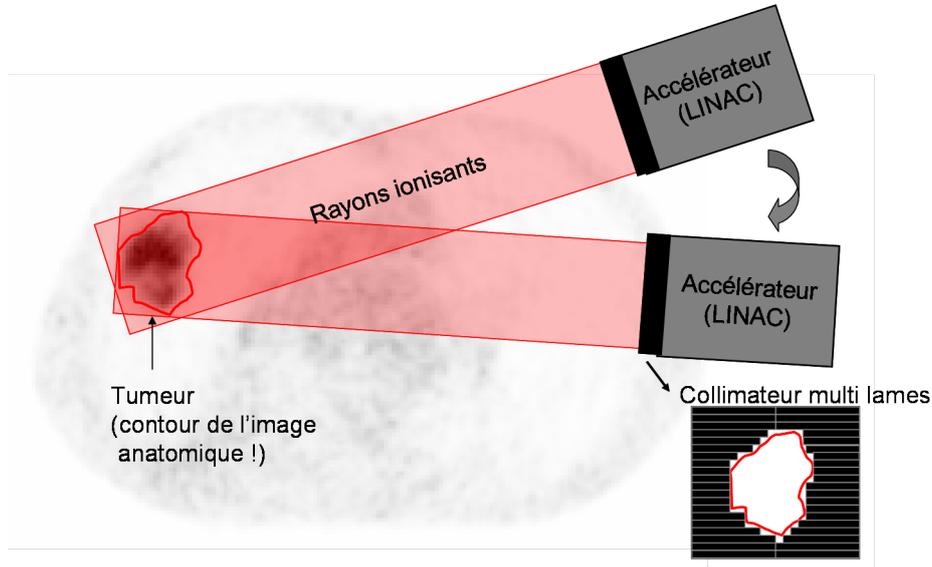


FIG. 6.1 – Illustration du fonctionnement de la radiothérapie externe conformationnelle.

de limiter significativement (diviser par 2 ou plus) les doses délivrées aux tissus environnants car ces derniers ont la propriété de déposer leur énergie en bout de course (la distance dépendant de l'énergie du proton), et non le long de leur trajet comme les photons ou les rayons X [120]. Malheureusement d'autres limitations sont associées à ce type de thérapie, comme par exemple l'hétérogénéité du parcours des protons, ce qui implique une plus grande sensibilité aux mouvements de la cible [120].

Une nouvelle approche et le principe de *dose painting*

La pratique actuelle de la radiothérapie vise une irradiation de la tumeur optimisée, préservant au maximum les organes à risques, comme la vessie par exemple, et les tissus sains environnants, pour limiter les effets secondaires indésirables. Dans le but d'améliorer la survie et le confort des patients, l'optimisation de la dosimétrie est essentielle, car il s'agit de délivrer des doses plus importantes dans la tumeur, et d'épargner dans le même temps au maximum les zones sensibles en limitant les doses qui leur sont délivrées. Pour cela, il est indispensable de prendre en compte l'information offerte par l'imagerie fonctionnelle. En effet, jusqu'à présent,

la pratique clinique consiste à délimiter les volumes d'irradiation en utilisant presque exclusivement l'image anatomique [13]. Or, les volumes anatomiques ne donnent pas l'information nécessaire pour optimiser la dosimétrie, car en effet, des tissus nécrosés ou dont l'activité cancéreuse est faible peuvent avoir la même densité que les tissus très actifs, et donc apparaître identiques sur l'image. La figure 6.2 illustre bien ce phénomène. Sur cette image, il apparaît clairement qu'utiliser le volume anatomique seul pour définir le volume d'irradiation conduit à irradier des régions dont l'activité cancéreuse est clairement limitée, voir nulle, au centre de la tumeur par exemple. Il s'agit donc de définir les volumes d'irradiation sur les images fonctionnelles, ce qui n'est pas fait actuellement, par manque d'outils adaptés et suffisamment précis, bien que de nombreuses études aient déjà souligné l'importance et le bénéfice de prendre en compte l'information offerte par la modalité TEP/TDM [3] [56]. Notre proposi-

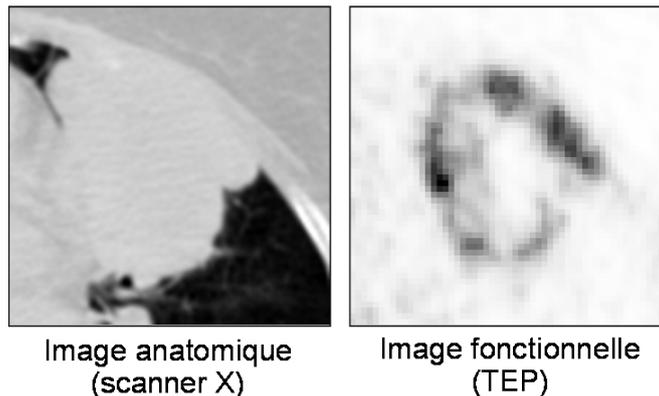


FIG. 6.2 – Illustration d'un cas où utiliser le volume anatomique pour irradier la tumeur ne prend pas en compte la réalité du métabolisme, que l'on peut observer sur l'image fonctionnelle.

tion consiste à combler ce manque en utilisant l'approche de segmentation automatique pour parvenir à l'objectif du *dose painting* [101]. Le *dose painting* consiste en la détermination des différentes zones à l'intérieur d'une masse tumorale, montrant une variabilité en termes de fonction, par exemple au niveau hypoxique, de prolifération cellulaire, de consommation de glucose ou la présence d'une nécrose. La détermination de ces différentes zones au sein même d'une tumeur permet une modulation de la dose à délivrer, menant à l'optimisation effective de la dose avec notamment la possibilité d'augmenter la dose sur les zones les plus actives tout en diminuant celle délivrée aux tissus sains. L'association du principe de dose painting aux

techniques innovantes de délivrance des rayons décrites ci-dessus devrait permettre d'atteindre une meilleure pratique de la radiothérapie, menant à des traitements plus efficaces tout en entraînant moins d'effets secondaires néfastes. Les résultats présentés dans le chapitre "performances des méthodes" sont encourageants en ce qui concerne la capacité de l'algorithme FLAB à correctement segmenter les tumeurs sur les images d'émission, et à fournir une carte de segmentation contenant éventuellement trois classes dans le cas de tumeurs présentant de une distribution d'activité fortement hétérogène. Ceci permet d'envisager l'utilisation de cette segmentation pour implémenter en pratique le principe de *dose painting* en radiothérapie.

Impact clinique : projet ANR

Le travail effectué au cours de ce doctorat a permis de mettre en lumière une piste de recherche jusqu'à présent négligée pour la détermination automatique des volumes fonctionnels en imagerie d'émission, qui présente un intérêt substantiel pour plusieurs applications pratiques de cette modalité d'imagerie. Toutefois, des études prolongées sur la précision et la robustesse de l'algorithme, ainsi que l'impact clinique de l'approche devraient être menées afin de confirmer les résultats encourageants de cette thèse. C'est ce que l'on se propose de mettre en oeuvre, notamment au travers du projet ANR 2008 SIFR.

L'usage de l'algorithme de segmentation FLAB dans le contexte de la radiothérapie peut être illustré par la figure 6.3. La tumeur étant sélectionnée par l'utilisateur dans l'image TEP corps entier du patient, la carte de segmentation de FLAB appliquée à cette sélection peut ensuite être utilisée pour définir une planification de traitement par radiothérapie en fonction des régions de la carte de segmentation, de façon à optimiser les doses. Pour intéresser les industriels du secteur, tant les fabricants de scanners TEP/TDM qui conçoivent également des stations d'analyse d'image associées, que les concepteurs et fabricants de stations d'analyse d'image, ainsi que les spécialistes des stations de planification en radiothérapie, il est nécessaire de valider à plus grande échelle l'approche de segmentation proposée. Ceci est prévu dans le cadre d'un appel à projet de l'Agence Nationale de la Recherche (ANR) pour 2008. Le projet que nous avons déposé et qui a été accepté en juillet 2008 (6 élus sur 42 propositions) est nommé SIFR (pour Segmentation d'Images Fonctionnelles pour la Radiothérapie), propose

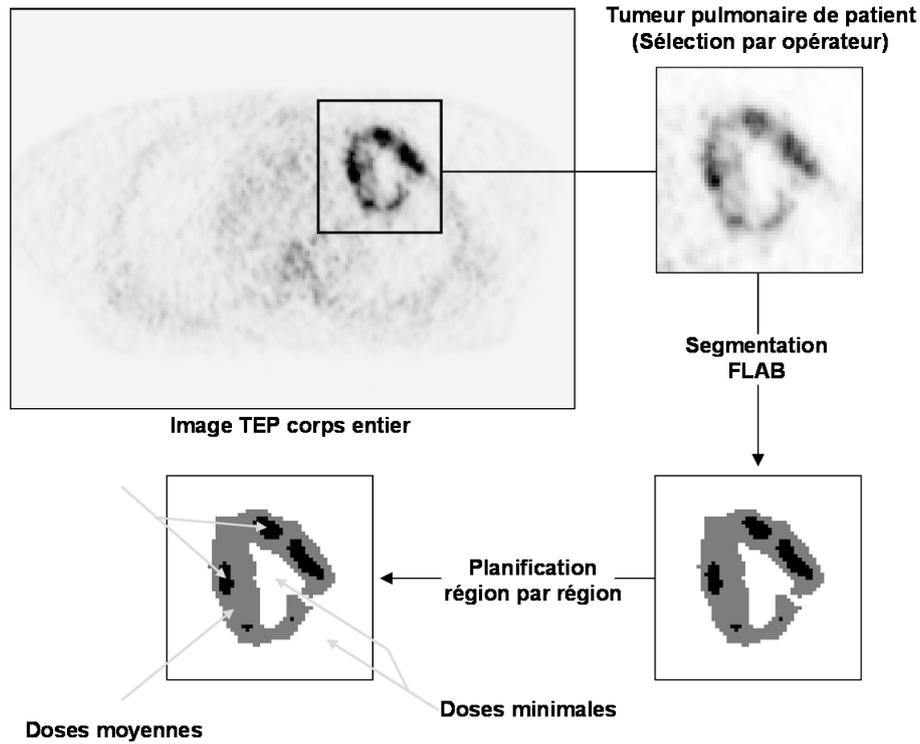


FIG. 6.3 – Positionnement de l’algorithme FLAB dans la planification de traitement par radiothérapie.

sur une durée de deux ans de valider et évaluer en profondeur l’approche de segmentation proposée dans cette thèse, en associant les industriels du secteur. Cette validation est proposée en trois tâches interdépendantes. La première (tâche 1) consiste à étendre la validation et l’étude de robustesse de l’algorithme de segmentation FLAB qui a été entamée dans cette thèse. Il faut pour cela compléter les résultats encourageants obtenus dans cette thèse, en étudiant la capacité de l’algorithme à segmenter correctement les images de fantômes issues d’autres scanners, reconstruits avec des algorithmes variés, aux propriétés de bruit distinctes. La deuxième étape (tâche 2) consiste au développement d’une plateforme logicielle intégrable à une station de planification de radiothérapie, de façon à pouvoir étudier et quantifier l’impact de l’utilisation de FLAB sur la dosimétrie dans un contexte clinique et une utilisation par les techniciens du service, par rapport aux méthodes utilisées habituellement. La phase finale (tâche 3) consiste en une validation complète sur des images simulées et réelles de tumeurs, associée avec une étude histologique permettant de comparer les résultats obtenus

par segmentation sur l'image à ceux obtenus par histologie post-opératoire. Dans le but de valoriser le travail réalisé, ainsi que de s'armer en vue de négociations nécessaires avec les industriels pour développer des solutions innovantes en radiothérapie, en diagnostic et en suivi thérapeutique, un dépôt de brevet a été proposé à la commission régionale de Valorisation Bretagne fin avril 2008. Cette dernière a donné un avis favorable pour le dépôt national du brevet de l'algorithme FLAB pour la radiothérapie. Le brevet a été déposé le 10 septembre 2008 sous la référence FR08/56089.

6.2.2 Application correction d'effets de volume partiel et quantification pour le suivi thérapeutique et le diagnostic

Les effets de volume partiel résultant des résolutions spatiales médiocres des modalités TEMP et TEP constituent une des limites les plus importantes nuisant à l'utilisation quantitative de l'imagerie d'émission (voir 1.6.2). Comme l'ont montré les résultats présentés dans la section 5.3, la segmentation par FLAB est celle permettant de suivre au mieux l'évolution du volume tumoral au cours du temps. Toutefois, la quantification reste un problème à cause de effets de volume partiel qui ont tendance à sous-estimer la valeur de l'activité dans une région d'activité supérieure à son environnement immédiat, ce qui entraîne le fait que disposer du volume correct ne permet pas pour autant d'en extraire directement une quantification correcte. L'idéal serait donc de disposer d'une approche permettant d'obtenir simultanément le volume le plus précis possible, et l'activité corrigée. Ainsi, l'extraction de paramètres quantitatifs thérapeutiques non biaisés et précis devrait permettre d'améliorer significativement le diagnostic ainsi que d'évaluer plus précisément la réponse thérapeutique. De nombreuses études se sont penchées sur la correction de ces effets de volume partiel. On en trouve une revue récente dans [137]. Une des méthodes a été développée au sein du LaTIM par Nicolas Bousson [15] [16], et se propose d'intégrer les détails d'une image haute résolution, acquise avec l'IRM ou la TDM par exemple, dans l'image basse résolution d'émission (TEMP ou TEP). Nous détaillons les principales caractéristiques de cette approche et ses limites dans la section suivante.

Analyse multi-résolutions MMA

L'analyse mutuelle multi-résolution (MMA) pour la correction de volume partiel en imagerie d'émission utilise une décomposition en ondelettes fondée sur l'algorithme "à trous" [10] sur les deux images dont le recalage est supposé parfait voxel à voxel. Ces décompositions offrent une structure hiérarchique des détails et des structures des images à différents niveaux de résolution. L'idée consiste à reconstruire les détails "manquants" dans l'image fonctionnelle à partir des détails de l'image anatomique, en établissant un modèle entre les décompositions en ondelettes de chaque modalité, à un niveau commun de résolution. Ce modèle est ensuite utilisé pour générer les niveaux de détails d'ondelettes manquants à l'image d'émission, à partir des niveaux de détails disponibles dans la décomposition de l'image anatomique. Par somme de ces niveaux recréés, on obtient une image correctrice qui est ajoutée à l'image fonctionnelle d'origine pour obtenir l'image fonctionnelle corrigée. Les figures 6.4 et 6.5 illustrent le fonctionnement de la méthode et le résultat obtenu sur une image du cerveau. L'intérêt de la méthode est

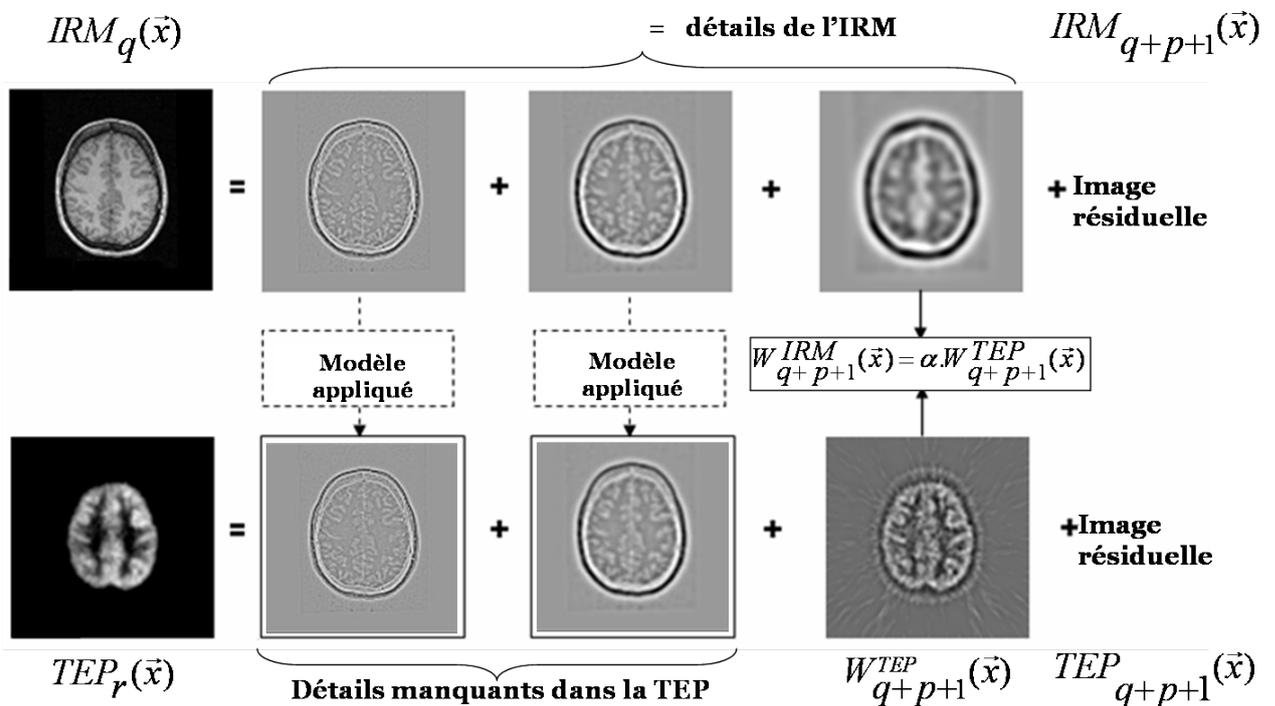


FIG. 6.4 – Principe de l'approche multirésolution développée par N. Boussion pour une image cérébrale.

