

Table des matières

Introduction générale	1
Chapitre 1 : La Tomographie par Emission de Positons (TEP)	
1.1 Introduction.....	3
1.2 L'imagerie conventionnelle.....	3
1.3 Principe d'utilisation de la TEP.....	4
1.3.1 Introduction.....	4
1.3.2 Qu'est ce qu'une TEP ?.....	4
1.3.3 Principe de l'imagerie TEP.....	5
1.4 Utilisation de la TEP en neurologie.....	14
1.4.1 Introduction.....	14
1.4.2 Petit rappel sur l'anatomie du cerveau.....	15
1.4.2.1 Les grandes régions anatomiques : face externe	15
1.4.2.2 Spécialisation du cortex.....	17
1.4.3 Quelques traceurs radioactifs pour la neurologie en TEP.....	18
1.4.4 Les tumeurs du cerveau.....	19
1.4.5 Les démences.....	21
1.4.6 Maladie d'Alzheimer	22
1.4.7 Maladie de Parkinson.....	23
1.5 Les limites de la TEP	25
1.6 Avenir de la TEP.....	25
1.7 Conclusion.....	26
Chapitre 2 : Les systèmes d'aide au diagnostic en imagerie médicale	
2.1 Introduction.....	27
2.2 Qu'est ce qu'un système CAD	27
2.3 Le système d'aide en imagerie Tep.....	28
2.3.1 Définition de base.....	28
2.3.2 Les systèmes CAD en TEP.....	29
2.4 La classification supervisée et la classification non supervisée.....	31
2.4.1 Introduction.....	31
2.4.2 La classification supervisée.....	32
2.4.3 La classification non-supervisée.....	32
2.5 Les systèmes CAD supervisés.....	33
2.5.1 Introduction	33
2.5.2 Les étapes d'un système CAD supervisé	33
2.5.3 L'identification initiale des anomalies.....	34

2.5.3.1	Le Prétraitement.....	34
2.5.3.2	Les caractéristiques descriptives.....	34
2.5.4	Les classifieurs.....	37
2.5.4.1	Introduction.....	37
2.5.4.2	Définition du classifieur.....	37
2.5.4.3	Les types de classifieur.....	39
2.6	Conclusion.....	41

Chapitre3 : les méthodes utilisées

3.1	Introduction.....	43
3.2	Théorie des ondelettes.....	43
3.2.1	Introduction.....	43
3.2.2	La transformation en ondelette.....	43
3.2.3	Analyse multi-résolution.....	44
3.2.3.1	Principe de l'analyse multirésolution.....	45
3.3	Types d'ondelettes.....	46
3.3.1	Ondelette de Haar.....	46
3.3.2	Ondelette de Daubechies.....	47
3.3.3	Ondelettes biorthogonales.....	48
3.4	Conclusion.....	49
3.5	Support vector machine ou séparateurs à vaste marge(SVM).....	49
3.5.1	Introduction.....	49
3.5.2	Le principe de fonctionnement général du SVM.....	49
3.5.2.1	Notions de base: Hyperplan, marge, et support vecteur.....	49
3.5.2.2	Pourquoi faut-il maximiser la marge?.....	51
3.5.2.3	Linéarité et non-linéarité.....	52
3.5.3	L'apprentissage.....	54
3.6	Noyaux.....	54
3.6.1	L'espace des caractéristiques.....	54
3.6.2	Quelles fonctions sont des noyaux.....	55
3.6.3	Conditions pour avoir un noyau.....	56
3.7	Conclusion.....	56

Chapitre 4 : Application du système CAD en imagerie TEP neurologique

4.1	Introduction.....	57
4.2	La base d'image PET-Sorteo.....	57
4.3	La lecture des images.....	59
4.3.1	Le format des images.....	59
4.3.2	Choix du langage de programmation.....	60
4.4	L'analyse fréquentielle.....	60
4.4.1	Extraction des caractéristiques.....	61
4.4.2	Algorithme de décomposition en ondelette.....	61
4.4.3	Utilisation de l'ondelette de Haar.....	67
4.4.4	Utilisation de l'ondelette coif2.....	68

4.4.5	Utilisation de l'ondelette bior4.4.....	69
4.5	La réduction des paramètres.....	70
4.5.1	Analyse en composantes principales.....	70
4.5.1.1	Algorithme de L'ACP	71
4.5.2	Calcul de la moyenne.....	73
4.6	Le choix du classifieur.....	75
4.6.1	Les séparateurs à vaste marge(SVM).....	75
4.6.1.1	Algorithme des SVM.....	75
4.6.1.1.1	Phase d'apprentissage.....	75
4.6.1.1.2	Phase de test.....	76
4.6.1.2	Conclusion.....	77
4.6.2	Régression linéaire.....	77
4.6.2.1	Conclusion.....	78
4.7	Recalage et fusion d 'images médicales.....	78
4.7.1	Introduction.....	78
4.7.2	Le recalage.....	79
4.7.2.1	Critère de ressemblance.....	80
4.7.2.2	Classe de transformation.....	80
4.7.2.3	Algorithmes d'optimisations.....	80
4.7.3	Principe de la fusion par ondelette.....	82
4.8	Justifications et discussions	84
4.8.1	Choix des caractéristiques.....	84
4.8.2	Réduction de la taille du vecteur caractéristique.....	85
4.8.3	Le choix des classifieurs.....	85
4.8.4	Le recalage et la fusion.....	86
7.9	Conclusion.....	86
	Conclusion générale.....	88
	Annexes.....	90
	Références bibliographiques.....	97

Table des figures

Figure 1.1 : Schéma du processus d'acquisition d'une TEP.....	5
Figure 1.2 : principe de l'imagerie TEP.....	6
Figure 1.3 : Désintégration du positon et réaction d'annihilation.....	7
Figure 1.4 : Les différentes étapes de la création d'une image TEP.....	8
Figure 1.5 : Les trois types d'acquisition : transmission, blanc et émission.....	9
Figure 1.6 : Principe de la détection en coïncidence de la TEP.....	10
Figure 1.7 : Nature des événements enregistrés en mode coïncidence.....	11
Figure 1.8 : Principe de fonctionnement d'un détecteur élémentaire.....	11
Figure 1.9 : reconstructions analytique et itérative.....	13
Figure 1.10 : Exemple de reconstruction tomographique.....	14
Figure 1.11 : Face externe des grandes régions anatomiques.....	16
Figure 1.12 : Spécialisation du cortex.....	16
Figure 1.13 : Aires primaires.....	17
Figure 1.14 : Aires associatives.....	18
Figure 1.15 : Scanner normal du cerveau au FDG-PET.....	20
Figure 1.16 : comparaison entre un sujet sain et un sujet atteint de la MA.....	23
Figure 1.17 : Le striatum.....	24
Figure 1.18 : La formation de la synapse dopaminergique dans le striatum chez un sujet sain, et un patient atteint de la maladie de Parkinson.....	24
Figure 1.19 : A : représente un examen normal au [18F]-Dopa, et B : patient atteint de la maladie de Parkinson.....	25
Figure2.1 : Schéma représentatif de la structure d'un système CAD.....	30
Figure2.2 : les différentes étapes de la méthode de Guan.....	31
Figure2.3 : Les différents types de classifications et les méthodes les plus utilisées pour chaque type.....	39
Figure2.4.Méthodologies d'une LDA a deux classes H1 et H0.....	40
Figure2.5 : Architecture et fonctionnement d'un réseau de neurones.....	41
Figure3.1: Deux exemples d'ondelettes mères utilisées pour la transformation en ondelette.....	44
Figure3.2:décomposition de l'image sur un niveau.....	46
Figure3.3:décomposition de l'image sur deux niveaux.....	46
Figure 4.1 : Description du système de détection.....	58
Figure 4.2 : les différentes images existantes dans la base de données simulées.....	59
Figure4.3 : Structure du signal X au troisième niveau de décomposition.....	61
Figure 4.4 : Première étape de décomposition.....	62
Figure 4.5 : la décomposition au $j^{\text{ème}}$ niveau de la 2^{nde} étape.....	63
Figure 4.6 : décomposition en ondelette au niveau $J=3$	63

Figure 4.7: la décomposition en ondelette de Haar selon plusieurs niveaux.....	64
Figure 4.8 : La décomposition en ondelette coif2 selon différents niveaux.....	65
Figure4.9: La décomposition en ondelette Bior4.4 selon différents niveaux	66
Figure4.10 : utilisation de l'ondelette de Haar lors de la décomposition en ondelette.....	67
Figure4.11: Les histogrammes du 3 ^{ème} niveau de décomposition avec ondelette de Haar.....	68
Figure4.12: images représentatives des résultats lors de l'utilisation de l'ondelette coif2.....	68
Figure4.13: Histogrammes représentant l'image après décomposition et l'image approximation obtenus avec l'ondelette coif2.....	69
Figure4.14 : images représentatives des résultats lors de l'utilisation de l'ondelette bior4.4.....	69
Figure4.15 : Histogramme de l'image après décomposition en ondelette et histogramme de l'image approximation obtenue avec l'ondelette de Bior4.4.....	70
Figure 4.16 : Histogrammes correspondants à une réduction des paramètres du vecteur de caractéristiques.....	74
Figure 4.17 : Fusion des images médicales.....	79
Figure4.18: Algorithme de recalage.....	81
Figure4.19:le recalage de deux images.....	82
Figure4.20 : algorithme de la fusion en utilisant la décomposition par ondelette.....	83
Figure 4.21 : fusion de deux images après recalage.....	83

Liste des abréviations

TEP : Tomographie par Emission de positons

CAD : Computer Aided Detection

CADi: Computer Aided Diagnosis

FDG : 2-fluoro-2-deoxy-D-glucose

18F: Fluor 18

TDM: Tomodensitométrie

IRM: Imagerie par Résonance Magnétique

3D : 3 dimensions

2D : 2 dimensions

SUV : Standardized Uptake Value

SVM : Séparateurs à Vaste Marge

VN : Vrais négatifs

VP : Vrais Positifs

FN : Faux Négatifs

FP : Faux Positifs

MA : Maladie d'Alzheimer

MC : Monte Carlo

KeV : kilo électronvolt

PM : Photomultiplicateur

AMR : Analyse Multi-Résolution

ACP : Analyse en Composantes Principales

Introduction générale

Introduction générale

Les progrès actuel de la médecine, de la biologie et des techniques d'imagerie médicale permettent de sauver la vie de nombreux malades, d'améliorer la prise en charge du patient et d'optimiser les traitements administrés. Ces avancées ont rendu possible le diagnostic de certaines pathologies et la détection d'anomalies liées au fonctionnement physiologique des différentes parties du corps humain. Parmi ces modalités, on retrouve l'imagerie conventionnelle, également appelée **anatomique** ou **structurale**, qui regroupe la radiologie traditionnelle dite radiologie X, le scanner, et l'Imagerie par Résonance Magnétique (IRM). Ces techniques permettent d'obtenir des informations sur la **structure** des organes, leur forme, leurs limites, et dans certains cas leur contenu (structures osseuses, calculs vésicaux).

La tomographie par émission de positons (TEP) est une technique appartenant à la nouvelle génération de l'imagerie médicale utilisée depuis les années 90 en clinique, cette modalité consiste à injecter un produit radiotraceur relatif aux types d'organes du corps humain, et qui a pour objectif de visualiser l'information fonctionnelle. Au début, son application était dédiée principalement pour la neurologie avant d'être généralisée à d'autres spécialités médicales telles que la cardiologie, l'oncologie, etc...

L'optimisation de l'utilisation de l'imagerie TEP en clinique passe par une amélioration de l'instrumentation des caméras et la mise en place d'un ensemble de méthodes optimisant l'interprétation des images provenant des examens cliniques. Malgré tous les efforts actuels pour faciliter l'utilisation de l'imagerie TEP d'un point de vue diagnostic, la sensibilité et la spécificité de cette imagerie pourraient encore être améliorées afin d'aider les cliniciens dans la difficile tâche du diagnostic.

L'objet de cette présente étude consiste à mettre en place un système d'aide au diagnostic (CAD) dont l'idée principale est d'aider les radiologues dans l'interprétation des images médicales en utilisant des systèmes informatiques dédiés à fournir une « seconde opinion ». Cependant, la décision finale revient aux radiologues. Ce système est fondé sur les techniques d'analyse et de traitement d'images.

En premier lieu, les images seront sujettes à une décomposition en ondelette puis à une réduction des données, afin d'utiliser les vecteurs caractéristiques obtenus comme paramètres d'entrée de notre classifieur. Ce dernier, représente la partie déterminante pour la prise de décision sur l'aspect pathologique de l'organe diagnostiqué. Enfin, le système CAD procède au recalage et à la fusion des images IRM/TEP pour exprimer simultanément les aspects anatomique et fonctionnel de l'organe.

Comme nous disposons d'une base de données d'images « PET-Sorteo », constituée de 134 images PET neurologiques saines et pathologiques, nous allons l'exploiter pour évaluer les performances du système CAD proposé.

Ce mémoire est organisé en quatre chapitres, répartis comme suit :

- ❖ Le premier chapitre est consacré à l'étude de la tomographie par émission de positons et à ses applications dans le domaine de la neurologie ;

- ❖ Le deuxième chapitre présente un état de l'art du système CAD pour l'imagerie TEP ainsi que pour les autres modalités développées. L'organisation de ces méthodes en étapes standardisées est ensuite précisée. Enfin, nous abordons les différents classifieurs existants dans la littérature ;
- ❖ Le troisième chapitre aborde les différents outils et méthodes utilisés par le système CAD, à savoir la théorie des ondelettes et le séparateur à vaste marge (SVM) ;
- ❖ Et enfin, dans le quatrième chapitre nous avons appliqué des méthodes d'analyse des images par ondelettes et d'extraction de caractéristiques, celles-ci seront envoyées à un classifieur SVM pour déterminer la nature des images (pathologiques ou saines).

Nous terminons notre travail par une conclusion générale et nous donnons quelques perspectives.

Chapitre 1 : La Tomographie par Emission de Positons (TEP)

1.1 Introduction

La tomographie par émission de positons (TEP) est une modalité d'imagerie médicale, pratiquée par les spécialistes en médecine nucléaire. Elle a été développée pour la première fois en 1975, mais son utilisation en recherche et clinique ne date que des années 1990. Elle permet de mesurer en trois dimensions l'activité métabolique d'un organe grâce aux émissions produites par les positons (ou positrons) issus de la désintégration d'un produit radioactif injecté au préalable. (Terminologie anglo-saxonne : « positron emission tomography » - PET ou PETscan-)

La TEP est utilisée en recherche biomédicale, par exemple en imagerie cérébrale où elle permet de révéler les régions actives du cerveau lors de telle ou telle activité cognitive de manière analogue à ce qui se fait avec l'imagerie par résonance magnétique fonctionnelle.

La TEP permet de visualiser les activités du métabolisme des cellules : on parle d'imagerie fonctionnelle par opposition aux techniques d'imagerie dite structurelle comme celles basées sur les rayons X (radiologie ou CT-scan) qui réalisent des images de l'anatomie. Elle est considérée comme un outil diagnostique qui permet de déceler certaines pathologies qui se traduisent par une altération de la physiologie normale comme les cancers.

Ce chapitre résume les principes physiques de la formation de l'image TEP et développe son application en neurologie.

1.2 L'imagerie conventionnelle

La tomodensitométrie (TDM) et l'imagerie par résonance magnétique (IRM) sont les piliers de tout travail-up complet de diagnostic d'une maladie neurologique. L'IRM permet de donner de très bons détails anatomiques, en permettant la résolution spatiale et la délimitation des structures de la matière grise et blanche. La modalité d'imagerie la plus sensible pour la détection des tumeurs cérébrales peut être trouvée aux normes en T1 et T2 des images IRM pondérées. L'administration de contraste par voie intraveineuse évalue l'intégrité de la barrière hémato-encéphalique et soulève également la sensibilité de l'IRM en raison de nombreuses tumeurs, en particulier celles qui sont de grade supérieur, montrent une augmentation qui correspond à la prise de contraste. Cependant, les tumeurs de bas grade ont une prise de contraste faible ou minime, mais, généralement, ne montrent pas que le signal a augmenté sur les séquences pondérées en T2. Souvent, l'IRM fournit des informations supplémentaires sur le plan clinique utile, telles que la présence et / ou le degré d'effet de masse, l'hémorragie, œdème et nécrose, qui accompagnent souvent les tumeurs cérébrales. L'IRM et la TDM fournissent des informations anatomiques. Le degré de rehaussement de contraste peut également fournir des informations sur le grade de la tumeur, mais elle est limitée. Parce que les tumeurs cérébrales sont hétérogènes et elles se développent souvent d'une manière infiltrante, des informations exactes concernant le grade tumoral est indispensable.

Les limitations de l'IRM peuvent survenir dans le contexte post-thérapeutique dans lequel les patients traités par résection et suivis par l'amélioration spectacle de la radiothérapie qui peut être

causée par un post-traitement (c-à-d changements, radionécrose) ou tumorales résiduelles. Pour remédier à ces limitations, beaucoup de travaux sont réalisés avec des techniques de médecine nucléaire, en particulier la tomographie par émission de positons (TEP), c'est un moyen de gagner de précieuses informations concernant le grade de la tumeur. Aussi, la précision du diagnostic augmente, et l'information anatomique fournie par l'IRM ou la TDM peut être combinée, ou enregistrée, avec l'information métabolique fournie par les techniques de médecine nucléaire. La plupart des centres ont l'habitude d'inscrire le 18F-2-fluoro-2-désoxy-D-glucose (FDG) en PET et des images IRM pour l'évaluation des patients atteints de tumeurs cérébrales.

1.3 Principe d'utilisation de la tomographie par émission de positons

1.3.1 Introduction

La TEP repose sur le principe général de la scintigraphie qui consiste à injecter un traceur dont on connaît le comportement et les propriétés biologiques pour obtenir une image du fonctionnement d'un organe. Ce traceur est marqué par un atome radioactif (carbone, fluor, azote, oxygène...) qui émet des positons dont l'annihilation produit elle-même deux photons. La détection de la trajectoire de ces photons par le collimateur de la caméra TEP permet de localiser le lieu de leur émission et donc la concentration du traceur en chaque point de l'organe. C'est cette information quantitative que l'on représente sous la forme d'une image faisant apparaître en couleurs les zones de forte concentration du traceur.

1.3.2 Qu'est-ce qu'une TEP ?

La scintigraphie en TEP est obtenue par injection d'un traceur faiblement radioactif par voie intraveineuse. Le marqueur le plus souvent utilisé est le fluor (^{18}F) incorporé dans une molécule de glucose formant le ^{18}F -fluorodésoxyglucose (en abrégé ^{18}F -FDG). Ce traceur est semblable au glucose : il se fixe au niveau des tissus qui consomment de grandes quantités de ce sucre comme les tissus cancéreux, le muscle cardiaque ou encore le cerveau. Le fluor 18, dont la demi-vie est inférieure à deux heures, émet ensuite de façon temporaire des rayonnements que l'on peut suivre dans l'organisme du patient grâce à une caméra TEP. Le fluor 18 ainsi que les autres isotopes pouvant être utilisés (oxygène (^{15}O), azote (^{13}N), carbone (^{11}C)) ont une courte demi-vie, jusqu'à 110 minutes pour le fluor. Ces isotopes de courte durée nécessitent pour leur production un cyclotron.

Une caméra TEP est un appareil qui a l'aspect d'un *scanner* mais son principe de fonctionnement est différent.

L'atome radioactif se désintègre en émettant un positon. Celui-ci va s'annihiler avec un électron du milieu, après un très court parcours (en général inférieur à 1 mm). Cette annihilation produit deux photons gamma de 511 keV qui partent sur une même direction mais dans un sens opposé, ce qui rend possible le traitement tomographique des données. En effet, les capteurs situés tout autour du patient détectent les photons d'annihilation en coïncidence (c'est-à-dire ceux qui arrivent en même temps), ce qui permet d'identifier la ligne sur laquelle se trouve l'émission des photons. Un système informatique reconstitue ensuite à l'aide d'un algorithme de reconstruction les images de la répartition du traceur au niveau d'une partie ou de la totalité

du corps sous la forme d'une image 2D ou d'un objet 3D. Les images ainsi obtenues sont dites "d'émission" (la radioactivité provient du traceur injecté au patient). La résolution spatiale de l'image ainsi obtenue est comprise entre 4 et 7 mm².

Il est possible d'améliorer la qualité des images en utilisant le principe de correction d'atténuation. Pour effectuer cette correction, on utilisait initialement des images de transmission obtenues grâce à une source radioactive qui tourne rapidement autour du patient ; mais aujourd'hui, la plupart des caméras TEP sont couplées à un tomodynamomètre à rayons X (système TEP/TDM ou PET/CT en anglais), ce qui permet de superposer l'image fonctionnelle (Image TEP) à sa localisation anatomique précise dans le corps (Image CT). La correction d'atténuation ainsi réalisée permet de réaliser l'examen beaucoup plus rapidement et permet d'obtenir des images de meilleure qualité.

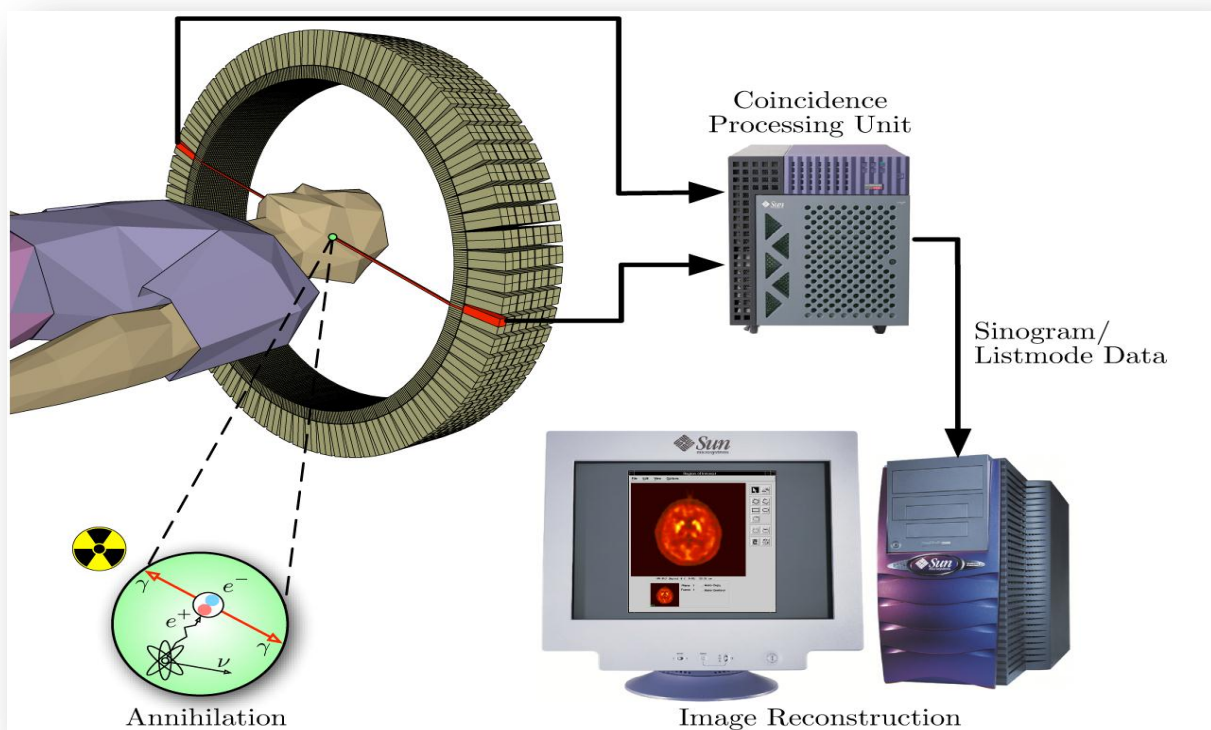


Figure 1.1 : Schéma du processus d'acquisition d'une TEP

1.3.3 Principe de l'imagerie TEP

De la même manière que pour les autres examens de médecine nucléaire, la réalisation d'un examen TEP nécessite l'injection d'un traceur radioactif.

Un traceur est constitué d'un vecteur moléculaire et d'un isotope radioactif qui permet de localiser la distribution de la molécule au sein de l'organisme.

Le choix d'un isotope est déterminé d'une part en fonction des propriétés chimiques qui conditionnent la possibilité de marquage et, d'autre part, selon le mode de désintégration qui permet une détection externe du rayonnement émis. La supériorité des émetteurs de positons par rapport aux émetteurs de photons gamma existe pour les deux critères précédents : marquage biochimique et détection externe.

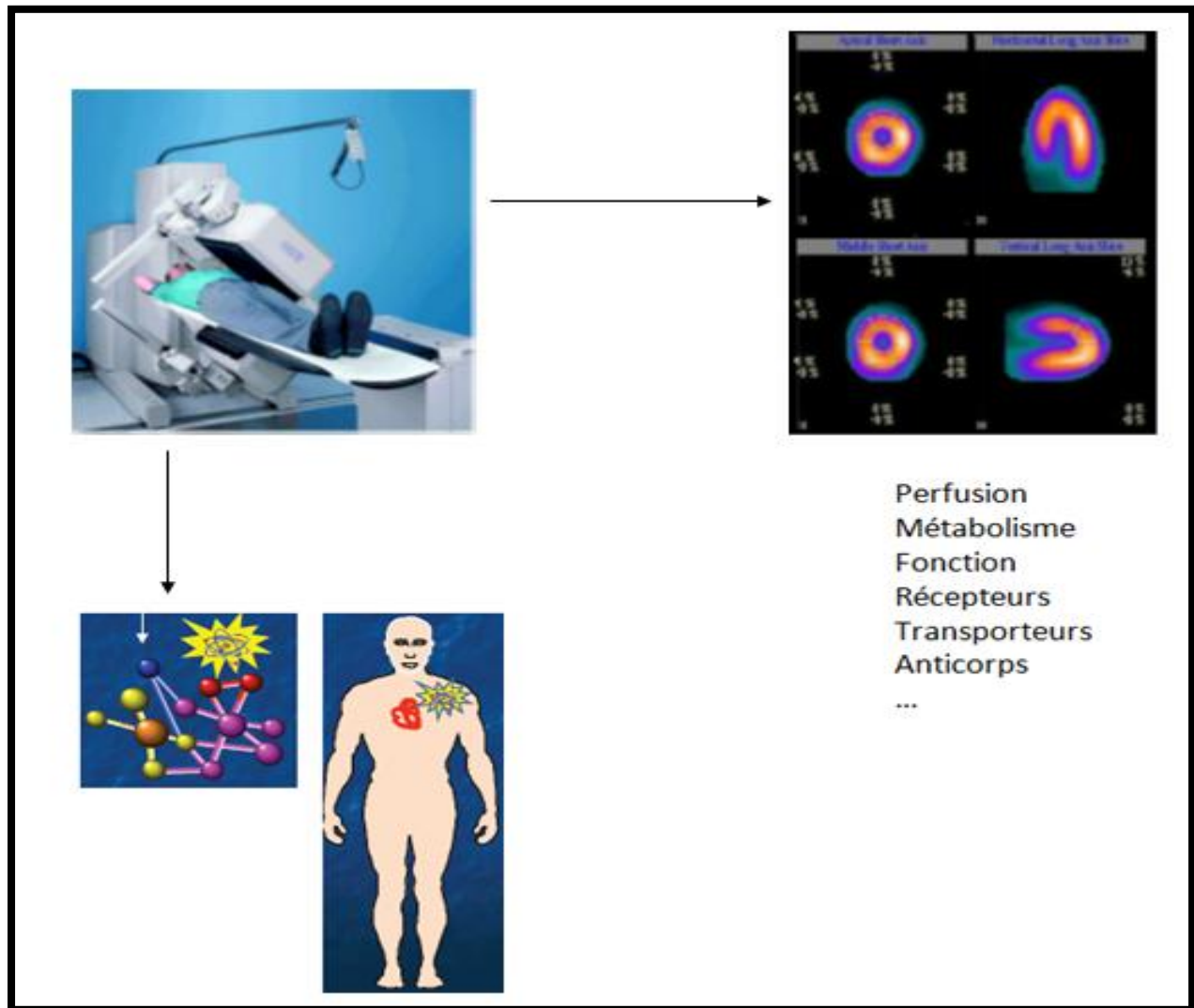


Figure 1.2 : principe de l'imagerie TEP

Radioactivité β^+ et détection des β^+

Les émetteurs de positons sont caractérisés par un excès de charge positive dans leurs noyaux. Ils se désintègrent vers un état stable, par une transformation d'un proton en un neutron qui conduit à l'émission d'un neutrino et d'un positon. Celui-ci est de masse égale à celle d'un électron mais de charge opposée. Une fois émis, le positon parcourt quelques millimètres dans les tissus, au cours desquels il perd toute son énergie cinétique. Quand le positon est au repos, il interagit avec un électron du milieu, suivant une réaction d'annihilation au cours de laquelle la masse des deux particules se transforme en deux photons gamma de 511 keV, émis dans des directions opposées (Fig. 1.3).

Le principe de base de la TEP consiste à détecter les deux photons de 511 keV pour déterminer le lieu de la réaction d'annihilation. L'information mesurée correspond au lieu d'annihilation et non à celui de l'émission β^+ . La distance entre ces deux lieux est appelée libre parcours moyen du positon, elle est déterminée par l'énergie d'émission des positons, et constitue une limite en termes de résolution

Spatiale. Le libre parcours moyen du positon est de 0,5 mm pour le ^{18}F et de 2,7 mm pour le ^{15}O .

Une autre limite intrinsèque, en termes de résolution spatiale en TEP, provient de la non-colinéarité des deux photons de 511 keV, résultant de l'annihilation d'un positon qui n'a pas intégralement perdu son énergie cinétique. Ce défaut de colinéarité peut être de l'ordre de $0,5^\circ$.

La combinaison de ces deux phénomènes constitue une limite intrinsèque en termes de résolution spatiale, propre à la détection en coïncidence, mais qui, pour le ^{18}F , est inférieure au millimètre. Cependant, la résolution spatiale d'une TEP dépend essentiellement de la taille des détecteurs élémentaires.

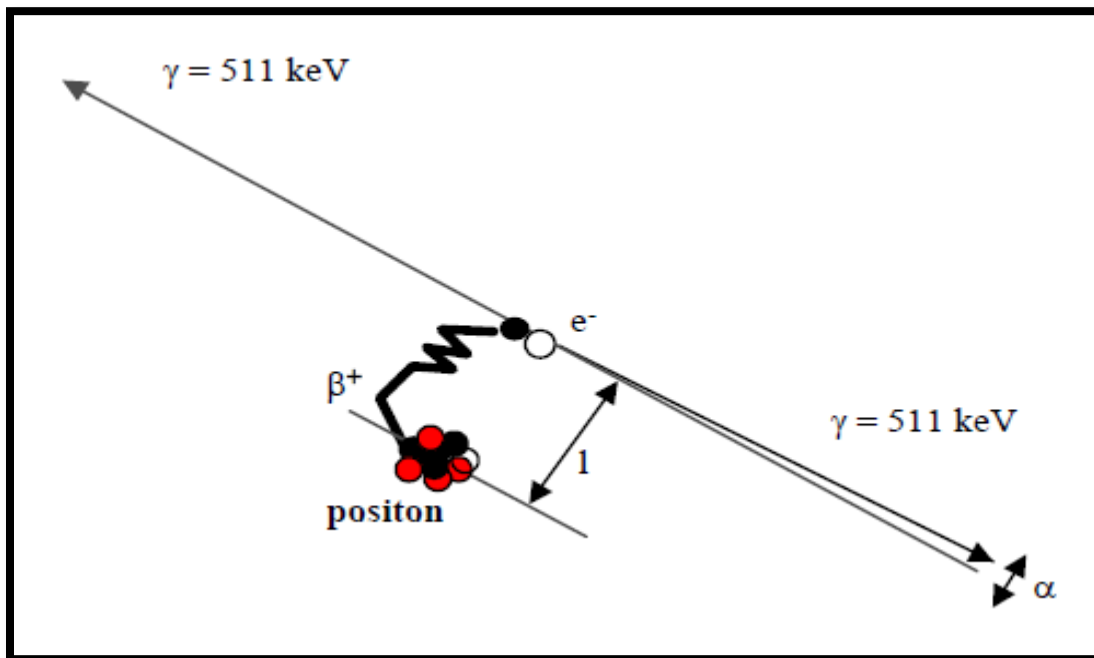


Figure 1.3 : Désintégration du positon et réaction d'annihilation. Une fois émis, le positon (+) parcourt quelques millimètres dans les tissus, durant lesquelles il perd toute son énergie cinétique. Quand le positon est pratiquement au repos, il interagit avec un électron (e^-) du milieu, suivant une réaction d'annihilation au cours de laquelle la masse des deux particules se transforme en deux photons gamma de 511 KeV, émis dans des directions opposées. α correspond à la non-colinéarité des deux photons γ émis.

L'obtention d'une image en TEP s'effectue en trois étapes selon la (fig. 1.4) :

- L'acquisition des données brutes et leurs corrections
- La mesure des projections de l'objet, en utilisant le principe de la détection en coïncidence
- La reconstruction tomographique, qui permet l'obtention de la distribution tridimensionnelle du traceur.

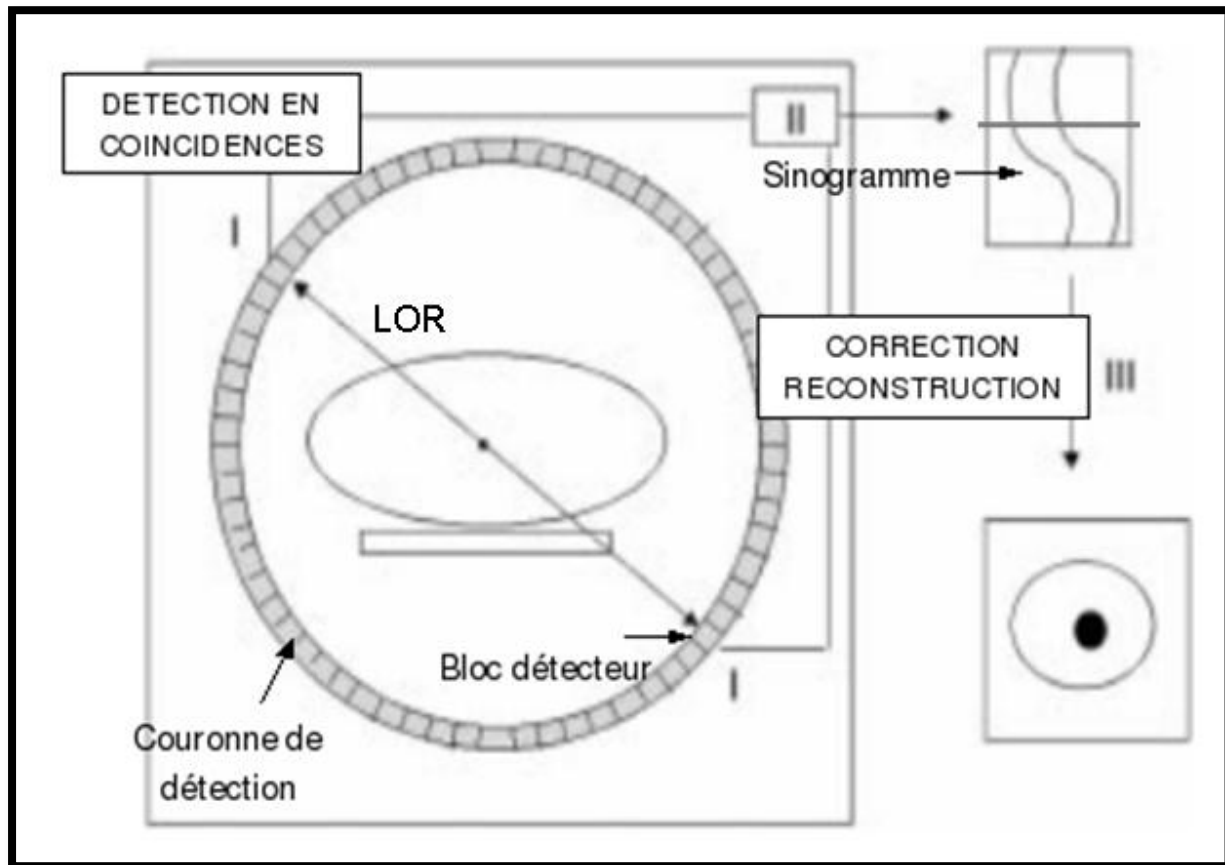


Figure 1.4 : Les différentes étapes de la création d'une image TEP

Il existe trois types d'acquisition TEP :

- ✓ L'acquisition en émission : le système de détection tourne autour du patient dont on veut extraire la distribution du radiotracteur injecté.
- ✓ L'acquisition en transmission comme l'illustre la Figure 1-5. Dans ce cas, l'acquisition est réalisée à partir d'une source externe, tournant autour du patient. Cette source peut être du germanium 68 (^{68}Ge), du césium 137 (^{137}Cs) ou encore des rayons X. Le temps de mesure en transmission est souvent inférieur au temps d'acquisition en émission. Cette acquisition permet de mesurer pour chaque LOR l'atténuation des photons dans le corps du patient afin de corriger ultérieurement les données de l'atténuation due aux événements atténués.
- ✓ L'acquisition « blanc » (Figure 1-5), elle est réalisée à vide, c'est-à-dire sans patient et suivant le même protocole que l'acquisition en transmission. Ces deux dernières acquisitions permettent de corriger les données de l'atténuation.

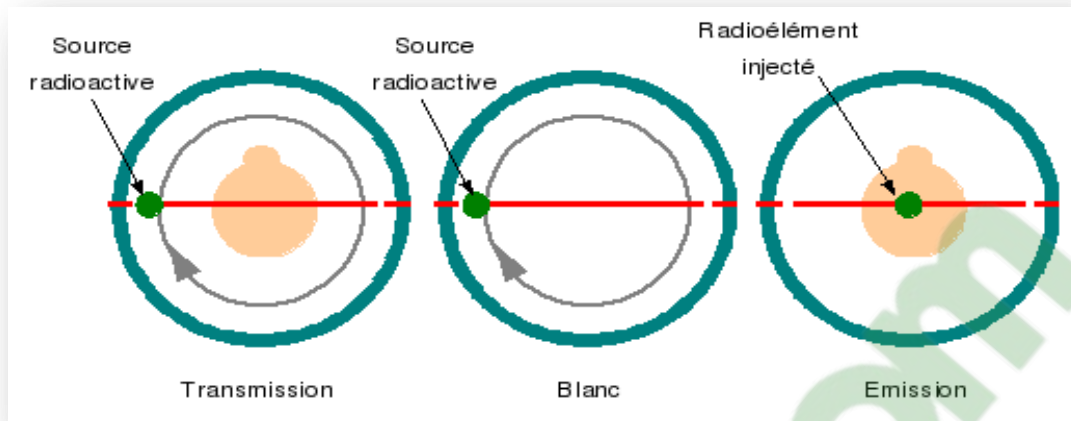


Figure 1.5 : Les trois types d'acquisition : transmission, blanc et émission

La plupart des caméras TEP sont constituées d'une série de couronnes de détecteurs élémentaires répartis en anneau autour du patient. Chaque détecteur est constitué d'un cristal scintillateur et d'un photomultiplicateur, optimisés pour la détection des photons gamma. Le circuit électronique, ou circuit de coïncidence, permet de déterminer au moyen de deux détecteurs élémentaires la projection du lieu d'annihilation. Le traitement de ces informations doit être extrêmement rapide, compte tenu des taux de comptage rencontrés en pratique. L'information mesurée par le circuit de coïncidence est alors positionnée dans une matrice, ou sinogramme, qui contient l'ensemble des éléments de projection d'une coupe de l'objet. Les images sont obtenues par reconstruction tomographique sur des stations informatiques puissantes équipées d'une application logicielle permettant l'analyse et la quantification des images.

Toute désintégration d'un positon se caractérise par l'émission de deux photons gamma de 511 keV dans des directions opposées. Cette propriété est intrinsèquement utilisée par les TEP pour déterminer la projection du lieu d'annihilation, au moyen de deux détecteurs indépendants qui mesurent quasi simultanément deux photons de 511 keV.

L'électronique de détection, ou circuit de coïncidence, est fondée sur deux critères : une fenêtre temporelle, de l'ordre de 4 à 20 ns, et une fenêtre en énergie. Ce système mesure les coïncidences vraies, qui correspondent à la détection des deux photons de 511 keV provenant d'une même annihilation (Fig. 1.6). En raison des imperfections de l'appareil de mesure en résolution en énergie et en résolution temporelle, le circuit de coïncidence enregistre de nombreux événements parasites qui génèrent du bruit de mesure.

Le rôle de chaque détecteur élémentaire consiste à arrêter le rayonnement, à déterminer l'énergie du photon incident et le moment où il a été détecté. La densité, la résolution en énergie et la résolution temporelle des détecteurs conditionnent les performances des différents instruments.

Les détecteurs sont constitués d'un ou plusieurs cristaux scintillateurs, couplés à un ou plusieurs tubes photomultiplicateurs (PM). Tout photon entrant dans la chaîne de détection est analysé individuellement (Fig. 1.8). Il interagit avec les électrons du cristal en cédant son énergie, soit en totalité par effet photoélectrique, soit en partie par effet Compton. Ces derniers vont exciter d'autres électrons du réseau cristallin, et générer des mécanismes de scintillation. L'énergie lumineuse ainsi créée est proportionnelle à l'énergie déposée dans le cristal. Elle est ensuite dirigée vers la photocathode du tube photomultiplicateur au travers d'un guide de lumière. Le rôle de la photocathode est de convertir l'énergie lumineuse en électrons. Ces derniers sont

accélérés et focalisés sur la première dynode, où le nombre d'électrons est multiplié par le facteur d'émission secondaire. Ce processus de multiplication est identique pour les dynodes suivantes. Le signal en sortie du PM fournit une impulsion électrique mesurable, dont l'intégrale est proportionnelle à l'énergie du photon déposée dans le cristal.

La durée d'intégration dépend de la rapidité de la décroissance de la lumière au sein du cristal.

Pendant cette durée, le détecteur élémentaire est dans l'impossibilité de mesurer un autre événement.

Ce phénomène, ou temps mort, est responsable des pertes en sensibilité pour des taux de comptage élevés.

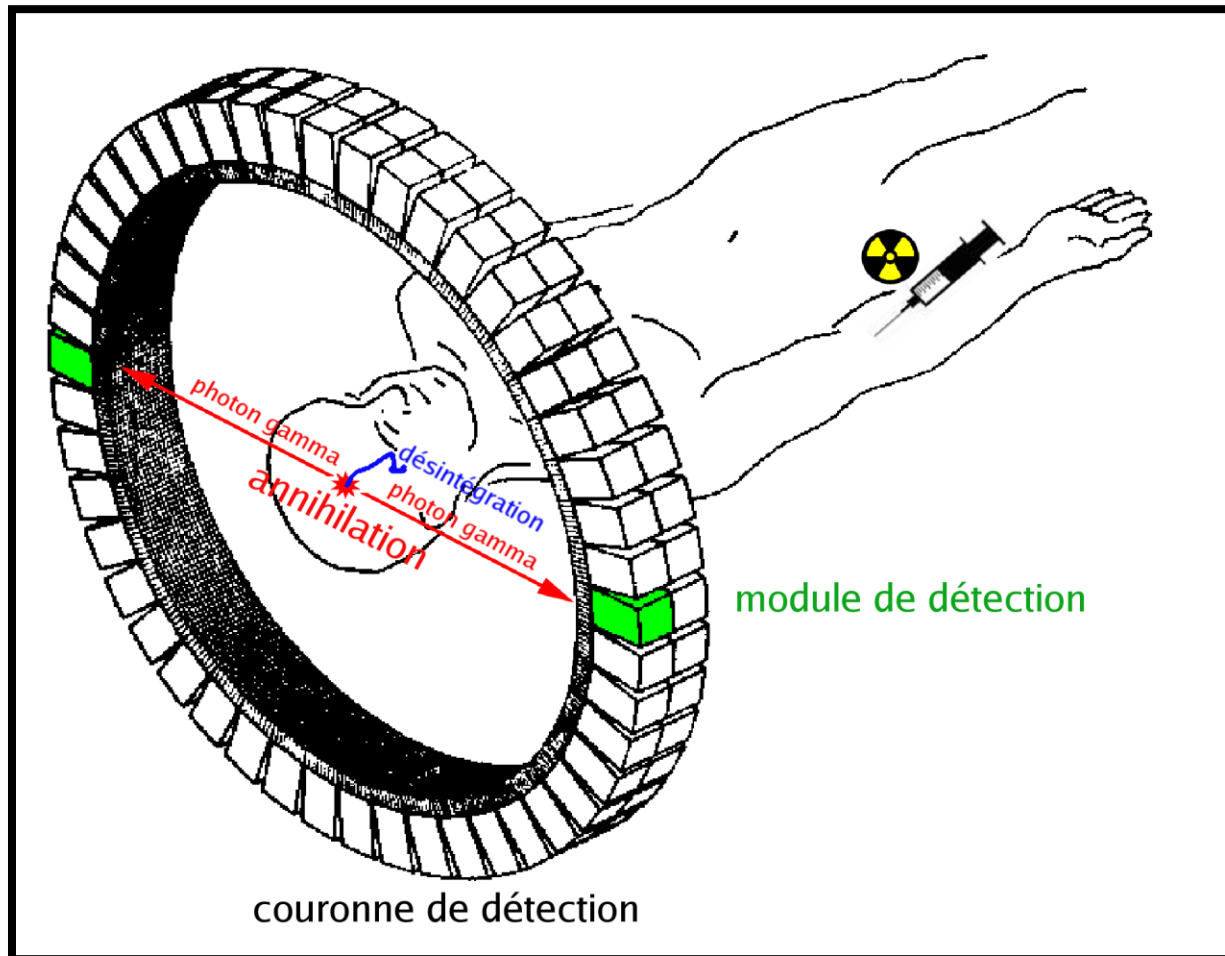


Figure 1.6 : Principe de la détection en coïncidence de la TEP

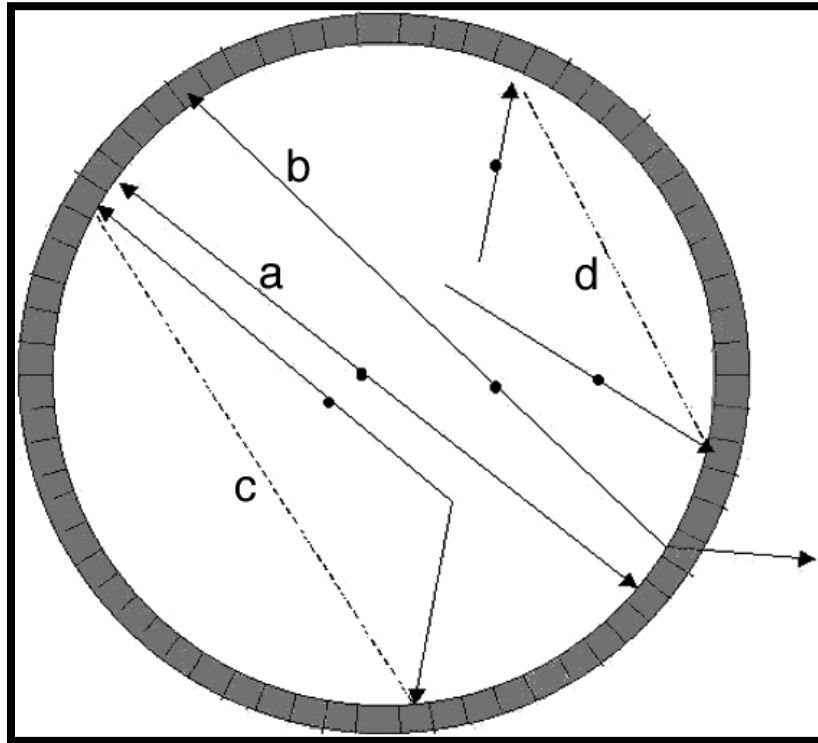


Figure 1.7 : Nature des événements enregistrés en mode coïncidence. La détection des deux photons γ de 511 keV résultant d'une annihilation correspond à une coïncidence vraie (annihilation a). L'absence de détection d'au moins un des deux photons entraîne l'absence de mesure d'une ligne de réponse, il en résulte une plus faible efficacité de détection (annihilation b). Du fait de la résolution en énergie limitée des scintillateurs, une fraction des photons ayant subi une diffusion Compton est acceptée, il en résulte l'enregistrement de coïncidences diffusées (annihilation c). La détection de deux photons c de 511 keV dans la fenêtre temporelle de coïncidence résultant de deux annihilations distinctes entraîne la mesure d'une coïncidence aléatoire (événement d). Les coïncidences aléatoires et diffusées sont des sources de bruit qu'il convient de limiter et de corriger.

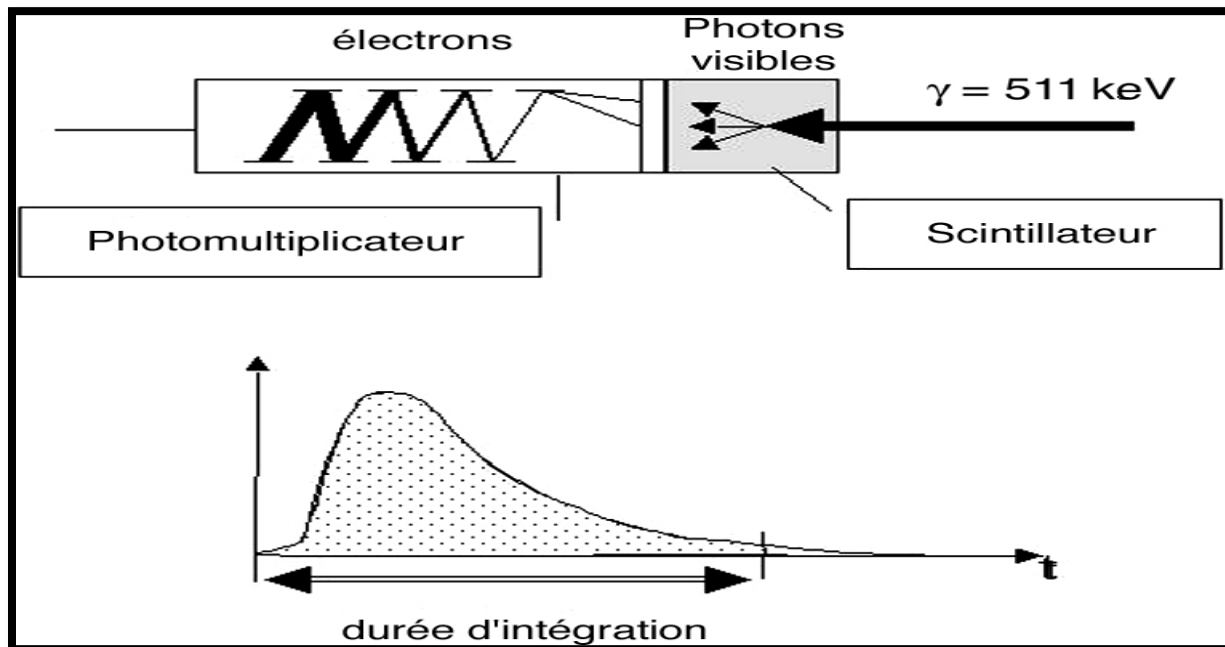


Figure 1.8 : Principe de fonctionnement d'un détecteur élémentaire

Choix du scintillateur

Le scintillateur idéal pour la TEP doit être caractérisé par :

- une densité élevée, pour offrir une efficacité de détection satisfaisante pour les photons de 511 keV (énergie trois à quatre fois supérieure à celle des gammas du technétium ^{99m} largement utilisé en médecine nucléaire) ;
- une efficacité de conversion lumineuse importante, c'est-à-dire un nombre de photons de scintillation important par rayonnement gamma détecté. Ce paramètre influe directement sur la résolution en énergie du scintillateur, et donc sur le taux de coïncidences diffusées ;
- une faible constante de décroissance, c'est-à-dire l'émission d'un grand nombre de photons lumineux par unité de temps ; elle conditionne la résolution temporelle du système de détection et réduit le temps mort relatif à chaque mesure ;
- une bonne transparence par rapport aux photons de scintillation, afin que ceux-ci ne soient pas atténués dans le milieu avant la surface d'entrée du PM. Par ailleurs, le spectre d'émission des photons de scintillation doit être compatible avec la réponse spectrale de la photocathode du PM ;
- des propriétés physiques favorables et, en particulier, il est préférable que les scintillateurs soient non hygroscopiques, insensibles à la température et facilement produits et usinés. Par ailleurs, le coût des matériaux influe directement sur le prix des appareils proposés.

Caractéristiques d'atténuation

Les paramètres d'atténuation du rayonnement de 511 keV dépendent de la densité du matériau et de la densité d'électrons estimée par le numéro atomique effectif (Z_{eff}). L'épaisseur du scintillateur peut être d'autant moins importante que le coefficient d'atténuation (μ) est élevé. Cette propriété permet de diminuer la taille des détecteurs élémentaires, et donc d'améliorer la résolution spatiale du système. La photofraction est le pourcentage d'effet photoélectrique par rapport aux interactions des photons de 511 keV dans le scintillateur.

Ce paramètre est essentiel, puisque seules les interactions photoélectriques sont caractérisées par un dépôt d'énergie de 511 keV dans le détecteur (annihilation (a) de la Fig. 1.4). Lorsque le photon incident subit une diffusion Compton dans le scintillateur, il en résulte un dépôt d'énergie inférieur à celui obtenu par effet photoélectrique (annihilation (b) de la Fig. 1.4). Cet événement est éliminé par la chaîne de mesure ; il en résulte une perte en efficacité. Parmi les différents scintillateurs utilisés en TEP, le germanate de bismuth (BGO) est l'un des matériaux qui possède le plus grand Z_{eff} . L'orthosilicate de lutétium (LSO) et l'orthosilicate de gadolinium (GSO) sont légèrement moins denses que le BGO. L'iodure de sodium dopé au thallium (Na [Tl]) est le scintillateur le moins dense, il est cependant utilisé pour certaines caméras TEP en raison de son faible coût de production.

Le *lutetium aluminium perovskite* (LuAp) est un scintillateur en développement, prometteur puisqu'il est caractérisé notamment par une densité élevée.

La reconstruction tomographique des images TEP

Quelle que soit la conception des différents instruments, la mesure des données correspond systématiquement à la projection du lieu des annihilations survenant au sein de l'objet dont on fait l'image. Une étape de reconstruction tomographique est nécessaire pour déterminer l'objet à

partir de ses projections. Les méthodes utilisées en TEP sont, d'un point de vue mathématique, similaires à celles de la TDM ou de la tomographie d'émission réalisée avec des émetteurs de photons gamma.

Il existe deux grandes familles d'algorithmes : la reconstruction analytique et la reconstruction itérative.

- Reconstruction analytique : elle est rapide et efficace pour des données peu bruitées (CT)
- Reconstruction itérative : elle est lente mais elle est efficace pour des données qui sont bruitées (TEP, SPECT)

Exemple de reconstruction tomographique :

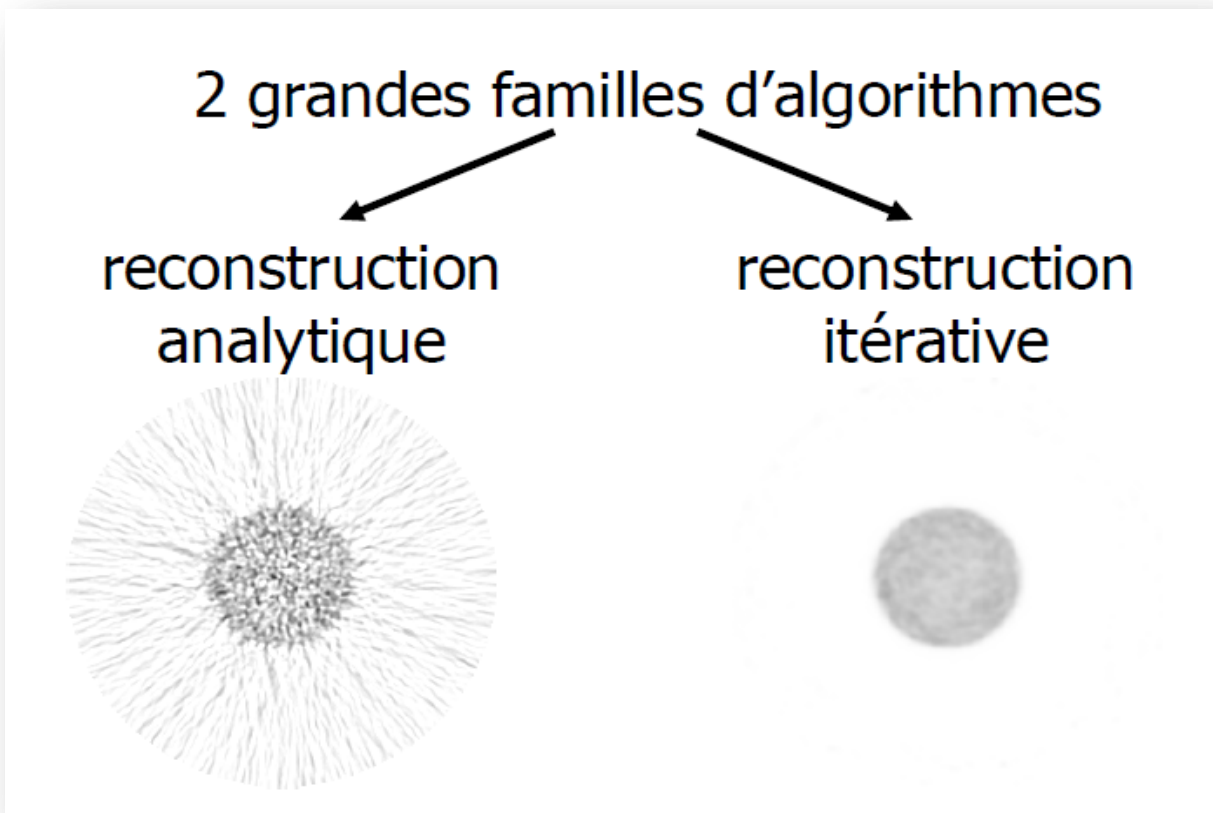


Figure 1.9 : reconstructions analytique et itérative

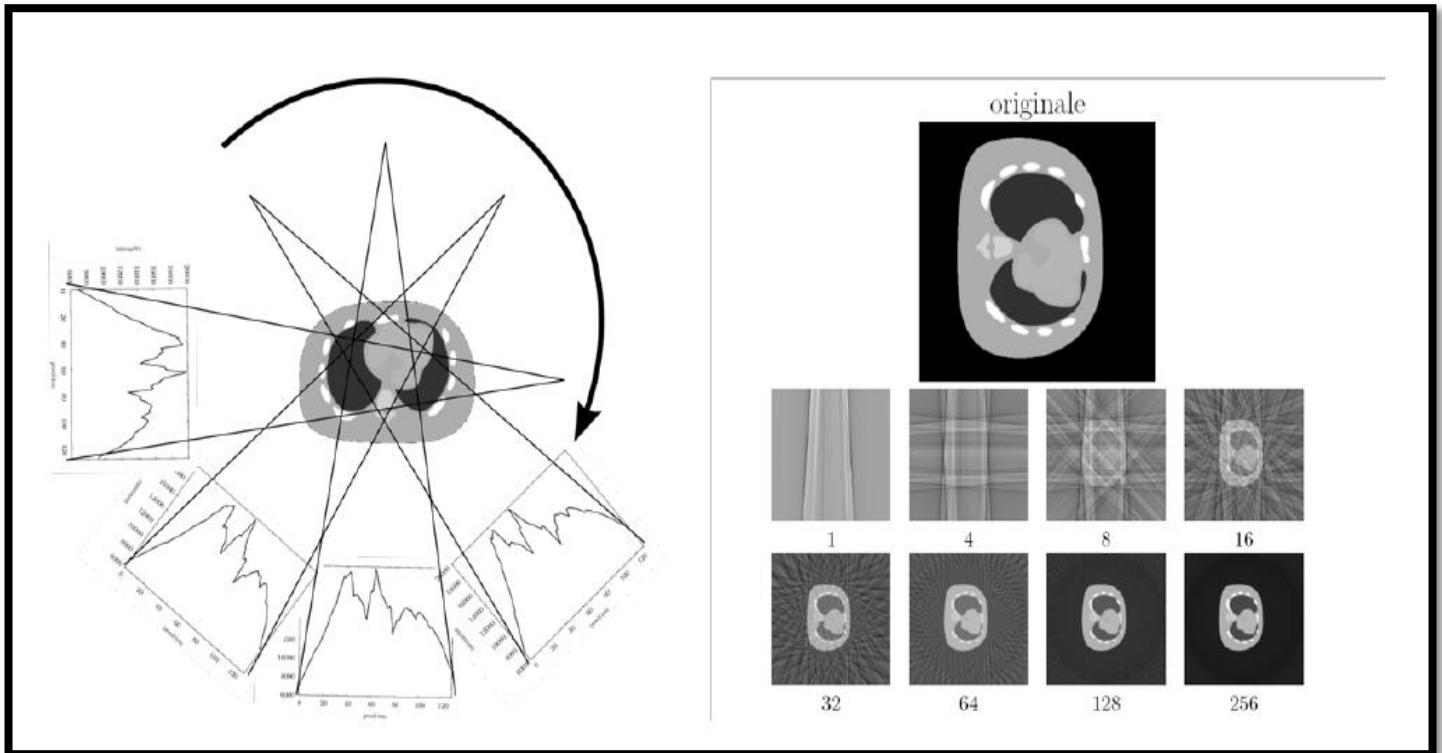


Figure 1.10 : Exemple de reconstruction tomographique

1.4 Utilisation de la TEP en neurologie

1.4.1 Introduction

Situé à l'intérieur du crâne, le cerveau gère et coordonne la plupart de nos fonctions :

- les fonctions internes du corps (battements du cœur, circulation du sang, contractions des muscles ou digestion),
- les fonctions dites supérieures (pensée, émotions, personnalité, capacité à communiquer ou à apprendre),
- et enfin, les cinq sens, qui nous permettent d'être en relation avec l'extérieur : la vue, l'ouïe, le toucher, l'odorat et le goût.

Le cerveau fonctionne comme un ordinateur central. Il reçoit en permanence une multitude d'informations sur ce qui se passe à l'intérieur et à l'extérieur du corps. Ces informations lui proviennent des nerfs rachidiens, qui parcourent l'ensemble du corps, et des nerfs crâniens, situés à l'intérieur du crâne. Chaque information est classée, analysée et traitée par des cellules nerveuses, qui sont appelés : « les neurones ». Ces derniers sont connectés entre eux, et sont capables de gérer plusieurs informations à la fois. C'est ce qui nous permet de voir, de parler, de marcher et de respirer, tout cela en même temps et sans avoir à y réfléchir. Ces milliards de neurones sont entourés et nourris par d'autres cellules appelées cellules gliales,

qui sont dix fois plus nombreuses que les neurones. Elles représentent 90 % de la composition du cerveau. Il en existe différents types : les astrocytes, les oligodendrocytes, les épendymocytes et les cellules microgliales.

Un peu plus de la moitié des tumeurs du cerveau se développent à partir de ces cellules. Ce type de tumeur est appelé gliome.

1.4.2 Petit rappel sur l'anatomie du cerveau

Le cerveau humain est composé de cellules nerveuses, les neurones (environ 10^{12}), qui échangent des informations par le biais de multiples fibres de connexions (les axones et dendrites) et de terminaisons (les synapses) existant entre elles (environ 10^4 par neurone). Lorsque le cerveau exécute une fonction, qu'il s'agisse d'actes moteurs ou de processus plus élaborés comme le langage ou la mémoire, certains groupes de neurones modifient leur dynamique d'échange d'informations. Ces modifications, qui peuvent être des augmentations, des diminutions voire des synchronisations avec d'autres groupes de neurones, sont rarement localisées en un endroit du cortex ; le plus souvent, elles sont distribuées dans un large réseau de régions cérébrales.

Les neurones génèrent des informations qui sont des séries de courants électriques, appelés « potentiels » d'action (PA), lesquels se propagent le long des fibres neuronales jusqu'aux synapses. A ce niveau, l'information électrique est convertie en information neurochimique : les potentiels d'action libèrent dans la fente synaptique des molécules particulières, les neurotransmetteurs. En se liant aux récepteurs du neurone postsynaptique, ces neurotransmetteurs engendrent soit une dépolarisation de sa membrane (potentiel postsynaptique excitateur ou PPSE), soit une surpolarisation (potentiel postsynaptique inhibiteur ou PPSI). Suivant l'amplitude de la somme de ces PPSE et PPSI, la génération d'un potentiel d'action au niveau du neurone postsynaptique est déclenchée ou inhibée.

1.4.2.1 Les grandes régions anatomiques : face externe

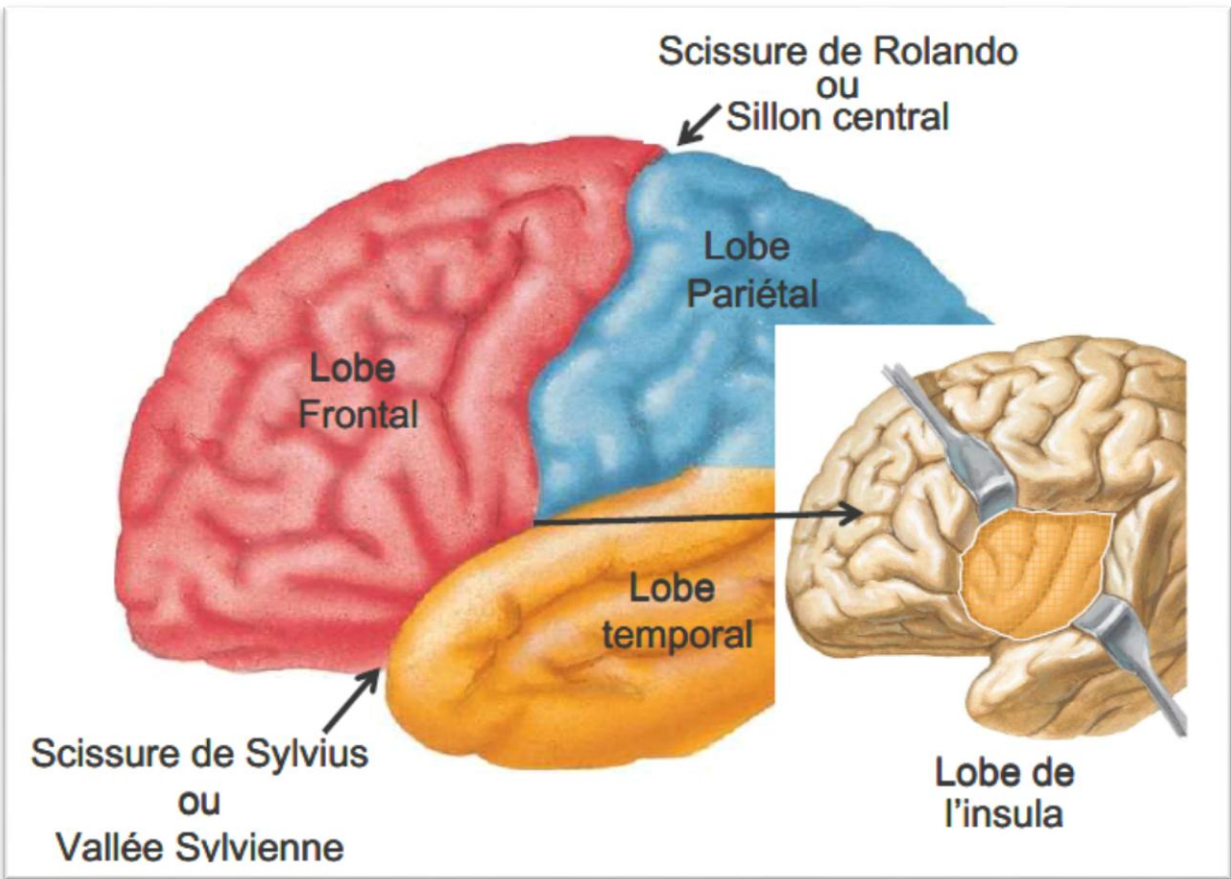


Figure 1.11 : Face externe des grandes régions anatomiques

1.4.2.2 Spécialisation du cortex

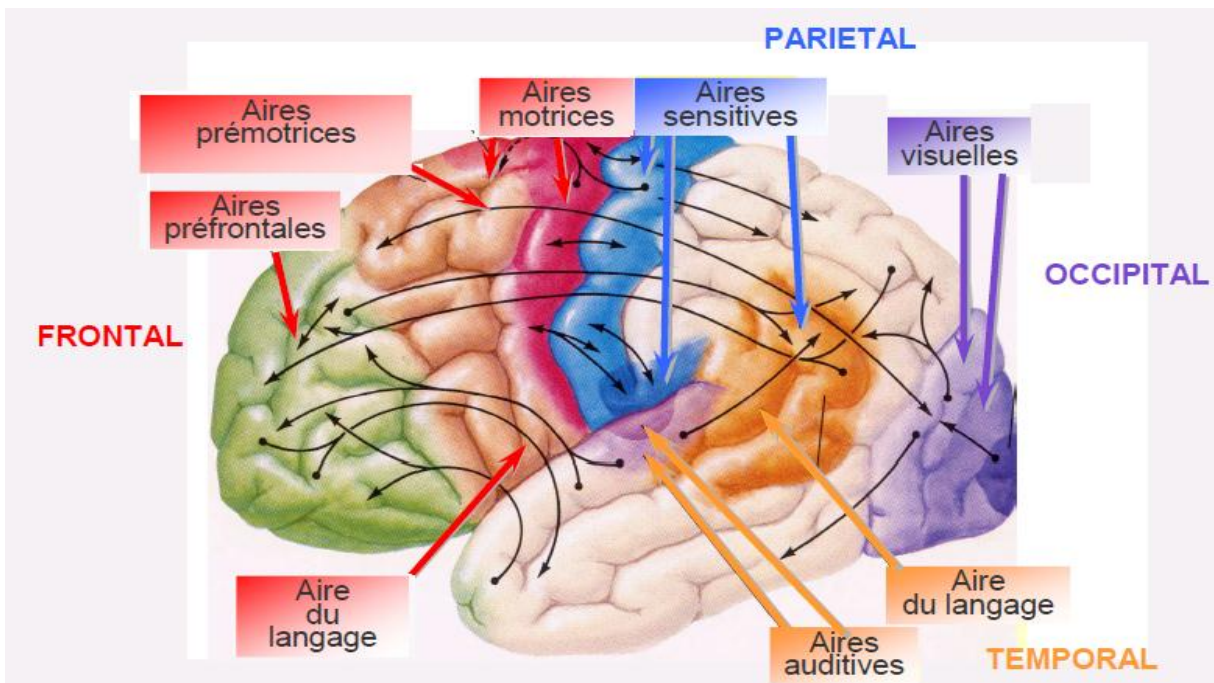


Figure 1.12 : Spécialisation du cortex

- **Aires primaires**
 - Premiers relais corticaux de l'information sensorielle (réception)
 - Assurent les fonctions élémentaires

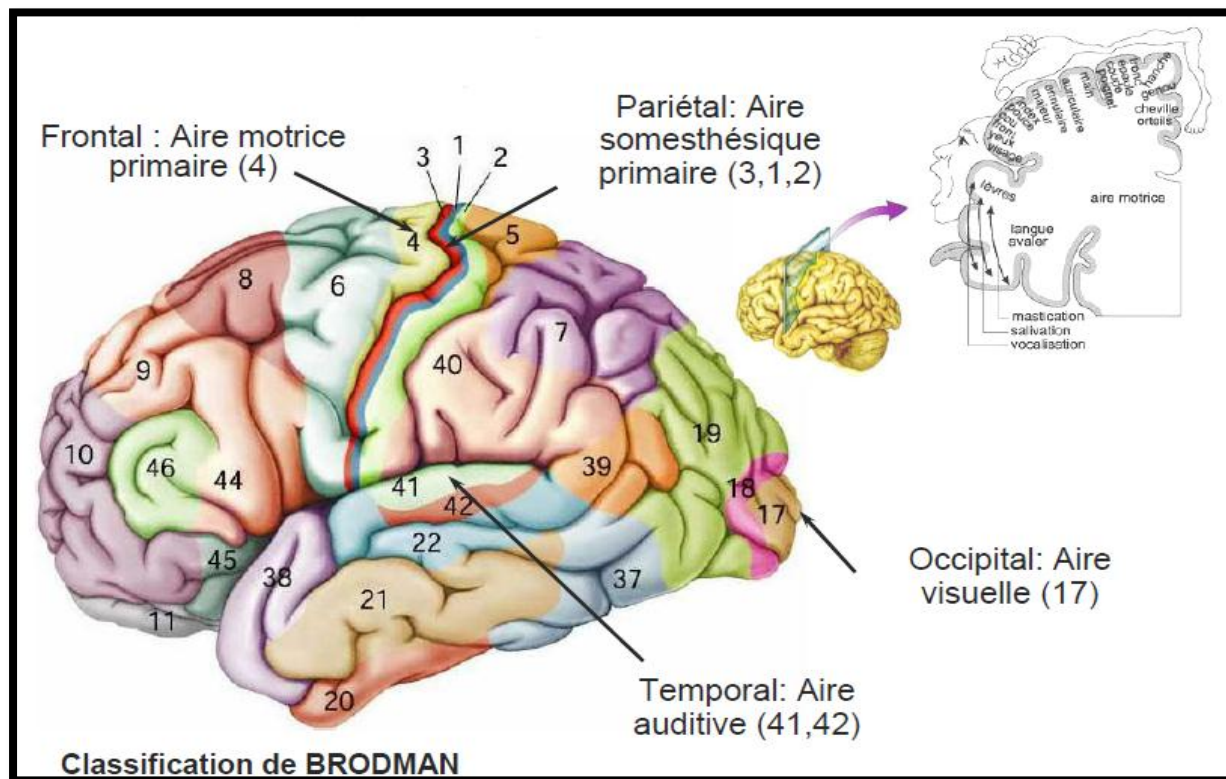


Figure 1.13 : Aires primaires

- **Aires associatives** : fonctions complexes
 - Unimodales (perception) :

Traitent l'information d'une aire primaire ou d'autres aires unimodales du même type

- Hétéromodales (interprétation) :

Traitent les informations provenant d'autres aires unimodales ou hétéromodales, quelque soit leur type.

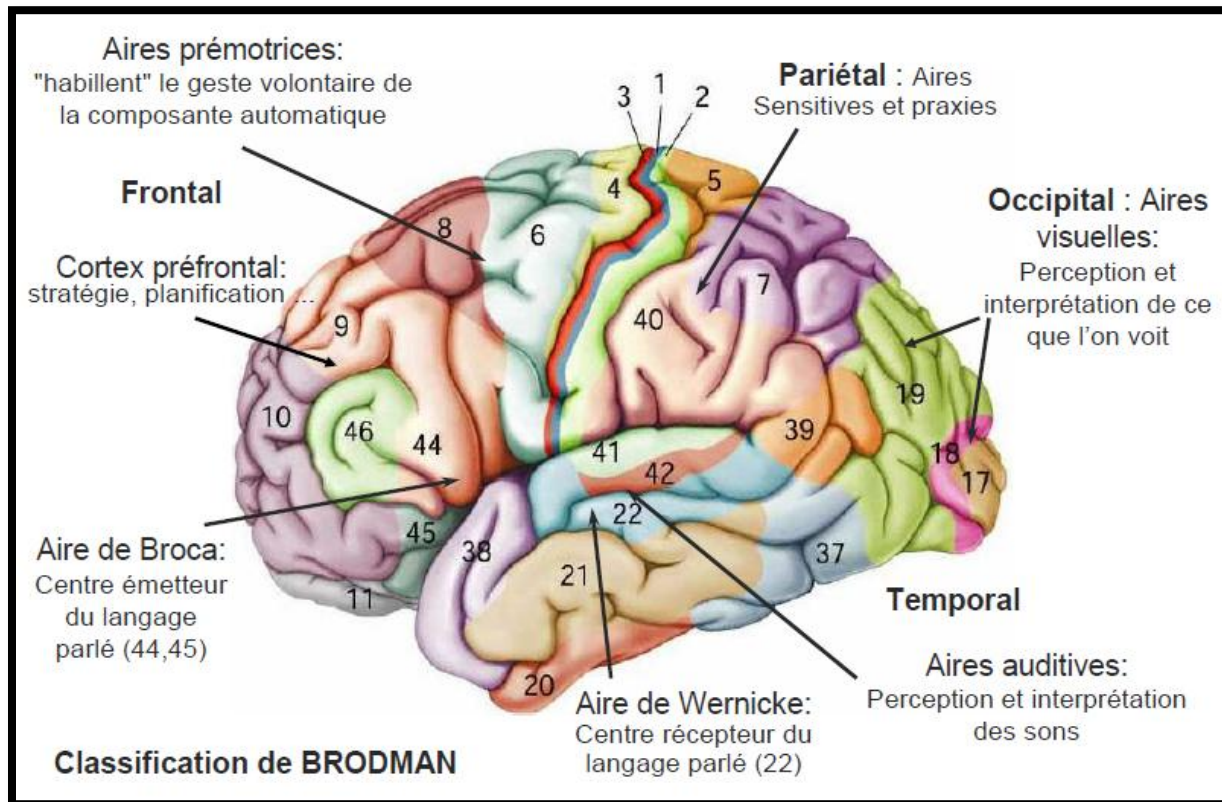


Figure 1.14 : Aires associatives

1.4.3 Quelques traceurs radioactifs pour la neurologie en TEP

Le traceur doit obéir à de nombreuses contraintes :

- il doit être spécifique de la fonction que l'on trace, c'est à dire s'accumuler le plus possible là où cette fonction est présente et avoir moins d'affinité là où elle ne l'est pas ;
- il doit également avoir une cinétique compatible avec la durée d'immobilisation du patient et de l'examen ;
- s'il est l'analogue d'une molécule naturellement présente dans l'organisme, le marquage ne doit pas modifier le comportement chimique que l'on cherche à étudier.

Le tableau 1.1 montre quelques traceurs utilisés en TEP.

Paramètre étudié	Traceurs	Demi-vie
Métabolisme cérébral	18F- déoxyglucose	110 min
Débit sanguin cérébral	15O-eau	2 min
Synthèse dopamine	18F-dopa	10 min
Transporteurs dopamine	11C-PE21	20 min
Récepteurs benzodiazépines	11C-flumazénil	20 min

Tableau 1.1 : quelques traceurs utilisés en TEP, leur paramètre étudié ainsi que leur demi-vie

Un radio-élément (tel que ^{18}F , ^{15}O , ^{11}C , ^{13}N), est classiquement produit dans un *cyclotron*, accélérateur de particules de type circulaire. Le cyclotron est composé d'un gros aimant circulaire ou règne un important champ magnétique (appliqué perpendiculairement dans une chambre vide) qui fait tourner les particules, et un champ électrique alternatif qui accélère leur mouvement à chaque tour : elles décrivent des cercles de plus en plus grands, suivant une trajectoire en forme de spirale. Lorsque le rayon de la trajectoire atteint la limite extérieure de l'aimant, les particules sont éjectées, guidées et localisées jusqu'à une cible où elles sont utilisées pour la recherche. L'accélération des particules peut aller, selon la taille de l'aimant, de quelques dizaines de MeV (millions électronvolts) à plusieurs centaines de GeV (giga électronvolts) dans le cas des synchrotrons, sorte de cyclotrons géants.

1.4. 4 Les tumeurs du cerveau

Depuis les années 1990, l'utilisation de la TEP dans les tumeurs cérébrales est devenue la norme de diligence dans de nombreux centres médicaux. La valeur de la TEP-FDG dans l'évaluation des tumeurs cérébrales, est la corrélation entre la fixation du FDG et le grade tumoral. Du point de vue physiologique, le FDG est faible dans la substance blanche, et élevé dans la matière grise. La figure 1.15 montre un scanner du cerveau FDG-PET normal.

Lors de l'évaluation de l'activité métabolique d'une tumeur cérébrale, la comparaison de la fixation du FDG au sein de la tumeur du cerveau controlatérale permet une évaluation relativement simple et rapide du grade. Les tumeurs de bas grade ont une accumulation du FDG similaire ou inférieure à la substance blanche normale, et les tumeurs de haut grade ont une activité qui est similaire ou supérieure à la matière grise normale.

La fixation du FDG dans les tumeurs à faible et à haute teneur par rapport aux deux matières blanche et grise normales, a également été étudiée quantitativement. En 1995, Delbeke et al. ont étudié 58 patients atteints de tumeurs cérébrales ; 32 ont fait une biopsie qui a révélé une tumeur de haut grade (stade III et IV), et 26 étaient à un faible grade de la tumeur (stade I et II). Les régions d'intérêt ont été utilisées pour attribuer une tumeur à la substance blanche (T / WM) et une tumeur à la matière grise (T / GM), ratios dans un effort pour déterminer une valeur seuil susceptible de faire la distinction entre le bas et le haut grade. Ils ont constaté que le rapport T / WM est supérieur à 1,5, et T / GM est supérieur à 0,6, ils sont indicatif de maladie de haut grade avec une sensibilité de 94% et une spécificité de 77%. Ces résultats soutiennent l'approche qualitative décrite ci-dessus dans laquelle la fixation du FDG dans les tumeurs de bas grade ressemble à la substance blanche et la consommation des tumeurs de haut grade ressemble à la matière grise.