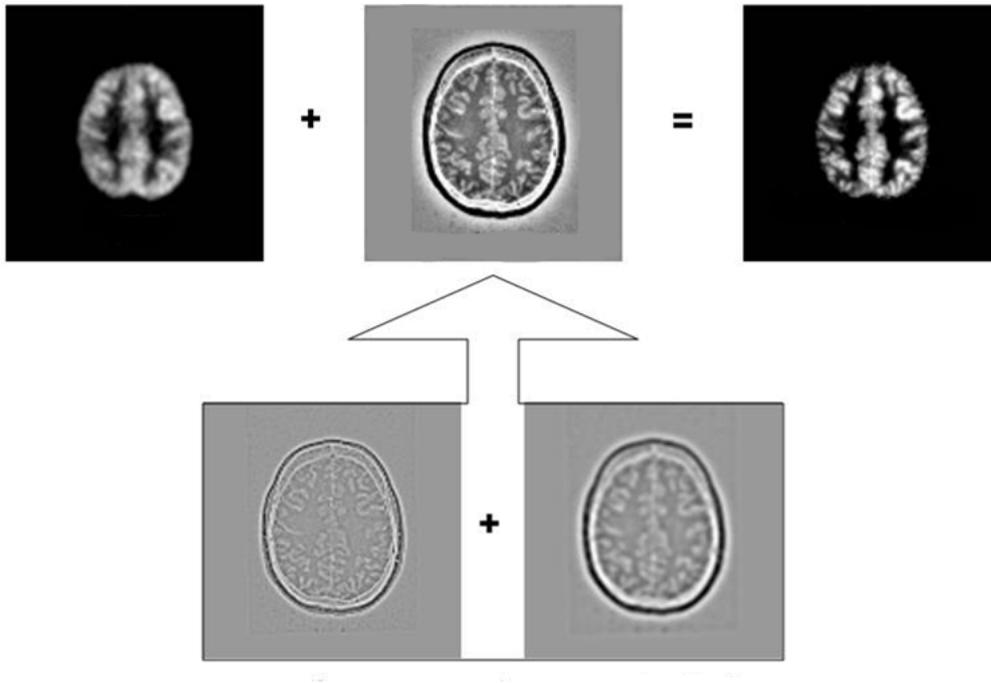


FIG. 6.5 – Image correctrice générée et résultat de correction sur la même image de cerveau.

de pouvoir générer des images corrigées qui restent des images fonctionnelles dont l'aspect quantitatif est préservé dans les régions homogènes, mais corrigé dans les zones de contours ou les effets de volume partiel sont importants. De plus, pour de toutes petites structures, dans lesquelles l'information quantitative est faussée même au centre, puisque les contours sont proches eu égard à la résolution spatiale du scanner, cette information est également correctement rehaussée. La limite de la méthode se situe dans le modèle dérivé des niveaux de résolution commun, qui est un modèle global. De plus, la correction est appliquée partout dans l'image, même si la corrélation entre les structures est faible ou inexistante, ce qui peut mener à des artefacts gênants dans l'image finale corrigée. Par exemple, les os présents dans l'image anatomique peuvent être insérés dans l'image TEP. Enfin, la méthode suppose que les deux images anatomiques et fonctionnelles soient parfaitement recalées voxel à voxel, ce qui est rarement le cas en réalité, et que la différence de résolution entre les deux modalités soit connue et de valeur entière, ce qui est également une approximation dans la plupart des cas réels. Plusieurs travaux ont été menés pour améliorer cette méthode et contourner ces

limites : une première approche consiste à utiliser une déconvolution filtrée dans le domaine des ondelettes permettant d'éviter la diminution du rapport signal sur bruit, phénomène généralement associé à la déconvolution itérative, pour se passer de l'image anatomique et de la nécessité du recalage parfait [17]. Malheureusement cette approche néglige le potentiel de l'imagerie TEP/TDM. Une autre approche consiste à modifier la méthode originale utilisant la modalité anatomique, soit en développant une MMA mettant en jeu une approche locale à base de fenêtre glissante [97] pour corriger le problème du modèle global, soit une correction conditionnelle fondée sur l'estimation d'arbres de Markov plaqués sur les décompositions en ondelettes des deux modalités qui sont intrinsèquement multi résolutions [98].

Combinaison de FLAB avec MMA

Pour s'affranchir de la nécessité de disposer d'une image anatomique recalée, ou pour pouvoir corriger les effets de volume partiel dans un cas où l'image anatomique n'offre que peu ou pas de corrélation de structure avec l'image fonctionnelle (voir l'exemple de la figure 6.2), une approche a été développée en collaboration avec N. Bousson : il s'agit d'utiliser la carte de segmentation obtenue par FLAB en lieu et place de l'image anatomique. Ce travail a fait l'objet d'une communication [18]. La figure 6.6 illustre le type de résultats obtenus pour la correction de volume partiel grâce à cette approche.

Utilisation dans le cadre du diagnostic et du suivi thérapeutique

Pour effectuer un diagnostic ou évaluer la réponse thérapeutique en utilisant plusieurs images consécutives dans le temps au cours du traitement, l'utilisation de l'approche combinant détermination de volume et correction de volume partiel permettrait d'obtenir automatiquement les informations de volume et de concentration d'activité pour un objet d'intérêt préalablement isolé manuellement dans les images.

Le cas le plus simple est celui d'un seul scan de patient, où le clinicien doit tout d'abord repérer l'anomalie et l'isoler dans une boîte d'intérêt grâce à la station de traitement des données. Ensuite, l'application de l'approche proposée précédemment permet de générer une carte de segmentation visuelle en deux ou trois classes, et les informations de volume et de

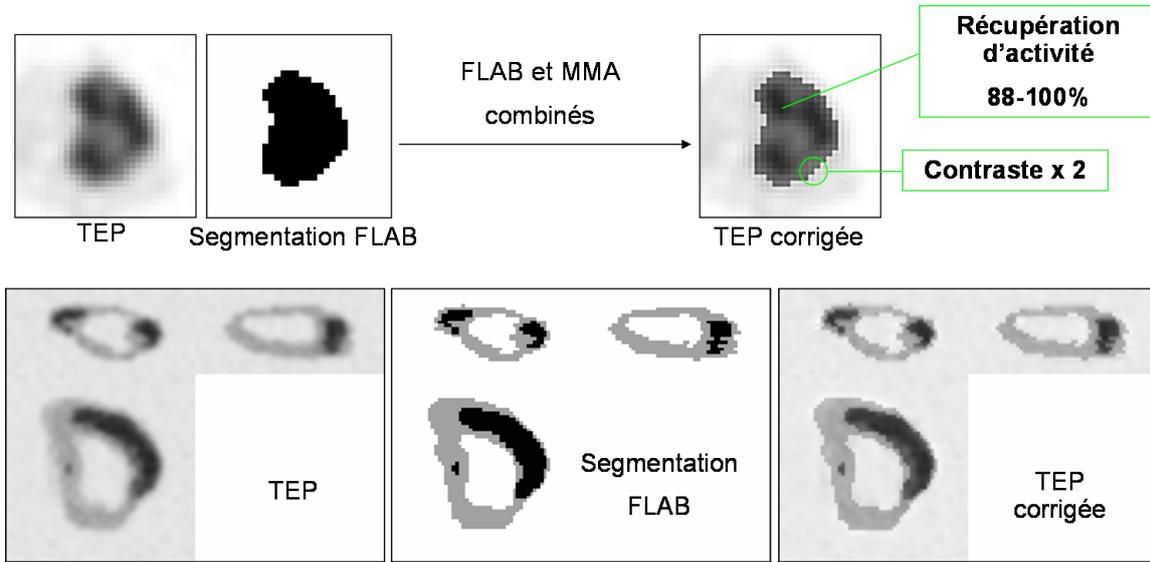


FIG. 6.6 – Résultats de correction de volume partiel en combinant la carte de segmentation de FLAB avec l’approche multirésolution développée par N. Bousson, dans le cas de deux classes (en haut) et trois classes (en bas).

quantification associées à ces segmentations. L’information quantitative ainsi obtenue est plus robuste et plus précise qu’en définissant une région d’intérêt à la main ou en utilisant des seuillages déterministes. De plus, dans le cas d’une tumeur hétérogène, la quantification serait disponible pour les différentes régions de la tumeur automatiquement, permettant de mieux poser le diagnostic. Enfin, l’information de quantification est donnée en complément de l’information du volume, lui aussi défini avec précision.

Dans ce cas il s’agit de mesurer l’évolution d’un objet d’intérêt sur plusieurs scans consécutifs. Etant donné la précision et la robustesse de l’approche proposée, les mesures indiquées précédemment (quantification et détermination du volume simultanés) peuvent être réalisées automatiquement sur chacun des scans indépendamment (après avoir pris soin de normaliser les images), puis de tracer des courbes d’évolution du volume et de l’activité de l’objet d’intérêt suivi, ceci permettant de mettre en évidence les changements ou la stabilité de l’un ou l’autre aspect de la tumeur. Ceci peut être envisagé pour un nombre de tumeurs réduit, et pour certains types de tumeurs seulement. Dans le cas de très nombreuses tumeurs dans un

même patient, il peut être nécessaire d'envisager d'autres approches car l'approche proposée nécessite d'isoler et de traiter chaque tumeur séparément.

6.2.3 Modèles de Markov couples et triplets

Comme cela est souvent observé dans ce domaine d'application, les approches issues d'autres domaines (ici, l'imagerie satellitaire et astronomique) trouvent des applications dans le domaine de l'imagerie médicale. Toutefois, les modèles considérés dans cette thèse sont relativement simples par rapport aux récents développements réalisés dans la famille des modèles utilisés en segmentation statistique d'images. L'utilisation de modèles de Markov couples [128] ou triplets [126] sont susceptibles de donner des résultats plus performants dans certaines situations, car il s'agit de généralisation des modèles standards, dans lesquelles certaines hypothèses peuvent être levées. Il reste à montrer que pour l'application en imagerie d'émission, ces modèles puissent effectivement améliorer les résultats par rapport aux modélisations adoptées dans le cadre de cette thèse, et ceci pourrait être un sujet à fort potentiel.

6.2.4 Etendre le champ d'application de FLAB

L'algorithme FLAB a été conçu pour segmenter des petites images, en 3D, fortement bruitées et floues, car c'est ce que l'on constate généralement pour les tumeurs imagées en imagerie d'émission. L'approche a été essentiellement validée pour la segmentation de tumeurs pulmonaires, abdominale et ORL en imagerie TEP par 18F-FDG, mais devrait pouvoir trouver une utilité également dans d'autres applications de la TEP, en neurologie (par exemple pour segmenter les noyaux codés dans des images de cerveau acquises par 18F-DOPA) et en cardiologie (pour la segmentation du coeur), et pourquoi pas également en imagerie TEMP ou du petit animal. Il existe donc de nombreuses applications dans le domaine de l'imagerie d'émission où l'algorithme FLAB pourrait potentiellement se révéler efficace et utile.

De plus, des caractéristiques proches de bruit et de flou se retrouvent dans d'autres modalités d'imagerie, de l'imagerie optique à l'imagerie par ultrasons, en passant par les images générées par bioluminescence.

Enfin, étudier les différents domaines d'applications de FLAB au delà de l'imagerie, pour la classification de données quelconques, pourrait également se révéler intéressant.

Chapitre 7

Annexes

7.1 Glossaire des termes et abréviations

3D-CRT : 3D Conformal radiothérapie ou radiothérapie conformationnelle. **ANR** : Agence Nationale pour la Recherche.

CHU : Centre Hospitalier Universitaire.

EC : Erreur de Classification (ou CE pour Classification Error). Mesure utilisée dans ce travail pour évaluer les performances des approches de segmentation en présence d'une vérité terrain voxel à voxel.

EV : Erreur de Volume (ou VE pour Volume Error). Mesure utilisée dans ce travail pour évaluer les performances des approches de segmentation en absence de vérité terrain voxel à voxel.

EM : Expectation Maximization ou maximisation d'espérance. Désigne un algorithme d'estimation de paramètres.

EVP : Effets de Volume Partiel. Désigne les effets de flou et de contamination croisée dus à la résolution spatiale réduite.

FBP : Filtered BackProjection : rétroprojection filtrée. Désigne un algorithme de reconstruction tomographique analytique des images TEP.

FCM : Fuzzy C-Means. Désigne l'algorithme classique de clustering flou.

FDG : FluoroDéoxyGlucose. Molécule semblable au glucose, utilisée en oncologie.

18F-FDG : FluoroDéoxyGlucose marqué au Fluor 18, l'isotope radioactif émetteur de positons.

FHMC : Fuzzy Hidden Markov Chains ou chaînes de Markov cachées floues. Désigne une des méthodes

développées pour segmenter les images TEP.

FLAB : Fuzzy Locally Adaptive Bayesian ou méthode bayésienne floue localement adaptative. Désigne une des méthodes développées pour segmenter les images TEP.

FWHM : Full Width at Half Maximum ou largeur à mi-hauteur. Désigne une mesure utilisée pour quantifier la réponse impulsionnelle ou résolution spatiale d'un imageur.

GATE : Geant 4 Application for Tomography Emission. Plateforme de simulation numérique pour l'imagerie d'émission.

GT : Ground-Truth. Voir VT (Vérité Terrain).

HMM : Hidden Markov Models. Voir MMC.

ICE : Iterative Conditional Estimation ou estimation conditionnelle itérative. Désigne un algorithme d'estimation de paramètres.

IEC : International Electrotechnical Commission.

IGRT : Image-guided RadioTherapy : radiothérapie guidée par l'image.

IMRT : Intensity Modulated RadioTherapy : radiothérapie à modulation d'intensité.

INSERM : Institut National de la Santé et de la Recherche Médicale. **IRM** : Imagerie par Résonance Magnétique.

KeV : Kilo-electronVolt : mesure d'énergie.

LaTIM : Laboratoire de Traitement de l'Information Médicale. Unité INSERM 650 dans laquelle ce travail a été mené.

LOR : Line Of Response ou ligne de réponse. Désigne la ligne virtuelle reliant deux détecteurs en coïncidence dans le scanner TEP.

ML-EM : Maximum Likelihood-Expectation Maximization. Algorithme itératif de reconstruction d'images en TEP.

MMA : Mutual Multiresolution Analysis. Désigne la méthode de correction des effets de volume partiel mettant en jeu l'analyse multirésolution des modalités anatomiques et fonctionnelles.

MMC : Modèles de Markov Cachés (ou HMM pour Hidden Markov Models). Désignation générale pour les modèles de Markov tels que les chaînes, les champs, les arbres. **NEMA** : National Electrical Manufacturers Association.

OPL-EM : One Pass Listmode-Expectation Maximization. Algorithme itératif de reconstruction

d'images en TEP.

ORL : Oto-Rhino-Laryngologie.

OSEM : Expectation Maximization. Algorithme itératif de reconstruction d'images en TEP.

PSF : Point Spread Function ou réponse impulsionnelle. Désigne la résolution spatiale d'un imageur.

RAMLA : Row-Action Maximum Likelihood Algorithm. Algorithme itératif de reconstruction d'images en TEP.

RTL : Relative Threshold Level : une approche de seuillage adaptative.

SEM : Stochastic Expectation Maximization ou maximisation d'espérance stochastique. Désigne un algorithme d'estimation de paramètres.

SUV : Standard Uptake Value. Index le plus communément utilisé pour caractériser la fixation du FDG. Il s'agit de la valeur des voxels de l'image normalisée par rapport au poids du patient, à la dose injectée et à la durée d'acquisition.

T42 ou **T50** : désigne les approches de segmentation d'image basées sur des seuillages fixes (à 42 ou 50 % de valeur maximum de l'objet d'intérêt).

Tbckg et **TSBR** : désigne les approches de segmentation d'image basées sur des seuillages adaptatifs prenant en compte la valeur du fond.

TDM : Tomodensitométrie (scanner à rayons X).

TEP : Tomographie par Emission de Positons.

TEMP : Tomographie par Emission MonoPhotonique.

TEP/TDM ou **TEMP/TDM** : Tomographie multimodalités associant un scanner fonctionnel TEP ou TEMP avec un scanner à rayons X dans le même statif.