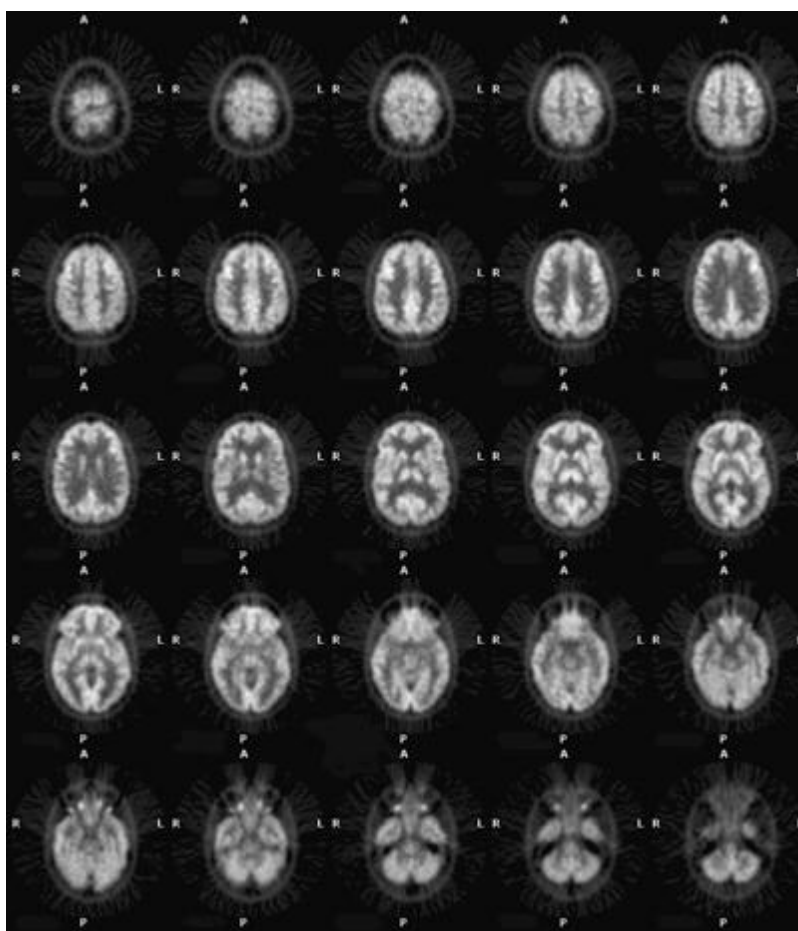


Figure 1.15 : scanner normal du cerveau au FDG-PET. Notez l'activité accrue de la matière grise, à la fois dans corticale et les structures profondes de gris, par rapport à la substance blanche relativement hypométabolique. Dans ce patient, vous pouvez également voir la physiologie du FDG dans les muscles extraoculaires.

Les tumeurs gliales sont des pathologies hétérogènes. Cela a été quantifié lorsque Paulus et Peiffer ont étudié les caractéristiques histologiques de 1000 échantillons provenant de 50 tumeurs cérébrales (20 échantillons par tumeur afin de simuler des biopsies multiples). Ils ont observé que les différents grades ont été détectés dans 82% des tumeurs, et une majorité (62%) des gliomes sont contenus à faible et haut grade (OMS grades II, III et IV). C'est ce qui explique l'hétérogénéité de l'erreur d'échantillonnage et la sous-mise en scène souvent rencontrés dans la gestion des tumeurs cérébrales. L'émergence de la maladie de haut grade au sein d'une lésion préalablement à faible teneur ne peut pas être détectée par IRM, mais avec la TEP-FDG, les zones hypermétaboliques sont plus susceptibles d'avoir le plus haut grade et peut être spécifiquement ciblées pour la biopsie stéréotaxique, améliorant ainsi les chances de guérison de la tumeur.

Il existe des dizaines de tumeurs cérébrales différentes, que l'on distingue en fonction de 3 caractéristiques :

- La localisation de la tumeur
- Le type de tumeur

- Le degré d'agressivité de la tumeur

1. La localisation de la tumeur

Le cerveau est organisé en plusieurs zones, qui gèrent chacune des activités spécifiques : le langage, l'équilibre du corps, les battements du cœur, la circulation du sang, la mémoire, etc. Une tumeur peut entraîner des troubles très différents selon la zone dans laquelle elle se développe.

La localisation de la tumeur est également un élément essentiel pour le choix des traitements. Une tumeur située à la surface du cerveau par exemple, est généralement plus facile à extraire qu'une tumeur située au centre du cerveau.

2. Le type de tumeur

Le cerveau est composé de différents types de cellules, qui peuvent chacune être à l'origine de tumeurs différentes.

Les tumeurs du cerveau portent généralement le nom des cellules à partir desquelles elles se développent : les gliomes se développent à partir des cellules gliales, qui nourrissent et soutiennent les neurones ; les méningiomes se développent à partir des cellules composant les méninges (enveloppes du cerveau), etc.

Selon leur type (les médecins parlent de type histologique), les tumeurs ne se comportent pas de la même manière. Certaines se développent plus vite que d'autres ou ont un risque accru de récurrence par exemple.

3. Le degré d'agressivité

Lorsqu'une tumeur du cerveau est découverte, il est indispensable de déterminer ses caractéristiques : sa localisation, son type et son degré d'agressivité. Pour cela, plusieurs examens doivent être réalisés. Plus la tumeur se développe rapidement, plus elle est jugée agressive. On parle de bas grade pour les tumeurs les moins agressives et de haut grade pour les tumeurs qui le sont davantage. Le grade peut aussi être exprimé par un chiffre romain allant de I à IV. Connaître le grade d'une tumeur est déterminant dans le choix des traitements et dans le pronostic. Plus le grade est bas, plus le pronostic est favorable.

Contrairement aux autres cancers, les tumeurs cancéreuses du cerveau n'entraînent pas de métastases à l'extérieur du cerveau. Elles ne s'étendent pas à d'autres organes.

1.4.4 Les démences

La démence est une condition de la détérioration des facultés mentales, en particulier la mémoire et la fonction cognitive, et peut également être associée à des changements de personnalité comme la paranoïa, l'agressivité ou la dépression. Le type le plus commun de la démence est la maladie d'Alzheimer (MA). Environ 70% des patients atteints de démence ont une maladie d'Alzheimer.

Le diagnostic de la démence s'effectue par :

- ✚ Altération de la mémoire à court et à long terme

- ✚ Et au moins l'une des manifestations suivantes :
 - Altération du raisonnement
 - Altération du jugement
 - Aphasie, apraxie, agnosie.
- ✚ Retentissement professionnel, social, et/ou relationnel indépendamment de toute confusion ou dépression.

La classification des démences est comme suit :

- ❖ Démences dégénératives (idiopathiques)
 - Maladie d'Alzheimer
 - Atrophies lobaires (incluant les démences fronto-temporales)
 - Démences associées à un syndrome parkinsonien.
- ❖ Démences secondaires
 - Démence vasculaire : accidents vasculaires multiples
 - Causes toxiques (alcool), métaboliques, nutritionnelles
 - Causes infectieuses (Creutzfeld-Jakob)
 - Tumeurs, traumatisme crânien.

1.4.6 Maladie d'Alzheimer

En 1907, Alois Alzheimer décrit « une maladie particulière du cortex cérébral ».

La MA est une maladie dégénérative du tissu cérébral qui entraîne de façon irréversible la perte de fonctions cérébrales, la prévalence augmente avec l'âge (10% après 65 ans et 47% après 85 ans).

La physiopathologie : 2 marqueurs histologiques associés à des dépôts de protéines

- Plaques séniles : dépôt de peptide bêta-amyloïde
- Dégénérescence neuro-fibrillaire : constituée de dépôts de protéines taux anormale
- Augmentation du stress oxydatif → mort neuronale

La MA est une démence progressive, dont la 1^{ère} étape c'est l'atteinte progressive de la mémoire hippocampique, puis l'apparition de :

- ✚ Troubles instrumentaux : « Aphasie, Apraxie, Agnosie »
 - Apraxie de l'habillage, constructive
 - Manque du mot, trouble de compréhension
 - Troubles visuo-spatiaux avec désorientation temporo-spatiale
- ✚ Troubles des fonctions exécutives (frontales)
 - Troubles de la planification, et du jugement
- ✚ Modification de la personnalité et du comportement
- ✚ Perte d'autonomie

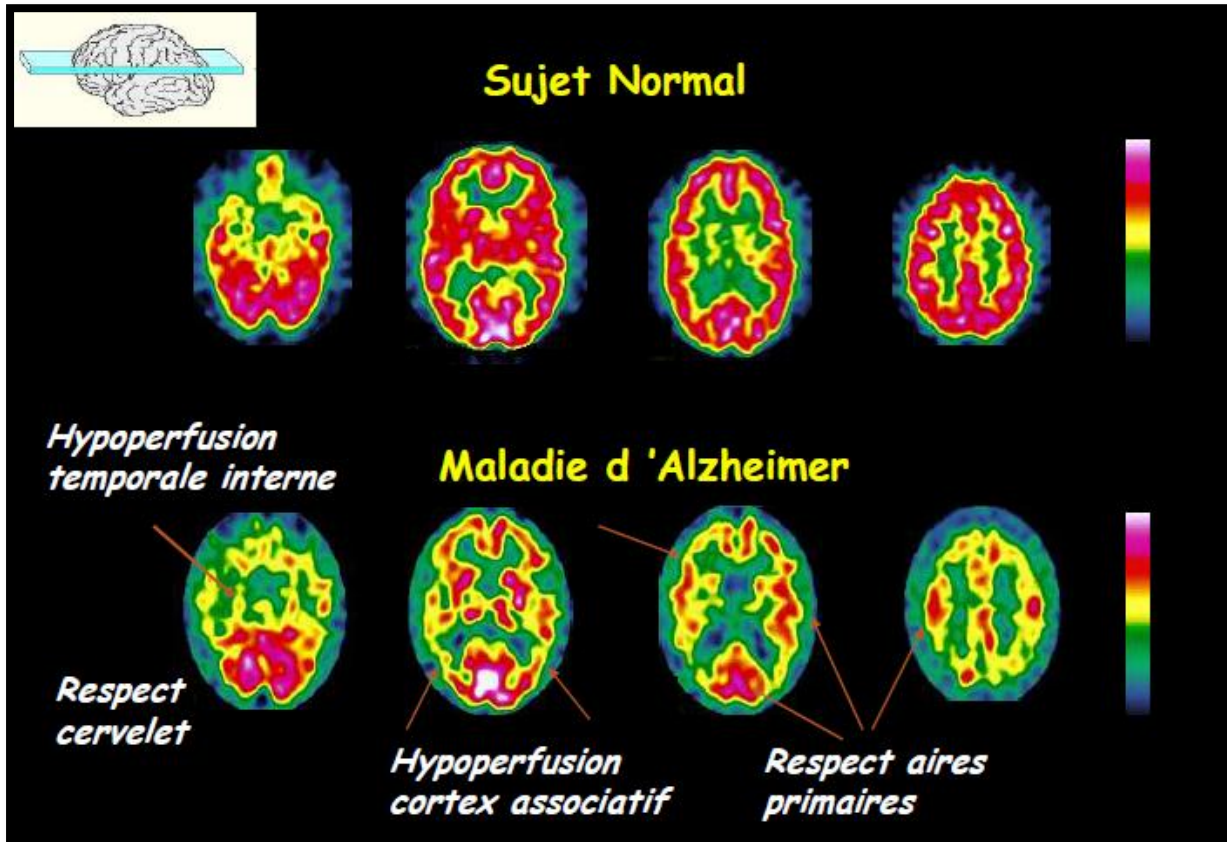


Figure 1.16 : comparaison entre un sujet sain et un sujet atteint de la MA

1.4.7 Maladie de Parkinson

Elle touche 1% de la population de plus de 55ans.

C'est une dégénérescence dopaminergique nigro-striatale. On remarque un déficit en dopamine striatale.

C'est aussi un handicap moteur :

- Akinésie
- Hypertonie
- Tremblement ± troubles cognitifs

Le striatum est formé du noyau caudé plus le putamen, comme illustré sur la figure 1.17

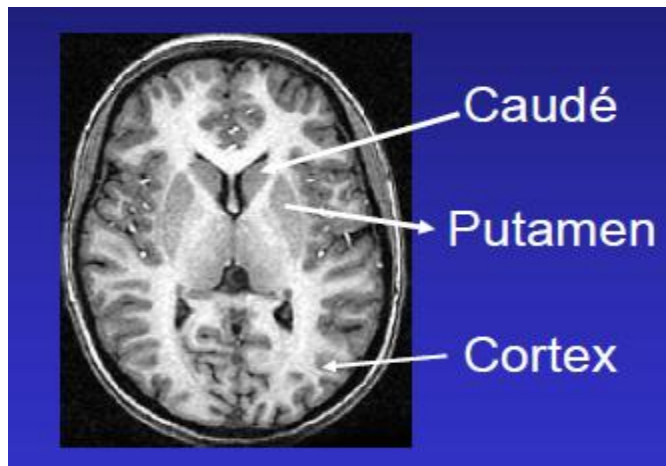


Figure 1.17 : Le striatum

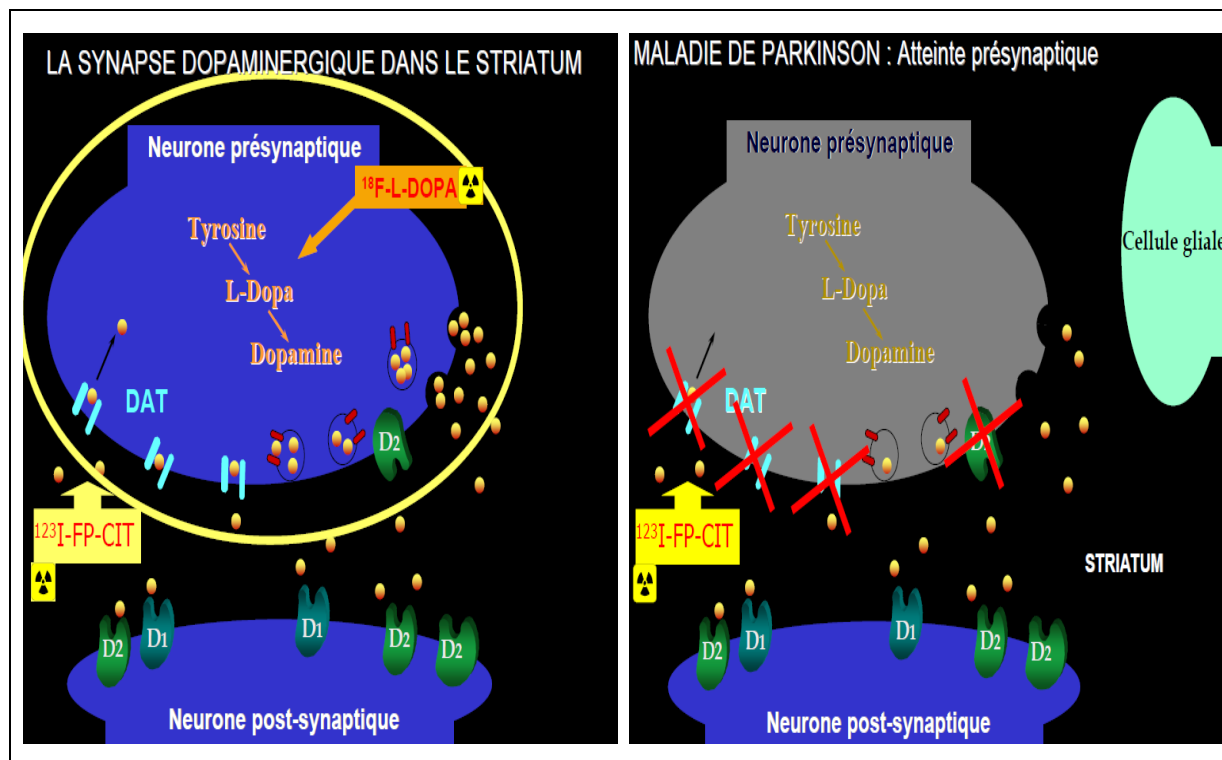


Figure 1.18 : La formation de la synapse dopaminergique dans le striatum chez un sujet sain, et un patient atteint de la maladie de Parkinson.

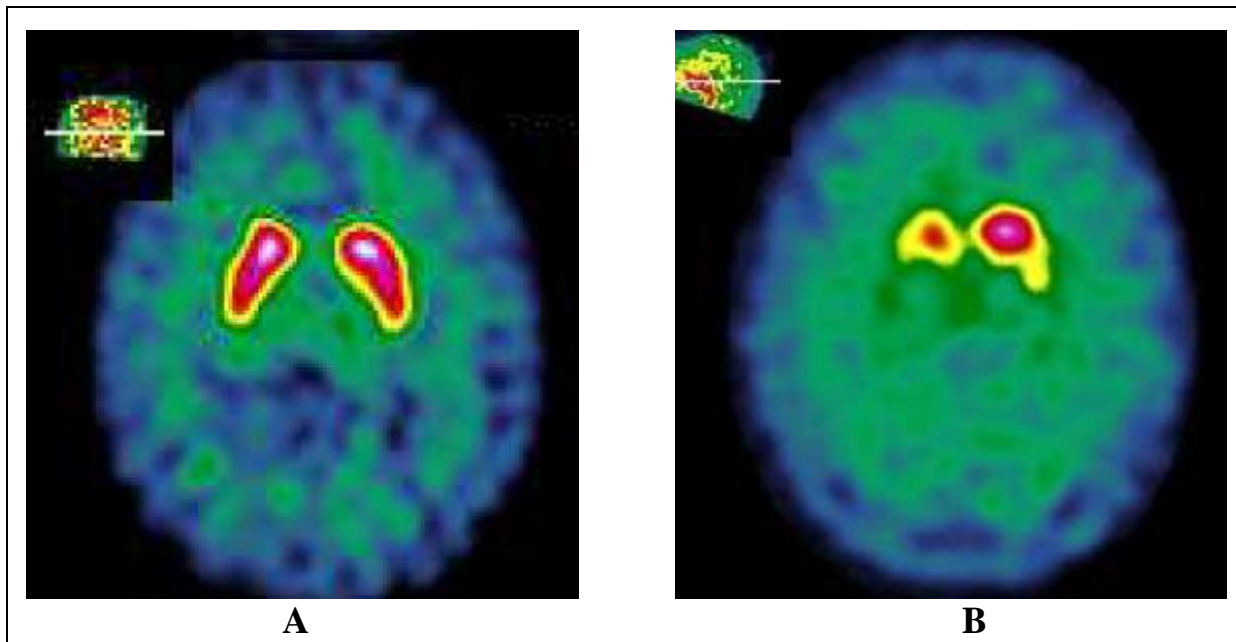


Figure 1.19 : A : représente un examen normal au [18F]-Dopa, et B : patient atteint de la maladie de Parkinson

1.5 Les limites de la TEP

La principale contrainte d'utilisation de la TEP est due à la courte demi-vie des isotopes radioactifs qui nécessitent la présence d'un Cyclotron à proximité du centre d'examen. Il existe un certain nombre de causes de faux positifs et de faux négatifs qui peuvent conduire à une mauvaise interprétation de l'examen.

Les faux positifs (examen positif alors qu'absence d'anomalie) sont essentiellement dus à des causes d'inflammation (granulome, anthracose, chirurgie et radiothérapie récentes, fracture récente...) ou d'infection (pneumopathie, tuberculose, ostéomyélite...)

Les faux négatifs (examen négatif alors que présence d'anomalies) sont essentiellement liés à la taille de la lésion (trop petite pour être détectée), à la localisation (proximité d'une région de fixation naturelle : cerveau, cœur, vessie...), à la nature de l'anomalie (certaines tumeurs présentent moins d'avidité au traceur : tumeur mucineuse...), à une glycémie élevée, ou à une chimiothérapie récente de moins de 15 jours.

1.6 Avenir de la TEP

L'évolution actuelle de la TEP est le couplage de cette technique d'imagerie fonctionnelle avec le scanner à rayons X (TEP/TDM) dans le but de permettre un repérage anatomique précis des anomalies métaboliques révélées par la TEP. L'imagerie multimodale est très performante dans la pratique clinique mais tous les services de médecine nucléaire n'en sont pas encore dotés. Cependant, les principaux constructeurs d'équipement médical cherchent d'ores et déjà à développer de nouveaux types d'appareils combinant l'examen IRM et la TEP. Aussi, l'arrivée de nouvelles molécules (NaF, Fes, FHLT, etc) permettront d'imager des nouveaux processus métaboliques.

1.7 Conclusion

La TEP est un outil puissant d'imagerie in vivo utilisé en clinique comme en recherche pour détecter et quantifier le fonctionnement cellulaire. Les applications cliniques les plus courantes aujourd'hui se trouvent en oncologie et en neurologie. Cependant, de nouvelles applications émergent. Les recherches tendant à améliorer la qualité des images obtenues sur deux fronts :

- instrumentation

Avec entre autres :

- le développement de nouveaux cristaux denses et rapides, ou d'autres matériaux détecteurs comme le gaz ou les semi-conducteurs [77, 78, 79, 80, 81].
- le développement de nouveaux photodétecteurs amplificateurs comme les diodes à avalanche [82].
- le développement de modules de détection permettant le décodage de la profondeur d'interaction du photon γ (amélioration de la résolution spatiale) [83, 84].
- le développement de tomographes dits multimodaux, permettant de réaliser sur le même appareil une image de type anatomique (CT, IRM) et une image fonctionnelle (PET). Ce type d'appareillage présente des avantages évidents pour l'analyse d'image, par exemple la localisation anatomique des structures sans faire intervenir des algorithmes de recalages multimodaux.

- traitement des données, avec entre autres :

- l'amélioration des méthodes de correction des diffusés.
- le développement de nouvelles méthodes de reconstruction qui incorporent une description fidèle du tomographe et de la nature statistique de l'émission radioactive ou bien qui utilisent l'information de temps de vol [85, 86].
- le développement de méthodes de correction de volume partiel, plus efficaces, et permettant une récupération du signal de manière plus robuste que les méthodes nécessitant l'information anatomique [87, 88, 89].
- la correction des mouvements du patient durant les acquisitions TEP (recalage algorithmique des frames d'émission, ou acquisition TEP synchronisée avec un appareil mesurant les mouvements du patient [90]).

Chapitre 2 : Les systèmes d'aide au diagnostic en imagerie médicale

2.1 Introduction

L'imagerie médicale, est maintenant considérée comme étant la première étape dans la détection et la prévention de certaines maladies. Et cela grâce aux développements considérables qu'elle a connu notamment l'intégration des systèmes d'aides à la détection et au diagnostic.

On peut définir un système d'aide en imagerie médicale comme étant un programme numérique développé spécialement afin de fournir aux cliniciens une aide dans l'interprétation et la classification des images pour améliorer la qualité des soins, la santé des patients et éviter certaines chirurgies exploratrices inutiles.

2.2 Qu'est ce qu'un système CAD

Le système d'aide ou bien comme on dit le système CAD peut être décomposé en deux applications principales :

- ✓ Les systèmes d'aide à la détection ou CAD (Computer Aided Detection) permettent la détection des différentes pathologies rencontrées en les différenciant les unes des autres.
- ✓ Les systèmes d'aide au diagnostic (Computer Aided Diagnosis), permettent de classer les pathologies des moins graves aux plus graves détectées par les systèmes d'aide à la détection ou bien observées par les cliniciens.

On peut dire qu'un système CAD est un observateur neutre pouvant analyser l'image selon des critères donnés permettant l'analyse de cette dernière et pouvant poser un diagnostic numérique destiné à aider le clinicien en complétant son diagnostic ou même parfois de lui montrer d'autres critères omis, débouchant sur d'autres pathologies probables. Un système CAD est une combinaison entre les méthodes du traitement des images et les éléments de la classification ou bien d'intelligence artificielle pour pouvoir détecter puis différencier entre les différentes pathologies.

On remarque que le développement de ces systèmes en imagerie médicale ne date pas d'il y a longtemps, ce n'est qu'à la fin des années 80 et avec les travaux de Chan et Giger [49, 50] qu'ils ont fait leur apparition. Au début, ils étaient destinés spécialement à la mammographie. Ces travaux ont été suivis par les études de Freer et Ulissey, Gilbert, Doi, Marten et Engelke [51, 52, 53, 54]. Ils ont permis de prouver que certaines méthodes existantes dans les systèmes CAD permettent une meilleure détection des pathologies en augmentant les performances en terme de détection et de stadification et encore plus, grâce aux éléments d'intelligence artificielle en diminuant la variation des diagnostics posés par les cliniciens que ça soit par un ou plusieurs cliniciens.

Plusieurs de ces méthodes ont été implantées sur des consoles cliniques comme pour Kodak mammography CAD engine et syngo mammoCAD (Siemens) et syngoCT (Siemens) pour le diagnostic du cancer du colon et des poumons pour l'imagerie TDM (<http://www.medical.siemens.com>).

Le Discovery PET/CT 600 avec Cristal BGO, le Discovery PET/CT 690 couplé au scanner Brightspeed Elite et le Discovery PET/CT 690 VCT couplé au scanner VCT proposé par GS pour l'imagerie TEP cérébrale, il est nécessaire de détecter correctement les différentes pathologies

existantes afin de statuer sur le diagnostic final du patient et cela pour éviter d'éventuelles chirurgies ou des pathologies inexistantes.

Pour cela, on se base précisément sur le diagnostic des pathologies en les classifiant selon leur existence ou non et non pas sur la détection des différentes pathologies existantes au niveau du cerveau.

2.3 Le système d'aide en imagerie TEP

2.3.1 Définition de base

Les systèmes de détection peuvent être évalués grâce aux notions suivantes:

- Les vrais positifs (VP).
- Les faux positifs (FP).
- Les vrais négatifs (VN).
- Les faux négatifs (FN).

Représentés dans le tableau suivant :

	Pathologie	Non pathologie
Détectée par le système	VP	FP
Pas détectée par le système	FN	VN

On parle de vrais positifs (VP) lorsque la pathologie est réelle et a été détectée par notre système alors que les faux positifs (FP) représentant une détection de pathologie non existantes apparaissant sur l'image, par contre on rencontre plusieurs fois des cas où la pathologie existe mais qui n'apparaît pas sur l'image requise ce qu'on appelle faux négatifs (VN), et enfin on parle de vrais négatif lors d'une non existence de la pathologie et qui n'a pas été détectée par notre système de détection.

Afin d'évaluer un système on doit calculer: la sensibilité, la spécificité et la précision.

La sensibilité, ou la probabilité que la pathologie soit détectée si la pathologie est présente, se mesure seulement lors de l'existence de celle-ci qui est définie par la relation suivante :

$$\text{Sensibilité} = \frac{VP}{VP + FN}$$

La mesure de la sensibilité est toujours accompagnée par la mesure de la spécificité sans oublier que celle-ci se mesure lors de la non existence de la pathologie seulement. Ainsi, la spécificité, ou la probabilité de non détection de la pathologie pour une non-existence de cette dernière est donnée par la relation suivante :

$$\text{Spécificité} = \frac{VN}{VN + FP}$$

Ensemble, la sensibilité et la spécificité d'un système de détection donnent une appréciation de sa validité intrinsèque. Prises séparément, elles ne veulent rien dire. Par exemple, un système de détection avec une sensibilité 95 % n'a aucune valeur si sa spécificité n'est que de 5 %. Dans ce cas, le système détecte des pathologies à tort et à travers sans tenir compte de leur existence.

2.3.2 Les systèmes CAD en TEP

Les systèmes d'aide au diagnostic (CAD) sont devenus récemment une tâche de routine pour la détection de différentes pathologies en imagerie TEP, donc on se permet de dire que les systèmes CAD sont une référence dans la détection et le diagnostic des différentes anomalies dans le corps humain invisibles à l'œil nu ou pour certaines techniques d'imagerie

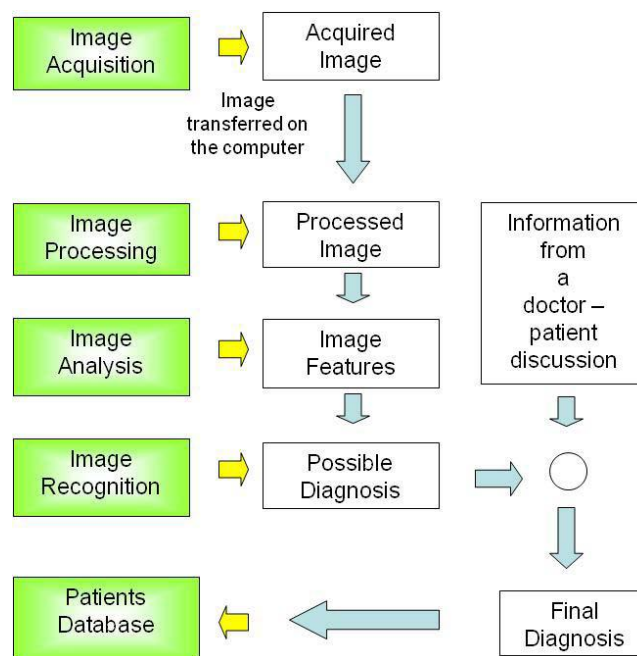


Figure2.1: Schéma représentatif de la structure d'un système CAD

La majorité des systèmes CAD présentés dans ce paragraphe se basent essentiellement sur le traitement d'image et la classification supervisée ou non supervisée, on peut citer par exemple la méthode de Tozaki [55, 56], la méthode de Guan [57] et bien d'autres.

Cette méthode a permis à TOZAKI une première extraction des tumeurs en se basant sur la sélection de pixels de courbures de Gauss2D maximales et qui sont donc obtenus par le produit des courbures principales. Ces courbures sont calculées par combinaisons linéaires des gradients des intensités dans les deux coordonnées de l'espace, une recherche de maxima sur des images de différence inter coupe permet de réduire les fausses détections. Cette méthode comporte un inconvénient majeur alors que les tumeurs ont un aspect tridimensionnel, cette méthode a un aspect bidimensionnel.

La méthode de Guan commence à l'aide des modèles de Markov cachés, par une labellisation des différentes parties du corps humain afin de faciliter l'extraction de ces derniers (le cou, les

poumons, l'abdomen et le cerveau) lors des quels les anomalies (Tumeurs candidates) peuvent être extraites grâce à un algorithme de compétition-diffusion.

Cet algorithme joue sur la minimisation d'une fonctionnelle dépendance des paramètres discriminants afin de diriger les deux coordonnées de l'espace vers deux états majoritaires (tumeur ou pas).

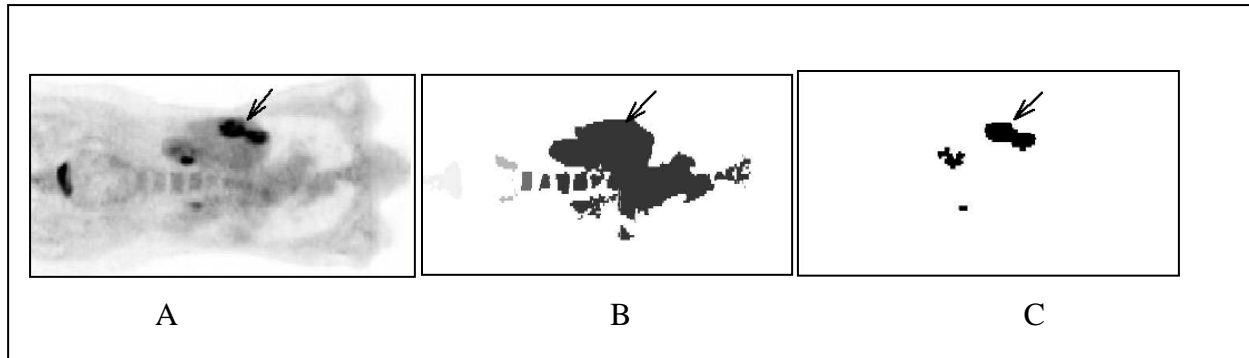


Figure 2.2 : les différentes étapes de la méthode de Guan. A: représente l'image TEP originale. B: la région d'intérêt. C: l'image segmentée.

On remarque que cette méthode est efficace pour des images de tumeurs moyennes par contre, elle a permis de détecter aussi les fixations physiologiques des reins.

En 2008, un autre système CAD basé sur la classification supervisée en utilisant des classifieurs supervisés a été développé par Kanakatte, un seuillage de l'image TEP d'origine constitue la première étape de ce système afin de sélectionner les voxels potentiellement pathologiques.

Un ensemble de caractéristiques statistiques de type moment du premier et second ordre est extrait des régions d'intérêt ROI de taille 32x32 voxels contenant les candidats initiaux. Des caractéristiques spatiales (transformées en cosinus discrète) et fréquentielles (moments statistiques de coefficients d'ondelette) sont aussi ajoutées pour chaque candidat. Les ondelettes sélectionnées sont celles de Daubechie d'ordre 2 et 3 (db2 et db3), la Coiflet (coif2, coif3) et la Biorthogonale (bior2.8, bio3.9). Ces caractéristiques sont ensuite envoyées aux classifieurs supervisés qui pourront déterminer l'existence des pathologies ou leurs absences.

2.4 La classification supervisée et la classification non supervisée

2.4.1 Introduction :

Les systèmes d'aide au diagnostic se basent à la fois sur les méthodes de traitement d'image et de classification.

La classification rassemble une famille de méthodes permettant l'automatisation du processus de reconnaissance, son but est de regrouper différentes observations en sous-ensembles (classes) selon des caractéristiques prédéfinies.

Le terme classification en français désigne à la fois les termes anglais *classification* (classification supervisée) et *clustering* (classification non supervisée).

- Classification supervisée : les classes sont connues à priori, elles ont en général une sémantique associée
- Classification non-supervisée: les classes sont fondées sur la structure des objets, la sémantique associée aux classes est plus difficile à déterminer.

2.4.2 La classification supervisée

La classification supervisée est une tâche de fouille comprenant deux phases :

L'apprentissage qui consiste à construire un classifieur (un modèle) à partir d'un ensemble d'exemples étiquetés par leurs classes, et le classement qui consiste à utiliser le modèle pour affecter une classe à un nouvel exemple.

Le problème de la classification est traité dans plusieurs communautés de recherche qui se découvrent et s'enrichissent mutuellement : statistiques, reconnaissances de formes, apprentissage automatique, réseaux de neurones et beaucoup d'autres. Il existe de nombreux domaines d'application de ce problème : l'attribution de crédit bancaire, la reconnaissance de gènes, la prédiction de sites archéologiques, le diagnostic médical, etc.

Plusieurs méthodes de classification supervisée publiées dans la littérature s'appuient sur des techniques différentes : inférence bayésienne, k plus proches voisins, arbres de décision, réseaux de neurones, machines à support de vecteurs (ou SVM), Programmation génétique ou treillis de concepts.

La plupart des algorithmes d'apprentissage supervisés tentent donc de trouver un modèle - une fonction mathématique - qui explique le lien entre des données d'entrée et les classes de sortie.

L'inconvénient principal de cette classification repose sur la nécessité d'avoir une base d'observations suffisamment large et représentative pour l'apprentissage et le test du classifieur.

2.4.3 La classification non-supervisée

Le but ultime de la classification non supervisée (ou regroupement) est d'identifier les différentes observations, et donc d'affecter à chaque observation une "étiquette de classe" qui matérialise l'appartenance de l'observation à l'une des classes. De plus, on souhaite pouvoir également affecter à toute nouvelle observation une étiquette de classe.

Cette situation n'est pas sans rappeler celle rencontrée en classification supervisée : les observations disponibles sont déjà affectées à une classe, et donc son objectif est d'affecter une nouvelle observation à une classe. Mais le problème est plus difficile puisque les observations disponibles ne sont pas initialement identifiées comme appartenant à telle ou telle classe: il faudra déduire cette information de la répartition spatiale des observations.

L'expression "non supervisée" fait donc référence au fait qu'aucun 'superviseur' n'est là pour nous dire à quelle classe appartient telle ou telle observation.

2.5 Les systèmes CAD supervisés

2.5.1 Introduction

En imagerie TEP et comme constaté lors de notre recherche, les méthodes d'aides au diagnostique ne sont pas très nombreuses et bien plus limitées par leurs critères de validation. En vue de cette limitation, on s'est donc basé sur les méthodes de développement des systèmes CAD déjà existantes dans d'autres modalités surtout en mammographie et en imagerie TDM d'où la grande révolution constitue l'introduction de la classification supervisée ayant permis, par disponibilité, d'images classées la création de superviseur efficace et paramétrable.

2.5.2 Les étapes d'un système CAD supervisé

L'identification initiale des anomalies et la réduction des faux négatifs sont considérés comme des étapes essentielles pour un système CAD.

La première étape a pour but d'identifier les zones suspectes de l'image contenant des pathologies ou anomalies en diminuant le plus possible les fausses détections qui peuvent être détections. Cette première étape peut être illustrée par plusieurs approches.

Les parties du corps sont séparées pour une meilleure détection lorsqu'on cible un organe ou régions précises, on peut donc procéder à une amélioration sur les images correspondantes, ce qu'on appelle une étape de prétraitement, lors de laquelle on utilise plusieurs méthodes selon l'image à traiter. Une segmentation par seuillage ou par LPE, une élimination du bruit, une détection des contours, un rehaussement et beaucoup d'autres.

Par la suite, on utilise des systèmes de classification optimisée par un apprentissage en se basant sur des informations de l'image ou des pathologies cibles, ce qui permet l'extraction des anomalies considérées comme suspectes. Cette classification a pour seul objectif la détermination de l'existence des pathologies ou leur absence dans l'image traitée.

La deuxième étape consiste à réduire le plus possible les fausses détections extraites lors de la première étape mais en gardant une sensibilité élevée.

Plusieurs méthodes peuvent être utilisées afin de trier les anomalies pour ne garder que les cas les plus pertinents.

2.5.3 L'identification initiale des anomalies

2.5.3.1 Le Prétraitement

Le prétraitement est une étape primordiale lors de l'analyse des images, cela pour améliorer et préparer ces dernières afin de faciliter l'extraction des anomalies.

On remarque que beaucoup de systèmes CAD sont conçus pour un organe précis surtout en imagerie TDM et mammographique dans lesquels on visualise un seul organe à la fois, par contre, l'image TEP possède une particularité qui se traduit par la visualisation du corps entier,

une partie du corps ou bien un organe spécifique en trois dimension sans changement de position pour le patient.

Les méthodes de classifications développées pour un seul organe permettent de mieux tenir compte de ses caractéristiques particulières en termes de bruit et de texture par exemple. Son inconvénient est la génération de systèmes à adapter à chaque nouvelle région cible et de ce fait, ne pouvant pas réaliser de traitement global sur les images du corps entier [S. Tomei, 2009].

Pour segmenter l'organe voulu on utilise généralement des méthodes simples comme des seuillages ou des méthodes de classification non-supervisées pour le cas des images anatomiques, mais dans le cas des images fonctionnelles et pas anatomiques comme pour l'imagerie TEP, la segmentation des organes est plus difficile et problématique.

L'amélioration des images ne constitue pas une étape très importante dont on peut s'en passer et on remarque que pour beaucoup d'études on utilise uniquement l'image brute, par contre l'amélioration des images tend à influencer les performances finales de classification. Un rehaussement en imagerie améliore généralement les résultats de classification [59]. Cependant le débruitage par des filtres inadaptés aux images utilisées baisse la sensibilité [60].

2.5.3.2 Les caractéristiques descriptives

Lors de la classification l'image brute n'est que très rarement traitée telle quelle. On préfère généralement des caractéristiques descriptives extraites de l'image. Ce qui rend ces images facilement interprétables compte tenu que ces caractéristiques portent un grand nombre d'informations.

Ces caractéristiques sont assimilées à ce que l'œil humain analyse avant de poser son diagnostic sur la présence ou pas de tumeurs. Elles sont l'interprétation de l'image d'un point de vue numérique et leur choix est primordial [58].

Les caractéristiques ont pour but de différencier entre les classes pour notre cas : différencier entre les cas pathologiques et les cas sains [61].

Plusieurs caractéristiques sont utilisées généralement classées selon leur type dans des familles appelées familles de caractéristiques, on distingue 5 grandes familles :

- Les caractéristiques statistiques.
- Les caractéristiques sémantiques.
- Les régions d'intérêt de l'image.
- Les caractéristiques fréquentielles.
- Les caractéristiques cliniques.

Chacune des ces caractéristiques est présentée ci-dessous lors du système CAD utilisé en différentes modalités d'imagerie :

Caractéristiques statistiques

Les propriétés de bruit et de texture des images à étudier sont toujours reflétées par des caractéristiques. Les statistiques liées au bruit représentent les niveaux de gris de l'image et donc on peut extraire la moyenne et l'histogramme sans aucune difficulté par contre les statistiques liées à la texture sont assimilées à l'apparence locale de l'image, c'est à dire l'organisation des détails d'une petite partie de l'image telle que le système visuel humain l'aperçoit.

Haralick [62] a proposé un ensemble de 14 caractéristiques qui peuvent être utilisées afin d'homogénéiser les mesures extraites pour les différentes modalités d'imagerie. Par contre, les études de [63, 64] proposent un ensemble de 48 caractéristiques statistiques en mammographie.

En imagerie TEP, la taille réduite des zones d'intérêt implique une caractérisation de leur texture limitée. Cette famille de caractéristiques peut donc parfois être rendue insuffisante ou inappropriée pour une tâche de discrimination.

Régions d'intérêt de l'image et intensités

Pratiquement dans toutes les modalités d'imageries médicales, les tumeurs apparaissent ou plus claires, ou plus foncées que l'organe à étudier et donc plus distinctes les unes des autres. En TEP au 18F-FDG et en mammographie les tumeurs sont beaucoup moins apparentes par rapport au tissu environnant. L'imagerie TEP associée à d'autres radiotraceurs que le 18F-FDG peut aussi être caractérisée par un hyposignal dans les tumeurs. Les tumeurs apparaissent sous forme de masses de forme spécifique et d'atténuation plus élevée en imagerie TDM. Dans certains cas la valeur même des pixels ou de voxels peut caractériser les tumeurs et donc beaucoup d'études l'utilisent en l'état comme caractéristique d'image.

Caractéristiques géométriques

Les critères de convexité expriment généralement les caractéristiques géométriques, pouvant aussi être discriminantes. Elles correspondent aux détails visuels sur lesquels se basent les cliniciens, ce qui rend leur interprétation plus simple. Ces caractéristiques sont spécifiques aux imageries anatomiques étant donné qu'elles reposent sur une délimitation précise des structures étudiées [58].

En imagerie TDM la majorité des CAD utilisés pour cette modalité se basent sur cette technique, les tumeurs initiales sont généralement pré-segmentées manuellement ou semi-manuellement et leur forme peut facilement être extraite.

Le système CAD ciblé sur le cancer du poumon développé par [65] se base sur l'extraction d'un ensemble de caractéristiques de forme discriminantes. On peut citer l'irrégularité de la tumeur pré segmentée, son degré de compacité et son diamètre effectif,

Lin Boroczky [64, 66] ajoutant aussi la notion de sphéricité pour des nodules pulmonaires. Il est également possible d'utiliser un critère de localisation, et notamment de symétrie afin de caractériser la présence de tumeurs thoraciques comme le fait [67] en imagerie TDM.

Cette famille de caractéristiques est aussi difficilement exploitable dans le cas de l'imagerie fonctionnelle.

Caractéristiques fréquentielles

L'étude de Sachs [68], a montré que l'information est traitée à travers des canaux sélectifs en fréquence pour le système visuel humain, donc par extraction des caractéristiques fréquentielles des images cibles on peut se rapprocher le plus possible d'un système visuel humain. Cette famille de caractéristiques se base sur la décomposition en ondelettes. Une ondelette est une petite onde (ou vague) qui a un début et une fin elle permet la représentation d'une fonction comme une somme pondérée de petites ondes translatées ou dilatées, de moyenne nulle et de support compact.

Le facteur de translation détermine sur quel intervalle du domaine de définition de la fonction on se place, alors que le facteur d'échelle calibre le domaine fréquentiel. Le positionnement conjoint dans les domaines spatial et fréquentiel et l'aspect multi-résolution de cette transformation est d'un attrait particulier dans la détection d'objets Gaussiens de taille réduite noyés dans du bruit haute fréquence.

Le choix de la fonction d'ondelette et son niveau de décomposition influence les images approximations et les images de détails ce qui constitue la plus grande difficulté de cette famille de caractéristiques.

Caractéristiques textuelles

La dernière grande famille de caractéristiques rassemble des informations qui ne sont plus directement extraites des images, mais du dossier médical du patient. Ce type de caractéristiques est communément appelé sémantique. L'ensemble des informations personnelles du patient se retrouve dans cette catégorie, allant de son âge à ses habitudes alimentaires. Cette famille de caractéristiques permet de tenir compte de l'influence de paramètres extérieurs à l'image, mais qui influencent néanmoins le diagnostic du clinicien au cours de l'interprétation des images [58]. Utilisé principalement dans les systèmes CAD implantés en imagerie TDM, ces données regroupent les informations sur les patients, leur antécédent, leur pathologie et leur évolution.

Cette famille de caractéristiques est la plus intéressante parmi les autres familles déjà étudiées précédemment, et cela en vue de son comportement de clinicien tenant compte des circonstances des patients et des images traitées. Par contre, les systèmes CAD utilisant cette famille de caractéristiques utilisent beaucoup plus d'informations supplémentaires que pour les autres familles sans les quelles les systèmes CAD ne peuvent fonctionner correctement.

2.5.4. Les classifieurs

2.5.4.1. Introduction

L'apprentissage sert à construire un modèle général à partir de données particulières. Son but est de prédire un comportement face à une nouvelle donnée et approximer une fonction ou une densité de probabilité. En apprentissage, le classifieur est une famille d'algorithmes de classement statistique. Le rôle d'un classifieur est de classer dans des groupes (des classes) les échantillons qui ont des propriétés similaires, mesurées sur des observations. Un classifieur linéaire est un type particulier de classifieur, qui calcule la décision par combinaison linéaire des échantillons.

2.5.4.2 Définition du classifieur

Pour un vecteur d'observation x , à valeur dans \mathbb{R}^N , la sortie du classifieur est donnée par:

$$g(x) = f(w^T x + w_0) = f\left(\sum_{j=1}^N w_j x_j + w_0\right),$$

Où \vec{w} est un vecteur de poids, w_0 est le biais, et f est une fonction qui convertit le produit scalaire des deux vecteurs dans la sortie désirée. Le vecteur de poids w est appris à partir d'un ensemble d'apprentissage étiqueté. La fonction f est souvent une simple fonction de seuillage, par exemple la fonction signe, la fonction de Heaviside, ou des fonctions plus complexes comme la tangente hyperbolique, ou la fonction sigmoïde. Une fonction de décision plus complexe pourrait donner la probabilité qu'un certain échantillon appartienne à une certaine classe.

Pour un problème de discrimination à deux classes, l'opération réalisée par un classifieur linéaire peut se voir comme la séparation d'un espace de grande dimension par un hyperplan: tous les points d'un côté de l'hyperplan sont classés en tant que 1 , les autres sont classés en tant que -1 . Cet hyperplan est appelé hyperplan séparateur, ou séparatrice.

Les classifieurs linéaires obtiennent souvent de bons résultats lorsque N , le nombre de dimension de l'espace des observations est grand.

Pour estimer le vecteur \vec{w} deux familles de méthode sont utilisées :

La première cherche à créer des modèles génératifs en se basant sur les probabilités conditionnelles. On peut citer les deux méthodes les plus utilisées pour cette première famille :

- ✓ Classifieur bayésien naïf.
- ✓ Analyse discriminante linéaire (LDA).

La deuxième famille permet une approche par modèle discriminant, qui en un premier temps maximise la qualité de la classification sur un jeu de test. Ensuite une fonction de coût va réaliser l'adaptation du modèle de classification final (en minimisant les erreurs).

On présente quelques méthodes parmi celles appartenant à cette famille :

- ✓ Régression logistique
- ✓ Perceptron.
- ✓ Machine à vecteurs de support (séparateur à vaste marge, SVM).

On considère généralement que les modèles entraînés par une méthode discriminante (SVM, Régression logistique) sont plus précis que ceux de type génératifs entraînés avec des probabilités conditionnelles (classifieur bayésiens naïfs ou linéaires). On considère que les classifieurs génératifs sont plus adaptés pour les processus de classification avec de nombreuses données manquantes (par exemple la classification de texte avec peu de données d'apprentissage).

Tous les classifieurs linéaires cités peuvent opérer sur des données non linéairement séparables en opérant sur un espace de représentation transformé $\varphi(\vec{x})$ avec un Kernel. Cette technique consiste à appliquer une transformation aux données d'entrées pour trouver dans un nouvel espace de grande dimension dans lequel elles sont projetées, un hyperplan séparateur optimal.

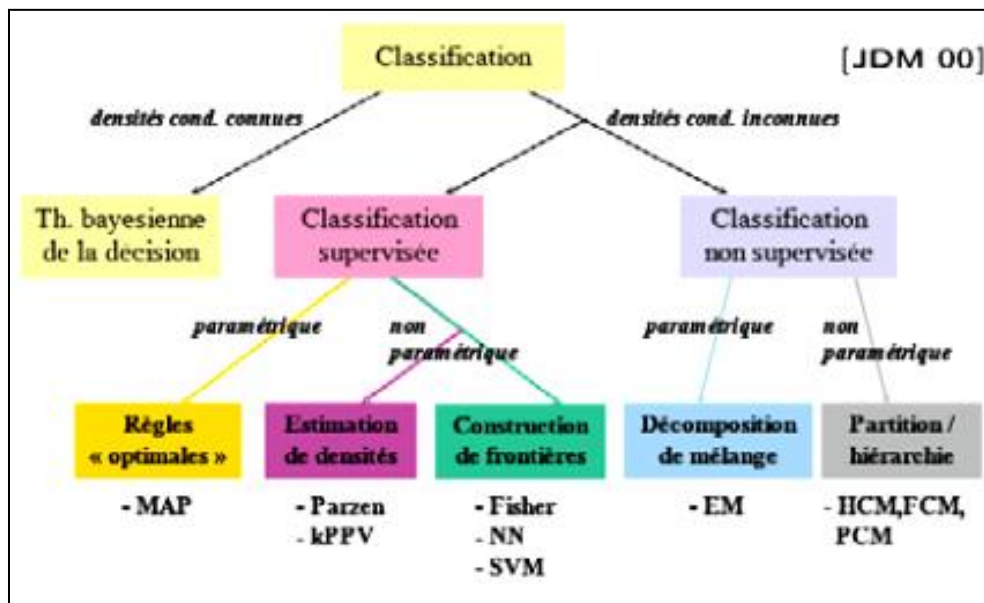


Figure2.3 : Les différents types de classifications et les méthodes les plus utilisées pour chaque type

2.5.4.3 Types de classifieurs

Le classifieur naïf de Bayes

Le classifieur bayésien naïf [69] est une méthode d'apprentissage supervisé qui repose sur une hypothèse simplificatrice forte : les descripteurs (X_j) sont deux à deux indépendants conditionnellement aux valeurs de la variable à prédire (Y). Pourtant, malgré cela, il se révèle robuste et efficace. Ses performances sont comparables aux autres techniques d'apprentissage.

Soient $\mathfrak{X} = (X_1, \dots, X_J)$ l'ensemble des descripteurs, Y la variable à prédire (l'attribut classe comportant K modalités).

En apprentissage supervisé, pour un individu w à classer, la règle bayésienne d'affectation optimale revient à maximiser la probabilité a posteriori d'appartenance aux classes donc :

$$\hat{y}(w) = y_{k^*} \Leftrightarrow y_{k^*} = \arg \max_k P[Y = y_k / \mathfrak{X}(w)]$$

En vue que la décision bayésienne se base sur une estimation fiable de la probabilité conditionnelle $P(Y/X)$. Et donc on peut écrire :

$$P[Y = y_k / \mathfrak{N}(\omega)] = \frac{P(Y = y_k) \times P[\mathfrak{N}(\omega) / Y = y_k]}{P[\mathfrak{N}(\omega)]}$$

Comme l'objectif est de détecter le maximum de cette quantité selon y_k , et que le dénominateur n'en dépend pas, nous pouvons ré écrire la règle d'affectation ci-dessus.

$$\hat{y}(\omega) = y_{k^*} \Leftrightarrow y_{k^*} = \arg \max_k P(Y = y_k) \times P[\mathfrak{N}(\omega) / Y = y_k]$$

L'analyse discriminante de Fisher

L'analyse discriminante linéaire selon Fisher [70] appelée LDA (linear discriminant analysis), basée sur la discrimination linéaire. Son but est de dériver une combinaison linéaire des caractéristiques des images de façon que les caractéristiques projetées dans le nouvel espace soient séparables de manière optimale.

Cette dérivation aboutit à un plan de projection ω et ne se base que sur le calcul de moments statistiques d'ordre un et deux de ces caractéristiques. Le plan de séparation optimale entre les classes correspond à la normale au plan ω . Cette dérivation aboutit à un plan de projection ω et ne se base que sur le calcul de moments statistiques d'ordre un et deux de ces caractéristiques. Le plan de séparation optimale entre les classes correspond à la normale au plan ω . ω étant calculé comme suit :

$$\omega = (\langle V_1 \rangle - \langle V_0 \rangle)^t \Sigma^{-1} \text{ avec } \Sigma = \frac{1}{2} \left((V - \langle V_0 \rangle)(V - \langle V_0 \rangle)^t + (V - \langle V_1 \rangle)(V - \langle V_1 \rangle)^t \right)$$

Il est nécessaire de connaître la moyenne des vecteurs de caractéristiques de chaque classe et de la covariance intra-classe pondérée ω , sans oublier que ces mesures sont extraites par apprentissage.

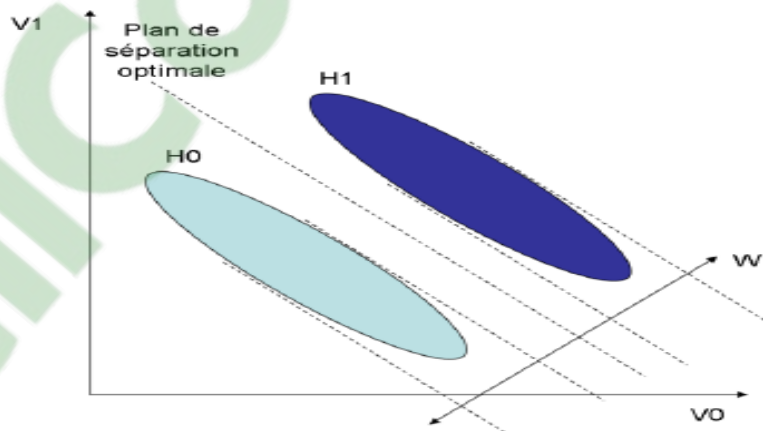


Figure 2.4. Méthodologies d'une LDA à deux classes H1 et H0

Les réseaux de neurones

On s'inspire du mode de raisonnement et de l'apprentissage de l'être vivant pour développer des modèles dit intelligents. Ces réseaux de neurones développés pour résoudre des problèmes de contrôle, de reconnaissances etc....

Le réseau de neurone est un classifieur non-linéaire se basant sur le fonctionnement des neurones humains de manière à imiter le fonctionnement de notre cerveau. Il est constitué de neurones artificiels interconnectés. Un neurone permet de définir une fonction discriminante dans l'espace d'entrée en réalisant une combinaison linéaire de son vecteur de caractéristiques. Les neurones sont arrangés par couches et comprennent une couche d'entrée, une couche de sortie et des couches cachées. Leur objectif est d'apprendre des caractéristiques d'entrée en utilisant les couches cachées.

Les réseaux de neurones sont assez difficile à mettre en place, il est en effet très difficile de définir le nombre de couche cachées du classifieur et le nombre de neurones par couche malgré leurs bonnes performances.

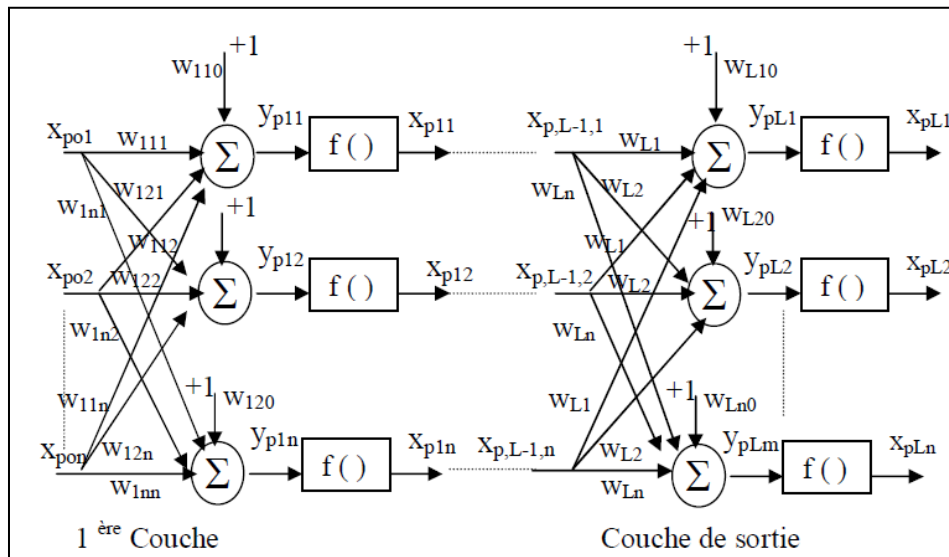


Figure 2.5 : Architecture et fonctionnement d'un réseau de neurones

Les séparateurs à vaste marge (SVM)

La méthodologie des SVM [71] repose directement sur la minimisation du risque théorique de mauvaise classification détaillé précédemment. L'idée basique des SVM est, comme leur nom l'indique, de trouver l'hyperplan optimal de séparation qui maximise la marge de séparation entre les deux classes H_1 et H_0 . Cette marge correspond en d'autres termes à la distance à l'hyperplan entre les plus proches vecteurs de caractéristiques appartenant à chacune des classes

Les SVM présentent généralement des performances proches de celles des réseaux de neurones. Ils sont connus pour être adaptés aux bases d'apprentissage de petite taille et des ensembles d'observations de grande dimension (nombre de caractéristiques élevé).

2.6 Conclusion

L'utilisation des différents classifieurs en mammographie et en imagerie TDM a été présentée dans de nombreux articles et revues. Ce qui montre qu'il n'est pas toujours possible de comparer entre les différents classifieurs. Il est donc quasiment impossible d'ordonner les classifieurs selon leur capacité à être ou pas les plus adaptés aux problématiques ciblées.

Campadelli [72] analyse la robustesse des SVM à un apprentissage à partir de classes de taille très inégale pour la détection de tumeurs pulmonaires en imagerie TDM. Boroczky Et Lu [66, 73] utilisent aussi les SVM sur des images TDM pulmonaires. En ce qui concerne les réseaux de neurones, Gletsos [74] il base son système CADi dédié aux tumeurs du foie en TDM sur une combinaison séquentielle de trois réseaux de neurones permettant de classifier les nodules.

D'autres systèmes reposent sur la combinaison de plusieurs classifieurs supervisés. Comme pour le CAD développé par Park [75] pour la mammographie. Celui-ci se base sur la comparaison et la combinaison des scores obtenus d'un classifieur par k-PPV et de réseaux de neurones.

Chapitre 3 : Les méthodes utilisées

3.1 Introduction

Dans ce chapitre, nous allons parler des outils utilisés dans le cadre de ce travail, et qui ont permis l'obtention d'un système CAD fiable. Les deux principaux outils utilisés sont l'analyse par ondelette et la classification à l'aide du séparateur à vaste marge (SVM).

3.2 La théorie des ondelettes

3.2.1 Introduction

L'imagerie médicale depuis son arrivée a permis des avancées révolutionnaires en pratiques médicales. Sauf que de nombreux problèmes restent posés et leur résolution peut améliorer l'interprétation de certaines pathologies. On peut citer par exemple plusieurs problèmes que rencontrent les cliniciens.

- La détection et l'analyse des structures malignes dans des images issues de mammographie de scanner ou d'IRM.
- La mauvaise résolution spatiale des images acquises avec la tomographie par émission de positons (TEP).

L'application des ondelettes en imagerie médicale ne date pas de longtemps (au début des années 90), et elle est plutôt très utilisée depuis. Les ondelettes ont été appliquées en premier lieu pour le débruitage et la compression mais par suite pour l'analyse fonctionnelle des données médicales.

3.2.2 La transformation en ondelette

L'analyse en ondelettes proposée initialement par J. Morlet au début des années 80, est considérée comme un outil très performant permettant pour la résolution des problèmes dans différents domaines d'application, dont le traitement d'image.

L'analyse par ondelette est basée sur un concept quelque peu différent du concept de fréquence: le concept d'échelle. Au lieu de considérer des fonctions oscillantes placées à l'intérieur d'une fenêtre, que l'on fait ensuite coulisser le long d'un signal à analyser, on utilise des fonctions mathématiques pour représenter le signal. Les ondelettes sont davantage des copies les unes des autres, donc un signal s'écrit sous la forme d'une superposition de telles ondelettes décalées et dilatées. Les poids de ces ondelettes dans la décomposition forment la transformée en ondelettes. la transformée en ondelette d'un signal $X(t)$ est donnée par :

$$g(a, b) = \frac{1}{\sqrt{a}} \int_{t=-\infty}^{t=\infty} x(t) \bar{\psi}_{a,b}(t) dt$$

La fonction $\psi_{a,b}(t)$ est obtenue par translation et dilatation d'une fonction particulière appelée ondelette mère.

$$\psi_{a,b}(t) = \Psi\left(\frac{t-b}{a}\right)$$

b : détermine la position, et a : donne l'échelle.

Pour le cas d'un signal : a représente la fréquence et b le temps.

La fonction Ψ doit être oscillante, intégrale nulle et de carré intégrable. Comme elle peut être à valeurs complexes.

La transformation en ondelette se rapproche le plus de la transformation de Fourier, et donc il existe de nombreuses ondelettes mères Ψ possibles.

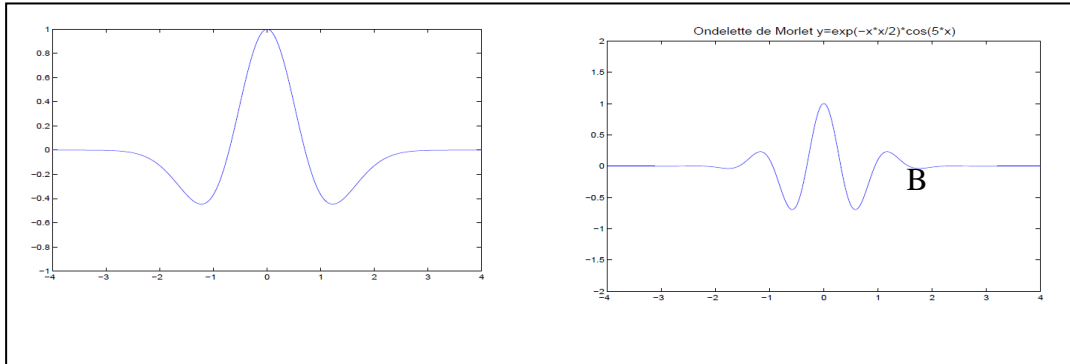


Figure 3.1 : deux exemples d'ondelettes mères utilisées pour la transformation en ondelette, où (A) est la dérivée seconde d'une gaussienne, et (B) est l'ondelette de Morlet.

Comme mentionné précédemment, l'analyse par ondelettes est un outil mathématique capable de transformer un signal dans le domaine spatial en un autre signal dans le domaine fréquentiel. Les composantes de ce nouveau signal sont appelées les coefficients d'ondelettes ces coefficients renseignent sur la variation locales des niveaux de gris au tour d'un pixel donné de l'image. Plus la variation est importante plus les coefficients sont grands.

3.2.3 Analyse multi-résolution

Le signal $x(t)$ est exprimé comme une superposition continue d'ondelettes avec la transformation continue par ondelettes. Cette reconstruction du signal est extrêmement *redondante* dans le sens où toutes les ondelettes sont utilisées et qu'elles sont « linéairement dépendantes ». Une première méthode de discrétisation, assez efficace bien que toujours redondante, consiste en une généralisation des théorèmes d'échantillonnage de Shannon. En 1986, une méthode radicalement différente est mise au point : l'*analyse multirésolution*. Elle est complètement récursive et donc idéale pour faire des calculs. De plus, elle engendre une famille d'ondelettes orthonormales, ce qui réduit la redondance à néant.

3.2.3.1 Principe de l'analyse multi-résolution

Plutôt que de commencer avec une ondelette-mère, on débute avec une *analyse multirésolution* (AMR). C'est-à-dire, une suite $(V_j)_{j \in \mathbb{Z}}$ de sous ensembles fermés de $L^2(\mathbb{R})$

Telle que :

- (i) $\forall j \in \mathbb{Z} : V_j \subset V_{j+1}$
- (ii) $\forall j \in \mathbb{Z} : f \in V_j \Leftrightarrow f(2 \cdot) \in V_{j+1}$
- (iii) $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$
- (iv) $\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R})$

(v) Il existe une fonction $\varphi \in V_0$ telle que $\{\varphi(\cdot - k) : k \in \mathbb{Z}\}$ est une base orthogonale de V_0 . Une telle fonction φ est appelée fonction d'échelle de l'AMR.

La fonction φ est en quelque sorte le potentiel qui permet de construire une ondelette mère ψ à partir de l'AMR.

Il est facile de prouver que $\forall j \in \mathbb{Z}, \{\varphi_{jk} := 2^{j/2} \varphi(2^j \cdot -k) : k \in \mathbb{Z}\}$ forme une base orthogonale de V_j .

Et donc l'intervalle d'échantillonnage dans le temps est fixé à 1 tandis que le facteur d'échelle à 2.

Cette analyse multi-résolution donne comme résultat un ensemble d'image de détails et une image d'approximation, plus on augmente le niveau de décomposition plus la tailles des images est réduite par un facteur P qui varie selon le type d'ondelette utilisée.

L'image approximation est obtenue par la convolution de l'image d'entrée avec un filtre passe-bas. Par contres les détails sont obtenus par convolution avec un filtre passe haut appliquée sur l'image d'entrée d'abord ligne par ligne puis colonne par colonne d'où l'obtention des détails horizontaux, verticaux et diagonaux.

Voici un exemple de décomposition d'une image au premier et au deuxième niveau représentés dans les figures 3.2 et 3.2.

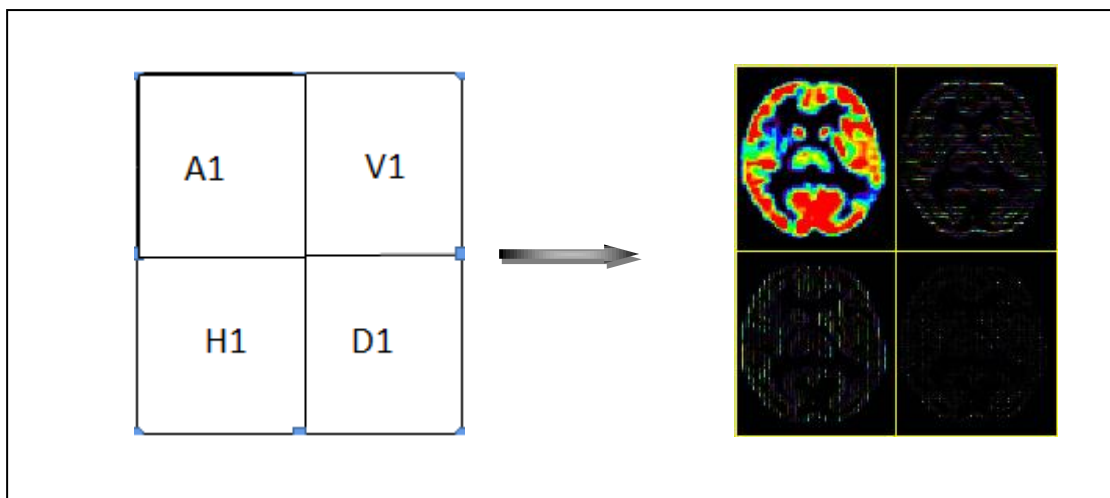


Figure3.2:décomposition de l'image sur un niveau

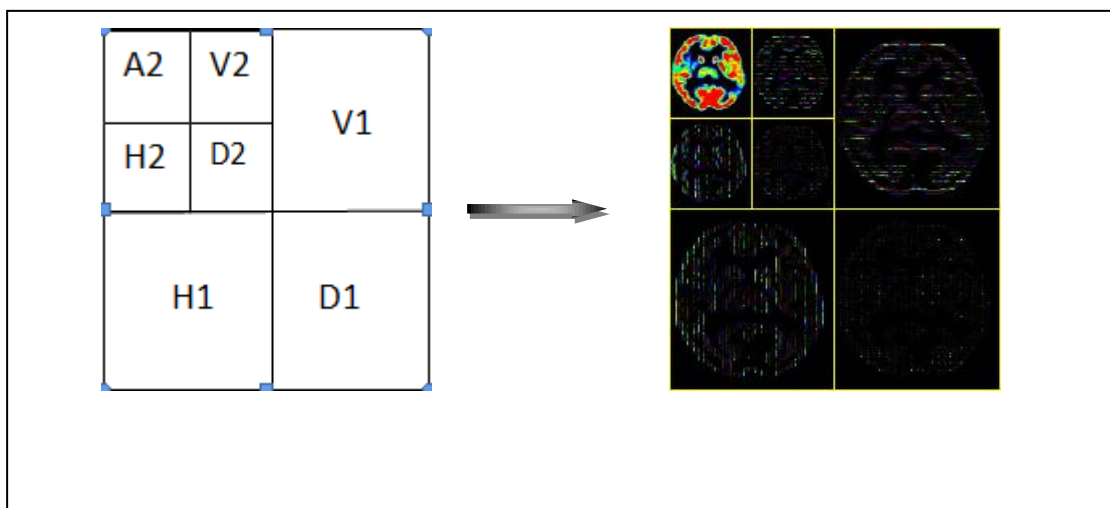


Figure 3.3: décomposition de l'image sur deux niveaux

3.2 types d'ondelettes

3.2.1 Ondelette de Haar

L'ondelette de Haar, ou fonction de Rademacher, est une ondelette créée par Alfred Haar en 1909. On considère que c'est la première ondelette connue. Elle est la plus simple à comprendre et à implémenter. C'est une fonction dilatée et/ou translatée de la fonction mère. Elle est définie comme une fonction réelle sur l'intervalle $[0, 1[$

Cette fonction appelée $x(t)$ est échantillonnée en N points avec $N=2^m$.

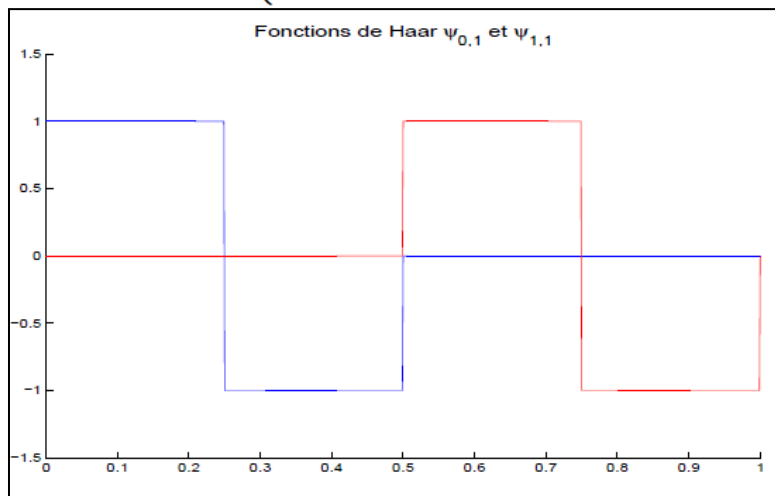
$$x_k = x(t_k) \text{ pour } t_k = 0, \frac{1}{N}, \dots, \frac{N-1}{N}$$

On définit la famille de fonction:

$$\phi(t) = \begin{cases} 1 & \text{pour } 0 \leq t < 1 \\ 0 & \text{autrement} \end{cases}$$

Et

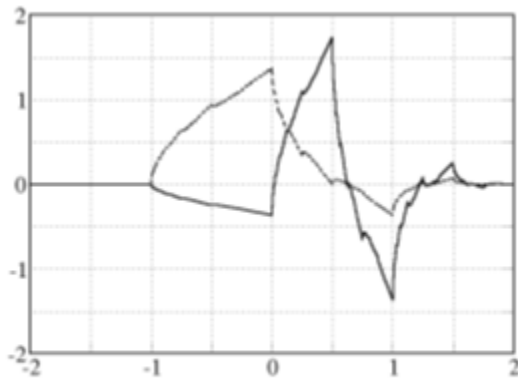
$$\phi_{j,k}(t) = \begin{cases} \phi(2^j t - k) & : t < \frac{k+1}{2^j} \\ 0 & \text{autrement} \end{cases}$$



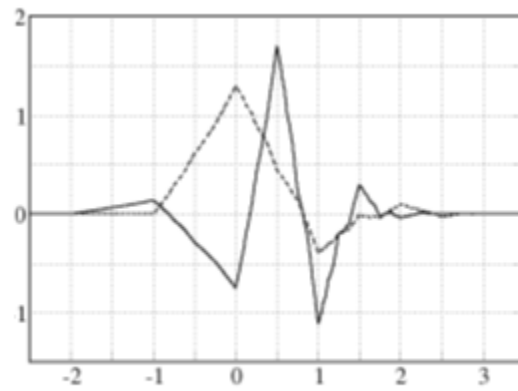
Ondelette de Haar.

3.2.2 Ondelette de Daubechies

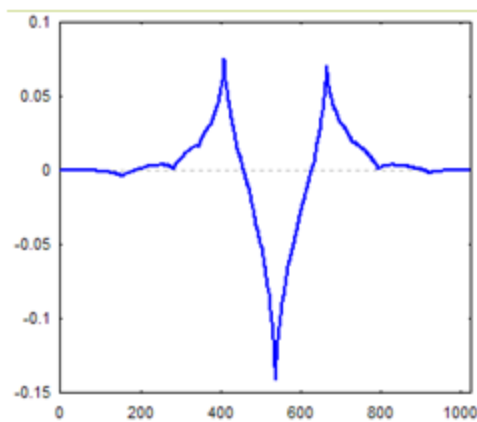
Nommées d'après leur créatrice Ingrid Daubechies, les ondelettes de Daubechies sont une famille d'ondelettes orthogonales définissant une transformée en ondelettes discrète, caractérisées par un nombre maximal de moments dissipants pour un support donné. Pour chaque type d'ondelette de cette classe, il existe une fonction d'échelle (appelée aussi ondelette mère) qui génère une analyse multirésolution orthogonale.



Ondelette D4(db2)



Ondelette D6(db3)



Ondelette coif6

3.3.3 Les ondelettes biorthogonales

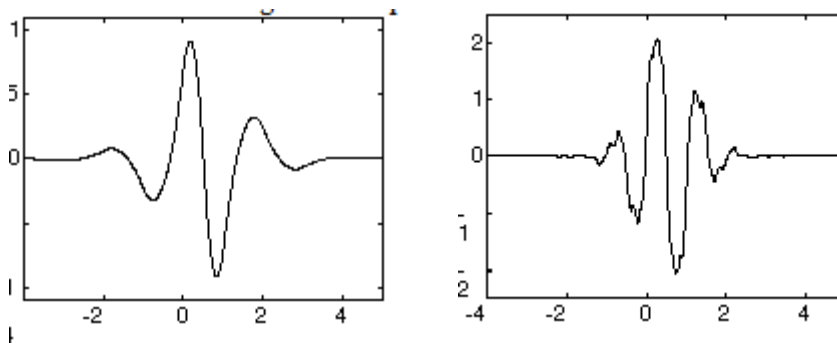
Les ondelettes biorthogonales se définissent de manière analogue aux ondelettes orthogonales, mais en partant de multirésolutions biorthogonales. On écrit les décompositions:

$$V_{j-1} = V_j \oplus W_j \text{ avec } W_j \subset (V_j^*)^\perp$$

$$V_{j-1}^* = V_j^* \oplus W_j^* \text{ avec } W_j^* \subset (V_j)^\perp$$

De manière analogue au cas orthogonal, un signal f de L^2 peut s'écrire comme suit :

$$\begin{aligned}
f(t) &= \sum_{j,n \in \mathbb{Z}} \langle f, \psi_{j,n}^* \rangle \psi_{j,n}(t) \\
&= \sum_{n \in \mathbb{Z}} \langle f, \phi_{j,n}^* \rangle \phi_{j,n}(t) \\
&+ \sum_{k \leq j, n \in \mathbb{Z}} \langle f, \psi_{k,n}^* \rangle \psi_{k,n}(t) \\
&= \sum_{j,n \in \mathbb{Z}} \langle f, \psi_{j,n} \rangle \psi_{j,n}^*(t) \\
&= \sum_{n \in \mathbb{Z}} \langle f, \phi_{j,n} \rangle \phi_{j,n}^*(t) \\
&+ \sum_{k \leq j, n \in \mathbb{Z}} \langle f, \psi_{k,n} \rangle \psi_{k,n}^*(t)
\end{aligned}$$



Ondelettes biorthogonales

3.4 Conclusion

L'analyse en ondelette est l'une des méthodes les plus efficaces utilisées pour le traitement des signaux et des images dans le domaine fréquentiel, visant à résoudre les problèmes et améliorer la qualité des images.

Cette méthode prend en considération plusieurs paramètres, comme le type d'ondelette utilisé et le niveau de décomposition, afin de traiter les informations contenues dans un signal ou une image de façon appropriée.

3.5 Support Vector Machine ou les séparateurs à vaste marge (SVM)

3.5.1 Introduction

Le SVM est une méthode de classification binaire par apprentissage supervisé, elle fut introduite par Vapnik en 1995. Cette méthode repose sur l'existence d'un classificateur linéaire dans un espace approprié. Puisqu'on possède deux classes, cette méthode fait appel à un jeu de données d'apprentissage pour apprendre les paramètres du modèle. Elle est basée sur l'utilisation de fonctions dites « noyaux » (kernel) qui permettent une séparation optimale des données.

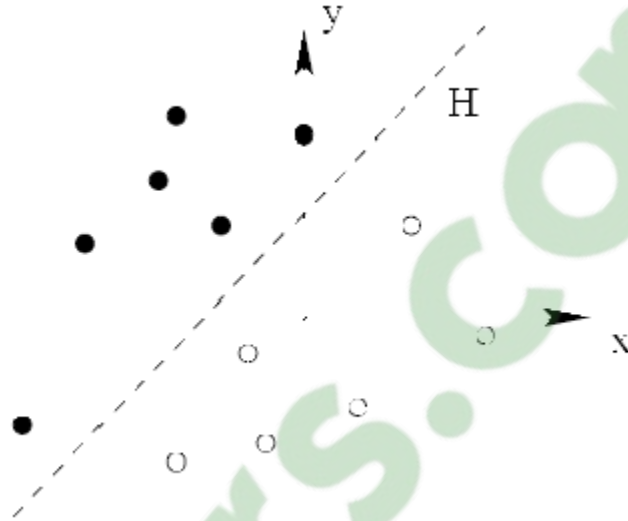
Le **SVM est un algorithme d'apprentissage** qui permet d'arriver à des conclusions par l'examen d'exemples particuliers. Il se divise en apprentissage supervisé et non supervisé. Le SVM utilise l'apprentissage supervisé. Les exemples particuliers sont représentés par un ensemble de couples d'entrée/sortie. Le but est d'apprendre une fonction qui correspond aux exemples vus et qui prédit les sorties pour les entrées qui n'ont pas encore été vues. Les entrées peuvent être des descriptions d'objets et les sorties la classe des objets donnés en entrée.

3.5.2 Le principe de fonctionnement général du SVM

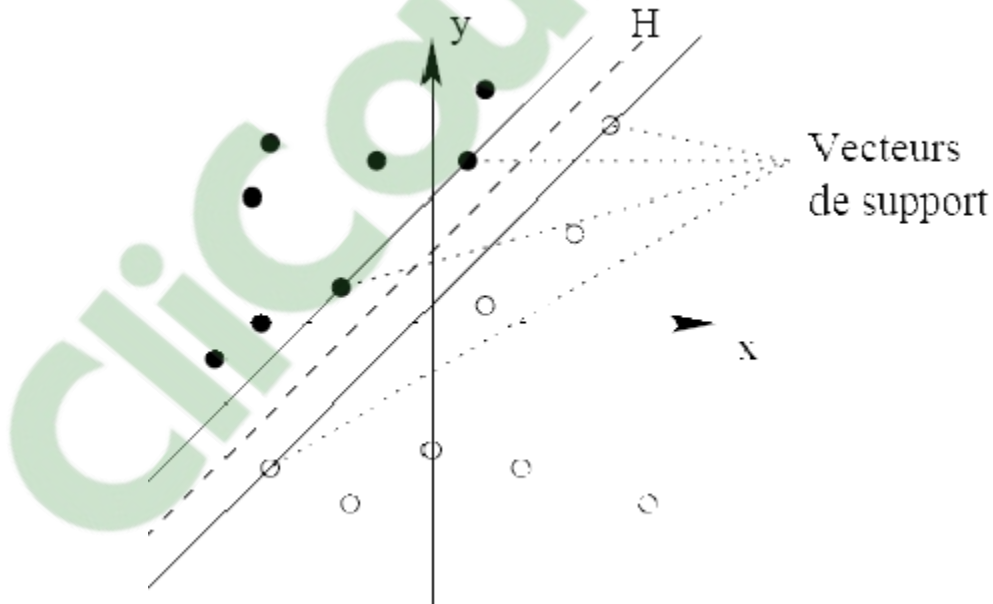
3.5.2.1 Notions de base : Hyperplan, marge et support vecteur

Pour deux classes d'exemples donnés, le but du SVM est de trouver un classificateur qui va séparer les données et maximiser la distance entre ces deux classes. Avec le SVM, ce classificateur est un classificateur linéaire appelé « hyperplan ».