TOF (ou **TF**) : Time Of Flight ou temps de vol. Désigne les systèmes TEP récents permettant d'affiner la position de l'annihilation le long de la ligne de réponse, par estimation du temps de vol des photons.

Voxel : élément de volume d'une image 3D. Il s'agit de l'équivalent du pixel pour les images tri-dimensionnelles.

VMAT : Volumetric intensity Modulated Arc Therapy.

VT : Vérité Terrain (ou GT pour Ground-Truth). Réalité simulée ou référence par rapport à laquelle se calculent les erreurs commises par les approches de segmentation.

7.2 Communications et publications dans le cadre de cette thèse

7.2.1 Publications sur le travail de thèse

Publications en revues à comité de lecture

M. Hatt, C. Cheze le Rest, P. Descourt, C. Roux, D. Visvikis, Accurate and robust automatic delineation of heterogeneous functional volumes in positron emission tomography for oncology applications, *Journal of Nuclear Medicine*, soumis, 2008.

M. Hatt, A. Turzo, C. Roux, D. Visvikis, A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET, *IEEE Transactions on Medical Imaging*, in press, 2008.

M. Hatt, F. Lamare, N. Boussion, A. Turzo, C. Collet, F. Salzenstein, C. Roux, K. Carson, P. Jarritt, C. Cheze-Le Rest, D. Visvikis, Fuzzy hidden Markov chains segmentation for volume determination and quantitation in PET, *Physics in Medicine and Biology*, Vol. 52, pp. 3467-3491, 2007.

Communications en conférences internationales à comité de lecture

M. Hatt, P. Bailly, A. Turzo, C. Roux, D. Visvikis, PET functional volume segmentation : a robustness study, *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC)*, Dresden, Germany, October 2008.

M. Hatt, A. Dekker, D. De Ruyscher, M. Oellers, P. Lambin, C. Roux, D. Visvikis, Accurate functional volume definition in PET for radiotherapy treatment planning, *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC)*, Dresden, Germany, October 2008.

M. Hatt, N. Boussion, C. Roux, O. Pradier, C. Cheze Le Rest, D. Visvikis, An automatic segmentation algorithm for accurate functional volume determination in radiotherapy treatment planning, *Annual Congress of the European Association of Nuclear Medicine*, Munich, Germany, October 2008.

M. Hatt, N. Boussion, O. Pradier, C. Roux, D. Visvikis, Automatic Segmentation of Functional Images for Radiotherapy Treatment Planning, *50th American Society for Therapeutic Radiology and Oncology annual meeting*, Boston, USA, September 2008.

M. Hatt, C. Cheze Le Rest, C. Roux, D. Visvikis, Automatic volume delineation of heterogeneous

tumours in PET : comparison of various methodologies, *Journal of Nuclear Medicine*, Vol. 49 (S1) pp. 381, 2008.

M. Hatt, C. Roux, D. Visvikis, A Segmentation Algorithm for Heterogeneous Tumor Delineation in PET, *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC) conference records*, Vol. 5 pp. 3939-3945, 2007.

M. Hatt, C. Roux, D. Visvikis, 3D Fuzzy Adaptive Unsupervised Bayesian Segmentation for Volume Determination in PET, *4th IEEE International Symposium on Biomedical Imaging : From Nano to Macro conference records*, pp. 328-331, 2007.

M. Hatt, C. Cheze-Le Rest, A. Turzo, F. Lamare, K. Carson, P. Jarritt, D. Visvikis, An automatic segmentation algorithm for functional volume and activity concentration determination in PET, *Annual Congress of the European Association of Nuclear Medicine*, Athens, Greece, September 2006.

M. Hatt, N. Bousson, F. Lamare, C. Collet, F. Salzenstein, C. Roux, Y. Bizais, C. Cheze-Le Rest, D. Visvikis, Fuzzy versus Hard Hidden Markov Chains Segmentation for Volume Determination and Quantitation in Noisy PET Images, *3rd IEEE International Symposium on Biomedical Imaging : From Nano to Macro conference records*, pp. 1376-1379, 2006.

M. Hatt, N. Bousson, K. Karson, P. Jarritt, F. Lamare, Y. Bizais, C. Cheze-Le Rest, D. Visvikis, Comparison of different methodologies for lesion volume determination in PET, *British Nuclear Medicine Society Spring Meeting conference records*, Vol. 27(3) pp. 283-284, 2006.

Communications en conférences nationales à comité de lecture

M. Hatt, D. Visvikis, N. Bousson, O. Pradier, Segmentation automatique d'images fonctionnelles TEP pour l'aide à la planification de traitement en radiothérapie, *47e journées scientifiques de la société française de physique médicale*, Marseille France, juin 2008.

M. Hatt, C. Roux, D. Visvikis, Evaluation de méthodes de segmentation bayésiennes pour l'imagerie TEP en oncologie, *XXIe colloque GRETSI - traitement du signal et des images*, Troyes France, septembre 2007.

M. Hatt, N. Bousson, C. Roux, D. Visvikis, Evaluation de méthodes de segmentation statistiques pour l'imagerie TEP en oncologie, *Journée de Recherche en Imagerie Médicale (JRIM) 2007*, Dijon France,

septembre 2007.

Présence à des workshops

M Hatt, A Turzo, F Lamare, P Jarritt, K Carson, C Roux, D Visvikis, Segmentation of functional volumes in PET/CT, *Quantitative imaging and dosimetry in Nuclear Medicine*, Berder island, 3-6 october 2006.

7.2.2 Publications annexes en co-auteur

Publications en revues à comité de lecture

N. Boussion, C. Cheze Le Rest, M. Hatt, D. Visvikis, Incorporation of wavelet based denoising in iterative deconvolution for partial volume correction in whole body PET imaging, *European Journal of Nuclear Medicine and Molecular Imaging*, soumis, 2008.

N. Boussion, M. Hatt, F. Lamare, C. Cheze Le Rest, D. Visvikis, Contrast enhancement in emission tomography by way of synergistic PET/CT image combination, *Computer Methods and Programs in Biomedicine*, Vol. 90(3), pp. 191-201, 2008.

F. Salzenstein, C. Collet, S. Lecam, M. Hatt, Non-stationary fuzzy Markov chain, *Pattern Recognition Letters*, Vol. 28(16), pp. 2201-2208, 2007.

N. Boussion, M. Hatt, F. Lamare, Y. Bizais, A. Turzo, C. Cheze-Le Rest, D. Visvikis, A multiresolution image based approach for correction of partial volume effects in emission tomography, *Physics in Medicine and Biology*, Vol. 51, pp. 1857-1876, 2006.

Communications en conférences internationales à comité de lecture

A. Le Pogam, M. Hatt, N. Boussion, F. Turkheimer, C. Prunier-Aesch, D. Guilloteau, J.-L. Baulieu, D. Visvikis, A wavelet-based hidden Markov model and multi-resolution approach for conditional partial volume correction in emission tomography, *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC)*, Dresden, Germany, October 2008.

- A. Le Pogam, M. Hatt, N. Bousson, D. Guilloteau, J-L. Beaulieu, C. Prunier, F. Turkheimer, D. Visvikis, A 3D multi resolution local analysis approach for correction of partial volume effects in emission tomography, *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC)*, Dresden, Germany, October 2008.
- A. Le Pogam, N. Bousson, M. Hatt, C. Prunier-Aesch, D. Guilloteau, J.-L. Baulieu, D. Visvikis, 3-D discrete ridgelet transform for emission tomography denoising, *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC)*, Dresden, Germany, October 2008.
- N. Bousson, C. Cheze-Le Rest, M. Hatt, D. Visvikis, Evaluation of resolution and quantitation preserving wavelet-based denoising in wholebody PET, *Journal of Nuclear Medicine*, Vol. 49 (Supplement 1), pp. 395, 2008.
- N. Bousson, M. Hatt, D. Visvikis, Partial volume correction in PET based on functional volumes, *Journal of Nuclear Medicine*, Vol. 49 (Supplement 1), pp. 388, 2008.
- A. Le Pogam, M. Hatt, N. Bousson, D. Guilloteau, J-L. Beaulieu, C. Prunier, F. Turkheimer, D. Visvikis, Conditional voxel-wise partial volume correction for emission tomography : combining a wavelet-based hidden Markov model with a mutual multi-resolution analysis approach, *Journal of Nuclear Medicine*, Vol. 49 (Supplement 1), pp. 62, 2008.
- A. Le Pogam, M. Hatt, N. Bousson, D. Guilloteau, J-L. Beaulieu, C. Prunier, F. Turkheimer, D. Visvikis, Conditional partial volume correction for emission tomography : a wavelet-based hidden Markov model and multi-resolution approach, *5th IEEE International Symposium on Biomedical Imaging conference records*, pp. 1319-1322, 2008.
- N. Bousson, M. Hatt, A. Reilhac, D. Visvikis, Fully automated partial volume attenuation in PET : a wavelet approach without anatomical information, *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC) conference records*, Vol. 4, pp. 2812-2816, 2007.
- N. Bousson, M. Hatt, F. Lamare, Y. Bizais, A. Turzo, C. Cheze-Le Rest, D. Visvikis, Generating resolution-enhanced images for correction of partial volume effects in emission tomography : a multi-resolution approach, *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS-MIC) Conference Records*, Vol. 4, pp. 2423-2427, 2005.
- N. Bousson, M. Hatt, F. Lamare, A. Turzo, Y. Bizais, C. Cheze-Le Rest, D. Visvikis, A multiresolution approach in partial volume correction for emission tomography imaging, *European Journal of Nuclear Medicine and Molecular Imaging*, Vol. 32(S1), S51, 2005.
- N. Bousson, M. Hatt, F. Lamare, A. Turzo, Y. Bizais, C. Cheze-Le Rest, D. Visvikis, A multiresolution approach in partial volume correction for emission tomography imaging. *Annual Congress of the European Association of Nuclear Medicine*, Istanbul, Turkey, 2005.

Communications en conférences nationales à comité de lecture

N. Boussion, M. Hatt, D. Wallach, O. Pradier, D. Visvikis, Quantification en TEP oncologique : les dernières avancées en traitement d'images et leur impact potentiel sur la planification de traitement en radiothérapie, *47e journées scientifiques de la société française de physique médicale*, Marseille France, juin 2008.

N. Boussion, M. Hatt, F. Lamare, D. Visvikis, Correction des effets de volume partiel en TEP sans a priori anatomique : déconvolution et filtrage à base d'ondelettes, *Journée de Recherche en Imagerie Médicale (JRIM) 2007*, Dijon France, septembre 2007.

7.3 Algorithme FLAB : schéma d'ensemble

La figure 7.1 présente le fonctionnement de l'algorithme FLAB sous forme de machine à états, avec les principales étapes.

7.4 Résultats complémentaires

7.4.1 Optimisation et validation

Résultats sur acquisition réelle du fantôme IEC

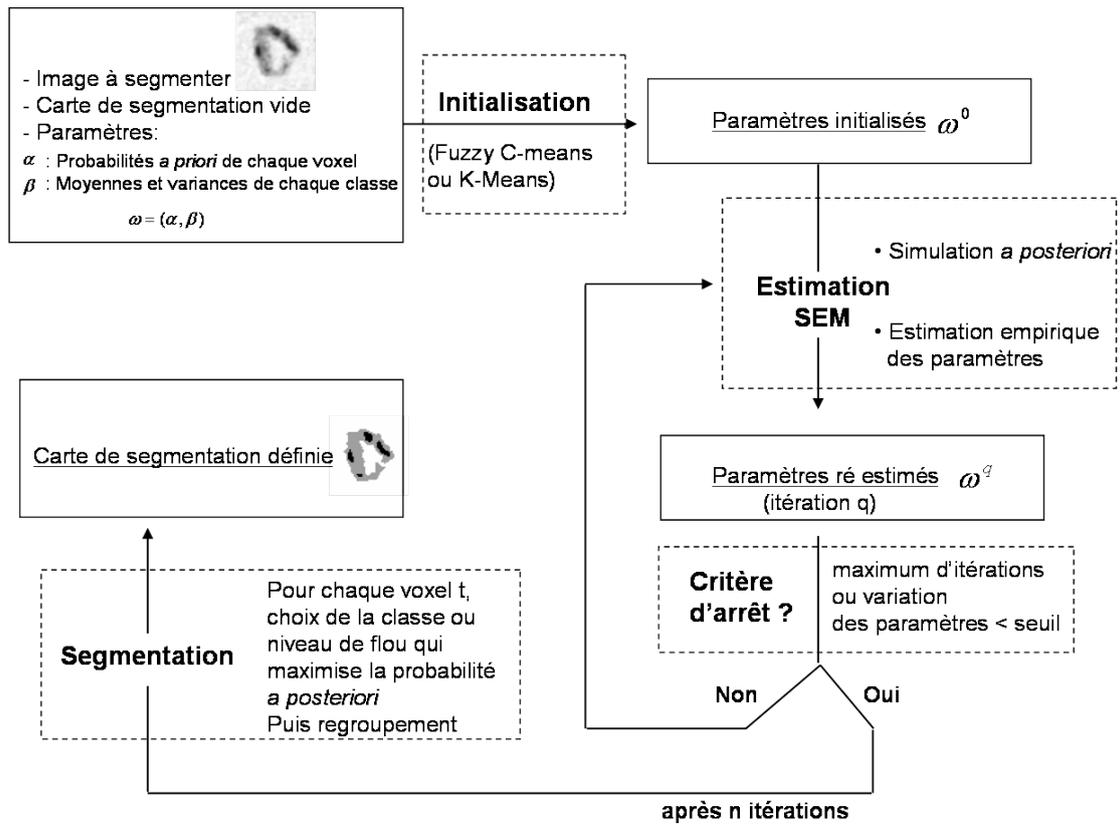


FIG. 7.1 – Schéma d'ensemble du fonctionnement de l'algorithme FLAB. Les cases à traits continus constituent les états des paramètres, et les cases à traits en pointillés indiquent des opérations de transitions entre états.

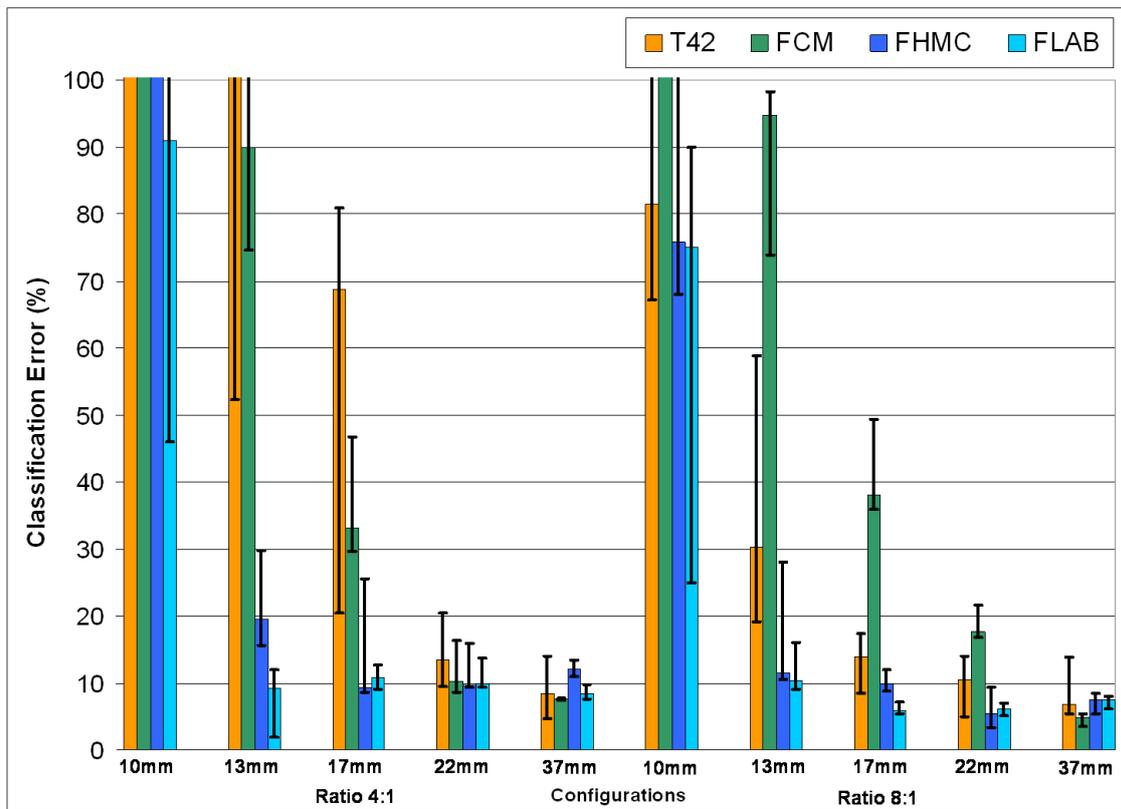


FIG. 7.2 – Résultats obtenus (erreur de classification) sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.