Dans le schéma qui suit, on détermine un hyperplan qui sépare les deux ensembles de points.

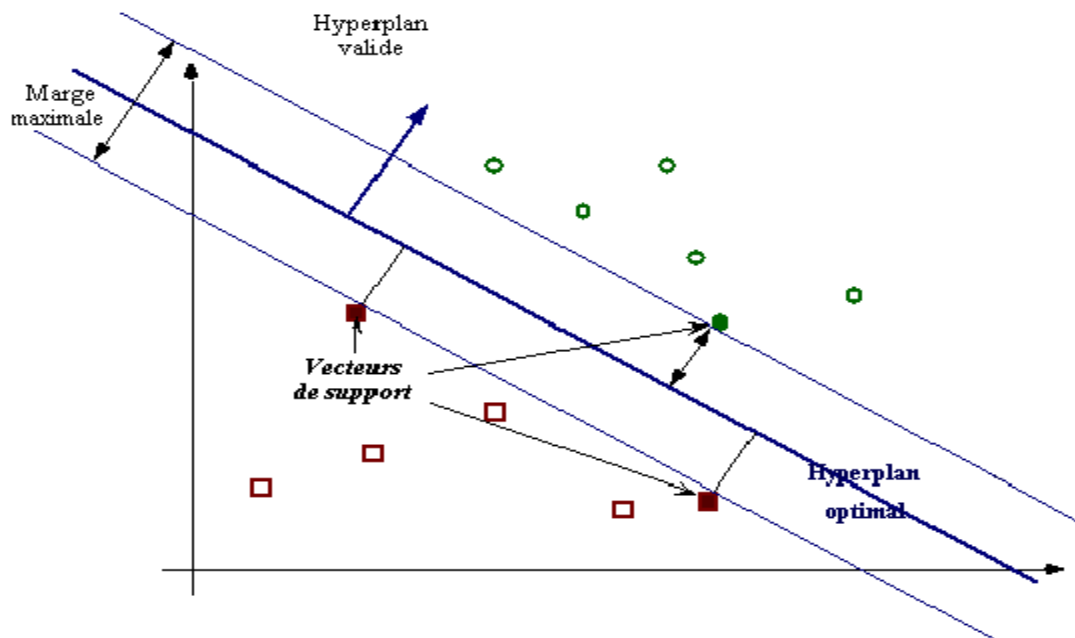


Les points les plus proches, sont les points les plus utilisés pour la détermination de l'hyperplan, ils sont appelés « vecteurs de support ».



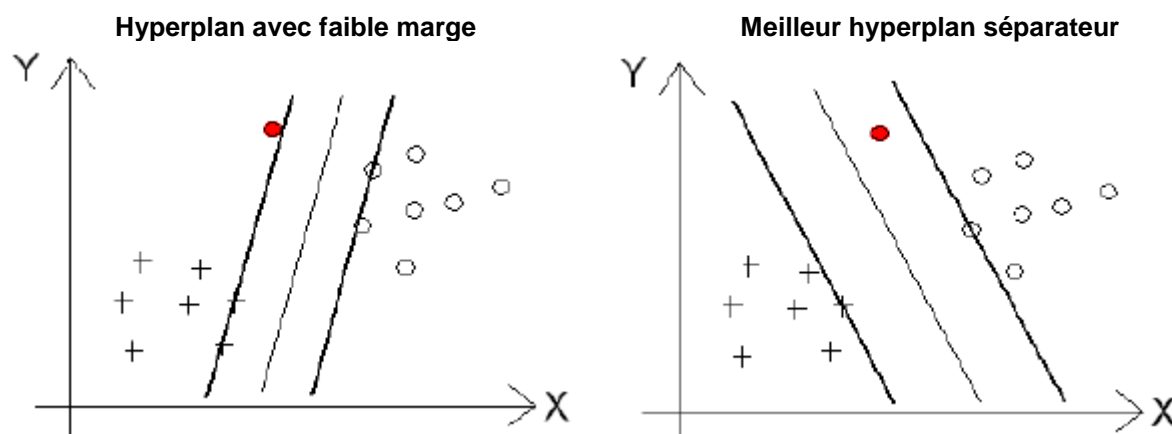
Il est évident qu'il existe une multitude d'hyperplan valide mais la propriété remarquable des SVM est que cet hyperplan doit être optimal. Nous allons donc en plus chercher parmi les hyperplans valides, celui qui passe « au milieu » des points des deux classes d'exemples. Intuitivement, cela revient à chercher l'hyperplan le « plus sûr ». En effet, supposons qu'un exemple n'ait pas été décrit parfaitement, une petite variation ne modifiera pas sa classification si sa distance à l'hyperplan est grande. Formellement, cela revient à chercher un hyperplan dont la distance minimale aux exemples d'apprentissage est maximale.

On appelle cette distance « marge » entre l'hyperplan et les exemples. L'hyperplan séparateur optimal est celui qui maximise la marge. Comme on cherche à maximiser cette marge, on parlera de séparateurs à vaste marge.

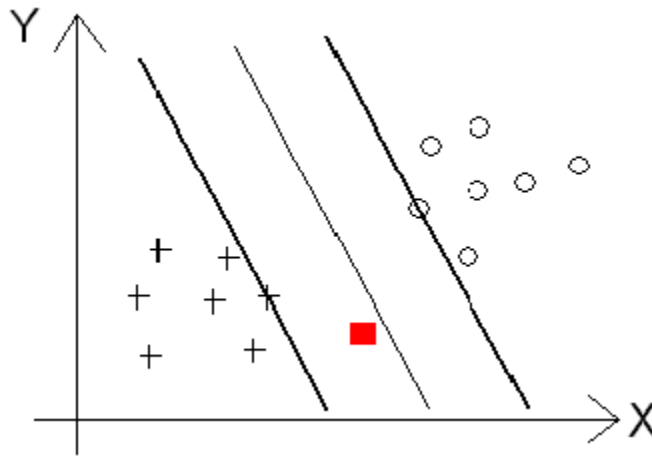


3.5.2.2 Pourquoi faut-il maximiser la marge ?

Intuitivement, le fait d'avoir une marge plus large procure plus de sécurité lorsque l'on classe un nouvel exemple. De plus, si l'on trouve le classificateur qui se comporte le mieux vis-à-vis des données d'apprentissage, il est clair qu'il sera aussi celui qui permettra au mieux de classer les nouveaux exemples. Dans le schéma qui suit, la partie droite nous montre qu'avec un hyperplan optimal, un nouvel exemple reste bien classé alors qu'il tombe dans la marge. On constate sur la partie gauche qu'avec une plus petite marge, l'exemple se voit mal classé. En résumé, la marge est la distance du point le plus proche à l'hyperplan.



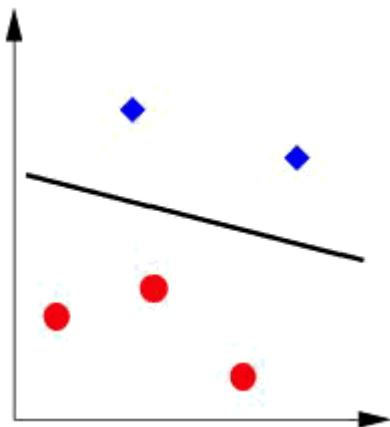
En général, la classification d'un nouvel exemple inconnu est donnée par sa position par rapport à l'hyperplan optimal. Dans le schéma suivant, le nouvel élément sera classé dans la catégorie des « + ».



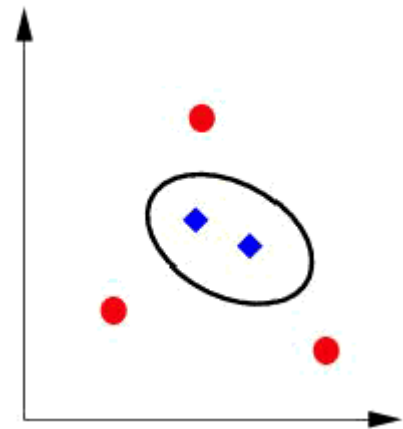
3.5.2.3 Linéarité et non-linéarité

Il existe deux types de séparation pour les modèles du SVM :

- Les cas linéairement séparable : ce sont les plus simples du SVM, car ils permettent de trouver facilement le classificateur linéaire
- et les cas non linéairement séparable, car dans la plupart des problèmes réels il n'y a pas de séparation linéaire possible entre les données, le classificateur de marge maximale ne peut pas être utilisé, il fonctionne seulement si les classes de données d'apprentissage sont linéairement séparables.



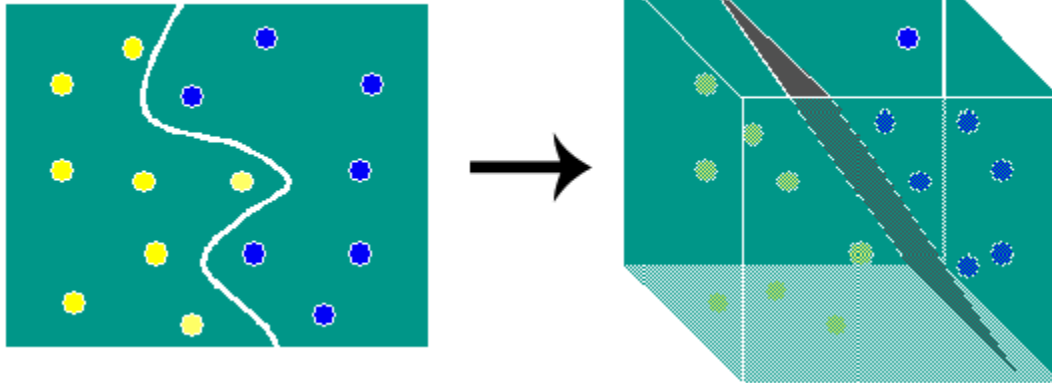
Cas linéairement séparable



Cas non linéairement séparable

Afin de remédier aux inconvénients des cas non linéairement séparable, les SVM pensent à changer l'espace des données. La transformation non linéaire des données peut permettre une séparation linéaire des exemples dans un nouvel espace. On va donc avoir un changement de dimension. Cette nouvelle dimension est appelé «espace de re-description».

En effet, intuitivement, plus la dimension de l'espace de re-description est grande, plus la probabilité de pouvoir trouver un hyperplan séparateur entre les exemples est élevée. Ceci est illustré par le schéma suivant :



Ce schéma nous montre la transformation d'un problème de séparation non linéaire dans l'espace de représentation en un problème de séparation linéaire dans un espace de re-description de plus grande dimension. Cette transformation non linéaire est réalisée via une fonction noyau. En pratique, quelques familles de fonctions noyau paramétrables sont connues et il revient à l'utilisateur de SVM d'effectuer des tests pour déterminer celle qui convient le mieux pour son application. On peut citer les exemples de noyaux suivants : polynomiale, gaussien, sigmoïde et Laplacien.

3.5.3 L'apprentissage

On s'intéresse à un phénomène f (éventuellement non déterministe) qui, à partir d'un certain jeu d'entrées x , produit une sortie $y = f(x)$.

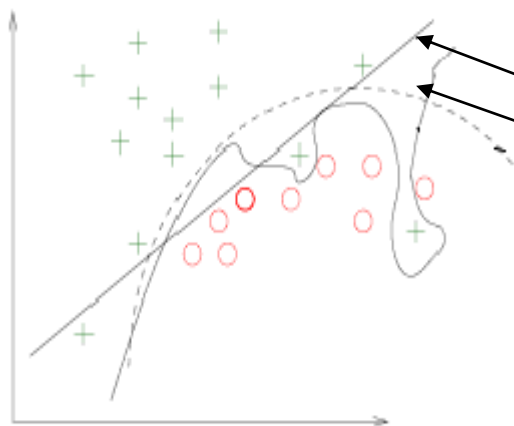
Le but est de retrouver cette fonction f à partir de la seule observation d'un certain nombre de couples entrée-sortie $\{(x_i; y_i) : i = 1, \dots, n\}$ afin de « prédire » d'autres événements.

On considère un couple (X, Y) de variables aléatoires à valeurs dans $X \times Y$.

Seul le cas $Y = \{-1, 1\}$ (classification) nous intéresse ici (on peut facilement étendre au cas $\text{card}(Y) = m > 2$ et au cas $Y = \mathbb{R}$). La distribution jointe de (X, Y) est inconnue.

Sachant qu'on observe un échantillon $S = \{(X_1, Y_1), \dots, (X_n, Y_n)\}$ de n copies indépendantes de (X, Y) , on veut: construire une fonction $h : X \rightarrow Y$ telle que $P(h(X) \neq Y)$ soit minimale.

Exemple de sur et de sous-apprentissage :



Si les données sont générées par un modèle quadratique :

- Le modèle linéaire est en situation de sous apprentissage
- Le modèle de haut degré est en situation de sur apprentissage (apprentissage par cœur)

Il faut donc trouver un compromis entre adéquation aux données et complexité pour pouvoir généraliser.

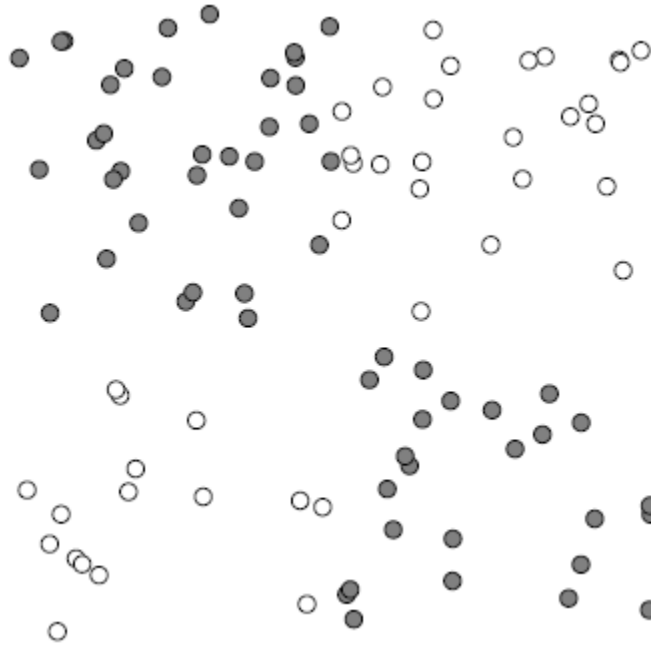
3.5.4 Noyaux

Le gros intérêt des fonctions noyaux est que tout ce qu'on vient de voir sur la séparation linéaire s'applique en fait très facilement à des séparations non linéaires, sous réserve de bien faire les choses.

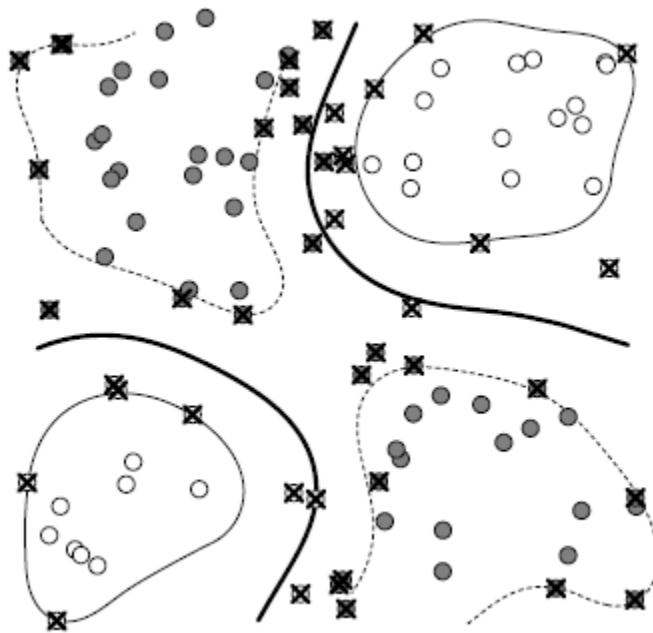
3.5.4.1 L'espace des caractéristiques

Imaginons un ensemble d'exemples x_i étiquetés par -1 ou 1 suivant la classe à laquelle ils appartiennent, qui ne soit pas du tout séparable linéairement. Par conséquent, la méthode de la séparation linéaire ne fonctionne pas, il faut faire appel à des fonctions noyaux.

Voici un exemple de deux classes qui ne peuvent pas être séparées linéairement :



Une solution pour mieux séparer les exemples, est de les projeter dans un espace différent, et de réaliser une séparation linéaire dans cet espace-là, où cette fois-ci elle devrait être plus adaptée. Pour cet exemple, on peut utiliser comme fonction noyau, la fonction gaussienne afin de résoudre le problème d'optimisation, la frontière de séparation sera la ligne épaisse, et les vecteurs supports seront marqués d'une croix.



3.5.4.2 Quelles fonctions sont des noyaux ?

Il serait bien sûr trop beau que n'importe quelle fonction de deux vecteurs puisse être un noyau, il faut en effet qu'il existe une projection dans un espace de caractéristiques dont la fonction donne le même résultat que le produit scalaire des projetés. Bien qu'on ne soit pas obligé de réaliser cette projection, il faut s'assurer de son existence. Il en va par exemple de la convergence de l'algorithme des SVM, car cette convergence n'est assurée que si le problème est convexe, ce qui requiert que le noyau ne soit pas n'importe quoi.

3.5.4.3 Conditions pour avoir un noyau

Il y a des conditions mathématiques, appelées théorème de Mercer, qui permettent de dire si une fonction est un noyau ou non, sans construire la projection dans l'espace des caractéristiques. En fait, il faut assurer que pour tout ensemble d'exemples de longueur l , la matrice suivante soit définie positive :

$$(k(\vec{x}_i, \vec{x}_j))_{1 \leq i, j \leq l}$$

3.6 Conclusion

Dans cette dernière partie du chapitre, nous avons tenté de présenter de manière simple et complète le concept du système d'apprentissage introduit par Vladimir Vapnik, les « Support Vector Machine ».

Cette méthode de classification est basée sur la recherche d'un hyperplan qui permet de séparer au mieux des ensembles de données. Nous avons exposé les cas linéairement séparables et les cas non linéairement séparables qui nécessitent l'utilisation des fonctions noyaux (kernel) pour changer d'espace. Cette méthode est applicable pour des tâches de classification à deux classes, mais il existe des extensions pour la classification multi classe.

Le champ d'application des SVM est donc large et représente une méthode de classification intéressante.

Chapitre 4 : Application du système CAD en imagerie TEP

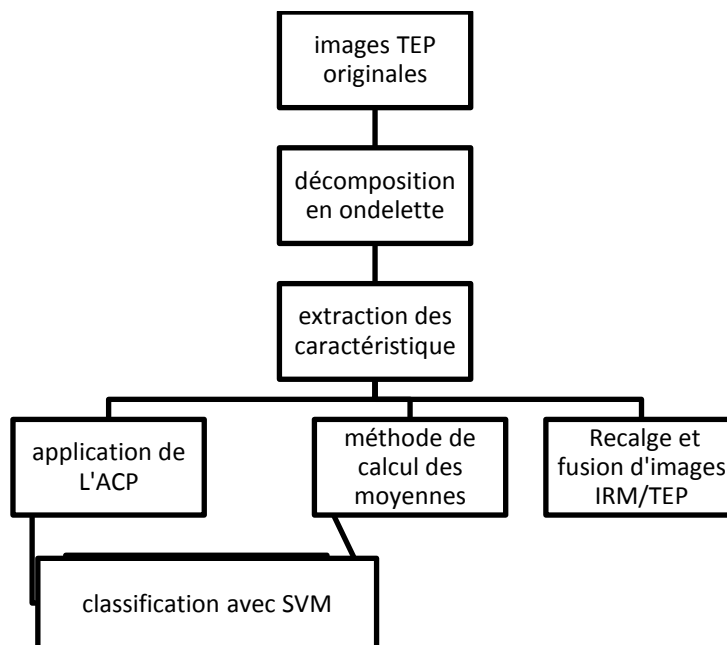
4.1 Introduction

Ce chapitre propose la méthodologie d'un système CAD pour les images TEP neurologiques. Ce système doit être capable de différencier entre les images saines et les images contenant des pathologies, de façon à aider le clinicien dans son diagnostic.

Un système CAD vise à extraire des informations pertinentes des images, ou cas traités, de manière à les envoyer à un classifieur supervisé. Celui-ci permet d'extraire une cartographie des pathologies contenant un nombre négligeable de faux positifs.

Dans un premier temps, un ensemble de caractéristiques de type ondelettes est extrait des images TEP originales. Celles-ci sont envoyées à un classifieur, qui est le SVM, après avoir réduit les données.

Les étapes du système CAD proposé, sont les suivantes :



4.2 La base d'image PET-SORTEO

Dans le cadre de ce mémoire, une base d'images simulées a été utilisée, elle est appelée : « **TEP SORTEO** ». Cette base contient des images TEP cérébrales destinées à la recherche. En l'absence d'une base d'images réelle cette base a permis aux chercheurs de bien avancer dans leurs travaux.

La raison principale du choix de cette base repose sur les temps de simulation affichés par PET-SORTEO qui est réduit par rapport aux autres bases.

SORTEO est un acronyme pour " Simulation Of Realistic Tridimensional Emitting Objects ". Cette plate-forme de simulation est née de la collaboration fructueuse entre le centre d'imagerie CERMEP à Lyon (France) et le Centre d'imagerie cérébrale McConnell à Montréal (Canada). Cet outil de simulation utilise des techniques de Monte Carlo pour produire des données réalistes PET à partir des descriptions de distributions voxelisées des traceurs, conformément à la géométrie et aux caractéristiques physiques du scanner.

PET-SORTEO, prend en compte pour toutes les principales sources de bruit et les préjugés qui peuvent survenir au sein de la formation des images TEP, et autorise la reproduction fidèle du processus de la formation d'image d'un tomographe moderne.

Le modèle de simulation comporte :

- La description numérique du traceur et l'atténuation des distributions.
- La modélisation de l'évolution de la concentration d'activité dans le temps au sein de chaque région émettrice.
- La nature de Poisson en émission de positons, la gamme de positons dans les tissus et les photons d'annihilation non-colinéaire.
- Interactions photoélectrique, Compton et Rayleigh dans les tissus et dans le matériau du scanner.
- Description précise du système de détection (voir fig. 4.1).
- Réglage fin de la sensibilité bloc pour englober les pertes générées par certains effets de détection qui ne sont pas modélisés dans le programme.
- Détecteur d'énergie et de la résolution spatiale de la modélisation en utilisant des fonctions gaussiennes floues.

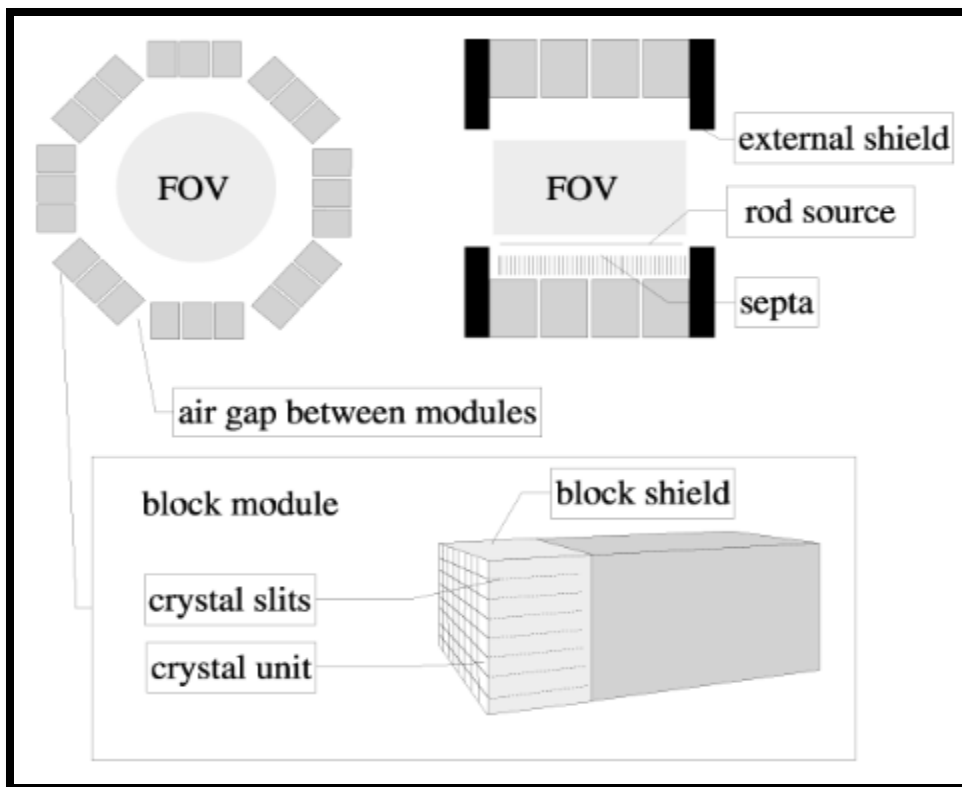


Figure 4.1 : Description du système de détection, le simulateur peut reproduire des données pour des géométries de détection cylindriques ou polygonales. Le simulateur permet de modéliser la géométrie du scanner et des modules.

Pour plus de détails sur la base, voir Annexe.

Cette base est constituée de 134 patients, et chacun possède cinq (5) images, à savoir : une image IRM, labels, [18F]FDG, [18F]Dopa, [11C]Raclopride. Ensuite, nous avons procédé à un

recalage et une fusion de l'image IRM sur l'image TEP marquée au $[18F]FDG$. La figure 4.2, montre les différentes images citées.

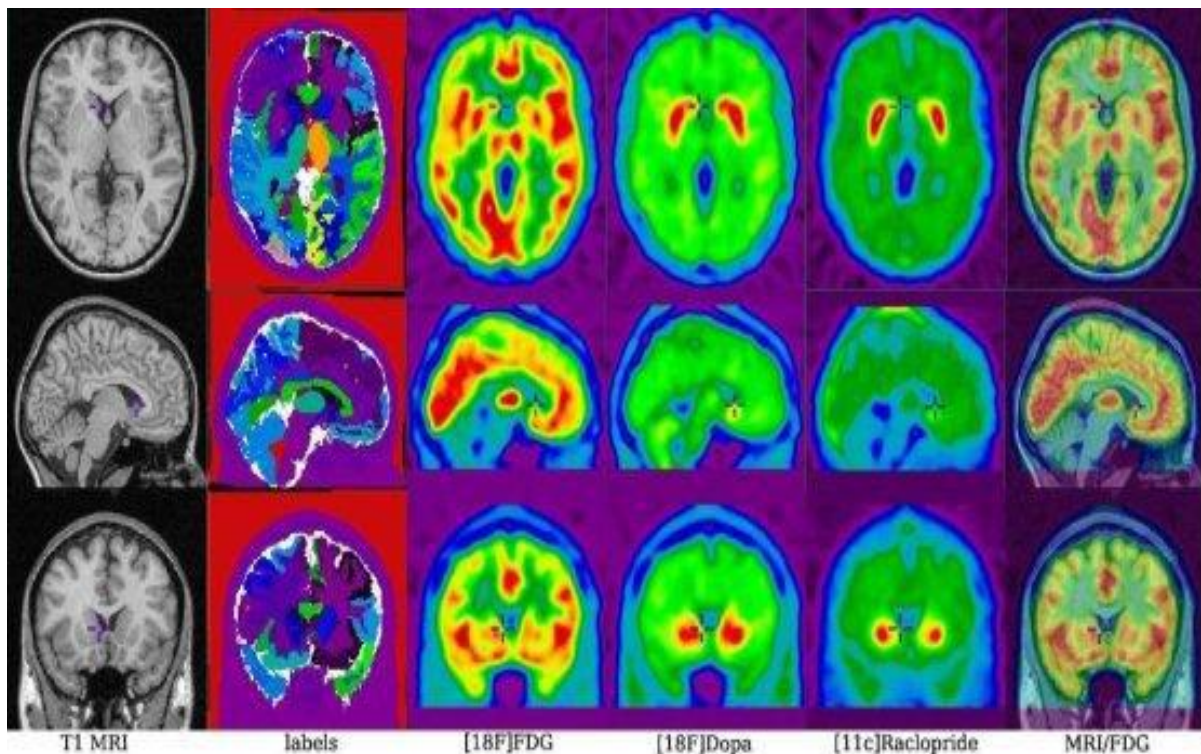


Figure 4-2 : De gauche à droite : coupes d'un modèles anatomique, du fantôme labellisé, d'images simulées au $18F$ -FDG, au $18F$ -Dopa, au $11C$ -Raclopride, et image recalée IRM clinique/TEP simulée.

4.3 La lecture des images

4.3.1 Le format des images

Les images de la base de données PET SORTEO sont stockées sous le format Minc, qui est utilisé pour les images médicales construites au dessus du format de données NetCDF généralisé. Le format est simple, auto-descriptif, extensible, portable et à N dimensions, avec des interfaces de programmation, à la fois, pour l'accès aux données de bas niveau et la manipulation de volume de haut niveau.

En plus des bibliothèques, c'est une suite d'outils génériques de manipulation d'image de fichiers. Les formats, les bibliothèques et les outils sont conçus pour une utilisation dans un environnement d'imagerie médicale de recherche: ils sont simples et puissants et fournissent une bonne interface pour les utilisateurs.

Le fichier MINC est analogue au fichier DICOM, ils sont utilisés tous les deux en médecine parce qu'ils préservent les propriétés des images et évitent leurs dégradations contrairement au format JPG ou JPEG. Aussi, ils contiennent un métafichier spécifique, c'est-à-dire, les images plus les données des patients en un seul fichier. Les images sont codées en binaire et ne sont généralement pas compressées. Quant aux données des patients, elles comprennent leurs noms, leurs âges, leurs antécédents, le type d'examen, le type d'acquisition etc...

4.3.2 Choix du langage de programmation

Dans ce travail, on a utilisé Matlab, qui est un langage de développement informatique dédié particulièrement aux applications scientifiques (imagerie, traitement du signal, etc...). Donc il va nous permettre de traiter les images avec efficacité et rapidité. Il est aussi doté d'un environnement simple et convivial, et il contient de nombreuses boites à outils : les ondelettes, les réseaux de neurones, les statistiques, etc...

Le format MINC peut être lu sous Matlab version 2011 et plus, dont il dispose de plusieurs commandes qui permettent d'extraire les informations contenues dans ces fichiers (les informations concernant le patient : nom, prénom, âge, etc... ; et les informations sur les images : type d'image, l'appareil utilisé, la dose injectée, etc...).

4.4 L'analyse fréquentielle

La TEP est une imagerie fonctionnelle permettant de visualiser les différentes parties du corps humain, les images obtenues ont une très faible résolution spatiale par rapport à d'autres techniques (IRM, TDM, Mammographie, etc...), ces dernières utilisent des méthodes classiques telles que : la segmentation, la morphologie mathématique, le débruitage, etc..., pour l'amélioration de la qualité des images.

Pour remédier à cette faible résolution en imagerie TEP, les méthodes classiques peuvent dégrader la qualité des images obtenues, et par conséquent, des informations peuvent être perdues, ce qui peut mener à une fausse interprétation. Pour éviter ce problème, nous avons opté pour une approche fréquentielle en se basant sur l'analyse par ondelettes.

La décomposition en ondelettes s'effectue en plusieurs niveaux, tout dépend du type d'ondelette utilisé. Cette décomposition permet d'extraire les approximations et les détails.

Nous avons appliqué la décomposition en ondelettes du signal X à un niveau N ($N=3$), en utilisant différents types d'ondelettes : de type Haar, daubechies, biorthogonales, etc... N doit être un entier strictement positif. La structure de décomposition en ondelettes de sortie contient le vecteur caractéristique C et le vecteur S de la comptabilité.

Le vecteur C est appelé vecteur caractéristique, il contient les coefficients d'approximation et les coefficients de détail. Les premiers paramètres du vecteur représentent l'approximation suivis par les détails verticaux, détails horizontaux et détails diagonaux.

Le vecteur S donne la taille de chaque matrice de coefficients, comme illustré sur la figure 4.3.

Dans notre étude, nous avons remarqué que l'analyse au troisième niveau de décomposition avec l'ondelette de Bior4.4, a permis la réduction de la taille du vecteur caractéristique sans perte d'information.

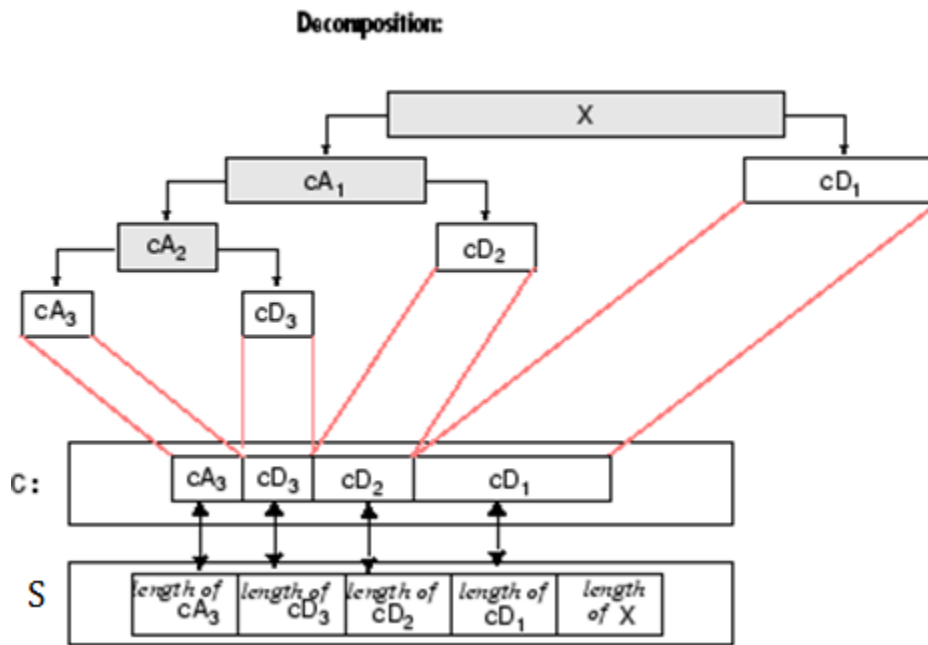


Figure 4.3 : Structure du signal X au troisième niveau de décomposition.

4.4.1 Extraction des caractéristiques

L'analyse des images se focalise généralement autour des attributs de bas niveaux (la texture, la forme, la couleur).

Il existe une approche de caractérisation **locale**, qui consiste à calculer des attributs sur des portions restreintes de l'image.

L'avantage de ces descripteurs locaux est de conserver l'information localisée dans l'image, en évitant que certains détails ne soient noyés. Et l'inconvénient, est que la quantité d'information produite est très grande.

On utilise dans ce travail cette approche locale, car elle est basée sur le choix de la texture et de la forme par la méthode d'ondelettes.

4.4.2 Algorithme de décomposition en ondelette

Pour un signal S de longueur N , la transformée en ondelette discrète se compose de $\log_2 N$ étages au maximum.

- ❖ Une première étape, engendre à partir du signal S deux ensembles de coefficients.

Les approximations cA_j et les détails cD_j . Ces coefficients sont obtenus par convolution du signal S avec un filtre passe-bas Lo_D .

La convolution du signal S avec un filtre passe-bas (low-pass filter) Lo_D permet l'obtention des coefficients d'approximation, et la convolution du signal S avec un filtre passe-haut (high-pass filter), suivie d'une décimation dyadique (sous-échantillonnage), permet l'obtention des détails.

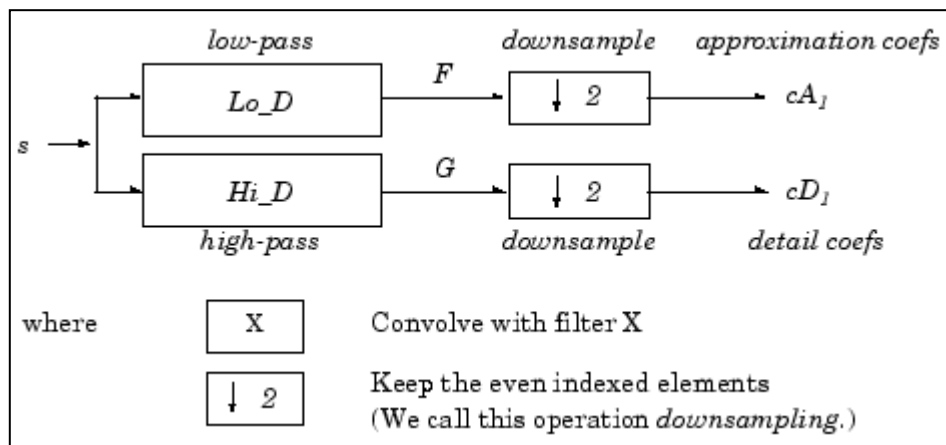


Figure 4.4 : Première étape de décomposition

La longueur de chaque filtre est égale à $2N$, si $n = \text{length}(S)$, alors il est considéré comme étant la longueur du signal S . Les deux signaux F et G de la fig. 4.4, ont une longueur de $n + 2(N - 1)$, et les coefficients cA_j et cD_j ont une longueur de :

$$\text{floor} \left(\frac{n-1}{2} \right) + N$$

- ❖ La seconde étape est la division des coefficients d'approximations cA_j en deux parties, selon la figure 4.4, en remplaçant cette fois le signal S par l'approximation cA_1 ce qui permet d'obtenir les coefficients d'approximations et les coefficients de détails au deuxième niveau. Cette étape est reproduite selon le niveau de décomposition voulu.

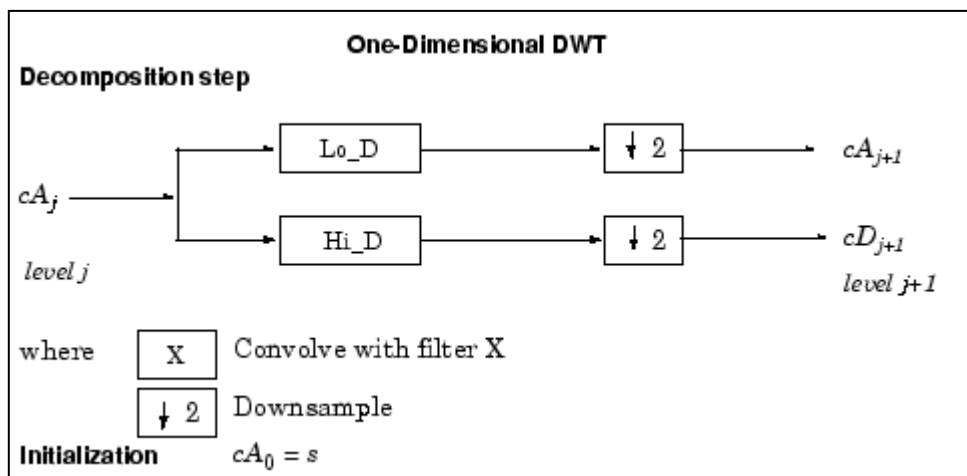


Figure 4.5 : la décomposition au j^{ème} niveau de la 2^{ème} étape

La décomposition en ondelette du signal S au niveau J est du type $[cA_j, cD_j, \dots, cD_1]$

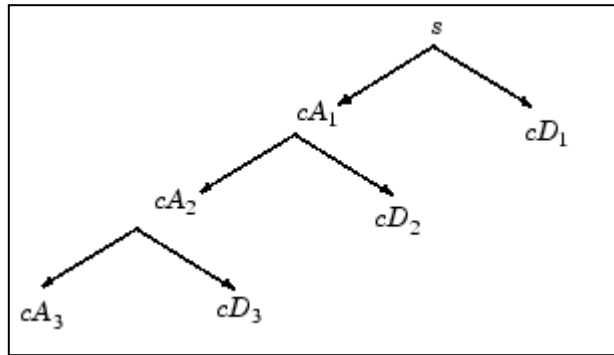


Figure 4.6 : décomposition en ondelette au niveau $J=3$

Dans un premier temps, nous avons décomposé l'image en utilisant différents types d'ondelettes. Nous avons varié les niveaux de décomposition (en allant du 1^{er} niveau jusqu'au 4^{ème}), afin de voir celui qui correspond le mieux à notre étude. Les résultats obtenus sont illustrés sur les figures 4.7, 4.8, et 4.9.

On a commencé par la variation du niveau de décomposition en utilisant l'ondelette de Haar.

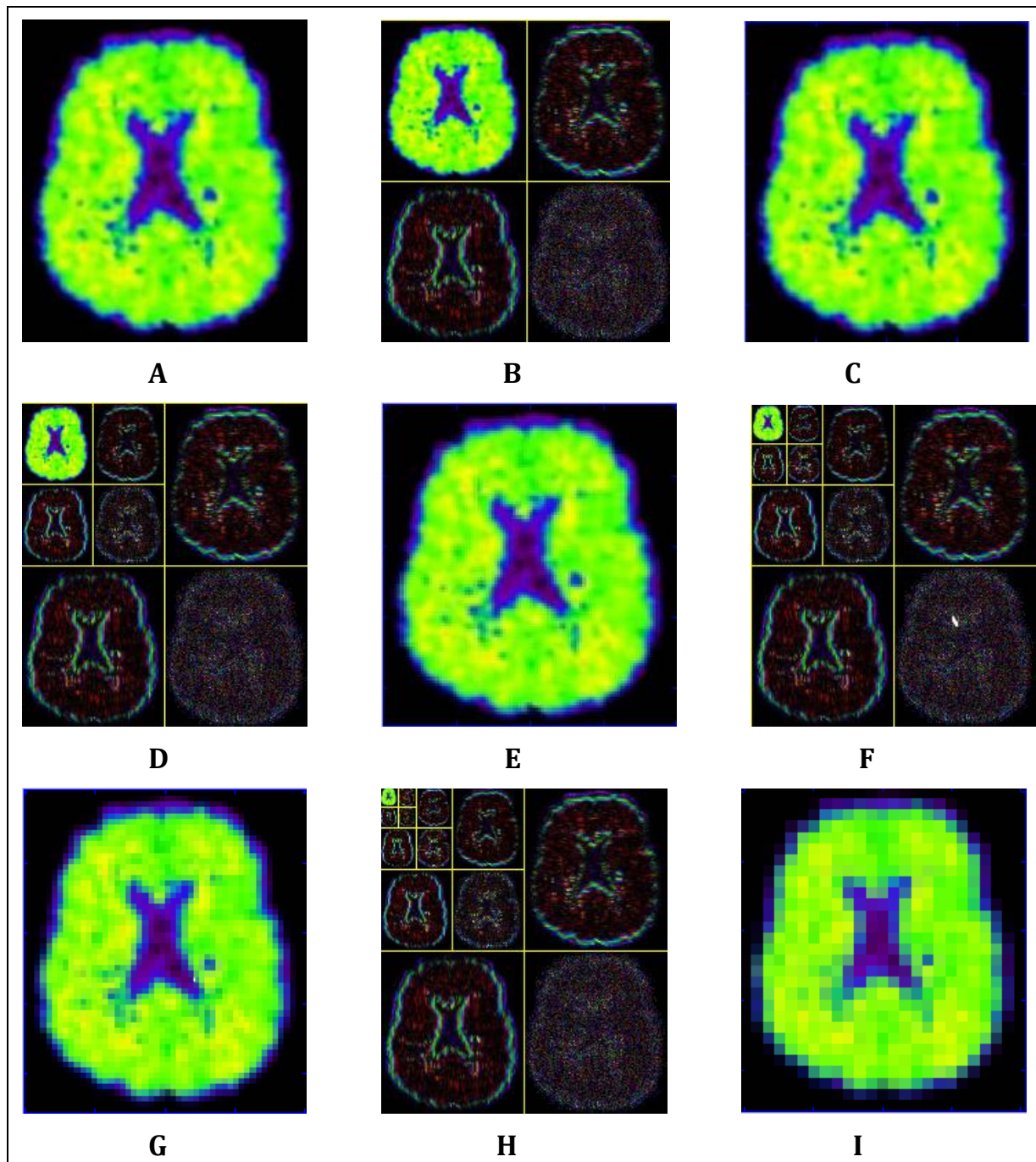


Figure 4.7: la décomposition en ondelette de Haar selon plusieurs niveaux. A : représente l'image originale. B: Le 1^{er} niveau de décomposition de l'image. C : L'image approximation du 1^{er} niveau. D: Le 2^{ème} niveau de décomposition de l'image. E: L'image approximation du 2^{ème} niveau de décomposition. F: Représente le 3^{ème} niveau de décomposition. G: L'image approximation du 3^{ème} niveau. H: Représente le 4^{ème} niveau de décomposition de l'image. I: L'image approximation du 4^{ème} niveau.

Nous avons par la suite utilisé l'ondelette « coif2 », et nous avons repris le même processus que pour l'ondelette de Haar.

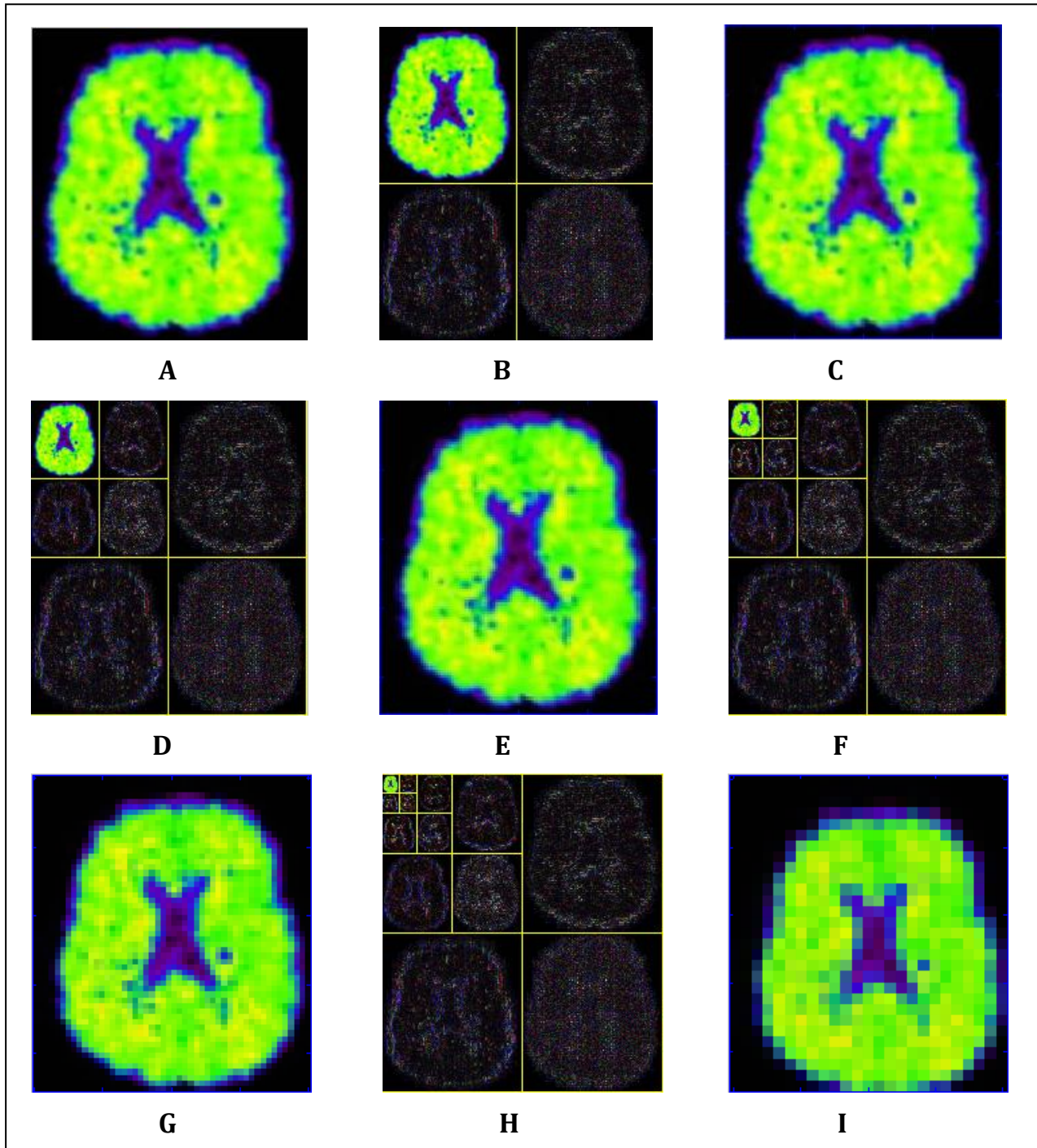


Figure 4.8 : La décomposition en ondelette coif2 selon différents niveaux. A: Représente l'image originale. B: La décomposition au 1^{er} niveau. C: L'image approximation du 1^{er} niveau. D: La décomposition au 2^{ème} niveau. E: L'image approximation du 2^{ème} niveau. F: La décomposition au 3^{ème} niveau. G: L'image approximation du 3^{ème} niveau. H: La décomposition au 4^{ème} niveau. I: L'image approximation du 4^{ème} niveau.

Et finalement, nous avons varié les niveaux de décomposition en utilisant l'ondelette de « Bior4.4 »

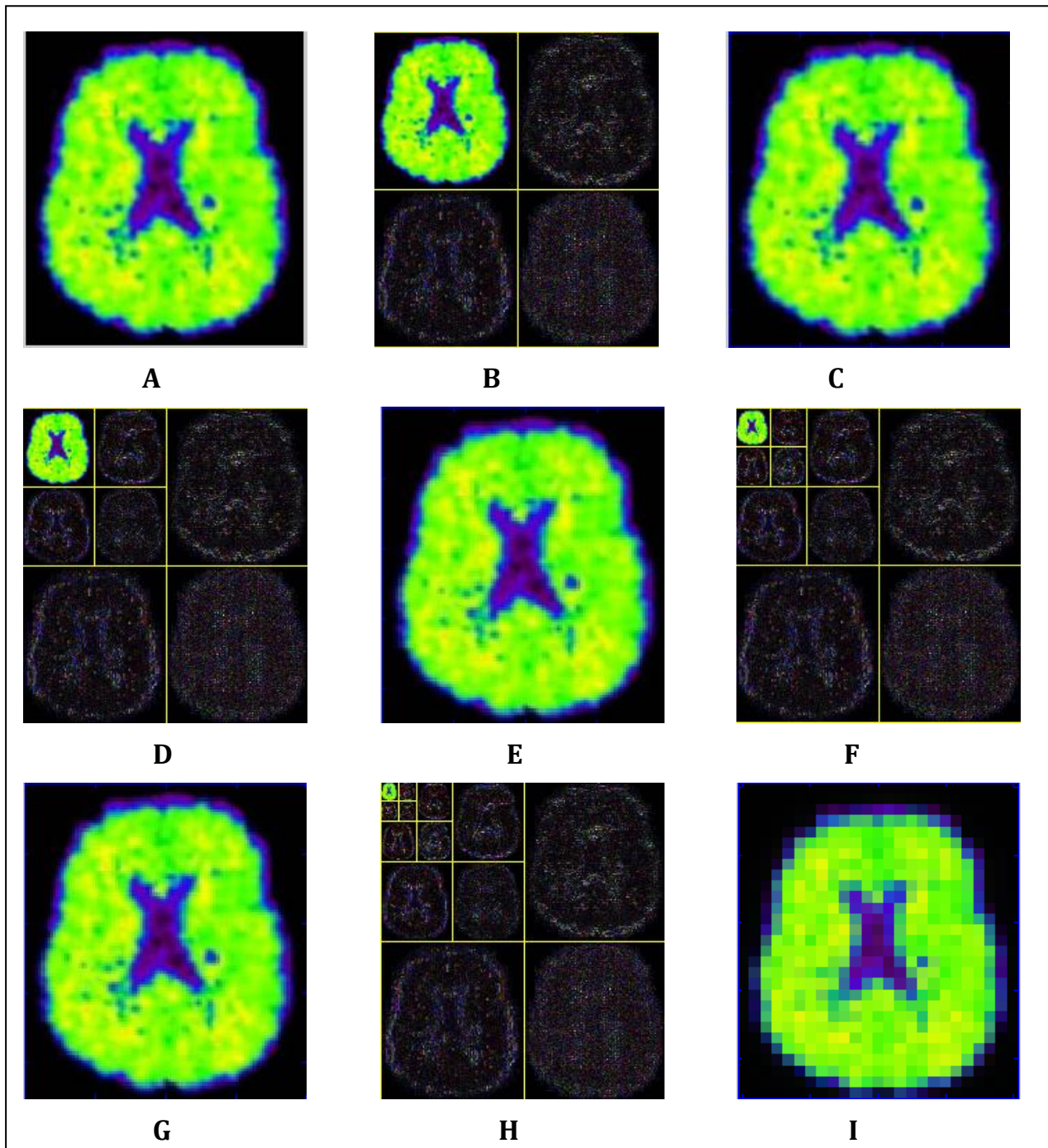


Figure 4.9: La décomposition en ondelette Bior4.4 selon différents niveaux de décomposition. A: Représente l'image originale. B : La décomposition au niveau 1. C: L'image approximation du 1^{er} niveau. D: La de l'image décomposition au 2^{ème} niveau. E: L'image approximation du 2^{ème} niveau de décomposition. F: La décomposition de l'image au 3^{ème} niveau. G : L'image approximation du 3^{ème} niveau. H: La décomposition de l'image au 4^{ème} niveau. I: Représente l'image approximation du 4^{ème} niveau de décomposition.

L'image originale représente une image cérébrale d'un patient atteint d'une démence. Nous avons appliqué à cette image une décomposition en ondelettes (Haar, coif2, et Bior4.4).

Nous avons commencé la décomposition par le niveau 1, on a obtenu une approximation et trois détails (horizontal, vertical, diagonal), les résultats obtenus montrent que l'approximation a presque la même résolution que l'image originale et que les détails montrent le contour de l'image et certaines parties du cerveau. Ensuite on est passée au 2^{ème} niveau de décomposition, elle consiste à décomposer l'approximation résultante du 1^{er} niveau en nouvelle approximation et d'autres détails, cette nouvelle approximation possède une faible résolution par rapport à celle du premier niveau (le nombre de coefficients diminue).

Plus on augmente le niveau de décomposition, plus la résolution devient faible. En arrivant au troisième niveau on a remarqué que pour les ondelettes de Haar et Coif2 les approximations obtenues sont de très faibles résolutions, ce qui mène à une très grande perte d'information. Par contre avec l'ondelette bior4.4 l'approximation du 3^{ème} niveau est de résolution moins faible par rapport aux autres approximations obtenues précédemment.

Afin de déterminer le type d'ondelette le plus approprié pour l'analyse des images TEP cérébrales, on a appliqué la décomposition au 3^{ème} niveau avec l'ondelette de Haar, l'ondelette de coif2 et l'ondelette de Bior4.4, sur une image d'un patient atteint de la maladie de parkinson.

4.4.3 Utilisation de l'ondelette de Haar

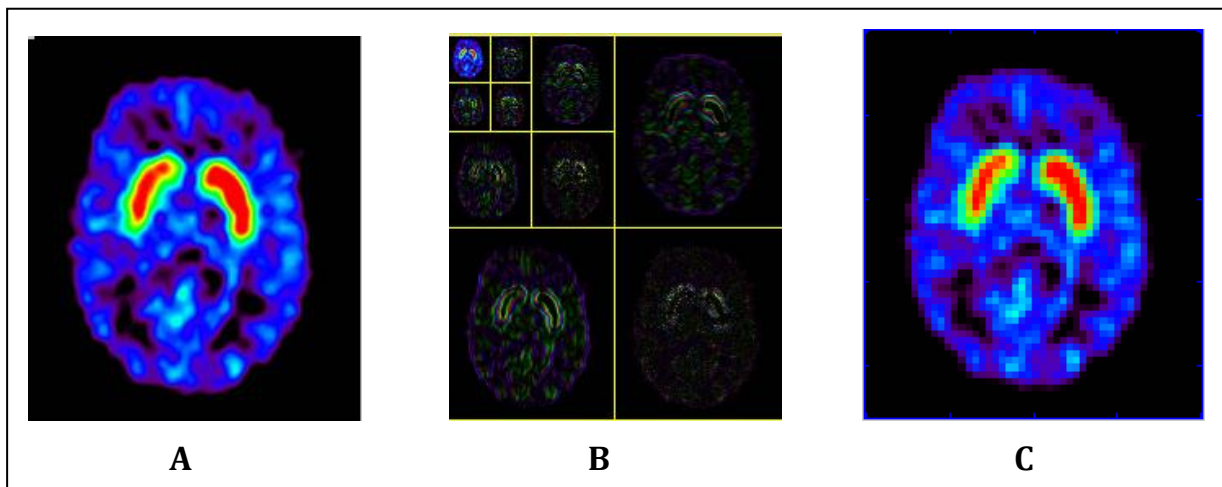


Figure 4.10 : utilisation de l'ondelette de Haar lors de la décomposition en ondelette. A : image originale, B : image décomposée en approximation et détails au 3^{ème} niveau, et C : image approximation extraite après la décomposition.

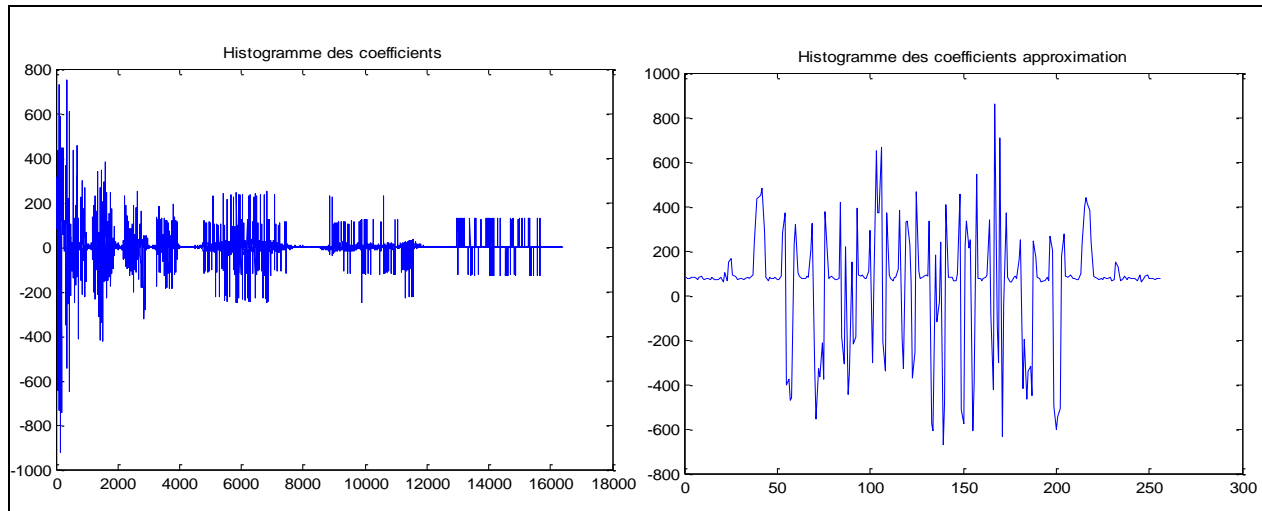


Figure 4.11 : Les deux histogrammes ci-dessus représentent respectivement l'histogramme de l'image après décomposition en ondelette de Haar au 3^{ème} niveau et l'histogramme de l'approximation de l'image résultante de la décomposition.

La décomposition avec l'ondelette de Haar donne une approximation floue et de très faible résolution, et les détails font apparaître les contours de l'image et les noyaux gris centraux.

4.4.4 Utilisation de l'ondelette coif2

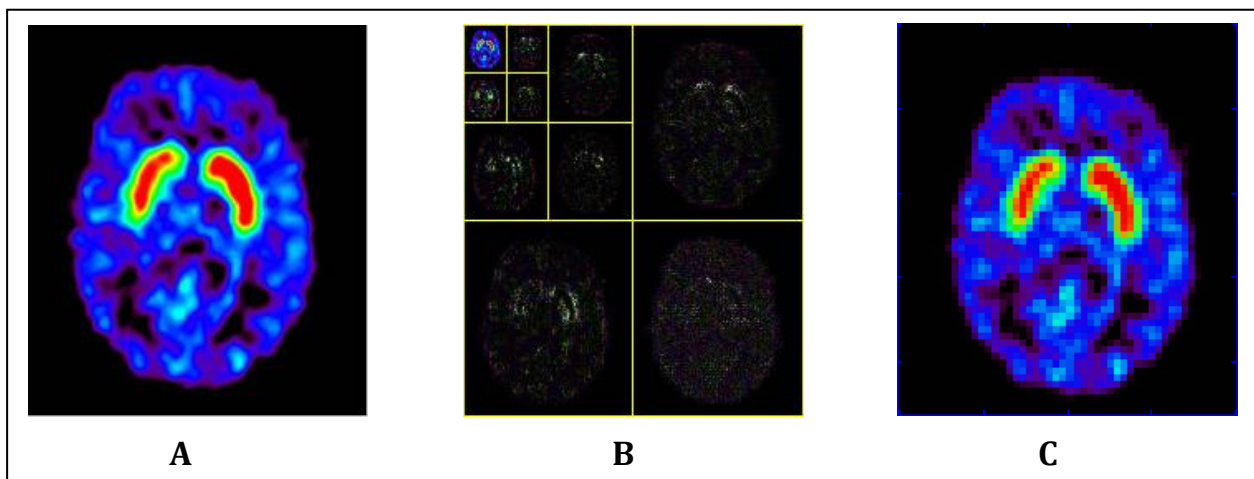


Figure 4.12: images représentatives des résultats lors de l'utilisation de l'ondelette coif2 pour la décomposition. *A:* image avant décomposition. *B:* image après décomposition en ondelette au 3^{ème} niveau. *C:* l'image approximation de la décomposition au 3^{ème} niveau.

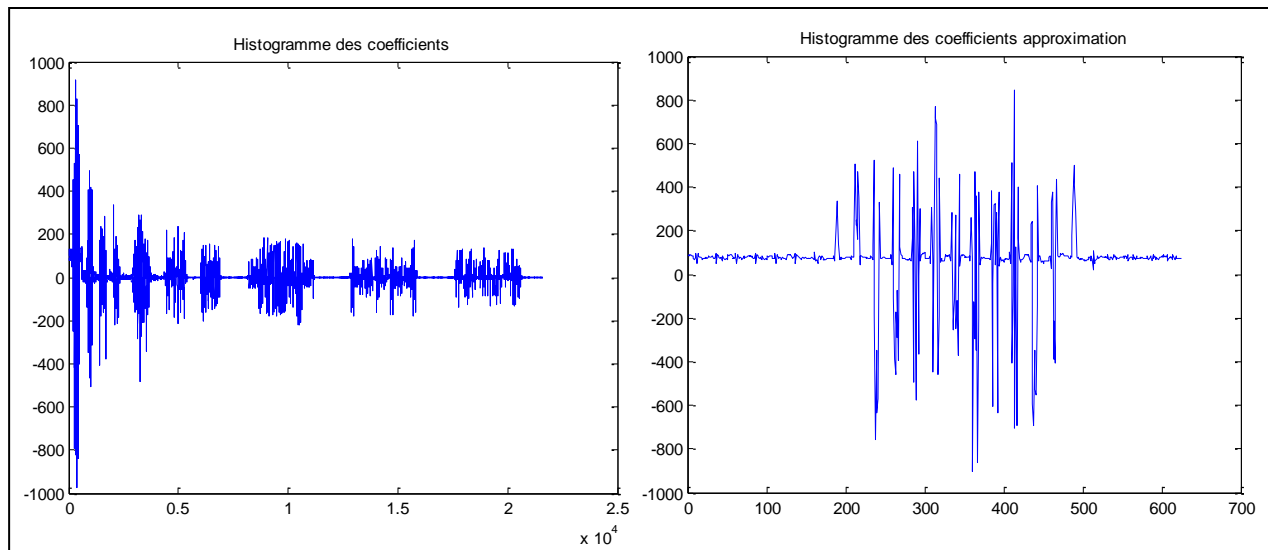


Figure 4.13 : Histogrammes représentant l'image après décomposition et l'image approximation obtenue avec l'ondelette *coif2*.

La décomposition avec l'ondelette *coif2* donne elle aussi comme résultat, une approximation de très faible résolution, et les détails ne font presque rien apparaître.

4.4.5 Utilisation de l'ondelette de *Bior4.4*

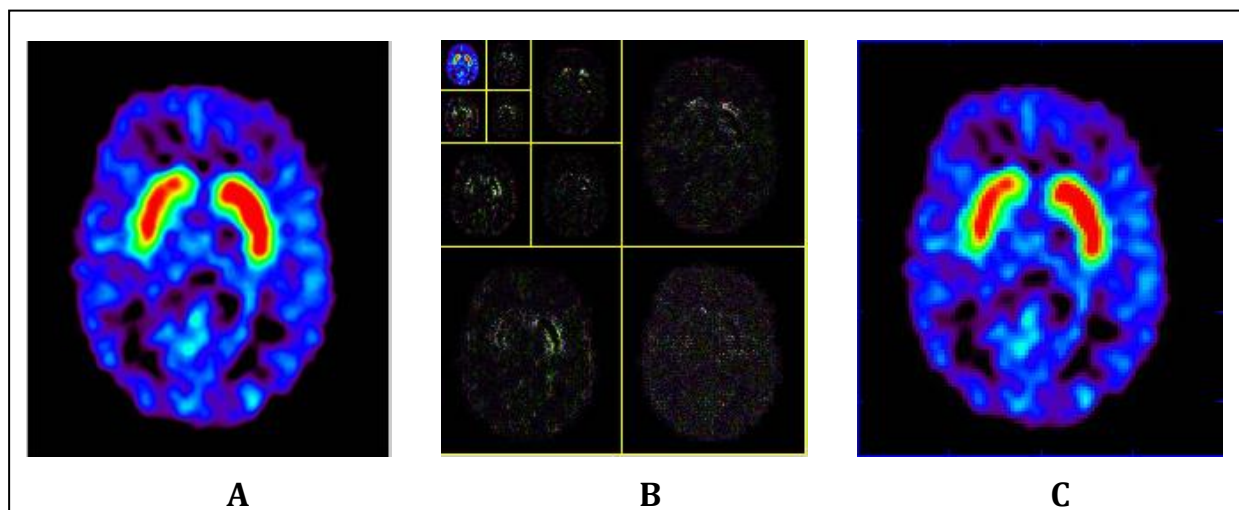


Figure 4.14 : images représentatives des résultats lors de l'utilisation de l'ondelette *Bior4.4* pour la décomposition. A: image avant décomposition. B: image après décomposition en ondelette au 3^{ème} niveau. C: l'image approximation de la décomposition au 3^{ème} niveau.

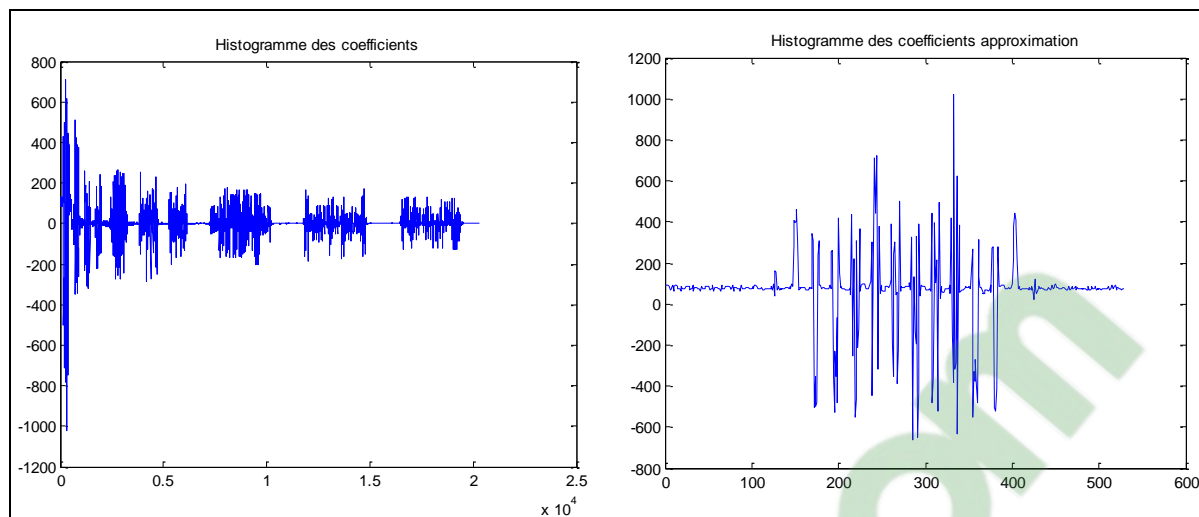


Figure 4.15 : Histogramme de l'image après décomposition en ondelette et histogramme de l'image approximation obtenue avec l'ondelette de Bior4.4.

La décomposition avec cette ondelette a donné un très bon résultat. On remarque que l'image de l'approximation est très claire, et elle garde le maximum d'information.

On peut prendre en considération que l'approximation, les détails peuvent être omis car ils ne contiennent pas d'informations utiles.

4.5 La réduction des paramètres

La décomposition en ondelette a permis l'extraction des caractéristiques fréquentielles dont on n'a retenu que l'image approximation. Les paramètres d'approximation obtenus composent les valeurs du vecteur caractéristique.

On se base souvent sur deux familles de méthodes lors du choix des caractéristiques à envoyer au classifieur. La première famille se base sur l'utilisation de certaines caractéristiques, par contre la deuxième regroupe des méthodes de réduction de la dimension.

L'analyse en composantes principales (ACP) est la méthode la plus utilisée, un nouvel espace est généré dans lequel les anciennes composantes du vecteur caractéristique sont projetées.

Cependant, on a utilisé une 2nde méthode qui consiste à calculer la moyenne.

On ne peut pas juger de la fiabilité des résultats de ces deux méthodes jusqu'à la dernière étape de notre travail, et qui se base sur la classification.

4.5.1 Analyse en composantes principales

L'Analyse en Composantes Principales consiste à transformer des variables liées entre elles (dites "corrélées" en statistique) en nouvelles variables décorréelées les unes des autres. Ces nouvelles variables sont nommées "composantes principales", ou axes. Elle permet d'extraire l'essentiel de l'information contenue dans une image et d'en fournir une représentation se prêtant plus aisément à l'interprétation.

Il s'agit d'une approche à la fois géométrique (représentation des variables dans un nouvel espace géométrique selon des directions d'inertie maximale) et statistique (recherche d'axes indépendants expliquant au mieux la variabilité - la variance - des données).

Cette méthode a été utilisée en 2008 par Papadopoulos, pour le développement des systèmes CAD en imagerie médicale. Son avantage c'est qu'elle est utilisable pour n'importe quel classifieur, elle implique l'extraction de l'ensemble des caractéristiques initiales à chaque étape du CAD.

4.5.1.1 Algorithme de L'ACP

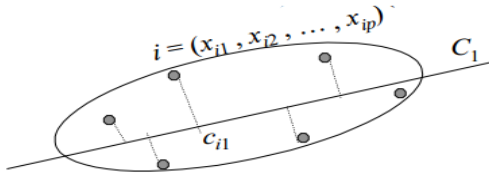
Soient X_1, X_2, \dots, X_p les variables d'origines

C_k , des nouvelles variables résultantes de la combinaison linéaire des variables d'origines.

$C_k = a_{1k} X_1 + a_{2k} X_2 + \dots + a_{pk} X_p$. Les coefficients a_{jk} sont à déterminer.

Les C_k doivent être 2 à 2 non corrélés, de variance maximale et d'importance décroissante C_1 la première composante principale doit avoir la plus grande variance.

Géométriquement : C_1 détermine une nouvelle direction dans le nuage de points qui suit l'axe d'allongement maximale du nuage.



c_{i1} = coordonnée du point i sur l'axe C_1
projection de x_i sur C_1

$$c_{i1} = \sum_{j=1}^p a_{1j} x_{ij}$$

Si C_1 est de variance maximale alors les projections c_{i1} sont les plus dispersées possibles.

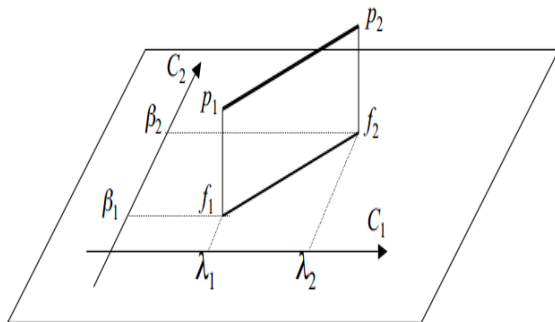
Pour fixer la droite, on impose qu'elle passe par le centre de gravité (g) sinon l'ensemble des droites parallèles convient.

C_1 : droite passant par le centre de gravité réalisant le meilleur ajustement possible du nuage c-à-d : qui conserve au mieux la distance entre les points (après projection) et assure une distorsion minimale.

C_2 est la 2^{ème} composante, orthogonale à C_1 et de variance maximale.

Géométriquement: C_2 détermine une droite perpendiculaire à C_1 (au point g), suivant un axe d'allongement maximum.

Les composantes C_1 et C_2 déterminent le plan principal : le meilleur plan de projection (de distorsion minimum).



C_1 est telle que la moyenne des $d^2(\lambda_i, \lambda_i)$ max.

C_2 est \perp à C_1 et telle que la moyenne des $d^2(\beta_i, \beta_i)$ max.

$\Rightarrow C_1$ et C_2 déterminent le plan tel que $d^2(f_i, f_i)$ soit maximum.

$\Rightarrow C_3$ est la droite \perp à C_1 et C_2 (par g) telle que la variance des coord. soit maximum ...

Décomposition de la variance

La variance se décompose de la manière suivante:

$$\begin{aligned}
\sigma^2 &= \frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{j=1}^n (\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j) \\
&= \frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{j=1}^n ((\mathbf{x}_i - \mathbf{g}) + (\mathbf{g} - \mathbf{x}_j))^T ((\mathbf{x}_i - \mathbf{g}) + (\mathbf{g} - \mathbf{x}_j)) \\
&= \frac{1}{2n(n-1)} \sum_{i=1}^n \sum_{j=1}^n \left((\mathbf{x}_i - \mathbf{g})^T (\mathbf{x}_i - \mathbf{g}) + (\mathbf{g} - \mathbf{x}_j)^T (\mathbf{g} - \mathbf{x}_j) + 2(\mathbf{x}_i - \mathbf{g})^T (\mathbf{g} - \mathbf{x}_j) \right) \\
&= \frac{1}{(n-1)} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{g})^T (\mathbf{x}_i - \mathbf{g})
\end{aligned}$$

Π est l'opérateur de projection orthogonale sur une droite de vecteur directeur unitaire \mathbf{v} s'écrit :

$$\Pi = \mathbf{v} * \mathbf{v}^T \quad \text{avec} \quad \mathbf{v}^T * \mathbf{v} = 1$$

La variance des observations projetées s'écrit alors :

$$\sigma_{\mathbf{v}}^2 = \frac{1}{(n-1)} \sum_{i=1}^n (\boldsymbol{\pi}(\mathbf{x}_i - \mathbf{g}))^T (\boldsymbol{\pi}(\mathbf{x}_i - \mathbf{g}))$$

Nous avons donc :

$$\begin{aligned}
\sigma_{\mathbf{v}}^2 &= \frac{1}{(n-1)} \sum_{i=1}^n (\boldsymbol{\pi}(\mathbf{x}_i - \mathbf{g}))^T (\boldsymbol{\pi}(\mathbf{x}_i - \mathbf{g})) \\
&= \frac{1}{(n-1)} \sum_{i=1}^n (\mathbf{v} \mathbf{v}^T (\mathbf{x}_i - \mathbf{g}))^T (\mathbf{v} \mathbf{v}^T (\mathbf{x}_i - \mathbf{g})) \\
&= \frac{1}{(n-1)} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{g})^T (\mathbf{v} \mathbf{v}^T \mathbf{v} \mathbf{v}^T) (\mathbf{x}_i - \mathbf{g}) \\
&= \frac{1}{(n-1)} \sum_{i=1}^n (\mathbf{x}_i - \mathbf{g})^T \mathbf{v} \mathbf{v}^T (\mathbf{x}_i - \mathbf{g})
\end{aligned}$$

Alors :

$$\begin{aligned}
\sigma_{\mathbf{v}}^2 &= \frac{1}{(n-1)} \sum_{i=1}^n ((\mathbf{x}_i - \mathbf{g})^T \mathbf{v}) (\mathbf{v}^T (\mathbf{x}_i - \mathbf{g})) \\
&= \frac{1}{(n-1)} \sum_{i=1}^n \mathbf{v}^T (\mathbf{x}_i - \mathbf{g}) (\mathbf{x}_i - \mathbf{g})^T \mathbf{v} \\
&= \frac{1}{(n-1)} \mathbf{v}^T \left[\sum_{i=1}^n (\mathbf{x}_i - \mathbf{g}) (\mathbf{x}_i - \mathbf{g})^T \right] \mathbf{v} \\
&= \mathbf{v}^T \boldsymbol{\Sigma} \mathbf{v}
\end{aligned}$$

$\boldsymbol{\Sigma}$ est positive et symétrique

Nous observons que la matrice $\boldsymbol{\Sigma}$ est la matrice variance-covariance qui devra être maximisée à cause du problème d'optimisation sous contrainte

$$\max_{\mathbf{v}} (\mathbf{v}^T \boldsymbol{\Sigma} \mathbf{v}) \quad \text{subject to} \quad (\mathbf{v}^T \mathbf{v} = 1)$$

Et pour cela il faudra former la fonction de Lagrange :

$$\mathcal{L} = \mathbf{v}^T \boldsymbol{\Sigma} \mathbf{v} + \lambda (1 - \mathbf{v}^T \mathbf{v})$$

Avec $\partial_{\mathbf{v}} \mathcal{L} = 0$ Condition d'optimisation.

On obtient alors l'équation aux valeurs propres $\boldsymbol{\Sigma} \mathbf{v} = \lambda \mathbf{v}$

Comme la matrice variance-covariance est positive et symétrique, les valeurs propres sont réelles et positives.

La variance s'écrit alors :

$$\begin{aligned}
\sigma_{\mathbf{v}}^2 &= \mathbf{v}^T \boldsymbol{\Sigma} \mathbf{v} \\
&= \mathbf{v}^T \lambda \mathbf{v} \\
&= \lambda
\end{aligned}$$

Et donc la solution est de projeter les données sur le vecteur propre ayant la valeur propre λ la plus élevée.

Afin de trouver le second axe de variance maximale, nous recherchons

$$\max_{\mathbf{v}} (\mathbf{v}^T \Sigma \mathbf{v}) \text{ subject to } (\mathbf{v}^T \mathbf{v} = 1) \text{ and } (\mathbf{v}^T \mathbf{v}_1 = 0)$$

Avec \mathbf{v}_1 étant le premier vecteur propre à valeur propre maximale.

Comme les vecteurs propres de Σ sont naturellement orthonormés, la solution est de choisir le deuxième vecteur propre de Σ (à deuxième valeur propre maximale).

Notons que si X est la matrice de données qui contient les vecteurs $(x_i - g)^T$ en ligne

La matrice $\Sigma = (n - 1)^{-1} X^T X$

La somme des valeurs propres correspond à la variance totale

$$\text{tr}(\Sigma) = \sigma^2 = \sum_{i=1}^n \lambda_i$$

Chaque valeur propre mesure la part de variance expliquée par l'axe factoriel correspondant.

4.5.2 Calcul de la moyenne

En raison de la dimension de la matrice caractéristique on a procédé à un calcul des moyennes de cette dernière afin d'éviter le plus possible, la perte d'information liée à d'autres méthodes.

Cette méthode est très simple et consiste à calculer la moyenne de chaque colonne de la matrice $[M \times N]$. Le vecteur obtenu garde le même aspect que la matrice, avec moins de paramètres et moins de perte. Ce vecteur a une dimension de $[1 \times N]$.

✚ Au départ la matrice obtenue était d'une dimension de 23×23 . Après application de l'ACP on a obtenu un vecteur de 23 colonnes et donc de dimension 1×23 et c'est d'ailleurs la méthode appliquant le calcul des moyennes.