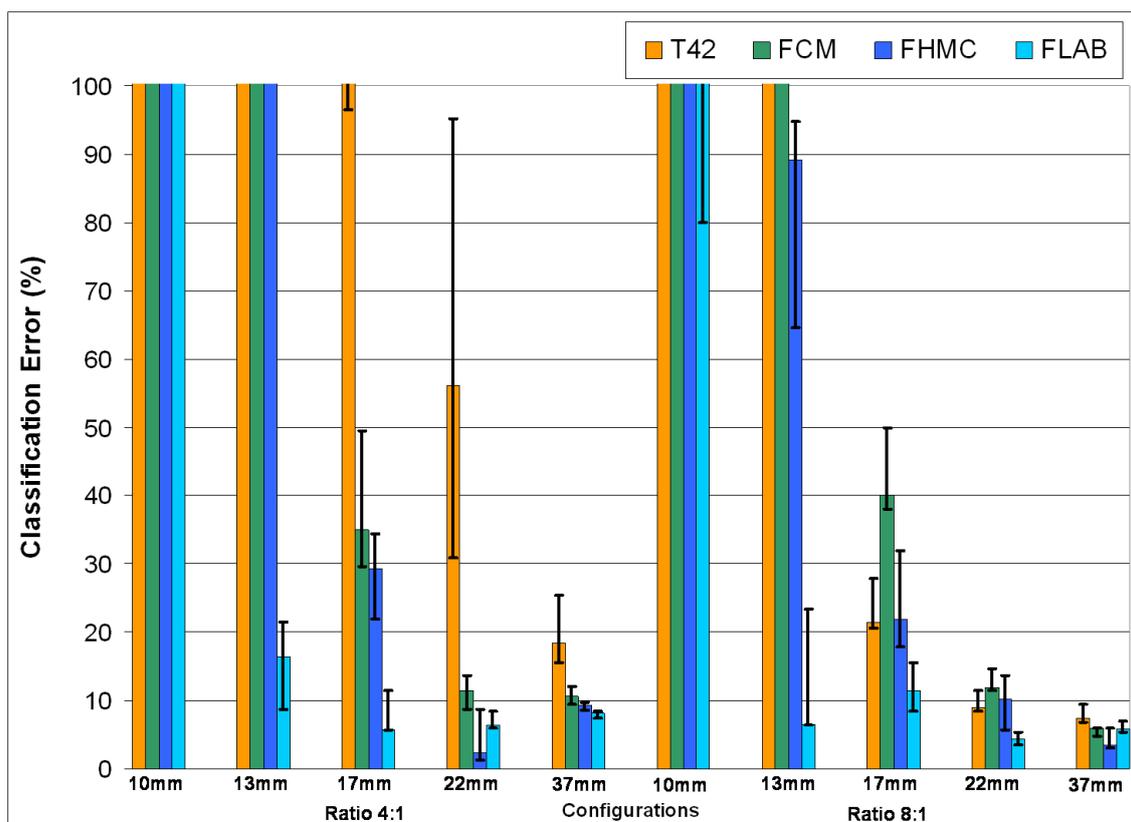


FIG. 7.3 – Résultats obtenus (erreur de classification) sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.

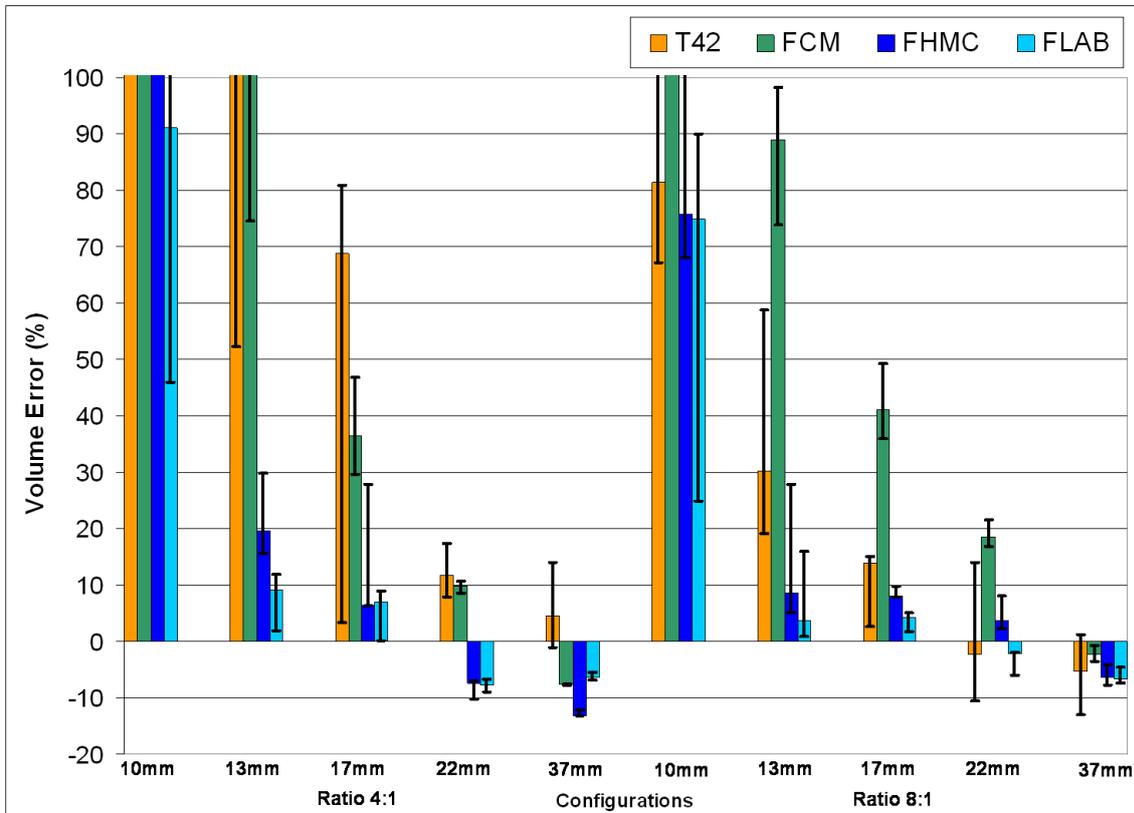


FIG. 7.4 – Résultats obtenus (erreur de volume) sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.

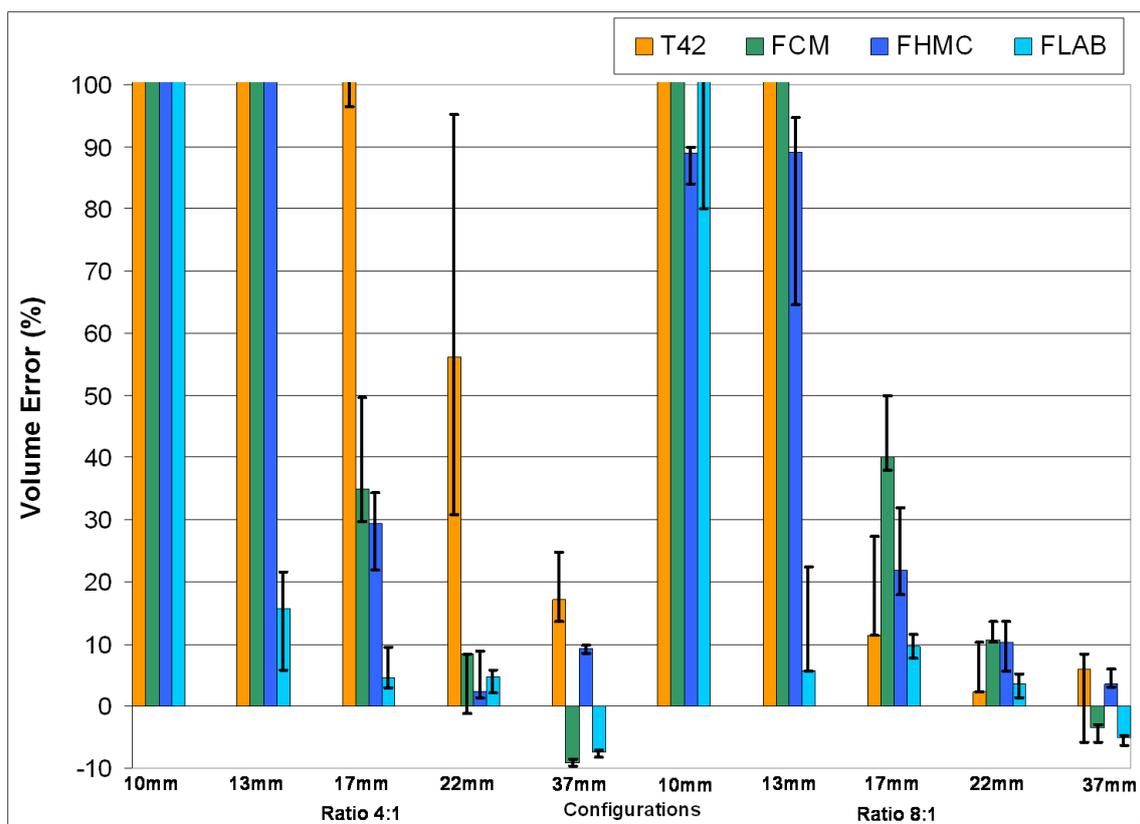


FIG. 7.5 – Résultats obtenus (erreur de volume) sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.

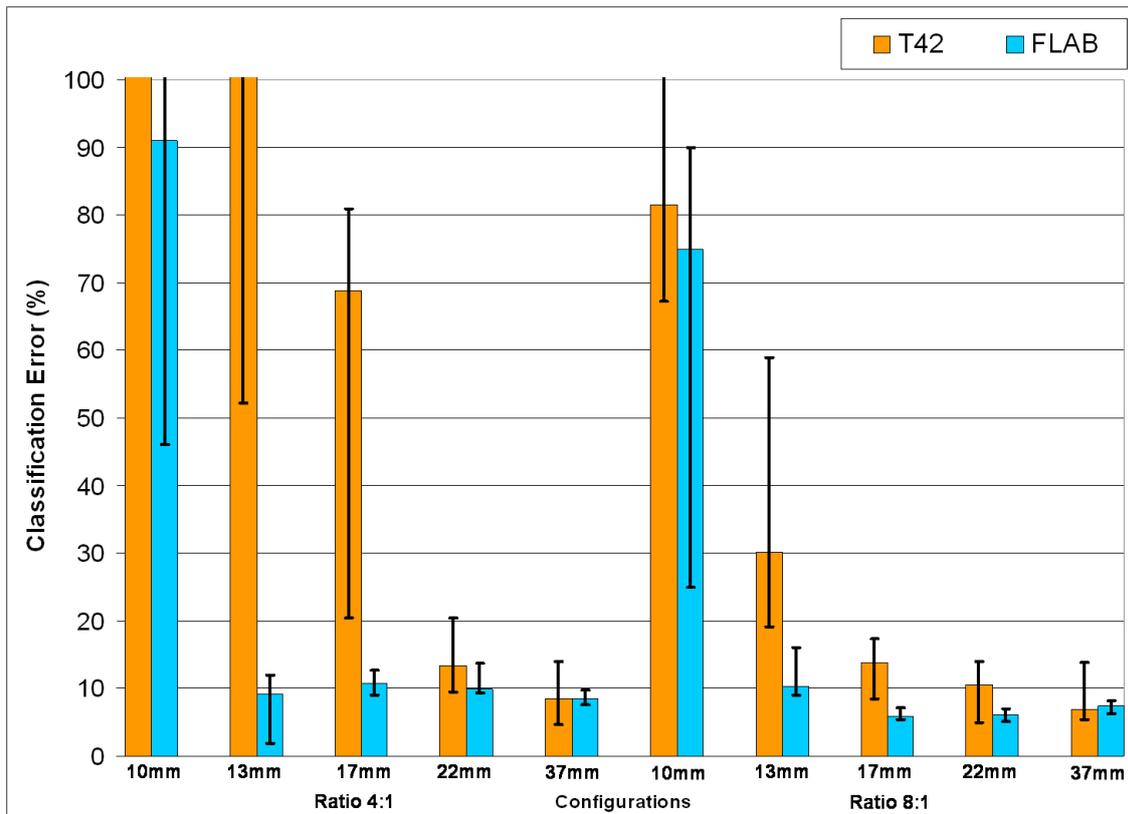


FIG. 7.6 – Comparaison des performances (erreur de classification) entre FLAB et T42 sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.

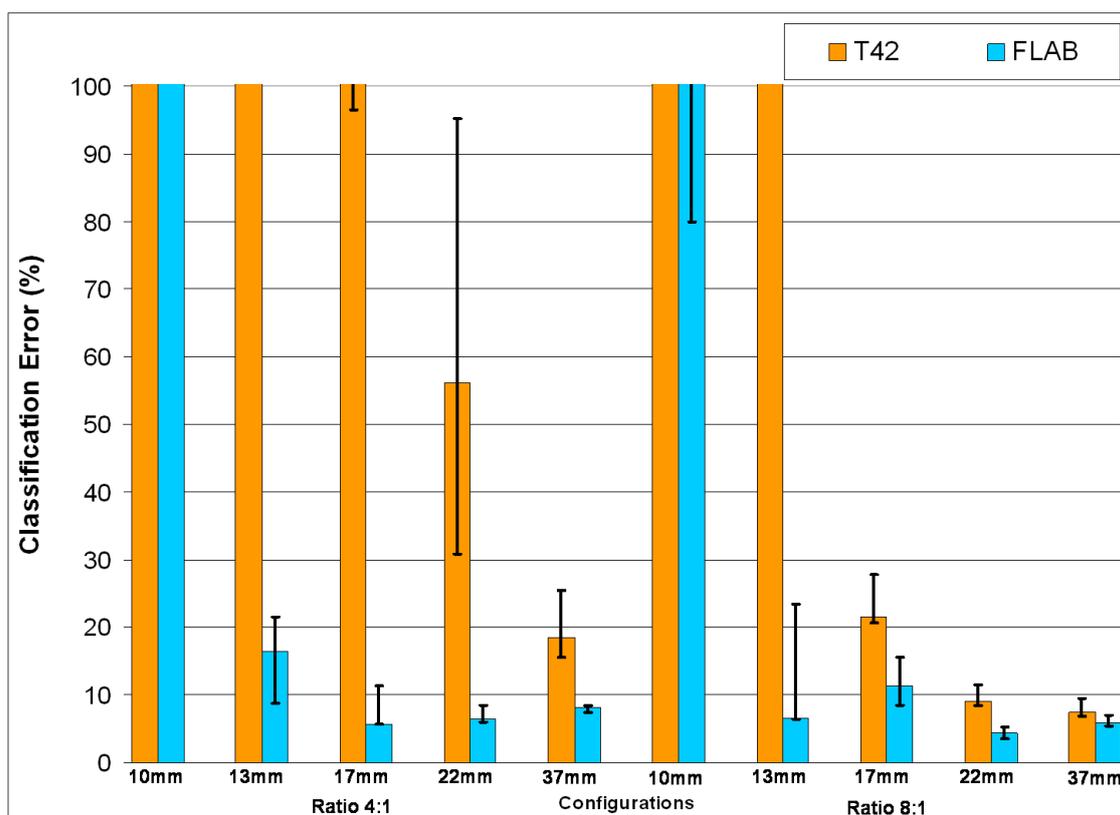


FIG. 7.7 – Comparaison des performances (erreur de classification) entre FLAB et T42 sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.