Nous avons appliqué ces deux méthodes pour réduire les paramètres du vecteur caractéristique obtenu après décomposition en ondelette Bior4.4. Ce vecteur caractéristique est représenté par les coefficients d'approximation de l'image. Voici quelques résultats obtenus :

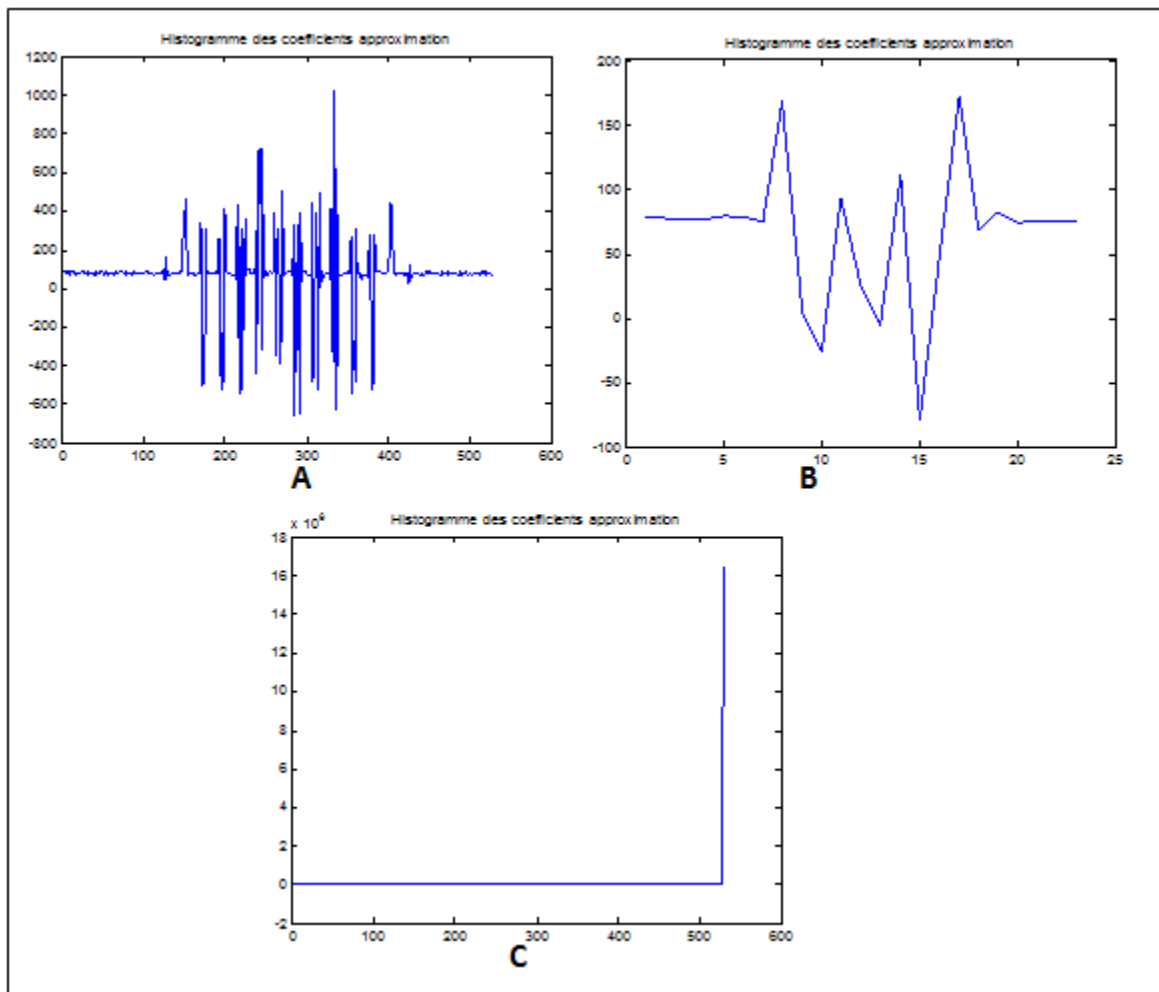


Figure 4.16 : Histogrammes correspondants à une réduction des paramètres du vecteur de caractéristiques. A: Histogramme de l'approximation avant réduction des paramètres. B: histogramme après application de la méthode de calcul des moyennes. C : histogramme obtenu après l'application de l'ACP.

L'application de la réduction des paramètres permet de garder les coefficients les plus pertinents de l'image d'approximation. Ainsi, on gagne en temps de calcul pour la classification.

Une application de L'ACP (Analyse en composantes principales) a réduit les paramètres de l'approximation le plus possible, mais l'aspect de l'histogramme a complètement changé.

L'application d'une deuxième méthode, nous a permis une réduction des paramètres en gardant le même aspect que celui de l'histogramme d'origine.

4.6 Choix du classifieur

Notre choix s'est porté sur la classification supervisée et cela pour des raisons de compatibilité avec notre étude. On sait que la classification non-supervisée classe les

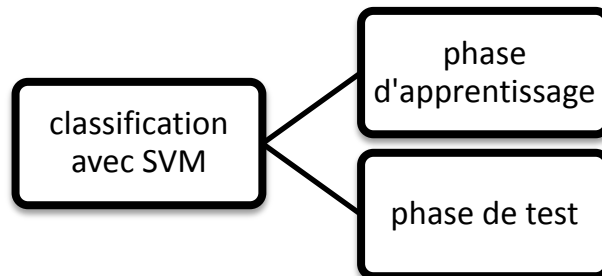
données selon des propriétés similaires ce qui mène parfois à une sur-classification. Pour éviter ce problème on a pensé à utiliser la classification supervisée dont les classes sont connues.

Dans le chapitre 2, on a vu les différents classifieurs supervisés existants dans la littérature. Il est assez difficile de sélectionner le classifieur idéal puisque leurs performances varient d'une problématique à une autre, et selon les caractéristiques reçues. Dans cette étude, on s'est basée sur l'utilisation du classifieur SVM (séparateurs à vaste marge) et afin de prouver ses performances, on l'a comparé à un autre du même type appelé régression linéaire multiple.

4.6.1 Les séparateur à vaste marge(SVM)

Les SVM reposent sur une théorie robuste et sont ajustés à travers différents paramètres totalement contrôlables. Leurs adaptations aux bases d'apprentissage de petite taille en font un classifieur attractif pour notre problématique. Les bonnes performances de détection obtenues lors des différentes études nous ont orienté à évaluer les SVM avec un deuxième classifieur.

4.6.1.1 Algorithme des SVM



4.6.1.1.1 Phase d'apprentissage

On utilise un procédé d'optimisation pour identifier les vecteurs supports « S_i », les poids α_i , et le biais « b » qui sont utilisés pour classer les vecteurs x conformément à l'équation suivante :

$$c = \sum_i \alpha_i k(s_i, x) + b,$$

Où k est une fonction noyau.

Dans le cas d'un noyau linéaire, k est le produit scalaire. Si $c \geq 0$, alors x est classé en tant que membre du premier groupe, sinon il est classé comme un membre du deuxième groupe.

4.6.1.1.2 Phase de test

Cette phase utilise les résultats de la phase d'apprentissage pour classer les vecteurs x conformément à l'équation suivante:

$$c = \sum_i \alpha_i k(s_i, x) + b,$$

Où les « Si » sont des vecteurs de support, α_i sont les coefficients de pondération, b est le biais, et k est une fonction noyau. Dans le cas d'un noyau linéaire, k est le produit scalaire. Si $c \geq 0$, alors x est classé en tant que membre du premier groupe, sinon il est classé comme un membre du deuxième groupe.

Avec la classification SVM, la base est divisée en deux parties, une partie est utilisée pour l'apprentissage et la deuxième pour le test.

On a varié le nombre d'images consacré à l'apprentissage et au test afin d'obtenir un bon taux de classification.

- Les résultats obtenus en utilisant les coefficients d'approximation traités avec la méthode de calcul de la moyenne, sont :

Variation du taux d'apprentissage (%)	Variation du taux de test (%)	Sensibilité (%)	Spécificité (%)	Résultats du taux de classification (%)
80	20	66	77,78	75.1852
90	10	50	57	51.5385
70	30	85,7	88,46	87
60	40	78,94	84	78.8889
50	50	81,2	83,33	82.0299
40	60	63,63	82,6	76.25
20	80	65,5	34,6	51.6822

Le taux total de classification est égal à : $VP+VN/VP+FP+FN+VN$

Remarque

D'après les résultats du tableau, on remarque que le classifieur SVM donne un très bon taux de classification (qui est de 87%), lorsqu'il utilise les 2/3 de la base d'images pour l'apprentissage et le 1/3 restant pour le test (c-à-d 70% pour l'apprentissage et 30% pour le test).

- Et les résultats obtenus utilisant avec l'ACP, sont :

Variation du taux d'apprentissage (%)	Variation du taux de test (%)	Sensibilité (%)	Spécificité (%)	Résultats du taux de classification (%)
80	20	9,09	81,25	51.6822
90	10	16,667	80	36.1538
70	30	4,76	89,47	45
60	40	15	79,4118	54.8148
50	50	14,28	93,75	52.6866
40	60	8,108	82,97	52.5
20	80	1,071	88,235	47.6636

4.6.2.2 Conclusion

La comparaison entre les résultats obtenus avec les deux méthodes étudiées précédemment, a montré que l'ACP élimine certaines caractéristiques essentielles, ce qui engendre une perte d'informations considérable, ceci mène le SVM à une classification peu fiable car il n'arrive pas à classer correctement toutes les images de la base. Par contre avec le calcul de la moyenne, seulement les informations de moindres importances sont supprimées afin de réduire la dimension du vecteur caractéristique sans pour autant perdre l'information utile, c'est pour cela que le SVM a donné un taux de classification élevé.

4.6.2 Régression Linéaire Multiple

Le premier travail sur la régression linéaire a été publié par Legendre en 1805, qui concerné la méthode des moindres carrés. Gauss prétendait la connaissance de cette méthode depuis 1795.

Legendre et Gauss ont appliqué cette méthode pour prédire l'orbite des planètes à partir des observations astronomiques.

Gauss a publié en 1821 une théorie sur la méthode des moindres carrés, incluant une version du théorème Gauss-Markov.

D'autres études ont été menées tout le 19ème et le début de 20ème siècle pour décrire des phénomènes biologiques et étendues à un contexte statistique général par Pearson, Yule (1877,1885) et Fisher (1922).

Les modèles de régression tentent de trouver une relation entre deux variables aléatoires x et y , ils cherchent à trouver une dépendance fonctionnelle entre les sorties réelles comme fonction des entrées et de prévoir la valeur de y connaissant celle de x .

En régression linéaire la forme de la dépendance fonctionnelle est une droite: $y=x'w+w_0$. Il s'agit d'estimer une variable réelle par une combinaison linéaire des caractéristiques d'entrée.

- en utilisant la base contenant les images traitées avec la méthode de calcul de la moyenne

Variation du taux d'apprentissage (%)	Variation du taux de test (%)	Résultats du taux de classification (%)
80	20	73.1776
90	10	43.3333
70	30	75
60	40	77.5926
50	50	80.0149
40	60	76.25
20	80	41.028

➤ en utilisant la base contenant les images traitées avec ACP

Variation du taux d'apprentissage (%)	Variation du taux de test (%)	Résultats du taux de classification (%)
80	20	71.4815
90	10	39.2308
70	30	67.5
60	40	49.2593
50	50	57.1642
40	60	55
20	80	38.271

4.6.2.1 Conclusion

La classification avec la méthode de la régression linéaire multiple (RLP) est moins fiable que la classification avec les SVM.

4.7 Le recalage et la fusion des images médicales

4.7.1 Introduction

La combinaison d'une imagerie morphologique à une imagerie métabolique, réalisée au mieux par des systèmes hybrides, a permis de faire un bond important dans la prise en charge des patients.

Jusqu'à une période récente, la Médecine Nucléaire détectait des anomalies fonctionnelles, sans pour autant pouvoir préciser leur situation anatomique exacte. L'imagerie par résonance magnétique nucléaire (IRM) et la tomodensitométrie X (TDM) permettent au contraire de caractériser des anomalies strictement morphologiques, en précisant leur rapport anatomique. Bien que d'utilisation pratique encore difficile, des logiciels permettent de combiner ces deux informations, en recalant et fusionnant des acquisitions faites sur deux imageurs différents.

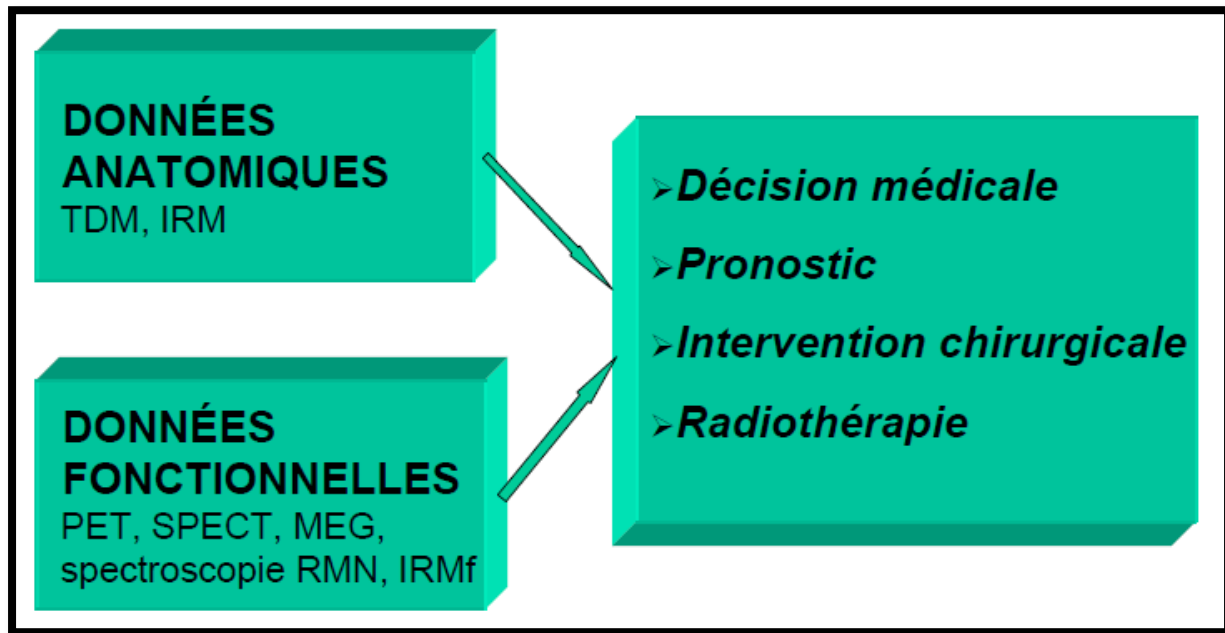


Figure 4.17 : Fusion des images médicales

4.7.2 Le recalage

Le recalage est l'étape préliminaire à la fusion, il se charge de définir les relations spatiales entre une image A et une image B pour compenser des variations dues à l'utilisation de modalités ou de conditions d'acquisitions différentes.

C'est une technique qui consiste en la mise en correspondance d'images, afin de comparer ou combiner leurs informations respectives. Cette mise en correspondance se fait par la recherche d'une transformation permettant de passer d'une image à une autre

Selon la formule suivante :

$$t = \underset{t \in T}{\operatorname{argmin}} f(I_1, t(I_2))$$

Avec :

- I1 et I2 : les images à recaler
- t : Transformation
- T : ensemble des transformations possibles/admissibles
- f : critère de dissimilarité (min) ou de similarité (max)

Les méthodes de recalage peuvent être classées selon trois critères principaux :

- le critère de ressemblance,
- la classe de transformation recherchée,
- l'algorithme d'optimisation

4.7.2.1 Critère de ressemblance

➤ Recalage géométrique

Cette méthode fut parmi les premières à être proposée de par sa similitude avec la méthode utilisée par un être humain pour mettre en correspondance deux images. Elle se base en effet sur l'extraction à partir de chacune des images de primitives géométriques (points saillants, coins, contours, ...). Par la suite, ces primitives sont appariées afin de déterminer la transformation entre les deux images.

➤ Recalage iconique

Cette méthode se base essentiellement sur l'utilisation des intensités des deux images via l'utilisation d'une mesure de similarité, sans prendre en compte l'information géométrique. De nombreuses mesures de similarité ont été développées à ce jour, supposant chacune une relation différente entre les intensités lorsque les images sont recalées.

4.7.2.2 Classe de transformation

➤ Transformation linéaire

La transformation géométrique recherchée peut être linéaire. Cette classe comprend notamment les transformations rigides (rotation, translation), les similitudes, les transformations affines et les transformations projectives.

Pour des raisons pratiques, ce type de transformation est par ailleurs souvent représenté sous forme matricielle. On parle alors de représentation en coordonnées homogènes.

➤ Transformation non-linéaire

La transformation peut également être non linéaire. Elle est alors d'ordre supérieur, par exemple les thin-plate splines, les combinaisons de B-Splines. Dans la littérature de traitement d'images, ces transformations peuvent également être nommées élastiques ou encore non rigides.

Le recalage rigide

Dans notre application, on s'est basée sur l'utilisation du recalage rigide, car cette opération est réalisée pour un même patient :

- La technique globale tend à estimer une transformation géométrique qui maximise la similarité entre les images données.

Le problème du recalage rigide repose sur les trois transformations suivantes: La translation, La rotation et l'homothétie.

4.7.2.3 Algorithmes d'optimisations

Les méthodes d'optimisation constituent le lien entre le critère de ressemblance et la transformation optimisée, utilisant les informations issues des critères choisis afin de

déduire la transformation entre les images. De ce fait, ces algorithmes dépendent de la transformation optimisée et du critère.

Les méthodes géométriques produisent souvent un ensemble d'appariements entre divers points des images. Dans ce cas, et si la transformation recherchée le permet, un algorithme dit des moindres carrés, ainsi que ses variantes robustes, peut être utilisé afin d'obtenir la transformation optimale.

Les critères iconiques sont quant à eux souvent optimisés via des méthodes de descente de gradient en fonction des paramètres de la transformation ou, lorsque le gradient ne peut être calculé, la méthode de Powell peut être utilisée afin d'optimiser la transformation.

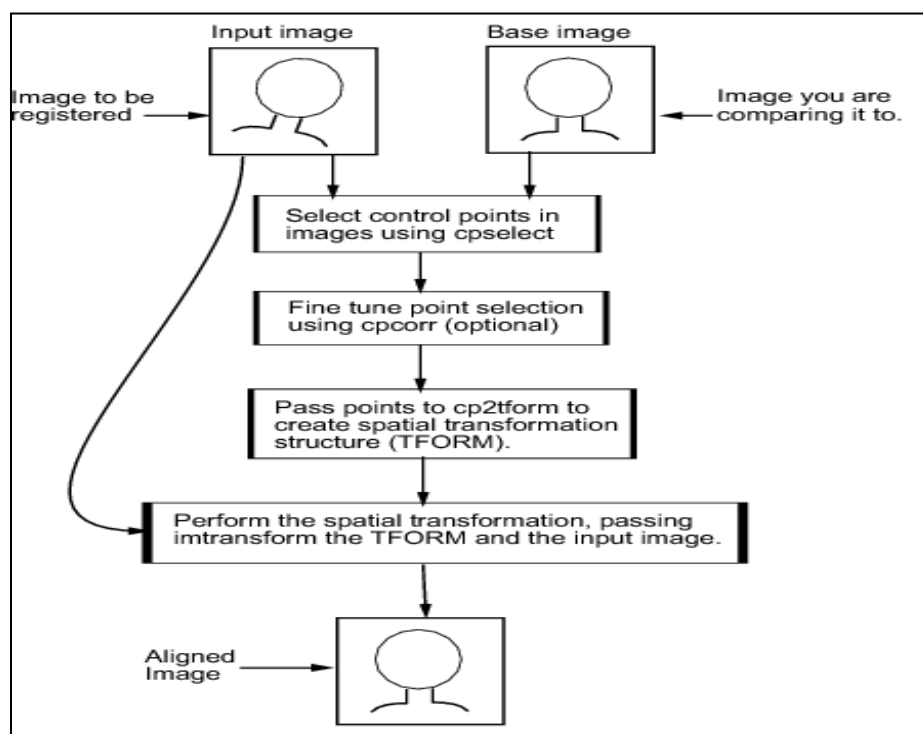
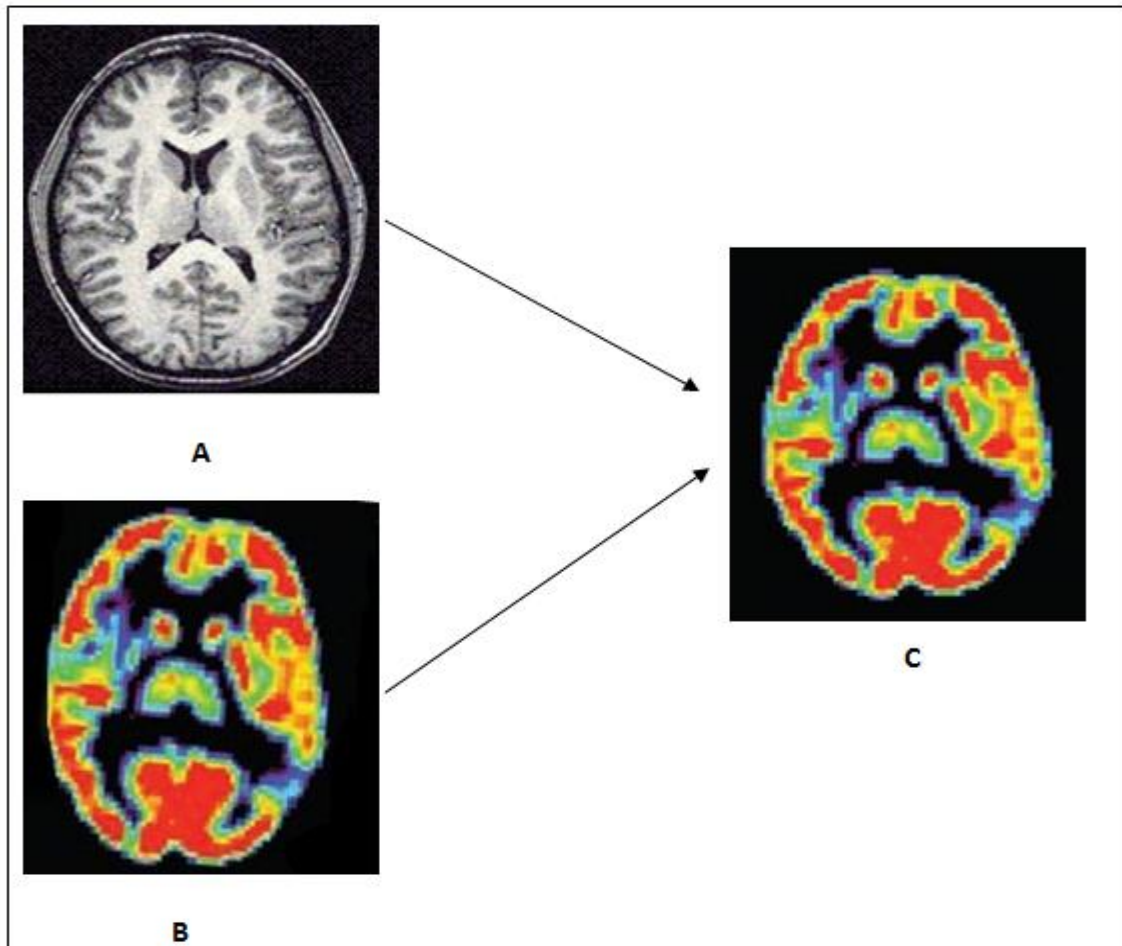


Figure 4.18 : Algorithme de recalage



Figure

4.19 : Le recalage de deux images. A : l'image anatomique (IRM) représentant l'image cible. B : l'image fonctionnelle (TEP) à recaler. C: le résultat du recalage de l'image B selon A.

4.7.3 La fusion

« La fusion d'informations consiste à combiner des informations issues de plusieurs sources afin d'améliorer la prise de décision », d'après Bloch [46].

4.7.3.1 Principe de la fusion par ondelette

Le principe de la fusion par ondelettes, est de combiner entre les décompositions en ondelettes de deux images différentes, ayant les mêmes dimensions, en utilisant des méthodes appliquant les coefficients d'approximations et les coefficients des détails [47].

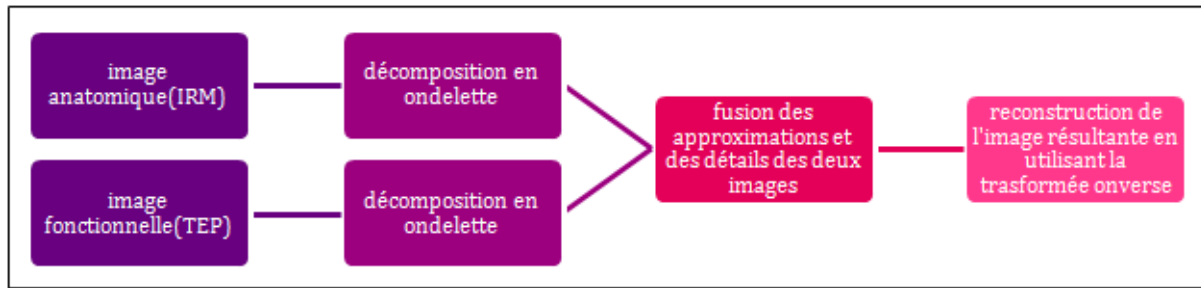


Figure4.20 : algorithme de la fusion en utilisant la décomposition par ondelette.

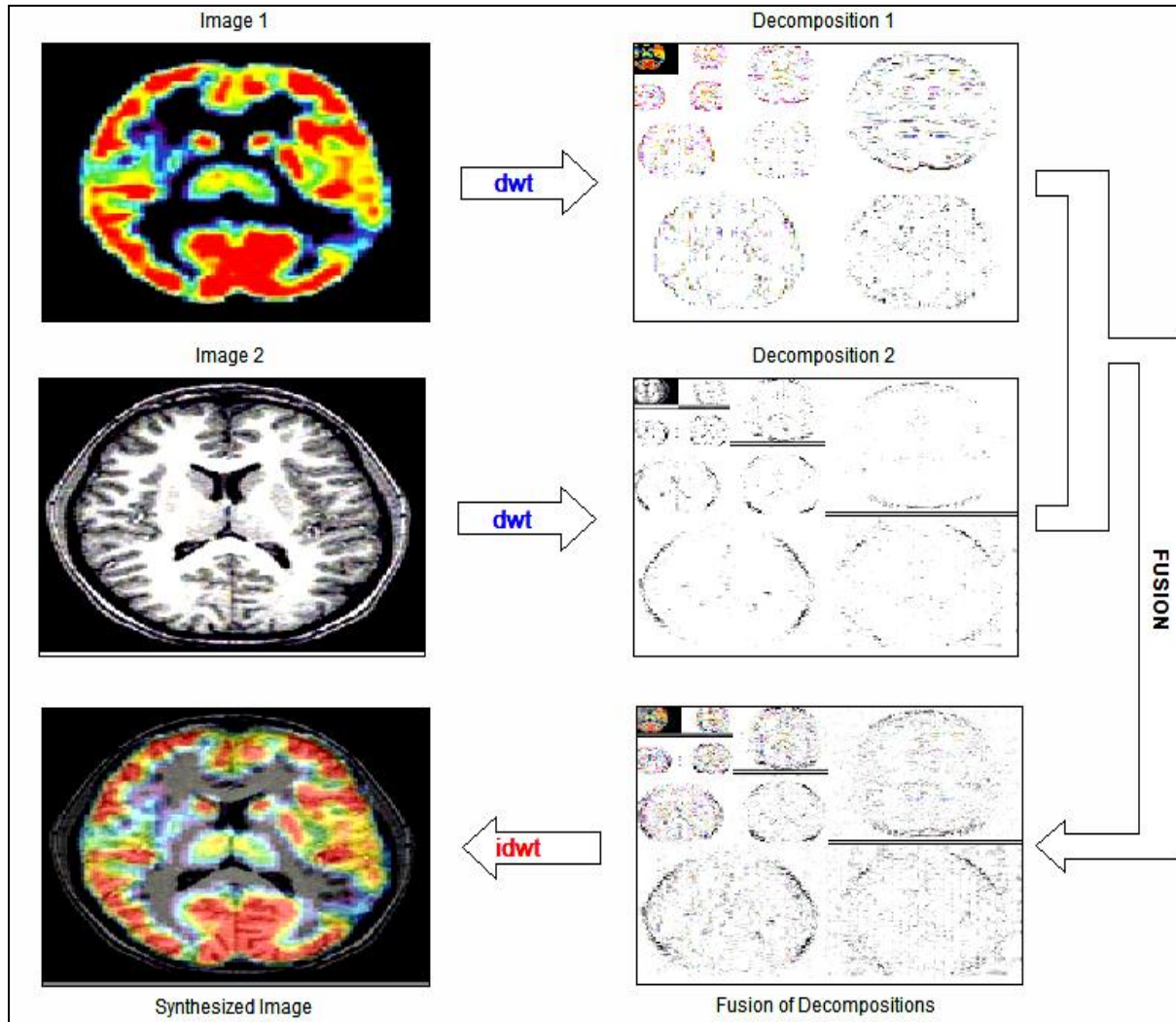


Figure 4.21 : fusion de deux images après recalage

4.8 Justifications et discussions

4.8.1 Choix des caractéristiques

Dans la 1^{ère} partie de notre travail, nous nous sommes intéressées aux caractéristiques fréquentielles qui permettent de synthétiser le contenu fréquentiel de l'image. Car l'extraction d'informations fréquentielles et plus précisément d'ondelettes permet de reproduire le système de perception humain traitant l'information à travers des canaux sélectifs en fréquences spatiales. La décomposition d'une image en ondelettes permet d'obtenir des composantes de l'image correctement localisées dans l'espace pour lesquelles les structures ou formes des signaux sont facilement accessibles (localisation dans le domaine fréquentiel). Cette transformation a l'avantage d'être flexible par rapport à la résolution de l'image et robuste à sa qualité variable. Elle permet d'extraire les informations fréquentielles, non seulement de la tumeur, mais aussi du bruit environnant. Celui-ci est ainsi plus facilement décorrélé de la tumeur cible.

L'utilisation d'une décomposition en ondelettes est optimale lorsque la fonction ondelette est judicieusement choisie. L'idéal est d'utiliser une ondelette dont la forme est proche de celle d'une tumeur ou d'une pathologie à détecter. Cela permet d'obtenir des images de coefficients d'ondelettes ayant leurs maximums localisés sur les pathologies.

En imagerie médicale, la plupart des études basées sur des caractéristiques discriminantes fréquentielles (Strickland et Hahn, Yu, Riyahi-Ala, Regentova), concernant le système CAD, ont exploité les ondelettes biorthogonale 4.4 (Bior4.4) [Cohen et al., 1992] et la Coiflet2 (Coif2) [Daubechies, 1992]. Le Tableau qui suit nous donne les caractéristiques des ondelettes citées ci-dessous.

	Orthogonale	Biorthogonale	Symétrique	Taille du filtre Passe-Bas	Taille du filtre Passe-Haut	Nombre de moments
Bior4.4	non	Oui	oui	9	7	4
Coif2	oui	Oui	presque	12	12	3

Tableau 4.1 : Principales caractéristiques des ondelettes Bior4.4, Coif2

Notre choix s'est porté sur l'utilisation de l'ondelette Bior4.4, car elle présente des caractéristiques permettant un bon compromis entre facilité de localisation et décorrélation des signaux, de plus sa forme semble plus adaptée à la problématique de ce travail.

Au cours de cette étude, d'autres ondelettes ont été testées : la Coif2, et Haar, mais les résultats ne sont pas apparus concluants en termes de faux positifs pour la première, et la seconde a mené à une dégradation de l'image qui a conduit à une perte d'information au 3^{ème} niveau de décomposition. Cette tendance est nettement observée sur les figures 4.10 et 4.12.

Dans la décomposition en ondelette, nous n'avons gardé que l'image « Approximation », les « détails » ne sont pas compris dans le vecteur caractéristique, car l'approximation de Bior4.4 contient pratiquement toutes les informations de l'image originale, les détails peuvent être supprimés du fait qu'ils ne contiennent pas d'informations pertinentes et qu'ils ne mènent à aucune perte.

4.8.2 Réduction de la taille du vecteur caractéristique

Après avoir extrait le vecteur caractéristique de chaque image de la base de données, on s'est retrouvée avec un grand nombre de coefficients, et pour mener à bien la classification, il faut réduire ce nombre.

Deux méthodes ont été utilisées :

- L'Analyse en Composantes Principales (ACP)
- Le calcul des moyennes

Avec l'ACP, on a une très grande perte d'informations, par contre avec la seconde méthode, c'est-à-dire, avec le calcul de la moyenne, seulement les coefficients qui ne comportent pas d'informations utiles ont été supprimés.

4.8.3 Le choix des classifieurs

Cette partie de notre travail se positionne dans une problématique de classification bi-classe, où on a attribué un « 1 » à la classe « non pathologique », et un « 2 » à la classe « pathologique ». Un classifieur supervisé a pour rôle de classer des observations inconnues dans une de ces deux classes après un apprentissage sur des observations d'étiquette connue. Nous avons choisi de tester et d'évaluer deux classifieurs supervisés, le SVM et la RLM. Ces deux classifieurs, se basant sur des approches différentes, sont ainsi exploités de manière complémentaire afin d'améliorer leurs résultats individuels.

Les résultats obtenus avec le SVM, ont montré une nette différence entre les deux méthodes citées auparavant, qui sont l'ACP et la Moyenne. Cette dernière, a donné un très bon taux de classification et qui correspond à 87%.

La phase d'apprentissage est la plus importante, mais la détermination des bons paramètres n'est pas toujours facile. Comme nous l'avons cité auparavant, nous nous sommes intéressées aux paramètres du classificateur SVM afin d'aboutir à une erreur de classification minimale. Pour fixer le modèle SVM approprié à notre classification, on a étudié les performances de trois types de modèles : la fonction linéaire, la fonction RBF à noyau gaussien, et la fonction Polynomiale (ordre 3). Ces fonctions ont été exploitées dans le but de savoir laquelle est apte à minimiser l'erreur du SVM sur les données d'apprentissage et à maximiser la marge.

C'est la fonction « Polynomiale » qui a été retenue, car elle a réussi à faire une bonne séparation des deux classes, en respectant les critères du SVM, ceci a conduit à un bon taux de classification en comparaison avec les deux autres fonctions utilisées, qui ont donné un taux inférieur.

4.8.4 Le recalage et la fusion

Par rapport aux images TDM, les images IRM présentent un bien meilleur contraste des tissus mous, ce qui est particulièrement utile lors de l'exploration du cerveau. En plus de l'anatomie macroscopique, l'IRM offre également la possibilité d'obtenir des informations sur les processus au niveau moléculaire. Ainsi, il est par exemple possible de mesurer la diffusion des molécules d'eau, ce qui est utile pour détecter des infarctus cérébraux, des abcès ou des tissus tumoraux.

La TEP et l'IRM sont complémentaires à bien des égards. D'une part, les altérations structurelles à l'IRM peuvent être mieux caractérisées par la TEP et d'autre part, les altérations à la TEP peuvent être détectées avant qu'elles ne donnent lieu à des altérations morphologiques à l'IRM. Ainsi, il est possible de détecter les dépôts conduisant à la maladie d'Alzheimer au moyen d'un traceur TEP spécifique avant que la dégénérescence des tissus cérébraux ne débute. Etant donné que les altérations au niveau moléculaire surviennent avant les altérations morphologiques, l'effet de médicaments ou d'agents chimiothérapeutiques peut souvent déjà être évalué peu après le début du traitement. De cette manière, il est possible d'arrêter rapidement les traitements inefficaces et d'administrer un autre traitement au patient.

Néanmoins, le couplage de la TEP et de l'IRM est assorti d'un certain nombre de défis techniques considérables.

La fusion d'images offre :

- Détection plus efficace d'anomalies fonctionnelles
- Meilleure compréhension des relations physio-pathologiques entre lésion et fonction
- Aide à l'interprétation des lésions structurelles ambiguës ou non spécifiques
- Aide à la prise de décisions thérapeutiques importantes
- Analyse quantitative

4.9 Conclusion

Le travail qui a été fait au cours de ce chapitre, a consisté à mettre en œuvre un système d'aide au diagnostic pour l'imagerie TEP afin d'aider les cliniciens dans la délicate tâche d'interprétation.

Une approche fréquentielle nous a permis d'extraire les caractéristiques des l'image en passant par une analyse par ondelette, le choix s'est porté sur l'ondelette Bior4.4 à cause des bons résultats obtenus, une réduction des données a été ensuite effectuée afin de minimiser le nombre de caractéristiques d'entrée du classifieur SVM.

Les résultats obtenus avec le SVM sont plus ou moins satisfaisants, vu qu'on a obtenu un taux de classification de 87%.

Un autre classifieur « **régression linéaire multiple** », a été utilisé pour effectuer une comparaison, cette dernière a montré que le classifieur SVM donne de meilleurs résultats.

Conclusion générale

Conclusion générale

L'imagerie TEP est une modalité qui permet de visualiser les différentes parties du corps humain, elle est considérée comme une révolution en imagerie médicale, elle informe sur l'activité fonctionnelle des organes à visualiser.

Les images issues de cette modalité sont de très faible résolution spatiale, ainsi, le traitement de ces images ne se fait pas de façon classique.

Nous avons procédé par approche fréquentielle, afin d'extraire les anomalies contenues dans les images.

On a donc proposé un système CAD fondé sur la combinaison de différentes techniques de traitement et d'analyse d'images, permettant non seulement d'extraire les pathologies, mais aussi de les classer en cas sains et pathologiques.

Les différentes simulations appliquées sur des images TEP ont permis de valider les résultats du système CAD proposé par rapport à la classification pathologique des démences. Ces simulations ont été accompagnées par des approches de recalage et de fusion des deux modalités IRM/TEP afin de guider les cliniciens.

L'imagerie TEP donne des images de faible résolution spatiale, ce qui ne nous a pas permis de procéder aux étapes habituelles (prétraitement, débruitage, segmentation, etc...), alors, pour le traitement de ces images, on s'est basé sur une approche fréquentielle (analyse par ondelette) pour pouvoir extraire les caractéristiques de l'image nous permettant ensuite la classification de cette dernière avec le classifieur SVM.

L'application intense de la tomographie par émission de Positons dans le diagnostic des démences cérébrales a permis de présenter les principales techniques liées aux outils numériques de diagnostic assisté par ordinateur et d'aide à la décision afin de faciliter l'interprétation pour les cliniciens

L'intérêt des systèmes CAD a été prouvé dans d'autres modalités d'imagerie médicale (mammographie, TDM)

Ils permettent en effet d'augmenter les performances en terme de détection et de stadification en diminuant les variations inter et intra-cliniques.

Les images de la base de donnée TEP-SORTEO, ciblées dans ce mémoire sont utilisées sans pré-traitement préalable. Les seules corrections exploitées sont généralement celles qui sont appliquées dans les protocoles cliniques (atténuation, diffusion, temps mort...).

Ce présent travail peut donc faire l'objet des perspectives suivantes :

- Simulations sur des images réelles,
- Utilisation d'autres approches de traitement d'image tel que le calcul des moments (HU),
- Exploitation d'autres types de classifieurs,
- Effectuer des simulations pour d'autres parties du corps humain,
- Classification des différentes pathologies neurologiques par type de démence,
- Il aurait pu être intéressant de tester l'influence de certaines méthodes de traitement d'images TEP faisant l'objet de développements importants dans la communauté scientifique et suscitant des débats. On peut citer par exemple les méthodes de débruitage ou de correction de volume partiel

Annexes

ANNEXES

Annexe A : La base d'images « TEP-Sorteo »

SORTEO signifie tirage aléatoire en Espagnol

La simulation TEP Monte Carlo

On appelle méthode de Monte-Carlo (MC) [Metropolis, 1947; Metropolis and Stanislas Ulam, 1949] toute méthode visant à calculer une valeur numérique, et utilisant des procédés aléatoires, c'est-à-dire des techniques probabilistes. Le nom de ces méthodes fait allusion aux jeux de hasard pratiqués à Monte-Carlo. L'imagerie TEP est particulièrement propice à l'utilisation de la simulation MC par la nature probabiliste des processus physiques qui la compose. En effet, la simulation MC permet de modéliser l'émission et le parcours du positon, l'annihilation et les interactions photon-matière jusqu'à la probabilité de détection des photons dans le détecteur.

Plusieurs simulateurs TEP MC rigoureusement validés sont actuellement disponibles. On peut par exemple citer SimSET [Harrison et al., 1993], EIDOLON [Zaidi and Morel, 1999], PET-EGS [Castiglioni et al., 1999], GATE [Jan et al., 2004] et PET-SORTEO [Reilhac et al., 2005].

Le simulateur GATE a été développé dans le cadre de la collaboration OpenGate et se base sur la librairie générique pour la simulation des particules, GEANT4 [Agostinelli et al., 2003]. Il est puissant et flexible et le processus de simulations des données inclut la modélisation de tous les phénomènes majeurs et permet de modéliser fidèlement la plupart des géométries existantes. Son principal inconvénient reste le temps de calcul nécessaire pour faire une simulation.

Le développement de PET-SORTEO a été initié avant celui de GATE dans le cadre de la thèse de [Reilhac, 2007] et bénéficie de temps de simulation plus réduits. PET-SORTEO a été validé pour différents micro-scanners [Lartzien et al., 2005] et pour la caméra TEP Siemens ECAT/EXACT HR+ (CTI Siemens) [Reilhac et al., 2005].

Une comparaison des codes de simulation en termes de réalisme, de temps de calcul et de flexibilité a été réalisée par [Buvat et al., 2005] et dans la thèse de [Reilhac, 2007].

Le développement de PET-Sorteo

Le développement de PET-SORTEO, entre 1999 et 2004, fût motivé par un constat assez similaire à celui étant à l'origine du développement de GATE : les limites des codes existants en 1999. Cependant, le cahier de charge étant différent (la rapidité était un critère essentiel dans le projet PET-SORTEO), les deux développements ont conduit à des outils différents répondant à deux besoins bien distincts. Le développement du code de simulation PET-SORTEO, a été réalisé et financé par l'institut neurologique de Montréal [Reilhac et al., 2004, Reilhac et al., 2005, SORTEO]. Ce code de simulation est écrit en langage C, et n'est basé sur aucune librairie

générique de simulation. Contrairement à GATE, il ne permet que la simulation de données TEP pour des scanners de géométries cylindrique ou polygonale. PET-SORTEO a été conçu pour les environnements multi-processeurs de type grilles de calcul, mais de par sa rapidité, il peut être utilisé sur une simple station de travail. Il doit être noté, qu'aucun effort n'a été réalisé pour distribuer le code source au grand public. De ce fait, ce code a été distribué que de façon ponctuelle, sur demande, ou à l'occasion de collaborations. Ceci est en contraste avec le projet GATE, pour lequel il existe une forte politique de promotion et de distribution.