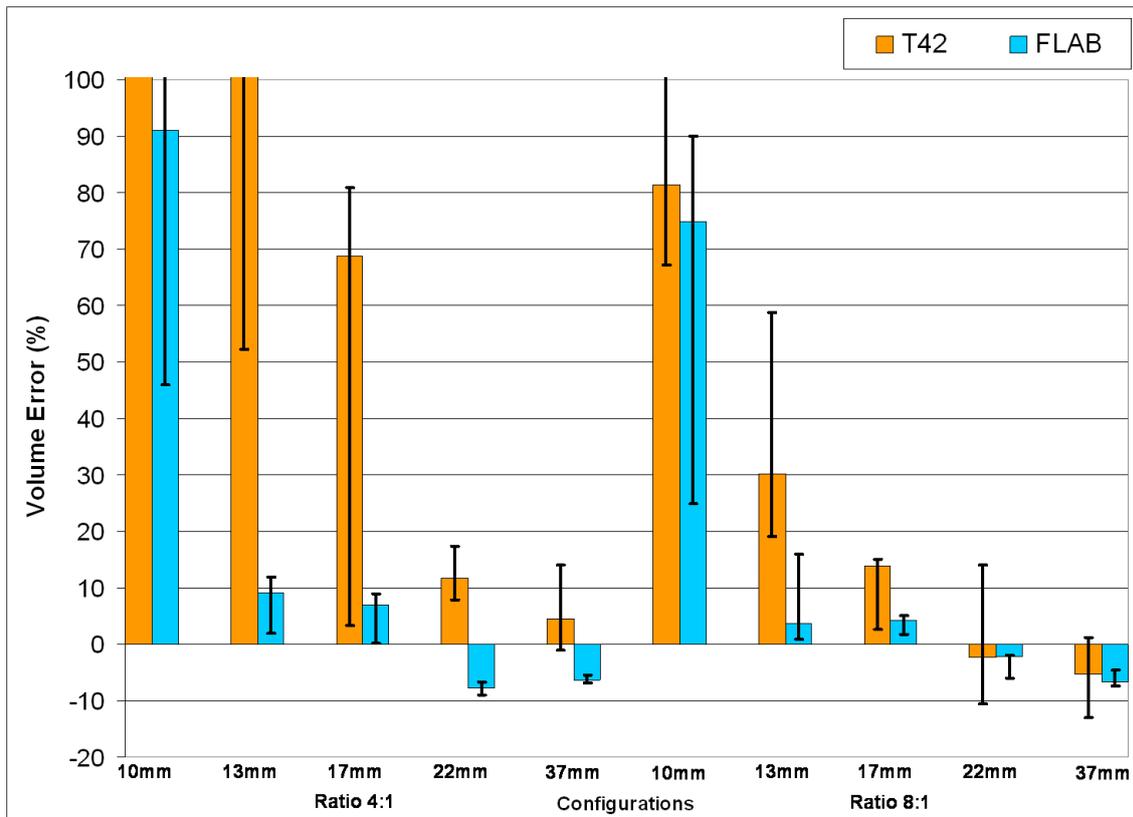


FIG. 7.8 – Comparaison des performances (erreur de volume) entre FLAB et T42 sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.

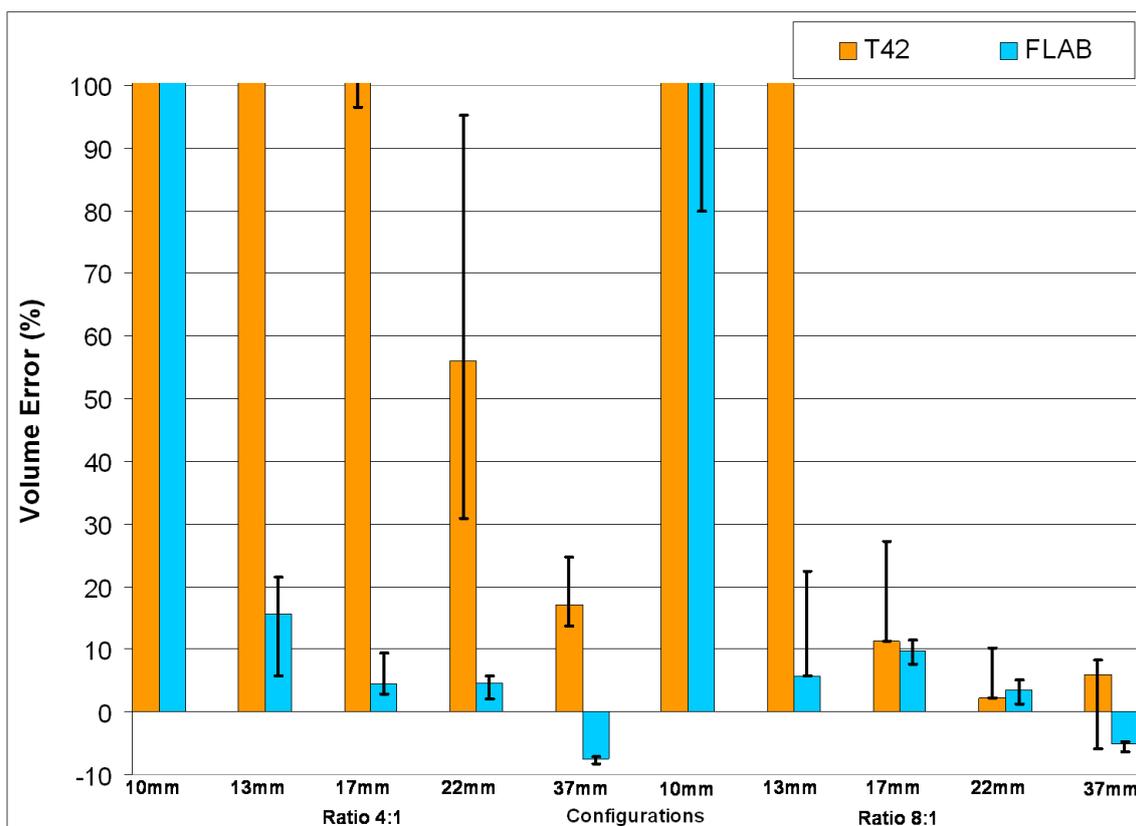


FIG. 7.9 – Comparaison des performances (erreur de volume) entre FLAB et T42 sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.

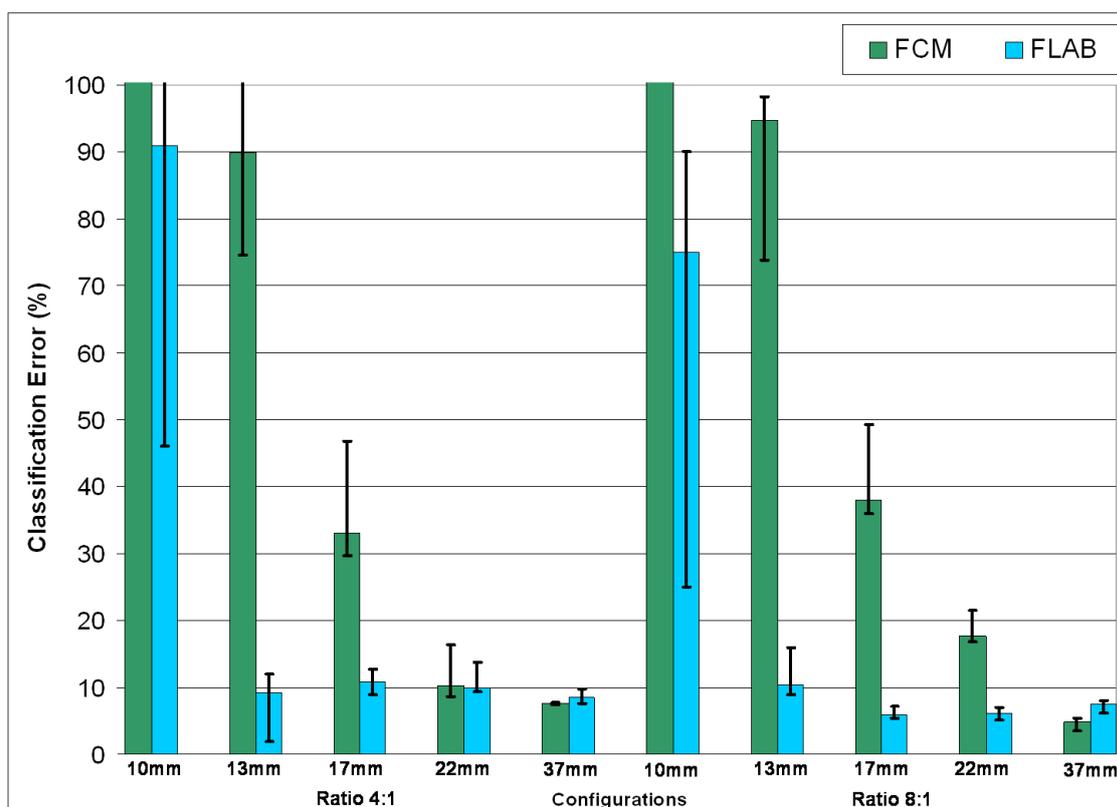


FIG. 7.10 – Comparaison des performances (erreur de classification) entre FLAB et FCM sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.

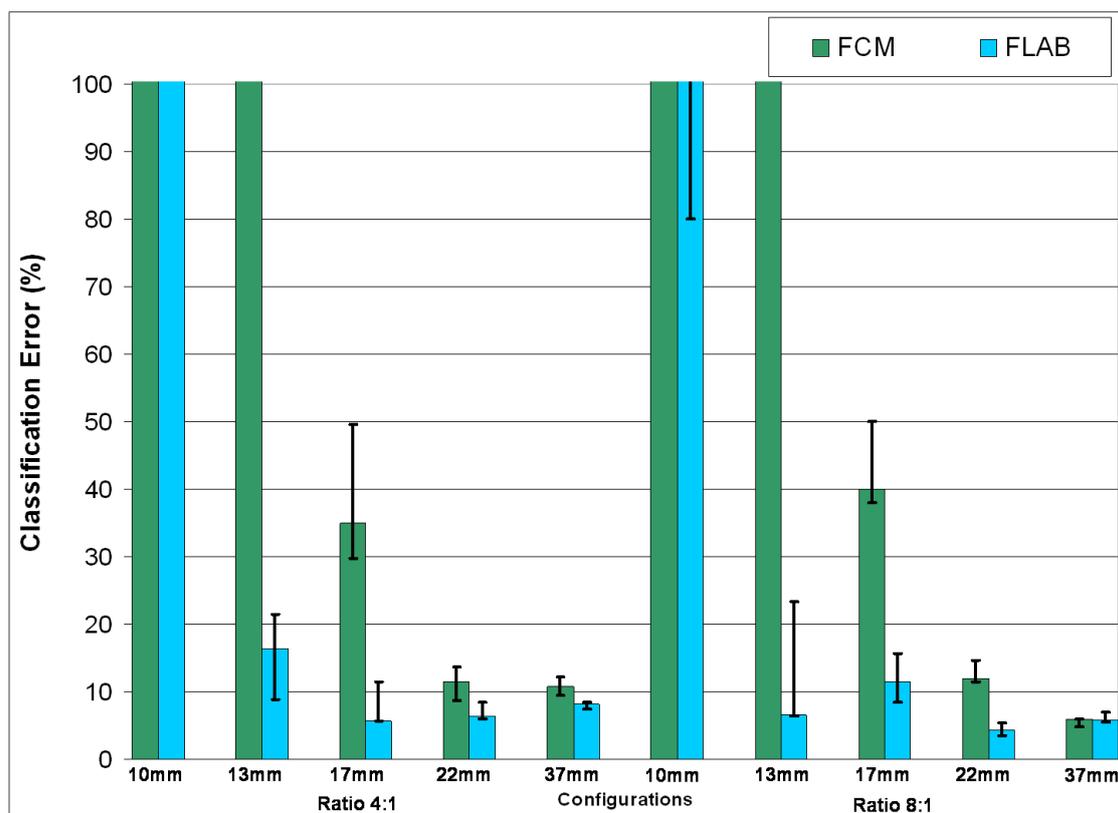


FIG. 7.11 – Comparaison des performances (erreur de classification) entre FLAB et FCM sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.

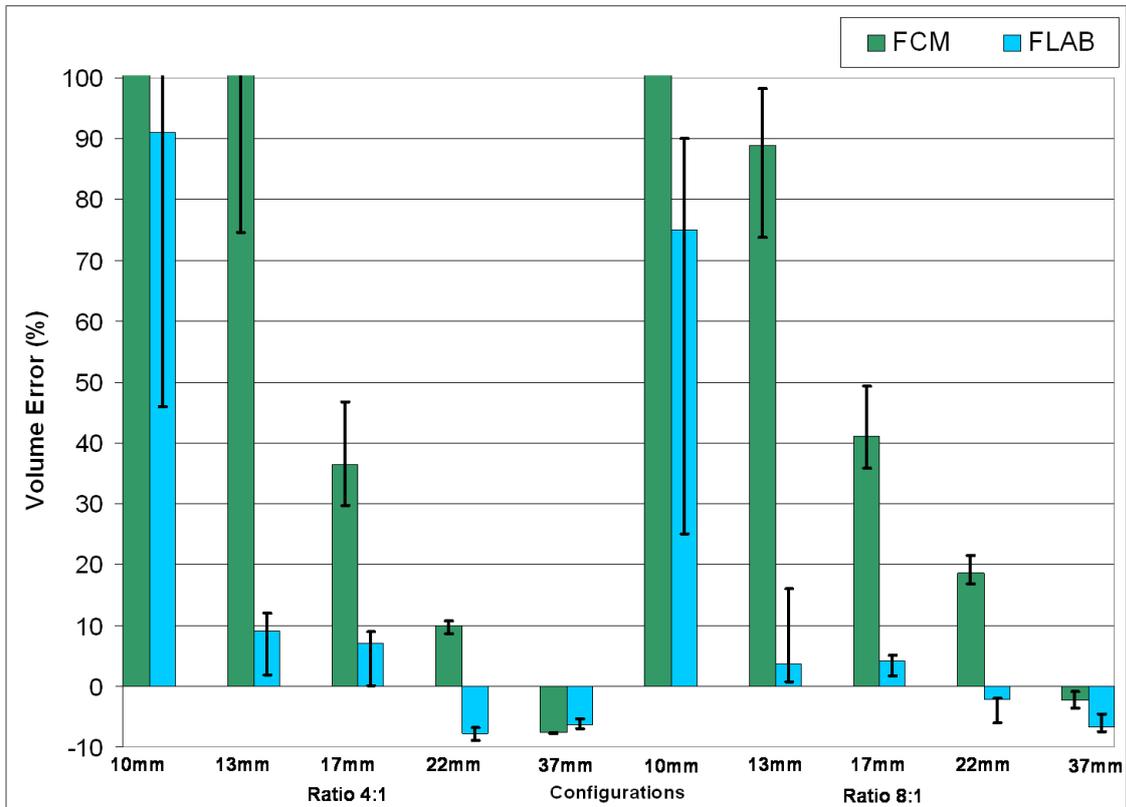


FIG. 7.12 – Comparaison des performances (erreur de volume) entre FLAB et FCM sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.

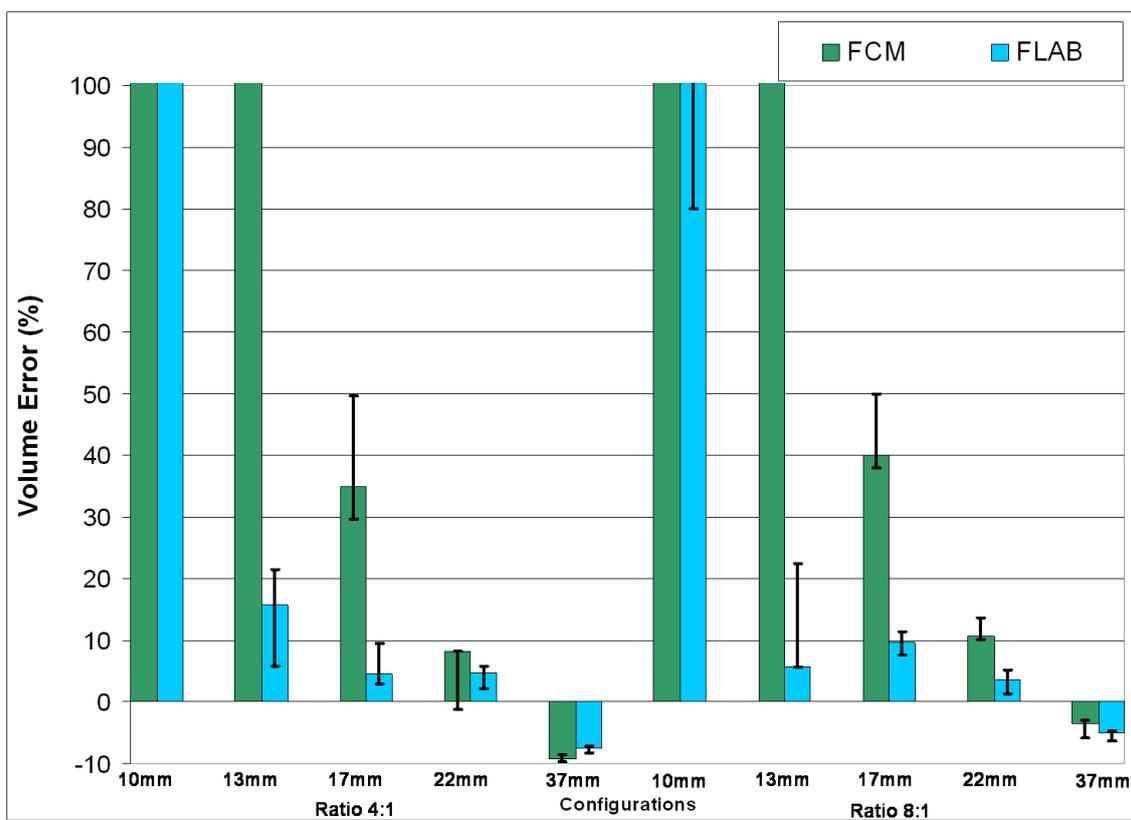


FIG. 7.13 – Comparaison des performances (erreur de volume) entre FLAB et FCM sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.

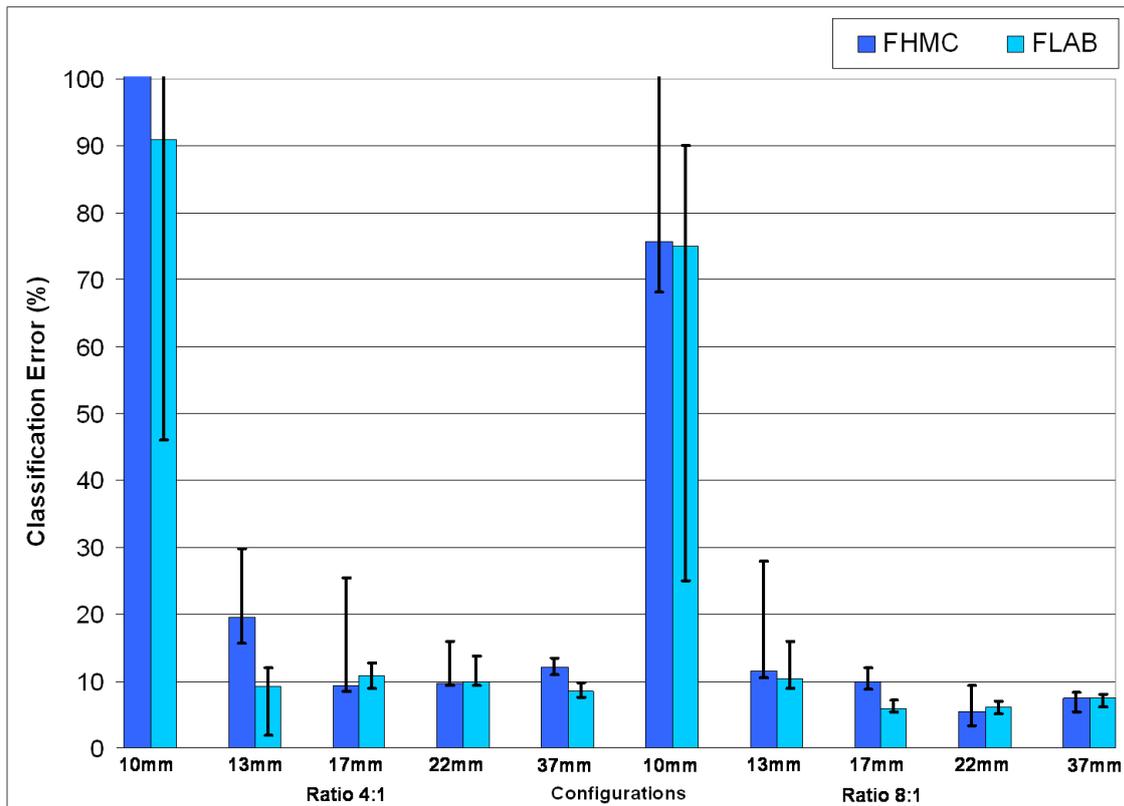


FIG. 7.14 – Comparaison des performances (erreur de classification) entre FLAB et FHMC sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.

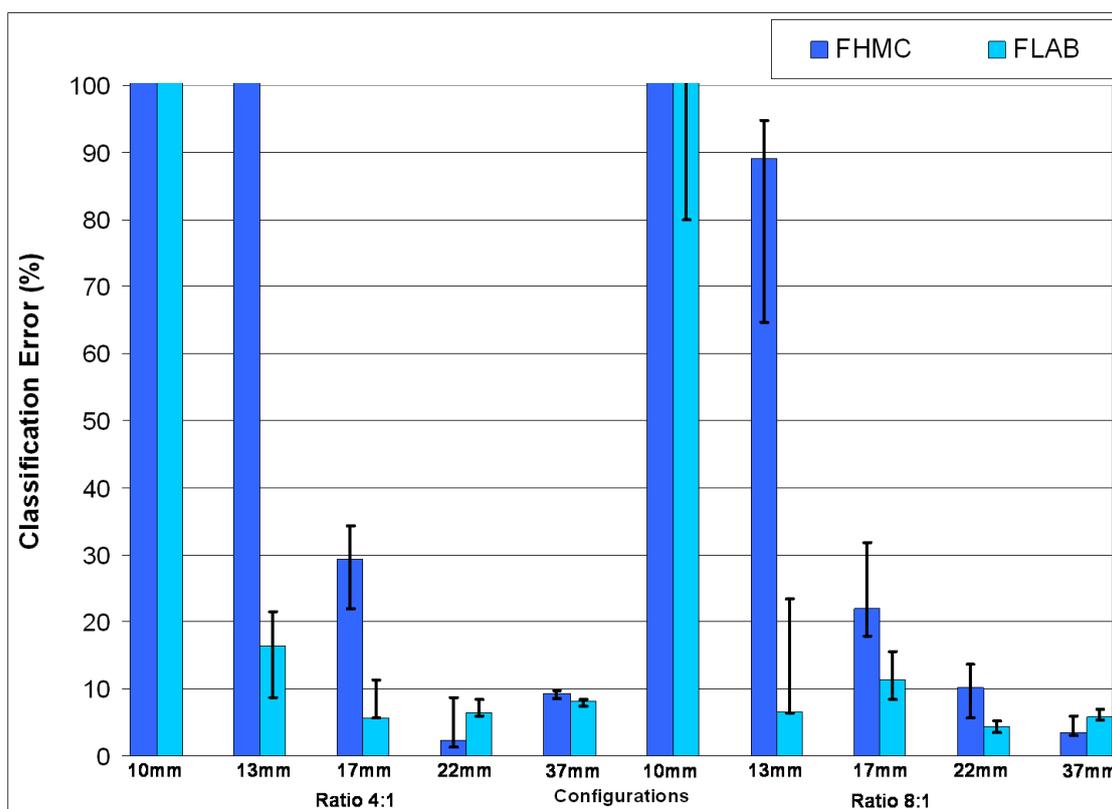


FIG. 7.15 – Comparaison des performances (erreur de classification) entre FLAB et FHMC sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.

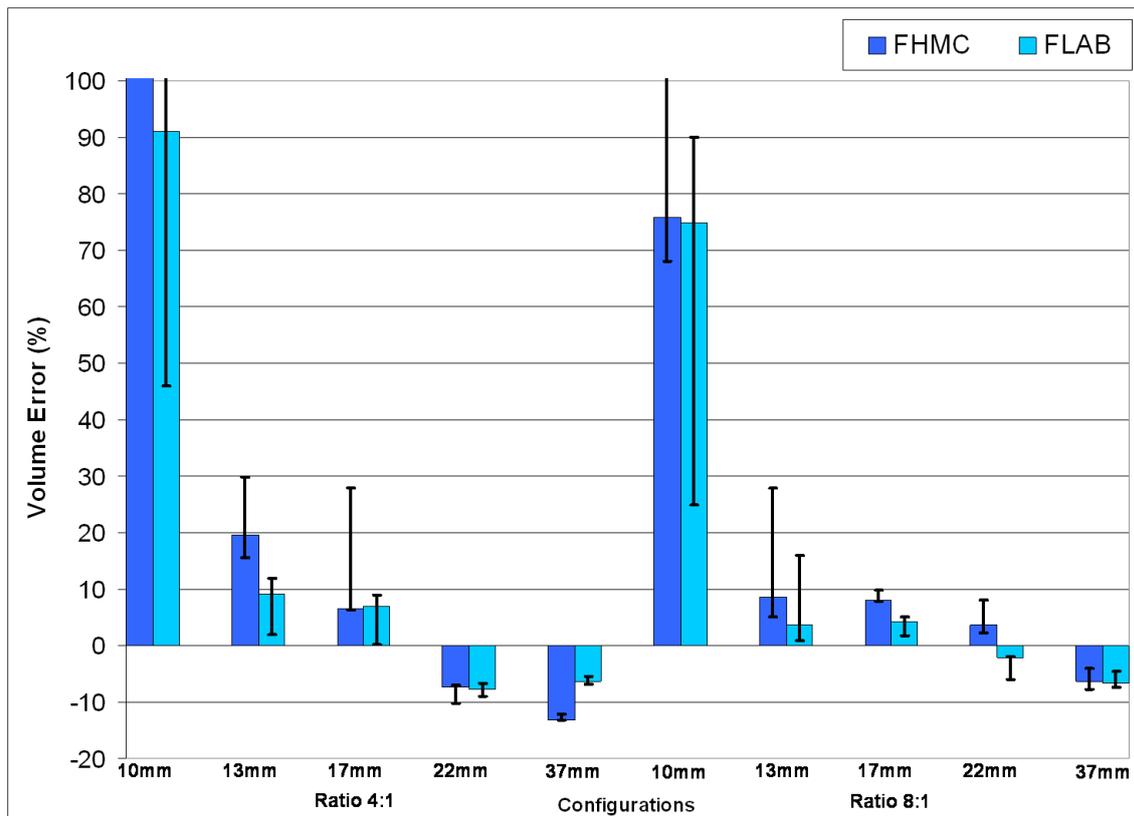


FIG. 7.16 – Comparaison des performances (erreur de volume) entre FLAB et FHMC sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 2x2x2 mm.

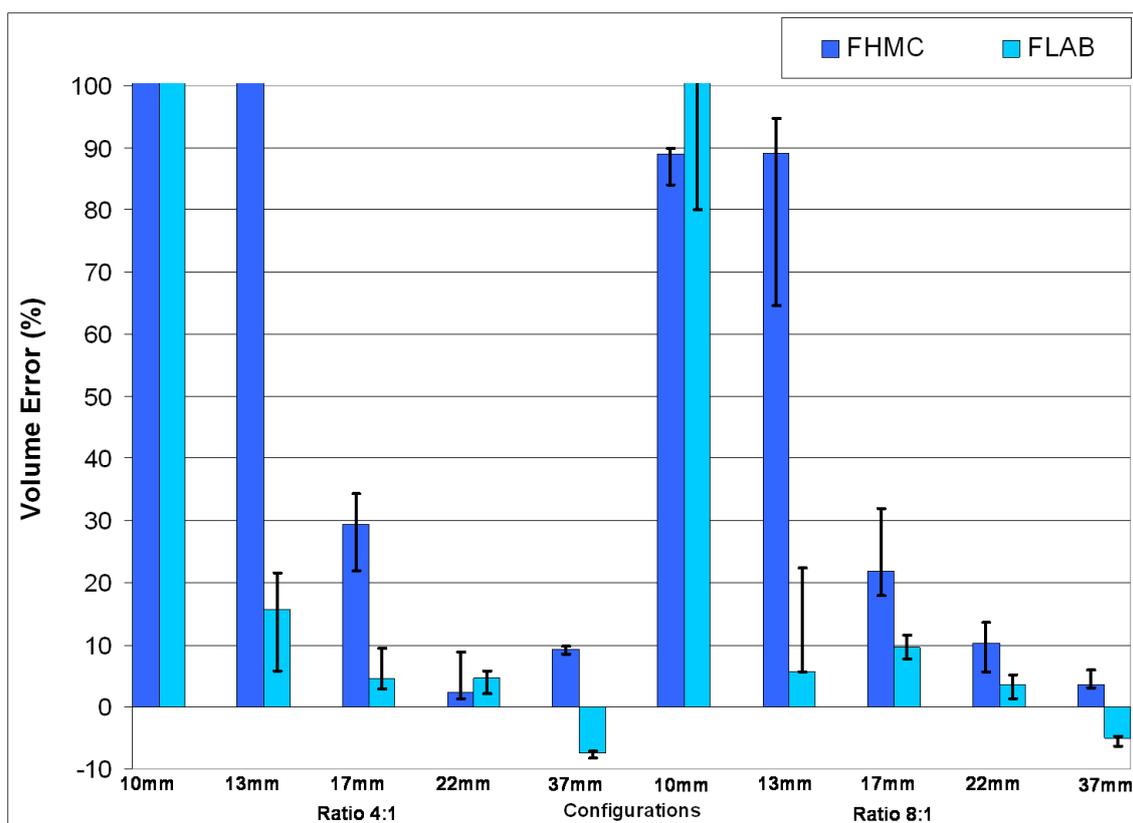


FIG. 7.17 – Comparaison des performances (erreur de volume) entre FLAB et FHMC sur les acquisitions réelles du fantôme IEC sur Philips Gemini et reconstruites par RAMLA 3D. Voxels de 4x4x4 mm.

Bibliographie

- [1] H. J. W. L. Aerts, G. Bosmans, A. W. van Baardwijk, A. L. A. J. Dekker, M. C. Oellers, P. Lambin, and D. de Ruyscher. Stability of ^{18}F -deoxyglucose uptake locations within tumor during radiotherapy for nscl: A prospective study. *International Journal of Radiology Oncology Biology Physics*, 71(5) :1402–1407, 2008.
- [2] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6) :716–723, 1974.
- [3] H. Ashamalla, S. Raffla, K. Parikh, B. Mokhtar, G. Goswami, S. Kambam, H. Abdel-Dayem, A. Guirguis, P. Ross, and A. Evola. The contribution of integrated PET/CT to the evolving definition of treatment volumes in radiation treatment planning in lung cancer. *International Journal of Radiology Oncology Biology Physics*, 63(4) :1016–1023, 2005.
- [4] R. Atcher, N. R. Dunnick, B. G. Haffty, and A. Padhani. Opportunities in imaging for radiation oncology. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [5] K. Avrachenkov and E. Sanchez. Fuzzy chains and decision-making. *Fuzzy Optimization and Decision Making*, 1(2) :143–159, 2002.
- [6] DL. Bailey and et al. *Bendriem B. and Townsend D.W. : The theory and practice of 3D PET*, chapter Quantitative procedures in 3D PET, pages 55–109. The Netherlands : Kluwer Academic, 1998.
- [7] A. D. Beggs, S. F. Hain, K. M. Curran, and M. J. O’Doherty. FDG PET as a ‘metabolic biopsy’ tool in non-lung lesions with indeterminate biopsy. *European Journal of Nuclear*

- Medicine and Molecular Imaging*, 29 :542–546, 2002.
- [8] J.C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York, 1981.
- [9] K. J. Biehl, F. Kong, F. Dehdashti, J. Y. Jin, S. Mutic, I. El Naqa, B. Siegel, and J. Bradley. 18F-FDG PET definition of gross tumor volume for radiotherapy of non-small cell lung cancer : Is a single standardized uptake value threshold approach appropriate? *Journal of Nuclear Medicine*, 47 :1808–1812, 2006.
- [10] A. Bijaoui, J.-L. Starck, and F. Murtagh. Restauration des images multiéchelles par l’algorithme à trous. *Traitement du Signal*, 3 :229–243, 1994.
- [11] Q.C. Black, I.S. Grills, L.L. Kestin, C-Y. O. Wong, J. W. Wong, A.A. Martinez, and D. Yan. Defining a radiotherapy target with positron emission tomography. *International Journal of Radiology Oncology Biology Physics*, 60(4) :1272–1282, 2004.
- [12] R. Boellaard, N.C. Krak, O.S. Hoekstra, and A.A. Lammertsma. Effects of noise, image resolution, and ROI definition on the accuracy of standard uptake values : a simulation study. *Journal of Nuclear Medicine*, 45 :1519–1527, 2004.
- [13] T. Bortfeld. IMRT : a review and preview. *Physics in Medicine and Biology*, 51 :R363–R379, 2006.
- [14] P. Bourguet and et al. Comprendre la tomographie par émission de positons au [18F]-FDG en cancérologie. Information à l’usage des personnes malades et de leurs proches. *Bulletin du cancer*, 92(7) :723–732, 2005.
- [15] N. Boussion, M. Hatt, F. Lamare, Y. Bizais, A. Turzo, C. Cheze-Le Rest, and D. Visvikis. A multiresolution image based approach for correction of partial volume effects in emission tomography. *Physics in Medicine and Biology*, 51 :1857–1876, 2006.
- [16] N. Boussion, M. Hatt, F. Lamare, C. Cheze-Le Rest, and D. Visvikis. Contrast enhancement in emission tomography by way of synergistic PET/CT image combination. *Computer Methods and Programs in Biomedicine*, 2008.
- [17] N. Boussion, M. Hatt, A. Reilhac, and D. Visvikis. Fully automated partial volume attenuation in PET : a wavelet approach without anatomical information. In *IEEE*

-
- Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, volume 4, pages 2423–2427, 2007.
- [18] N. Bousson, M. Hatt, and D. Visvikis. Partial volume correction in PET based on functional volumes. In *Society of Nuclear Medicine (SNM) annual meeting*, 2008.
- [19] J. D. Bradley, C. A. Perez, F. Dehdashti, and B. A. Siegel. Implementing biologic target volumes in radiation treatment planning for non-small cell lung cancer. *Journal of Nuclear Medicine*, 45 :96–101, 2004.
- [20] D. Brasse, P.E. Kinahan, C. Lartizien, C. Comtat, M. Casey, and C. Michel. Correction methods for random coincidences in fully 3D whole-body PET : Impact on data and image quality. *Journal of Nuclear Medicine*, 46(5) :859–867, 2005.
- [21] J.A. Browne and A.R. Pierre. A row-action alternative to the EM algorithm for maximizing likelihoods in emission tomography. *IEEE Transactions on Medical Imaging*, 15 :687–699, 1996.
- [22] G.L. Brownell. A history of positron imaging. Présentation faite à l’occasion des célébrations de la 50ème année de service de l’auteur auprès du Massachussetts General Hospital, october 1999.
- [23] G.L. Brownell, C.A. Burnham, S. Wilensky, S. Aranow, H. Kazemi, and D. Strieder. New developments in positron scintigraphy and the application of cyclotron-produced positron emitters. In *Proceedings of the Symposium on Medical Radioisotope Scintigraphy*, pages 163–176, 1968.
- [24] G.L. Brownell and W.H. Sweet. Localization of brain tumors with positron emitters. *Nucleonics*, 11 :40–45, 1953.
- [25] H. Caillol, A. Hillion, and W. Pieczynski. Fuzzy random Markov fields and unsupervised image segmentation. *IEEE Transactions on Geoscience Remote Sensing*, 4 :801–810, 1993.
- [26] H. Caillol, W. Pieczynski, and A. Hillion. Estimation of fuzzy gaussian mixture and unsupervised statistical image segmentation. *IEEE Transactions on Image Processing*, 6(3) :425 – 440, 1997.