Le modèle de simulation tient compte de tous les phénomènes majeurs, et entre autres du temps mort et de la contamination des données par les événements aléatoires. Cependant, afin que la prise en compte de ces phénomènes n'augmente pas de façon intolérable le temps de calcul, une méthode originale a été développée, permettant la simulation de ces phénomènes sans avoir à simuler tous les photons simples. Ce modèle proposé est certainement la partie la plus innovante du code. Par ailleurs, le code de simulation du transport des particules est similaire aux autres codes, avec certains processus simulés analytiquement (transport du positon, angulation des photons d'annihilation,..) et d'autres entièrement selon des méthodes de Monte Carlo (transport des photons). PET-SORTEO offre également la flexibilité souhaitée. Il permet entre autre la simulation de données dynamiques en 2D, 3D ainsi que la simulation de données de transmission. Enfin, le code permet l'utilisation de plusieurs processeurs pour simuler une acquisition.

Déroulement d'une simulation

PET-SORTEO utilise un protocole pour réaliser une simulation. Le protocole de simulation décrit l'acquisition TEP à simuler. Il précise entre autres le nombre et la durée des frames temporelles et les courbes d'activité temporelle pour chaque structure émettrice. C'est aussi lors de l'établissement de ce protocole que l'utilisateur définit les cartes d'émission et d'atténuation ainsi que les propriétés géométriques et physiques du tomographe. La simulation d'une acquisition se déroule à partir du protocole en trois étapes :

1. Une première étape consiste à simuler rapidement une version simplifiée du protocole d'acquisition. Au lieu de simuler une acquisition dynamique avec des courbes d'activité temporelle, le simulateur se contente de simuler un certain nombre de désintégrations par structure. Cette première étape, qu'on appellera acquisition simplifiée, permet de dériver certaines grandeurs spécifiques à la simulation en cours. C'est entre autres durant l'acquisition simplifiée que certaines probabilités de détection sont déterminées. Elles permettront la prise en compte des événements aléatoires et du temps mort de chaque module de détection durant la simulation des coïncidences sans avoir à simuler tous les photons simples. D'autres grandeurs ayant pour but de réduire le temps de simulation des coïncidences sont également calculées pendant cette étape.
2. La deuxième étape consiste à simuler les coïncidences vraies (diffusées et non diffusées), une à une, selon le protocole défini, et en s'aidant des probabilités calculées préalablement afin de raccourcir le temps de calcul.
3. Enfin, la troisième étape consiste à simuler les événements aléatoires enregistrés dans la fenêtre de coïncidence ainsi que les événements mesurés dans la fenêtre temporelle retardée.

❖ Étape 1 : L'acquisition simplifiée

L'objectif de cette étape est de déterminer les grandeurs suivantes :

$P_{wi,k}$: probabilité qu'un photon émis de la structure i soit détecté dans la fenêtre énergétique, par le module de détection k (temps mort non considéré).

$P_{Ei,k}$: probabilité qu'un photon émis de la structure i soit détecté avec une énergie supérieure au seuil de déclenchement d'une intégration du signal, par le module de détection k (temps mort non considéré).

$P_{u+s,i}$: probabilité qu'une paire de photons émis de la structure i soit enregistré en coïncidence par le tomographe (temps mort non considéré).

L'intervalle $\theta_{i,min}$, $\theta_{i,max}$: Un photon émis de la structure i selon un angle polaire θ n'appartenant pas à cet intervalle a une probabilité quasi-nulle d'être détecté.

$E_{i,min}$: Énergie en dessous de laquelle un photon émis de la structure i n'a plus de chance d'être détecté dans la fenêtre énergétique.

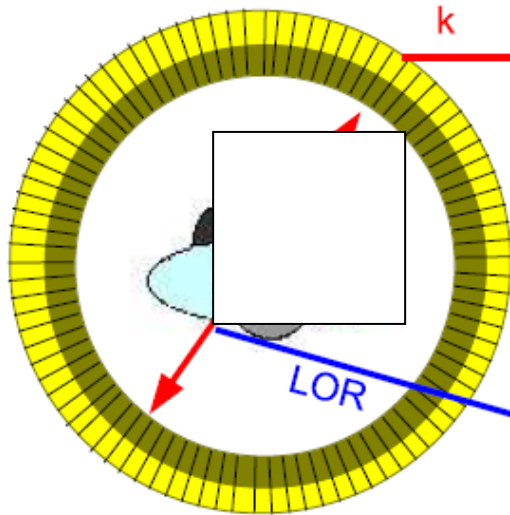
Ces grandeurs sont déterminées durant l'étape de pré-simulation d'une version simplifiée du protocole d'acquisition, en utilisant les modèles précédemment définis.

Durant cette étape, aucune source de temps mort ne doit être considérée. Pour chaque structure, N_i désintégrations sont simulées. Les événements produits sur le système de détection sont enregistrés et classés dans des matrices, selon la structure émettrice i , le module impliqué k et l'énergie enregistrée. Un photon émis par la structure i et enregistré dans le module k avec une énergie supérieure au seuil de déclenchement de l'intégration E_{trig} incrémente la matrice $C_{Ei,k}$. S'il est aussi enregistré dans la fenêtre énergétique, il incrémente en plus la matrice $C_{wk,i}$. Les coïncidences vraies (diffusées et non diffusées) enregistrées sont également classées dans une matrice $C_{u+s,i}$ selon leur structure d'origine i . Une fois la simulation simplifiée pour la structure i terminée, les probabilités $P_{wi,k}$ (respectivement $P_{Ei,k}$) sont déduites par le rapport de $C_{wi,k}$ (respectivement $C_{Ei,k}$) et le nombre total de photons émis par la structure i durant la simulation. De la même manière, les facteurs $P_{u+s,i}$ sont donnés par le rapport de $C_{u+s,i}$ et le nombre total de paires de photons générées (voir figure 1). Une approche similaire, permet aussi de déterminer les valeurs de $\theta_{i,min}$, $\theta_{i,max}$, $E_{i,min}$. La valeur de N_i , nombre de désintégrations simulées pour la structure i n'est pas fixé ni connu a priori. Il dépend fortement des caractéristiques de l'objet imagé, du tomographe modélisé et des paramètres d'acquisition.

Tout comme pour les autres grandeurs, ce nombre est déterminé lors de la simulation simplifiée. Il correspond concrètement au nombre de désintégrations qu'il a été nécessaire de simuler à partir de la structure i pour que plus de 50000 photons simples soient détectés avec une énergie valide par module de détection.

Ainsi, avec le tomographe Ecat Exact HR+, la simulation simplifiée pour la structure i s'arrête lorsque au minimum 14,4 millions photons simples provenant de la structure i ont été enregistrés par les 288 modules du tomographe. La valeur de 50000 photons détectés par module a été jugée suffisante pour que les différentes grandeurs dérivées de la simulation simplifiée soient déterminées avec une bonne précision.

Simulation de Monte Carlo



$$C_{\epsilon_{i,k}} = C_{\epsilon_{i,k}} + 1 \text{ si } E_w > \epsilon_{trig}$$

$$C_{w_{i,k}} = C_{w_{i,k}} + 1 \text{ si } lld < E_w < uld$$

Si les deux photons sont détectés

$$C_{u+s,i} = C_{u+s,i} + 1$$

Avec

E_w : Energie du photon incident

$C_{\epsilon_{i,k}}$: Nombre de photons issus de la région i et détectés par le module k avec une énergie supérieure au seuil de déclenchement d'une intégration

$C_{w_{i,k}}$: Nombre de photons issus de la région i et détectés par le module k dans la fenêtre énergétique

$C_{u+s,i}$: Nombre de coïncidences vraies ($u + s$) détectées de photons issus de la région i

N_i : Nombre de désintégrations simulées pour la région i lors de la pré-simulation.

Figure 1 : Les trois étapes d'une simulation. D'abord une première simulation du protocole d'acquisition est réalisée rapidement. Un certain nombre de désintégrations sont simulées pour chaque région émettrice i . Les événements enregistrés (photons simples, coïncidences non diffusées, coïncidences diffusées) incrémentent des matrices. Une fois cette première étape achevée, des probabilités de détection sont dérivées à partir de ces matrices. Les probabilités servent ensuite, durant la simulation des coïncidences et des événements aléatoires, à évaluer les taux de photons simples entrant dans un module de détection (et donc le temps mort), ainsi que le taux de coïncidences à traiter par le système.

❖ Étape 2 : Simulation des coïncidences vraies

Les événements vrais (diffusés et non diffusés) sont simulés un par un, de l'émission du positon jusqu'à éventuellement la détection en coïncidence de la paire de photons γ . L'histoire de chacun des deux photons est simulée séparément.

Un photon n'est simulé que si son angle polaire d'émission, θ , appartient à l'intervalle $[\theta_{min}, \theta_{max}]$. En effet, en dehors de cet intervalle, la probabilité qu'il soit détecté est infime. Aussi, si l'un d'eux est perdu, absorbé, ou possède une énergie inférieure au seuil $E_{i,min}$ alors le programme arrête la simulation de l'évènement en cours et passe directement à la désintégration

suivante. Ainsi, si le premier photon simulé issu de la désintégration n'est pas détecté, le deuxième photon n'est pas simulé. Si les deux photons sont enregistrés avec une énergie appartenant à la fenêtre imposée et si la ligne de réponse correspondante est autorisée par les paramètres de l'acquisition, alors l'élément correspondant du sinogramme est incrémenté d'une unité. Durant cette étape de simulation, les facteurs de perte de comptage due au temps mort au niveau de chaque module et du système de traitement des coïncidences évoluent en fonction de l'activité ; ils devraient calculer à chaque désintégration. Cependant, afin d'accélérer le processus de simulation, ce calcul n'est effectué que cycliquement (typiquement toutes les secondes). La détermination des facteurs est réalisée comme suit :

- Les différents taux de photons, $S_{\epsilon k}(t)$ et $S_{wk}(t)$ incidents aux modules sont calculés.
- Les taux d'évènements à traiter par les processeurs sont calculés.
- Les probabilités de survie des photons $P_k(t)$ et $P_l(t)$ et des coïncidences $P_{pc}(t)$ et $P_{fe}(t)$ sont calculées.

Les facteurs de pertes sont appliqués durant la simulation comme suit :

- Les pertes de comptage au niveau des photons sont appliquées à tout photon pénétrant un module en comparant chaque probabilité de survie à un nombre aléatoire uniformément distribué sur $[0,1]$.
- Les pertes affectant des évènements en coïncidence liées à leur traitement (tri et stockage) ne dépendent pas de la saturation des modules impliqués dans leur détection, mais seulement des taux de coïncidences à traiter.

Ainsi, ces pertes peuvent être appliquées avant même que l'on connaisse les modules à l'origine de la détection, et donc avant même la simulation du transport des deux photons. Les pertes sont appliquées exactement après l'émission du positon. Cette simplification représente un important gain en temps de calcul, puisqu'elle nous évite de simuler inutilement le parcours des photons d'évènements qui seraient ensuite écartés.

❖ Étape 3 : Simulation des évènements aléatoires et retardés

L'objectif de cette étape est de simuler la distribution des évènements aléatoires et de l'ajouter au sinogramme d'évènements vrais. Le taux d'évènements aléatoires, $E_{al}(t)$, est fonction du taux de photons simples enregistrés par le tomographe $D(t)$ (temps mort au niveau des modules inclus), de la longueur de la fenêtre de coïncidence τ_{coin} et des facteurs de temps morts intervenant au niveau du traitement des coïncidences $P_{pc}(t)$ et $P_{fe}(t)$.

$$E_{al}(t) = P_{pc}(t)P_{fe}(t)\tau_{coin}D(t)^2/2$$

La distribution spatiale de ces évènements est intimement liée à la distribution spatiale $dk(t)$ des photons simples enregistrés par les N_{mod} modules du tomographe.

$$E_{al}(t) = P_{pc}(t)P_{fe}(t)\tau_{coin}D(t)^2/2$$

Chacun des $E_{al}(t)$ évènements fortuits est simulé comme suit :

1. Deux modules impliqués dans la détection des deux photons non corrélés sont sélectionnés à partir de la distribution $dk(t)$.
2. Pour chaque module, une unité de détection est uniformément tirée au sort.
3. Si la ligne de réponse formée par les deux cristaux possède un élément correspondant dans le sinogramme alors ce dernier est incrémenté d'une unité.

La même logique est utilisée pour la génération des événements retardés puisque ces derniers ont la même distribution que les fortuits. Cependant, les événements retardés sont soustraits du sinogramme.

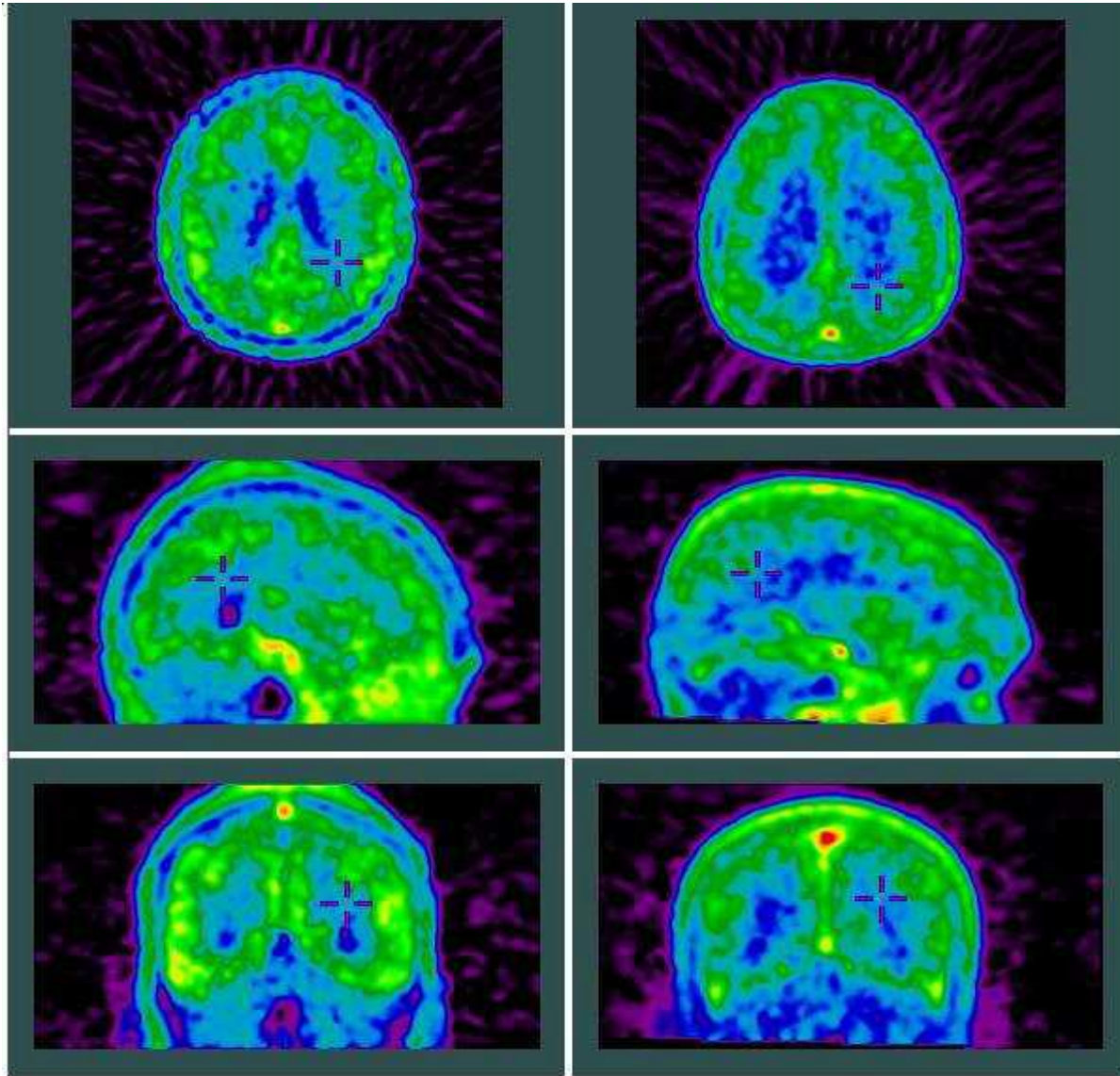


Figure 2 : Images TEP au $[^{18}\text{F}]\text{MPPF}$ acquises au CERMEP (droite) et simulées avec les mêmes conditions avec PET-SORTEO (gauche).

L'image de gauche a été entièrement simulée par PET-SORTEO, à partir d'un fantôme numérique de cerveau et de courbes d'activité temporelle reflétant l'évolution temporelle de la molécule injectée dans chaque structure. L'image de droite est une image réelle obtenue dans les mêmes conditions. La comparaison permet de juger du réalisme obtenue avec le simulateur. Cependant, une validation plus minutieuse du modèle a été réalisée. Les résultats sont présentés dans le chapitre suivant.

Bases de données de volumes TEP simulées

Simuler une acquisition TEP n'est pas une opération triviale. Outre la complexité des programmes de simulation TEP et la puissance de calcul nécessaire, la simulation TEP demande aussi :

- la définition du protocole d'acquisition (nombre et durée des frames)
- la création d'un fantôme réaliste du point de vue anatomique
- la définition de courbes d'activités temporelles reflétant la dynamique réelle du traceur dans chaque organe et non une courbe simplement mesurée sur des données TEP réelles, affectée par le bruit d'acquisition et par les effets de volume partiel

Enfin, dans une étape de validation, il est préférable de tester la méthode de correction ou de traitement proposée à l'aide de plusieurs volumes simulés. Pour toutes ces raisons, les bases de données de volumes simulés, accessibles aux chercheurs, présentent un intérêt grandissant. L'article suivant décrit la méthodologie employée pour générer une base de données d'acquisition TEP cérébrales incluant une variabilité anatomique inter-individuelle et cela pour différents protocoles d'acquisition.

Cette base contient les données brutes avant reconstruction et les volumes reconstruits selon la méthode conventionnelle (rétro-projection filtrée). Bien qu'il soit possible avec PET-SORTEO de sauvegarder à part les sinogrammes des différentes sources de contamination (événements fortuits et diffusés), la base ne propose que les sinogrammes de type "Net Trues" contenant les données acquises pré-corrigées des événements fortuits (soustraction des événements retardés). Cette base est destinée au chercheurs oeuvrant dans des thématiques allant de la reconstruction tomographique aux méthodes de traitement post reconstruction (recalage, effets de volume partiel). A ma connaissance, cette base est la première du genre avec la base proposée la même année par [Castiglioni et al., 2005]. Cette dernière, plus large, regroupe des données simulées de nature plus diverses, incluant des acquisitions simulées de fantôme géométriques, de corps entiers et de cerveaux, et obtenues pour différentes géométries de détection. Plus d'une centaine de laboratoires ont à ce jour téléchargé des données de la base PET-SORTEO pour la validation de méthodes de traitement allant de la reconstruction tomographique à la modélisation compartimentale (voir figure 4.4). Il est difficile de suivre précisément l'utilisation de la base, et surtout de recenser les articles qui en ont découlé. J'ai récemment proposé une amélioration de base qui inclue à présent une variabilité anatomique et fonctionnelle inter-individuelle [Reilhac et al., 2006]. Cette nouvelle base offre de nouvelles perspectives, comme la possibilité d'optimiser le traitement pour la détection de variations biologiques entre groupes de sujets.

Références bibliographiques

Références bibliographiques

- [1] Fluorine-18 radiopharmaceuticals beyond [18F]-FDG for use in oncology and neurosciences *Nuclear Medicine and Biology, Volume 37, Issue 7, October 2010, Pages 727-740* H.H. Coenen, P.H. Elsinga, R. Iwata, M.R. Kilbourn, M.R.A. Pillai, M.G.R. Rajan, H.N. Wagner Jr., J.J. Zaknub
- [2] **A Broad Overview of Positron Emission Tomography Radiopharmaceuticals and Clinical Applications: What Is New?**
Seminars in Nuclear Medicine, Volume 41, Issue 4, July 2011, Pages 246-264
Shankar Vallabhajosula, Lilja Solnes, Brigitte Vallabhajosula
- [3] Current themes in neuroimaging of epilepsy: Brain networks, dynamic phenomena, and clinical relevance
Clinical Neurophysiology, Volume 121, Issue 8, August 2010, Pages 1153-1175
Mark Richardson
- [4] Imaging and biomarkers for Alzheimer's disease
Maturitas, Volume 65, Issue 2, February 2010, Pages 138-142
Charlotte L. Allan, Claire E. Sexton, David Welchew, Klaus P. Ebmeier
- [5] **PET** radiotracers for molecular imaging in the brain: Past, present and future
NeuroImage, In Press, Corrected Proof, Available online 28 December 2011
Luc Zimmer, André Luxen
- [6] Biomarkers for Alzheimer's disease therapeutic trials
Progress in Neurobiology, Volume 95, Issue 4, December 2011, Pages 579-593
Harald Hampel, Gordon Wilcock, Sandrine Andrieu, Paul Aisen, Kaj Blennow, K. Broich, Maria Carrillo, Nick C. Fox, Giovanni B. Frisoni, Maria Isaac, Simon Lovestone, Agneta Nordberg, David Prvulovic, Christina Sampaio, Philip Scheltens, Michael Weiner, Bengt Winblad, Nicola Coley, Bruno Vellas, for the Oxford Task Force Group, *et al*
- [7] **Outils de quantification pour l'imagerie TEP au FDG**
Innovations Technologiques en Médecine Nucléaire, Saclay, 29 septembre 2011.
- [8] **Recent advances in PET Monte Carlo simulations and PET image quantification towards enhancement of the role of PET in cancer patients**
Invited seminar at the Maastricht clinics, Maastricht, February 8th, 2012.
- [9] **Rétroprojection filtrée et reconstruction itérative : rappels théoriques et propriétés des deux approches.**
Réunion thématique de l'APRAMEN, janvier 2002, Paris.
- [10] **Modèles numériques anthropomorphiques pour les simulations en TEMP et TEP**
Journée « Partage de modèles d'objets anatomiques et physiologiques pour la simulation d'images médicales cliniques » du GDR Stic Santé, Paris, 25 janvier 2011.

[11] **De la conception d'imageurs à l'estimation de paramètres physiologiques**

Journées de réflexion CNRS/IN2P3, Saint-François, 8 décembre 2010.

[12] **Quantification en Tomographie d'Emission.**

Séminaire invité chez la société Guerbet, Aulnay-sous-Bois, 5 octobre 2010.

[13] **SUV in PET: Silly or Smart Uptake Values?**

15th BHPA Symposium. Brussels, 5 février 2010.

[14] **Comment évaluer les méthodes de détermination de volumes métaboliquement actifs en Tomographie par Emission de Positons (TEP) ?**

Journée Volumes cibles biologiques en Radiothérapie. GDR STic-Santé. 9 décembre 2009.

[15] **Evaluation de tâches d'estimation sans gold standard**

Réunion de travail IMPEIC - GDR Stic Santé, 5 mars 2009, Paris.

[16] **Reconstruction tomographique** (Master de Physique Médicale Lyon)

[17] L'imagerie en Médecine Nucléaire

[18] Multi-resolution Bayesian regression in PET dynamic studies using **wavelets**

NeuroImage, Volume 32, Issue 1, 1 August 2006, Pages 111-121

F.E. Turkheimer, J.A.D. Aston, M.-C. Asselin, R. Hinz

[19] Development of wavelet de-noising technique for PET images

Computerized Medical imaging and Graphics, Volume 29, Issue 4, June 2005, Pages 297-304

Yen-Yu Shih, Jyh-Cheng Chen, Ren-Shyan Liu

[20] Wavelet variance components in image space for spatiotemporal neuroimaging data

NeuroImage, Volume 25, Issue 1, March 2005, Pages 159-168

John A.D. Aston, Roger N. Gunn, Rainer Hinz, Federico E. Turkheimer

[21] Quantification and wavelet-aided parametric imaging of cerebral amyloid using the HRRT PET - system and [18F]AZD4694

NeuroImage, Volume 52, Supplement 1, August 2010, Page S142

Zsolt Cselényi, Maria Erikdotter Jönhagen, Anton Forsberg, Christer Hallidin, Per Julin, Magnus Schou, Peter Johnström, Katarina Varnäs, Samuel Svensson, Lars Farde

[22] Wavelets and functional magnetic resonance imaging of the human brain
NeuroImage, Volume 23, Supplement 1, 2004, Pages S234-S249
Ed Bullmore, Jalal Fadili, Voichita Maxim, Levent Şendur, Brandon Whitchee, John Suckling, Michael Brammer, Michael Breakspear

[23] A Haar-wavelet-based Lucy-Richardson algorithm for positron emission tomography image restoration
Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 648, Supplement 1, 21 August 2011, Pages S122-S127
Naomi W.P. Tam, Jih-Shian Lee, Chi-Min Hu, Ren-Shyan Liu, Jyh-Cheng Chen

[24] Functional and structural synergy for resolution recovery and partial volume correction in brain PET
NeuroImage, Volume 44, Issue 2, 15 January 2009, Pages 340-348
Miho Shidahara, Charalampos Tsoumpas, Alexander Hammers, Nicolas Bousson, Dimitris Visvikis, Tetsuya Suhara, Iwao Kanno, Federico E. Turkheimer

[25] Optimization of dynamic measurement of receptor kinetics by wavelet denoising
NeuroImage, Volume 30, Issue 2, 1 April 2006, Pages 444-451
Nathaniel M. Alpert, Anthonin Reilhac, Tat C. Chio, Ivan Selesnick

[26] Analysis of brain electrical topography by spatio-temporal wavelet decomposition
Mathematical and Computer Modelling, Volume 49, Issues 11-12, June 2009, Pages 2224-2235
Adil Deniz Duru, Ahmet Ademoglu, Tamer Demiralp

[27] Classification of schizophrenic patients and healthy controls using [18F] fluorodopa PET **imaging**
Schizophrenia Research, Volume 106, Issues 2-3, December 2008, Pages 148-155
Subrata K. Bose, Federico E. Turkheimer, Oliver D. Howes, Mitul A. Mehta, Rhian Cunliffe, Paul R. Stokes, Paul M. Grasby

[28] Classification of human brain tumours from MRS data using Discrete wavelet Transform and Bayesian Neural Networks
Expert Systems with Applications, Volume 39, Issue 5, April 2012, Pages 5223-5232
Carlos Arizmendi, Alfredo Vellido, Enrique Romero

[29] Exploring feature-based approaches in PET images for predicting cancer treatment outcomes
Pattern Recognition, Volume 42, Issue 6, June 2009, Pages 1162-1171
I. El Naqa, P.W. Grigsby, A. Apte, E. Kidd, E. Donnelly, D. Khullar, S. Chaudhari, D. Yang, M. Schmitt, Richard Laforest, W.L. Thorstad, J.O. Deasy

[30] Computer-assisted detection of infectious lung diseases: A review
Computerized Medical Imaging and Graphics, Volume 36, Issue 1, January 2012, Pages 72-84
Ulaş Bağcı, Mike Bray, Jesus Caban, Jianhua Yao, Daniel J. Mollura

[31] ^{18}F -FDG PET- imaging analysis for computer aided Alzheimer's diagnosis
Information Sciences, Volume 181, Issue 4, 15 February 2011, Pages 903-916
I.A. Illán, J.M. Górriz, J. Ramírez, D. Salas-Gonzalez, M.M. López, F. Segovia, R. Chaves, M. Gómez-Rio, C.G. Puntonet, the Alzheimer's Disease Neuroimaging Initiative

[32] La tomographie par émission de positons (TEP) hors oncologie : l'exploration du cerveau
Médecine Nucléaire, Volume 32, Issue 10, October 2008, Pages 502-510
F. Lomeña

[33] PET radiotracers for molecular imaging in the brain: Past, present and future
NeuroImage, In Press, Corrected Proof, Available online 28 December 2011
Luc Zimmer, André Luxen

[34] **Brain effective connectivity modeling for alzheimer's disease by sparse Gaussian Bayesian network**

Shuai Huang, Jing Li, Jieping Ye, Adam Fleisher, Kewei Chen, Teresa Wu, Eric Reiman

August 2011

[35] PET imaging using ^{64}Cu -labeled sulfophthalocyanines: Synthesis and biodistribution
Bioorganic & Medicinal Chemistry Letters, Volume 21, Issue 24, 15 December 2011, Pages 7470-7473
Elena R. Ranyuk, Nico le Cauchon, Hasrat Ali, Roger Lecomte, Brigitte Guérin, Johan E. van Lier

[36] Simultaneous PET/MR imaging in a human brain PET/MR system in 50 patients Current state of image quality
European Journal of Radiology, In Press, Corrected Proof, Available online 17 January 2012
N.F. Schwenger, L. Stegger, S. Bisdas, C. Schraml, A. Kolb, A. Boss, M. Müller, M. Reimold, U. Ernemann, C.D. Claussen, C. Pfannenbergl, H. Schmidt

[37] Personalized mode transductive spanning SVM classification tree
Information Sciences, Volume 181, Issue 11, 1 June 2011, Pages 2071-2085
Shaoning Pang, Tao Ban, Youki Kadobayashi, Nikola Kasabov

[38] A multi-class SVM classification system based on learning methods from indistinguishable chinese official documents
Expert Systems with Applications, Volume 39, Issue 3, 15 February 2012, Pages 3127-3134
JuiHsi Fu, SingLing Lee

[39] Classification of hyperspectral remote-sensing data with primal SVM for small-sized training dataset problem

Advances in Space Research, Volume 41, Issue 11, 2008, Pages 1793-1799
Mingmin Chi, Rui Feng, Lorenzo Bruzzone

[40] Document classification Algorithm Based on MMP and LS-SVM
Procedia Engineering, Volume 15, 2011, Pages 1565-1569
Ziqiang Wang, Xia Sun

[41] Wavelet SVM in Reproducing Kernel Hilbert Space for hyperspectral remote sensing image classification
Optics Communications, Volume 283, Issue 24, 15 December 2010, Pages 4978-4984
Peijun Du, Kun Tan, Xiaoshi Xing

[42] Generalized N -dimensional principal component analysis (GND-PCA) and its application on construction of statistical appearance models for medical volumes with fewer samples
Neurocomputing, Volume 72, Issues 10–12, June 2009, Pages 2276-2287
Rui Xu, Yen-Wei Chen

[43] **Machines à vecteurs supports Didacticiel**
Hervé Frezza-Buet, Supélec, 25 Janvier 2012

[44] **Support Vector Machines ou séparateurs à vaste marge**
A. Cornuéjols et L. Miclet. Eyrolles 2007.

[45] « **SVM** » par Mohamadally Hasan, Fomani Boris

[46] Bloch, "**Fusion numérique d'informations en traitement du signal et des images**", Hermès Science Publication, collection IC2, Paris, France, 2003. ISBN 2-7462-0628-5.

[47] Zeeuw, P.M. (1998), "**Wavelet and image fusion**" CWI, Amsterdam, March 1998.

[48] « **Caractérisation des tumeurs gliales en TEP/TDM à la 18F-Dopa et en IRM de perfusion** » par Christophe NIOCHE, Juin 2011

[49] "**Image feature analysis and computer-aided diagnosis in digital radiography. 1. Automated detection of microcalcifications in mammography.**" CHAN, H. P., DOI, K., GALHOTRA, S., VYBORNÝ, C. J., MACMAHON, H. AND JOKICH, P. M. (1987).

[50] "**Computerized detection of lung nodules in digital chest radiographs.**" Giger, M. L., Doi, K. and MaxMahon, H. (1987).

[51] "**Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center.**" Freer, T.W. and Ulissey, M. J. (2001).

[52] "**Single Reading with Computer-aided Detection and Double Reading of Screening Mammograms in the United Kingdom National Breast Screening Program.**" Gilbert, F. J., Astley, S. M., McGee, M. A., Gillan, M. G. C., Boggis, C. R. M., Griffiths, P. M. and Duffy, S. W. (2006).

[53] "**Computer-aided diagnosis in medical imaging: historical review, current status and future potential.**" Doi, K. (2007).

[54] "**Computer-aided detection and automated CT volumetry of pulmonary nodules.**" Marten, K. and Engelke, C. (2007).

[55] **True positive reduction method for cancer diagnosis based on FDG-PET images.**" Tozaki, T., Fujimoto, T., Sakamoto, S., Matsumoto, K. and Senda, M. (2004).

[56] **"Computer assisted diagnosis method of whole body cancer using FDG-PET images".** Tozaki, T., Senda, M., Sakamoto, S. and Matsumoto, K. (2003).

[57] **"Automatic hot spot detection and segmentation in whole body FDG-PET images."** Guan, H. Y., Kubota, T., Huang, X. L., Zhou, X. S. and Turk, M. (2006)

[58] **" Détection automatique de fixations radioactives dans les images TEP oncologiques "** S. Tomei, (2009).

[59] **"Surface normal overlap: A computer-aided detection algorithm with application to colonic polyps and lung nodules in helical CT."** Paik, D. S., Beaulieu, C. F., Rubin, G. D., Acar, B., Jeffrey, R. B., Yee, J., Dey, J. and Napel, S. (2004)

[60] **"Computer-aided detection of microcalcification clusters on full-field digital mammograms: Multiscale pyramid enhancement and false positive reduction using an artificial neural network".** Ge, J., Wei, J., Hadjiiski, L. M., Sahiner, B., Chan, H.-P., Helvie, M. A. and Zhou, C. (2005).

[61] **"Approaches for automated detection and classification of masses in mammograms."** Cheng, H. D., Shi, X. J., Min, R., Hu, L. M., Cai, X. P. and Du, H. N. (2006).

[62] **"Textural features for image classification."** [Haralick et al., 1973] Haralick, R. M., Shanmugam, K. and Dinstein, I. (1973).

[63] IRENE BUVAT

[64] **"Autonomous detection of pulmonary nodules on CT images with a neural network-based fuzzy system."** Lin, D.-T., Yan, C.-R. and Chen, W.-T. (2005).

[65] **"Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans."** Li, Q., Sone, S. and Doi, K. (2003).

[66] **"Feature subset selection for improving the performance of false positive reduction in lung nodule CAD."** Boroczky, L., Zhao, L. and Lee, K. P. (2006).

[67] **"Local contralateral subtraction based on bilateral symmetry of lung for reduction of false positives in computerized detection of pulmonary nodules."** Yoshida, H. (2004).

[68] **"Spatial-Frequency Channels in Human Vision."** Sachs, M. B., Nachmias, J. and ROBSON, J. G. (1971).

[69] **"An Essay towards solving a Problem in the Doctrine of Chances."** Bayes, T. (1763).

[70] **"The use of multiple measurements in taxonomic problems."** Fisher, R. A. (1936).

[71] **"A training algorithm for optimal margin classifiers."** Boser, B. E., Guyon, I. M. and Vapnik, V. N. (1992).

[72] **"Support vector machines for candidate nodules classification."** Campadelli, P., Casiraghi, E. and Valentini, G. (2005).

[73] **"Automated detection of pulmonary nodules in CT images with support vector machines".** Lu, L., Wanyu, L. and Xiaoming, S. (2009).

[74] **"A Computer-Aided Diagnostic System to Characterize CT Focal Liver Lesions: Design and Optimization of a Neural Network Classifier."** Gletsos, M., Mougiakakou, S. G., Matsopoulos, G. K., Nikita, K. S., Nikita, A. S. and Kelekis, D. (2003).

[75] **"Improving performance of computer-aided detection scheme by combining results from two machine learning classifiers."** Park, S. C., Pu, J. and Zheng, B.(2009).

Rafecas et al., 2001, Karp et al., 1999, Moses et al., 1998, Shlichta, 2001, Lecoq and Korzhik, 2002, Aykac et al., 2006, Bercier et al., 2004

[76] **"Use of a Monte-Carlo based probability matrix for 3D iterative reconstruction of MADPET-II data"**. Magdalena Rafecas , Brygida Mosler , Melanie Dietz , Markus Pögl , Ros Stamatakis , David P. Mcelroy , Sibylle I. Ziegler Member (2001).

[77] **" Integrated pathway-genome databases and their role in drug discovery."** Karp PD, Krummenacker M, Paley S, Wagg J(1999).

[78] **" Early onset group B streptococcal neonatal infection in Oxford 1985"**, Moses, L. M., Heath, P. T., Wilkinson, A. R., Jeffery, H. E. & Isaacs, D. (1998). (1998).

[79] **"Variation, selection and evolution of function-valued traits."**Kingsolver, J. G., R. Gomulkiewicz, Shlichta, and P. A. Carter. 2001.

[80] **"Evolution of photoluminescence as a function of the structural order or disorder in CaMoO₄ nanopowders"**. Lecoq and Korzhik, 2002,

[81] **"New developments in image-based characterization of coated-particle nuclear fuel"** J Price, D Aykac, J Hunn, A Kercher, R Morris (2006)

[82] **"Reply to the Comment on "The importance of high-order correlation effects for the CO-CO interaction potential"** Michał Rode, Joanna Sadlej, Robert Moszynski^a, Paul E.S. Wormer, Ad van der Avoird(2001).

[83] **"An inter-comparison of three spectral-deconvolution algorithms for gamma-ray spectroscopy"** Meng and Ramsden, 2000

[84] **"System Design for Accurately Estimating the Spectral Reflectance of Art Paintings"** Hideaki Haneishi, Takayuki Hasegawa, Asako Hosoi, Yasuaki Yokoyama, Norimichi Tsumura, and Yoichi Miyake.2000

[85] **"Mast cells as targets of corticotropin-releasing factor and related peptides"** Theoharis C. Theoharides, Jill M. Donelan, Nikoletta Papadopoulou, Jing Cao, Duraisamy Kempuraj, Pio Cont.2004

[86] **“Fully 3-D PET reconstruction with system matrix derived from point source measurements”** Panin et al., 2006

[87] **“Deconvolution-Based Partial Volume Correction in Raclopride-PET and Monte Carlo Comparison to MR Based Method”**. J. Tohka and A. Reilhac. 2008

[88] **“A large fraction of unclassified variants of the mismatch repair genes *MLH1* and *MSH2* is associated with splicing defects”** Isabelle Tournier¹, Myriam Vezain¹, Alexandra Martins¹, Françoise Charbonnier¹, Stéphanie Baert-Desurmont^{1,2}, Sylviane Olschwang³, Qing Wang⁴, Marie Pierre Buisine⁵, Johann Soret⁶, Jamal Tazi⁶, Thierry Frébourg, Mario Tosi¹⁽²⁰⁰⁷⁾

[89] **“A multiresolution image based approach for correction of partial volume effects in emission tomography”** N Boussion, M Hatt, F Lamare, Y Bizais, A Turzo, C Cheze-Le Rest and D Visvikis 2006

[90] **“Regulation of endothelium-derived nitric oxide production by the protein kinase Akt.”** Fulton D, Gratton JP, McCabe TJ, Fontana J, Fujio Y, Walsh K, Franke TF, Papapetropoulos A, Sessa WC (1999)

Résumé

Le diagnostic précoce et précis des pathologies neurologiques reste une tâche difficile pour les spécialistes.

Parmi les modalités d'imagerie fonctionnelle les plus performantes, on retrouve la tomographie par émission de positons, qui est souvent utilisée dans le but de parvenir à un diagnostic précoce. Toutefois, l'évaluation classique des images TEP repose souvent sur la réorientation manuelle, la lecture visuelle de coupes tomographiques et de l'analyse semi-quantitative de certaines régions d'intérêt. Ces mesures prennent du temps et sont subjectives et sujettes à erreur.

L'objectif de notre travail est de proposer un système d'aide au diagnostic assisté par ordinateur (CAD), afin d'aider les cliniciens dans leur interprétation.

La base d'images utilisée dans ce mémoire est : « PET-SORTEO ». C'est une plateforme de simulation, permettant la génération de données TEP brutes réalistes en un temps de calcul acceptable. Cette plateforme est issue d'une collaboration impliquant d'une part le CERMEP à Lyon et d'autre part le McConnell Brain Imaging Centre de l'Institut Neurologique de Montréal (université McGill). La validation du modèle de génération des données TEP, incluant la simulation des projections brutes, la correction et la reconstruction.

L'approche proposée est basée sur la sélection des paramètres des images et sur la classification réalisée par les machines à vecteurs de support (SVM). Une étude est réalisée afin de trouver par une approche fréquentielle par ondelettes, les paramètres discriminants à partir des images TEP. Ces caractéristiques seront alors envoyées à un classifieur SVM après la réduction du vecteur par deux approches : l'ACP et le calcul des moyennes.

Le système proposé donne une précision de 87% dans la classification des images saines et pathologiques.

Mots clés : Tomographie par Emission de Positons (TEP), Computer Aided Diagnosis (CAD), transformée en ondelette, Machines à Vecteurs de support (SVM).