- [27] C. Carincotte and J-M. Boucher. Chaînes de Markov cachées floues et segmentation non supervisée d'images. In *Logique Floue et Applications (LFA'04)*, 2004.
- [28] G. Celeux and J. Diebolt. L'algorithme SEM : un algorithme d'apprentissage probabiliste pour la reconnaissance de mélange de densités. *Revue de statistique appliquée*, 34(2) :35–52, 1986.
- [29] J.L. Chen, S.R. Gunn, and M.S. Nixon. Markov random field model for segmentation of PET images. *Lecture Notes on Computer Science*, 2082 :468–474, 2001.
- [30] N. Christian, A. Bol, M. de Bast, D. Labar, J. Lee, P. Mahy, and et al. Determination of tumour hypoxia with the PET tracer [(18F)]EF3 : improvement of the tumour-to-background ratio in a mouse tumour model. *European Journal of Nuclear Medicine and Molecular Imaging*, 34(9) :1348–1354, 2007.
- [31] C. Collet and F. Salzenstein. Fuzzy Markov random fields versus chains for multispectral image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 28(11) :1753–1767, 2006.
- [32] A. H. Compton. A quantum theory of the scattering of X-rays by light elements. *The physical review*, 25(5) :483–502, 1923.
- [33] O. Couturier, G. Jerusalem, J. M. N'Guyen, and R. Hustinx. Sequential positron emission tomography using [18F]fluorodeoxyglucose for monitoring response to chemotherapy in metastatic breast cancer. *Clin. Cancer Res.*, 12(21) :6437–6443, 2006.
- [34] O. Couturier, F. Leost, M. Campone, T. Carlier, J. F. Chatal, and R. Hustinx. La 3'-déoxy-3'-[18f] fluorothymidine ([18f]-ft) est-elle le prochain traceur utilisé en routine clinique pour la tep après le [18f]-fdg? *Bulletin du Cancer*, 92 :789–798, 2005.
- [35] O. Couturier, A. Luxen, J. F. Chatal, J. P. Vuillez, P. Rigo, and R. Hustinx. Fluorinated tracers for imaging cancer with positron emission tomography. *European Journal of Nuclear Medicine and Molecular Imaging*, 31(8) :1182–1206, 2004.
- [36] J. Dai. Hybrid approach to speech recognition using hidden Markov models and Markov chains. *Proc. Inst. Elect. Eng : Vision Image Signal Processing*, 141(5) :273–279, 1994.

-
- [37] J.F. Daisne, M. Sibomana, A. Bol, and et al. Tri-dimensional automatic segmentation of PET volumes based on measured source-to-background ratios : influence of reconstruction algorithms. *Radiotherapy Oncology*, 69 :247–250, 2003.
- [38] J. B. Davis, B. Reiner, M. Huser, C. Burger, G. Szekely, and I.F. Ciernik. Assessment of ^{18}F PET signals for automatic target volume definition in radiotherapy treatment planning. *Radiotherapy and Oncology*, 80 :43–50, 2006.
- [39] M. Defrise. Reconstruction d’image en tomographie par émission. *Médecine nucléaire*, 31 :142–152, 2007.
- [40] Y. Delignon, A. Marzouki, and W. Pieczynski. Estimation of generalized mixtures and its application in image segmentation. *IEEE Transactions on image processing*, 6(10) :1364–1375, 1997.
- [41] O. Demirkaya. Lesion segmentation in wholebody images of PET. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, pages 2873–2876, 2003.
- [42] A.P. Dempster, N.M. Laird, and D.B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. B*, 39 :1–38, 1977.
- [43] Y. Demura, T. Tsuchida, T. Ishizaki, S. Mizuno, Y. Totani, S. Ameshima, and et al. ^{18}F -FDG accumulation with PET for differentiation between benign and malignant lesions in the thorax. *Journal of Nuclear Medicine*, 44 :540–548, 2003.
- [44] M. Desco, S. Reig, A. Santos, J. Pascau, N. Malpica, and P. Garcia-Barreno. Statistical segmentation of multidimensionnal brain datasets. *SPIE medical imaging*, 2001.
- [45] P. Devijver. Baum’s forward backward algorithm revisited. *Pattern Recognition Letters*, 3 :369–373, 1985.
- [46] K. Dhital, C. A. Saunders, P. T. Seed, M. J. O’Doherty, and J. Dussek. ^{18}F Fluorodeoxyglucose positron emission tomography and its prognostic value in lung cancer. *European Journal of Cardiothoracic Surgery*, 18 :425–428, 2000.
- [47] IMV Medical Information Division. 2007 radiation oncology market summary report. *IMV*, 2008.

-
- [48] R.K. Doot, P.E. Christian, D.A. Mankoff, and P.E. Kinahan. Reproducibility of quantifying tracer uptake with PET/CT for evaluation of response to therapy. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, 2007.
- [49] J.C. Dunn. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3 :32–57, 1973.
- [50] I. El Naqa, D. Yang, A. Apte, D. Khullar, S. Mutic, J. Zheng, J.D. Bradley, P. Grigsby, and J.O. Deasy. Concurrent multimodality image segmentation by active contours for radiotherapy treatment planning. *Medical Physics*, 34(12) :4738–4749, 2007.
- [51] Y.E. Erdi, O. Mawlawi, S.M. Larson, M. Imbriaco, H. Yeung, R. Finn, and J.L. Humm. Segmentation of lung lesion volume by adaptative positron emission tomography image thresholding. *Cancer*, 80(S12) :2505–2509, 1997.
- [52] P. M. Evans. Anatomical imaging for radiotherapy. *Physics in Medicine and Biology*, 53 :151–191, 2008.
- [53] L. Fin, P. Bailly, J. Daouk, and M. E. Meyer. A practical, semi-experimental system matrix for 2-D PET image reconstruction : Comparison with a geometrical model. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, 2008.
- [54] R. Fjortoft, Y. Delignon, W. Pieczynski, M. Sigelle, and F. Tupin. Unsupervised classification of radar images using hidden Markov chains and hidden Markov random fields. *IEEE Transactions on geoscience and remote sensing*, 41(3) :675–686, 2003.
- [55] E. Ford. Pancreatic tumor delineation with fdg-pet/ct : smaller volumes and regions not identified with ct alone. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [56] J.L. Fox, R. Rengan, W. O’meara, E. Yorke, Y. Erdi, S. Nehmeh, S.A. Leibel, and K.E. Rosenzweig. Does registration of PET and planning CT images decrease interobserver and intraobserver variation in delineating tumor volumes for non-small-cell lung cancer ? *International Journal of Radiology Oncology Biology Physics*, 62(1) :70–75, 2005.

- [57] N.M.T. Freedman, S.K. Sundaram, K. Kurdziel, J.A. Carrasquillo, and et al. Comparison of SUV and Patlak slope for monitoring of cancer therapy using serial PET scans. *European Journal of Nuclear Medicine*, 30 :46–53, 2003.
- [58] X. Geets, J.A. Lee, A. Bol, M. Lonneux, and V. Grégoire. A gradient-based method for segmenting FDG-PET images : methodology and validation. *European Journal of Nuclear Medicine and Molecular Imaging*, 34 :1427–1438, 2007.
- [59] F. Giammarile. Tep et optimisation de la radiothérapie : comment ? quels traceurs ? pour quoi faire ? *Médecine nucléaire*, 31 :516–520, 2007.
- [60] J. S. Greenberg, A. Kalend, F. Sciurba, J. R. Jett, R. Landreneau, and C. Belani. Developement of multifield three-dimensional conformal radiotherapy of lung cancer using a total dose-volume histogram. *Radiotherapy Oncology Investigation*, 3 :243–255, 1996.
- [61] V. Gregoire. Is there any future in radiotherapy planning without the use of PET : unraveling the myth. *Radiotherapy and Oncology*, 73 :261–263, 2004.
- [62] H. Guan, T. Kubota, X. Huang, X. S. Zhou, and M. Turk. Automatic hot spot detection and segmentation in whole body FDG-PET images. In *International conference on image processing (ICIP)*, 2006.
- [63] A. Hafiane, B. Zavidovique, and S. Chaudhuri. A modified FCM with optimal Peano scans for image segmentation. *IEEE conférence ?*, 2005.
- [64] S. M. Hahn, J. Karp, and D. Mankoff. Joint ASTRO/SNM : advances in molecular imaging for radiation oncology. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [65] S. F. Hain, K. M. Curran, A. D. Beggs, I. Fogelman, M. J. O’Doherty, and M. N. Maisey. FDG PET as a ‘metabolic biopsy’ tool in thoracic lesions with indeterminate biopsy. *European Journal of Nuclear Medicine*, 28 :1336–1340, 2001.
- [66] L.M. Hamberg, G.J. Hunter, N.M. Alpert, N.C. Choi, J.W. Babich, and A.J. Fischman. The dose uptake ratio as an index of glucose metabolism : useful parameter or oversimplification. *Journal of Nuclear Medicine*, 35 :1308–1312, 1994.

- [67] G. Hanna. Does PET-CT in the radiotherapy treatment planning process for non-small cell lung cancer alter treatment volumes in already PET-CT staged patients? In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [68] S. Hanna. PET/CT in radiotherapy treatment planning : An unistitutional series. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [69] L.B. Harrison, M. Chadha, R.J. Hill, K. Hu, and D. Shasha. Impact of tumor hypoxia and anemia on radiation therapy outcomes. *Oncologist*, 7 :492–508, 2002.
- [70] M. Hatt, F. Lamare, N. Bousson, A. Turzo, C. Collet, F. Salzenstein, C. Roux, K. Carson, P. Jarritt, C. Cheze-Le Rest, and D. Visvikis. Fuzzy hidden Markov chains segmentation for volume determination and quantitation in PET. *Physics in Medicine and Biology*, 52 :3467–3491, 2007.
- [71] M. Hatt, A. Turzo, C. Roux, and D. Visvikis. A fuzzy locally adaptive Bayesian segmentation approach for volume determination in PET. *IEEE Transactions on Medical Imaging*, 2008.
- [72] S-C. Huang. Anatomy of SUV. *Nuclear Medicine Biology*, 27 :643–646, 2000.
- [73] H.M. Hudson and R.S. Larkin. Accelerated image reconstruction using ordered subsets of projection data. *IEEE Transactions on Medical Imaging*, 13 :601–609, 1994.
- [74] D. Huglo. *Détermination du volume fonctionnel des tumeurs en Tomographie par Emission de Positons*. PhD thesis, Faculté de médecine Henri Warembourg, 2005.
- [75] D. Huglo, M. Vermandel, J. Rousseau, C. Vasseur, and M. Steinling. Segmentation of lung lesion volume by adaptative positron emission tomography image thresholding. *Médecine nucléaire*, 28(4) :155–162, 2004.
- [76] R. Hustinx, F. Bénard, and A. Alavi. Whole-body FDG-PET imaging in the management of patients with cancer. *Seminars in nuclear medicine*, 21(1) :35–46, 2002.
- [77] R. Hustinx, R. J. Dolin, F. Bénard, A. Bhatnagar, D. Chakraborty, R. J. Smith, S. Jang, and A. Alavi. Impact of attenuation correction on the accuracy of FDG-PET in patients

- with abdominal tumors : a free-response ROC analysis. *European Journal of Nuclear Medicine*, 27(9) :1365–1371, 2000.
- [78] R. Hustinx, C. Lemaire, G. Jerusalem, P. Moreau, D. Cataldo, B. Duysinx, J. Aerts, M-F. Fassotte, J. Foidart, and A. Luxen. Whole-body tumor imaging using PET and 2-18F-Fluoro-L-Tyrosine : Preliminary evaluation and comparison with 18F-FDG. *Journal of Nuclear Medicine*, 44 :533–539, 2003.
- [79] R. Hustinx, R.J. Smith, F. Benard, D.I. Rosenthal, M. Machtay, L. A. Farber, and A. Alavi. Dual time point fluorine-18 fluorodeoxyglucose positron emission tomography : a potential method to differentiate malignancy from inflammation and normal tissue in the head and neck. *European Journal of Nuclear Medicine*, 26(10) :1345–1348, 1999.
- [80] Leon Jaroff. Time magazine’s medical invention of the year 2000. *Time Magazine*, november, thursday the 30th 2000.
- [81] H. Jarrit, K. Carson, A.R. Hounsel, and D. Visvikis. The role of PET/CT scanning in radiotherapy planning. *British Journal of Radiology*, 2005.
- [82] W. Jentzen, L. Freudenberg, E.G. Eising, M. Heinze, W. Brandau, and A. Bockisch. Segmentation of PET volumes by iterative image thresholding. *Journal of Nuclear Medicine*, 48 :108–114, 2007.
- [83] G. Jerusalem, R. Hustinx, Y. Beguin, and G. Fillet. The value of positron emission tomography (PET) imaging in disease staging and therapy assessment. *Ann. Oncol.*, 13(S4) :227–234, 2002.
- [84] G. Jerusalem, N. Withofs, A. Rorive, and R. Hustinx. Positron emission tomography in oncology : an update. *La Revue du praticien*, 57(17) :1864–1870, 2007.
- [85] N. L. Johnson and S. Kotz. *Distributions in Statistics : Continuous Univariate Distributions*, volume 1. New York : Wiley, 1970.
- [86] K. Jordan. IEC emission phantom appendix performance evaluation of positron emission tomographs. *Medical and Public Health Research Programme of the European Community*, 1990.

- [87] S. Kamata, R.O. Eason, and Y. Bandou. A new algorithm for N-dimensional Hilbert scanning. *IEEE Transactions on Image Processing*, 8 :964–973, 1999.
- [88] J. S. Karp, S. Surti, M. E. Daube-Witherspoon, and G. Muehllehner. Benefit of Time-of-Flight in PET : Experimental and clinical results. *Journal of Nuclear Medicine*, 49(3) :462–470, mars 2008.
- [89] J.T. Kent and K.V. Mardia. Spatial classification based on fuzzy membership. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5) :659–671, 1991.
- [90] J.W. Jr. Keyes. SUV : Standard uptake or silly useless value? *Journal of Nuclear Medicine*, 36 :1836–1839, 1995.
- [91] N.C. Krak, R. Boellaard, and et al. Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *European Journal of Nuclear Medicine and Molecular Imaging*, 32 :294–301, 2005.
- [92] M. Krengli. (18)F-FDG-PET/CT imaging for staging and radiotherapy treatment planning of patients affected by carcinoma of the anal canal. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [93] F. Lamare, M. Busch, R. Brinks, J. Claire, P. Bruyant, and D. Visvikis. Local respiratory motion correction (LRMC) for PET/CT imaging. *Journal of Nuclear Medicine*, 48(S2) :196–200, 2007.
- [94] F. Lamare, M. Ledesma-Carbayo, T. Cresson, G. Kontaxakis, A. Santos, C. Cheze-Le Rest, A. Reader, and D. Visvikis. List mode based image reconstruction for respiratory motion correction in PET using non-rigid body transformations. *Physics in Medicine and Biology*, 2007.
- [95] F. Lamare, M. Ledesma-Carbayo, A. Reader, V. Bettinardi, O. Malwawi, G. Kontaxakis, A. Santos, C. Cheze-Le Rest, and D. Visvikis. Performance of an image reconstruction algorithm incorporating continuous B-spline functions in the system matrix for respiratory motion correction in 4D PET/CT. *Journal of Nuclear Medicine*, 42(2) :197–200, 2007.

-
- [96] F. Lamare, A. Turzo, Y. Bizais, C. Cheze-Le Rest, and D. Visvikis. Validation of a Monte Carlo simulation of the Philips Allegro/Gemini PET systems using GATE. *Physics in Medicine and Biology*, 51 :943–962, 2006.
- [97] A. Le Pogam, N. Bousson, M. Hatt, G. Fahd, and D. Visvikis. A 3D multi resolution local analysis approach for correction of partial volume effects in emission tomography. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, 2008.
- [98] A. Le Pogam, M. Hatt, N. Bousson, D. Guilloteau, J-L. Beaulieu, C. Prunier, F. Turkeheimer, and D. Visvikis. A wavelet-based hidden Markov model and multi-resolution approach for conditional partial volume correction in emission tomography. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, 2008.
- [99] H. Li, W. L. Thorstad, K. J. Biehl, R. Laforest, Y. Su, K. Shoghi, E. D. Donnelly, D. A. Low, and W. Lu. A novel PET tumor delineation method based on adaptive region-growing and dual-front active contours. *Medical Physics*, 35(8) :3711–3721, 2008.
- [100] Z. Liang, R. J. Jaszczak, and R. E. Coleman. Parameter estimation of finite mixture using the EM algorithm and information criteria with application to medical image processing. *IEEE Transactions on Nuclear Sciences*, 39(4) :1126–1133, 1992.
- [101] C.C. Ling, J. Humm, S. Larson, and et al. Towards multi-dimensional radiotherapy (MD-CRT) : biological imaging and biological conformality. *International Journal of Radiology Oncology Biology Physics*, 47 :551–560, 2000.
- [102] C-Y. Liu and C.Y. Hsu. Automatic segmentation of PET images. In *conference ? ?, ???*
- [103] J. P. Loannidis and J. Lau. 18F-FDG PET for the diagnosis and grading of soft-tissue sarcoma : a meta-analysis. *Journal of Nuclear Medicine*, 44 :717–724, 2003.
- [104] M. A. Lodge, V. Dilsizian, and B. R. Line. Performance assessment of the philips GEMINI PET/CT scanner. *Journal of Nuclear Medicine*, 45 :425, 2004.
- [105] M. Macedon. Contribution of PET/CT in radiation therapy treatment planning for rectal cancer in the neoadjuvant setting. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.

-
- [106] J.B. MacQueen. Some methods for classification and analysis of multivariate observations. In University of California Press, editor, *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297, 1967.
- [107] K. Mah. Variability in identification of positive nodes for head and neck cancers : Comparison of CT alone with PET-CT. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [108] K. Mah and J. D. Bourland. Advances in CT and PET imaging for radiation treatment planning. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [109] T.P.J. Maroquin and S. Mitte. Probabilistic solution of ill-posed problems in computational vision. *Journal of the american statistical association*, 82 :75–89, 1987.
- [110] M. Mohammed and P. Gader. Generalized hidden Markov models-part I : theoretical frameworks. *EEE Transaction on Fuzzy Systems*, 8(1) :67–81, 2000.
- [111] D. W. G. Montgomery, A. Amira, and H. Zaidi. Fully automated segmentation of oncological PET volumes using a combined multiscale and statistical model. *Medical Physics*, 34(2) :722–736, 2007.
- [112] Y. Morimoto, Y. Ueno, K. Tsuchiya, S. Kojima, T. Ishitsu, I. Takahashi, W. Takeuchi, and K. Amemiya. Performance of a prototype brain PET scanner based on semiconductor detectors. In *Society of Nuclear Medicine (SNM) annual meeting*, june 2008.
- [113] D.C. Munson. A note on lena. *IEEE Transactions on Image Processing*, 1996.
- [114] S. A. Nehmeh, Y. E. Erdi, A. Pevsner, K. E. Rosenzweig, E. Yorke, G. S. Mageras, H. Schoder, P. Vernon, O. Squire, H. Mostafavi, S. M. Larson, and J. L. Humm. Four dimensional (4D) PET/CT imaging of the thorax. *Medical Physics*, 31 :3179–3186, 2004.
- [115] U. Nestle, A. Schaefer-Schuler, C. Sebastian-Welch, D. Hellwig, C. Rube, and C-M. Kirsch. Comparison of different methods for delineation of 18F-FDG PET-positive tissue for target volume definition in radiotherapy of patients with non-small cell lung cancer. *Journal of Nuclear Medicine*, 46(8) :1342–1348, 2005.

- [116] A. Obata, S. Kasamatsu, J. S. Lewis, T. Furukawa, S. Takamatsu, J. Toyohara, and et al. Basic characterization of ^{64}Cu -ATSM as a radiotherapy agent. *Nuclear Medicine Biology*, 32 :21–28, 2005.
- [117] J. Okada, K. Yoshikawa, M. Itami, K. Imaseki, K. Uno, J. Itami, J. Kuyama, A. Mikata, and N. Arimizu. Positron emission tomography using fluorine-18-fluorodeoxyglucose in malignant lymphoma : a comparison with proliferative activity. *Journal of Nuclear Medicine*, 32 :333–339, 1992.
- [118] M. Öllers, G. Bosmans, A. vann Baardwijk, A. Dekker, and et al. The integration of PET-CT scans from different hospitals into radiotherapy treatment planning. *Radiotherapy and Oncology*, 2008.
- [119] D. Palma, E. Vollans, K. James, S. Nakano, V. Moiseenko, R. Shaffer, M. McKenzie, J. Morris, and K. Otto. Volumetric modulated arc therapy for delivery of prostate radiotherapy : Comparison with intensity-modulated radiotherapy and three-dimensional conformal radiotherapy. *International Journal of Radiology Oncology Biology Physics*, 2008.
- [120] J. R. Palta. Understanding the uncertainties in proton therapy planning and delivery. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [121] N. Paquet, A. Albert, J. Foidart, and R. Hustinx. Within-patient variability of $(^{18}\text{F})\text{-FDG}$: standardized uptake values in normal tissues. *Journal of Nuclear Medicine*, 45(5) :784–788, 2004.
- [122] W. Paul Segars, D. S. Lalush, and B. M. W. Tsui. Modeling respiratory mechanics in the MCAT and spline-based MCAT phantoms. *IEEE Transactions on Nuclear Sciences*, 48(1) :89–97, 2001.
- [123] A. Peng and W. Pieczynski. Adaptive mixture estimation and unsupervised local Bayesian image segmentation. *Graph. Models Image Processing*, 57 :389 – 399, 1995.
- [124] W. Pieczynski. Statistical image segmentation. *Machine graphics and vision*, 1(1/2) :261–268, 1992.

-
- [125] W. Pieczynski. Champs de Markov cachés et estimation conditionnelle itérative. *Traitement du Signal*, 11(2) :141–153, 1994.
- [126] W. Pieczynski. Chaines de Markov triplet. *Comptes rendus de l'académie des sciences - mathématiques, Paris Ser. I*, 335 :275–278, 2002.
- [127] W. Pieczynski. Modèles de Markov en traitement d'images. *Traitement du Signal*, 20(3) :255–277, 2003.
- [128] W. Pieczynski. Pairwise Markov chains. *IEEE Transactions on pattern analysis and machine intelligence*, 25(5) :634–639, 2003.
- [129] J. Qi. A unified noise analysis for iterative image estimation. *Physics in Medicine and Biology*, 48 :3505–3519, 2003.
- [130] J. Qi. Noise propagation in iterative image reconstruction with line searches. *IEEE Transactions on Nuclear Sciences*, 52(1) :57–62, 2005.
- [131] J. Qi and R. M. Leahy. Iterative reconstruction techniques in emission computed tomography. *Physics in Medicine and Biology*, 51 :R541–R578, 2006.
- [132] J. Rasey, W. Koh, M. Evans, L. Peterson, T. Lewellen, M. Graham, and K. Krohn. Quantifying regional hypoxia in human tumors with positron emission tomography of 18ffluoromisonidazole : a pretherapy study of 37 patients. *International Journal of Radiology Oncology Biology Physics*, 24 :417–428, 1996.
- [133] A.J. Reader, S. Ally, F. Bakatselos, R. Manavaki, R.J. Walledge, A.P. Jeavons, and et al. One-pass list-mode EM algorithm for high-resolution 3D PET image reconstruction into large array. *IEEE Transactions on Nuclear Sciences*, 49(3) :693–699, 2002.
- [134] A. Reilhac, C. Lartizien, N. Costes, S. Sans, C. Comtat, R. N. Gunn, and A. C. Evans. PET-SORTEO : A Monte Carlo-based simulator with high count rate capabilities. *IEEE Transactions on Nuclear Sciences*, 51(1) :46–52, 2004.
- [135] C. Riddell, R.E. Carson, and S.L. Bacharach. The watershed algorithm : A method to segment noisy PET transmission images. *IEEE Transactions on Nuclear Sciences*, 1999.

-
- [136] W. Römer, A. R. Hanauske, S. Ziegler, R. Thodtmann, W. Weber, C. Fuchs, and et al. Positron emission tomography in non-Hodgkin's lymphoma : assessment of chemotherapy with fluorodeoxyglucose. *Blood*, 91 :4464–4471, 1998.
- [137] O. Rousset, A. Rahmim, A. Alavi, and H. Zaidi. Partial volume correction strategies in PET. *PET clinics*, 2(2) :235–249, 2007.
- [138] F. Salzenstein, C. Collet, S. Lecam, and M. Hatt. Non-stationary fuzzy Markov chain. *Pattern Recognition Letters*, 28(16) :2201–2208, 2007.
- [139] F. Salzenstein and W. Pieczynski. Parameter estimation in hidden fuzzy Markov random fields and image segmentation. *CVGIP : Graphical Models and Image Processing*, 59(4) :205–220, 1997.
- [140] F. Salzenstein and W. Pieczynski. Sur le choix de méthode de segmentation statistique d'images. *Traitement du Signal*, 15(2) :119–127, 1998.
- [141] C. A. B. Saunders, J. E. Dussek, M. J. O'Doherty, and M. N. Maisey. Evaluation of fluorine-18 fluorodeoxyglucose whole body positron emission tomography imaging in the staging of lung cancer. *Ann. Thorac. Surg.*, 67 :790–797, 1999.
- [142] A.M.B. Seminario. *Segmentation et modélisation dynamiques. Application à la reconstruction 3D d'images échocardiographiques*. PhD thesis, ENST Bretagne, Rennes I : Brest, 2003.
- [143] J. Seppala. Biologic target volume definition with [11C]acetate PET/CT in IMRT of patients with prostate cancer. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [144] L.A. Shepp and Y. Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, MI-1 :113–122, 1982.
- [145] A. Shields, J. Grierson, B. Dohmen, and et al. Imaging proliferation *in vivo* with F-18 FLT and positron emission tomography. *Nature Medicine*, 4 :1334–1336, 1998.
- [146] W. Skarbek. Generalized Hilbert scan in image printing. *Theoretical foundations of computer vision*, R. Klette and W.G. Kropetsh, editors, Akademik Verlag, pages 45–57, 1992.

-
- [147] M. Soret, S.L. Bacharach, and I. Buvat. Partial-volume effect in PET tumor imaging. *Journal of Nuclear Medicine*, 48 :932–945, 2007.
- [148] C.W. Stearns and J. A. Fessler. 3D PET reconstruction with FORE and WLS-OS-EM. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, volume 2, pages 912–915, 2002.
- [149] A. Sun. A pilot prospective study of metabolic and anatomic response using FDG PET CT before, during and after radiotherapy in lung cancer. In *American Society for Therapeutic Radiology and Oncology (ASTRO) annual meeting*, 2008.
- [150] S. Surti, A. Kuhn, M. E. Werner, A. E. Perkins, J. Kolthammer, and J. S. Karp. Performance of philips gemini TF PET/CT scanner with special consideration for its time-of-flight imaging capabilities. *Journal of Nuclear Medicine*, 48 :471–480, 2004.
- [151] W.H. Sweet. The use of nuclear disintegration in the diagnosis and treatment of brain tumor. *New England Journal of Medicine*, 245 :875–878, 1951.
- [152] J. A. Thie. Understanding the standardized uptake value, its methods, and implications for usage. *Journal of Nuclear Medicine*, 45(9) :1431–1434, 2004.
- [153] J. Tohka, A. Reilhac, J. Mykkänen, and U. Ruotsalainen. Brain surface extraction from PET images with deformable model : assessment using Monte Carlo simulator. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, 2003.
- [154] S. Tomei, C. Odet, and C. Lartizien. Détection d’hyper signaux en tomographie par émission de positons 3D : approche statistique basée sur la combinaison d’une transformée en ondelettes discrète avec un test de Hotelling. In *Colloque GRETSI - traitement du signal et des images*, pages 869–872, 2007.
- [155] D. W. Townsend, T. Beyer, and T. M. Blodgett. PET/CT scanners : a hardware approach to image fusion. *Semin Nucl Med*, 33 :193–204, 2003.
- [156] P. Tylski, G. Bonniaud, E. Decencièrre, J. Stawiaski, J. Coulot, D. Lefkopoulos, and Ricard M. 18F-FDG PET images segmentation using morphological watershed : a phantom study. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, pages 2063–2067, 2006.

- [157] A. Van Baardwijk and et al. PET-CT-based auto-contouring in non-small-cell lung cancer correlates with pathology and reduces interobserver variability in the delineation of the primary tumour and involved nodal volumes. *International Journal of Radiology Oncology Biology Physics*, 68(3) :771–778, 2007.
- [158] A. W. van Baardwijk, G. Bosmans, L. Boersma, S. Wanders, A. L. A. Dekker, A. M. C. Dingemans, G. Bootsma, W. Geraedts, C. Pitz, J. Simons, Lambin P., and D. de Ruyscher. Individualized radical radiotherapy of non-small-cell lung cancer based on normal tissue dose constraints : A feasibility study. *International Journal of Radiology Oncology Biology Physics*, 71(5) :1394–1401, 2008.
- [159] J. A. Van Dalen, A. L. Hoffman, V. Dicker, W.V. Vogel, B. Wiering, T.J. Ruers, N. Karssemeijer, and W.J.G. Oyen. A novel iterative method for lesion delineation and volumetric quantification with FDG PET. *Nuclear Medicine Communications*, 28 :485–493, 2007.
- [160] J. F. Vansteenkiste, S. G. Stroobants, P. J. Dupont, P. R. De Leyn, E. K. Verbeke, G. J. Deneffe, and et al. Prognostic importance of the standardized uptake value of (18)F-fluoro-2-deoxy-glucose positron emission tomography scan in non-small-cell lung cancer : an analysis of 125 cases. *Journal of Clinical Oncology*, 17 :3201–3206, 1999.
- [161] Y. Venel, H. Garhi, A. de Muret, J-L. Baulieu, I. Barillot, and C. Prunier-Aesch. Comparaison de six méthodes de segmentation du volume tumoral sur la 18F-FDG TEP-TDM avec le volume de référence anatomopathologique dans les cancers bronchopulmonaires non à petites cellules. *Médecine nucléaire*, 32 :339–353, 2008.
- [162] D. Visvikis, O. Barret, T.D. Fryer, A. Turzo, Y. Bizais, and C. Cheze-Le Rest. Evaluation of respiratory motion effects in comparison with other parameters affecting PET image quality. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, 2004.
- [163] D. Visvikis, C. Cheze-Le Rest, and P. Jarritt. PET technology : current trends and future developments. *British Journal of Radiology*, 77 :906–910, 2004.

- [164] D. Visvikis, C. Cheze Le Rest, and F. Lamare. Le mouvement respiratoire en imagerie fonctionnelle du cancer : une revue des effets et des méthodes de correction. *Traitement du Signal*, 23(3) :1–11, 2003.
- [165] D. Visvikis, A. Turzo, S. Gouret, P. Damine, F. Lamare, Y. Bizais, and C. Cheze-Le Rest. Characterisation of SUV accuracy in FDG PET using 3D RAMLA and the Philips Allegro PET scanner. *Journal of Nuclear Medicine*, 45(5), 2004.
- [166] O. Warburg. On the origin of cancer cells. *Science*, 123 :309–314, 1956.
- [167] M.B. White. A semi-automatic approach to the delineation of tumour boundaries from PET data using Level Sets. In *Society of Nuclear Medicine (SNM) annual meeting*, 2005.
- [168] B. Yarenko, T. Riauka, D. Robinson, B. Murray, A. Alexander, A. McEwan, and W. Roa. Thresholding in pet images of static and moving targets. *Physics in Medicine and Biology*, 50 :5969–5982, 2005.
- [169] H. Young, R. Baum, U. Cremerius, K/ Herholz, O. Hoekstra, A. A. Lammertsma, and et al. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography : review and 1999 eortc recommendations. *European Journal of Cancer*, 35 :1773–1782, 1999.
- [170] H. Zaidi, M. Diaz-Gomez, A. Boudraa, and D.O. Slosman. Fuzzy clustering-based segmented attenuation correction in whole-body PET imaging. *Physics in Medicine and Biology*, 47 :1143–1160, 2002.
- [171] W. Zhu and T. Jiang. Automation segmentation of PET image for brain tumours. In *IEEE Nuclear Sciences Symposium and Medical Imaging Conference (NSS-MIC)*, volume 4, pages 2627–2629, 2003.