



N°d'ordre NNT : 2021LYSEI092

THESE de DOCTORAT DE L'UNIVERSITE DE LYON
opérée au sein de
(Institut National des Sciences Appliquées, INSA - Lyon)

Ecole Doctorale N° 160
(ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE)

Spécialité/ discipline de doctorat :
Traitement du Signal et de l'Image

Soutenue publiquement le 08/12/2021, par :
Suzanne Bussod

**Développement d'algorithmes
d'apprentissage profond pour la
tomographie spectrale :
Application à l'arthrose du genou**

Devant le jury composé de :

Douek, Philippe	Professeur des universités, Université Lyon 1 Praticien hospitalier, HCL	Président du jury
Rodet, Thomas	Professeur des Universités, ENS Cachan	Rapporteur
Visvikis, Dimitris	Directeur de recherche, INSERM, Brest	Rapporteur
Chappard, Christine	Chargée de recherche, INSERM, Paris	Examinatrice
Boursier, Yannick	Maître de conférence, Université Aix/Marseille	Examineur
Peyrin, Françoise	Directrice de recherche, INSERM, Lyon	Directrice de thèse
Ducros, Nicolas	Maître de conférence, INSA, Lyon	Co-encadrant
Abascal, Juan	Ingénieur R&D, Carestream dental, Paris	Invité

Département FEDORA – INSA Lyon - Ecoles Doctorales

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	<u>CHIMIE DE LYON</u> https://www.edchimie-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage secretariat@edchimie-lyon.fr	M. Stéphane DANIELE C2P2-CPE LYON-UMR 5265 Bâtiment F308, BP 2077 43 Boulevard du 11 novembre 1918 69616 Villeurbanne directeur@edchimie-lyon.fr
E.E.A.	<u>ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE</u> https://edeea.universite-lyon.fr Sec. : Stéphanie CAUVIN Bâtiment Direction INSA Lyon Tél : 04.72.43.71.70 secretariat.edeea@insa-lyon.fr	M. Philippe DELACHARTRE INSA LYON Laboratoire CREATIS Bâtiment Blaise Pascal, 7 avenue Jean Capelle 69621 Villeurbanne CEDEX Tél : 04.72.43.88.63 philippe.delachartre@insa-lyon.fr
E2M2	<u>ÉVOLUTION, ÉCOSYSTÈME, MICROBIOLOGIE, MODÉLISATION</u> http://e2m2.universite-lyon.fr Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 secretariat.e2m2@univ-lyon1.fr	M. Philippe NORMAND Université Claude Bernard Lyon 1 UMR 5557 Lab. d'Ecologie Microbienne Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69 622 Villeurbanne CEDEX philippe.normand@univ-lyon1.fr
EDISS	<u>INTERDISCIPLINAIRE SCIENCES-SANTÉ</u> http://ediss.universite-lyon.fr Sec. : Sylvie ROBERJOT Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 secretariat.ediss@univ-lyon1.fr	Mme Sylvie RICARD-BLUM Institut de Chimie et Biochimie Moléculaires et Supramoléculaires (ICBMS) - UMR 5246 CNRS - Université Lyon 1 Bâtiment Raulin - 2ème étage Nord 43 Boulevard du 11 novembre 1918 69622 Villeurbanne Cedex Tél : +33(0)4 72 44 82 32 sylvie.ricard-blum@univ-lyon1.fr
INFOMATHS	<u>INFORMATIQUE ET MATHÉMATIQUES</u> http://edinfomaths.universite-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage Tél : 04.72.43.80.46 infomaths@univ-lyon1.fr	M. Hamamache KHEDDOUCI Université Claude Bernard Lyon 1 Bât. Nautibus 43, Boulevard du 11 novembre 1918 69 622 Villeurbanne Cedex France Tél : 04.72.44.83.69 hamamache.kheddouci@univ-lyon1.fr
Matériaux	<u>MATÉRIAUX DE LYON</u> http://ed34.universite-lyon.fr Sec. : Yann DE ORDENANA Tél : 04.72.18.62.44 yann.de-ordenana@ec-lyon.fr	M. Stéphane BENAYOUN Ecole Centrale de Lyon Laboratoire LTDS 36 avenue Guy de Collongue 69134 Ecully CEDEX Tél : 04.72.18.64.37 stephane.benayoun@ec-lyon.fr
MEGA	<u>MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE</u> http://edmega.universite-lyon.fr Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bâtiment Direction INSA Lyon mega@insa-lyon.fr	M. Jocelyn BONJOUR INSA Lyon Laboratoire CETHIL Bâtiment Sadi-Carnot 9, rue de la Physique 69621 Villeurbanne CEDEX jocelyn.bonjour@insa-lyon.fr
ScSo	<u>ScSo*</u> https://edsciencessociales.universite-lyon.fr Sec. : Mélina FAVETON INSA : J.Y. TOUSSAINT Tél : 04.78.69.77.79 melina.faveton@univ-lyon2.fr	M. Christian MONTES Université Lumière Lyon 2 86 Rue Pasteur 69365 Lyon CEDEX 07 christian.montes@univ-lyon2.fr

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

Résumé

La tomographie spectrale est une nouvelle modalité d'imagerie à rayons X qui permet d'acquérir des données avec une dimension énergétique. Cela est possible grâce aux détecteurs à comptage de photons qui classent les photons en fonction de leur énergie. En exploitant cette dimension énergétique nous pouvons estimer les matériaux constituant l'objet ou reconstruire des images mono-énergétiques. Dans le domaine médical, la décomposition de matériaux est souvent opérée sur des bases de matériaux anatomiques tels que les tissus mous et l'os, combinés à des produits de contraste.

D'un point de vue mathématique, la décomposition en matériaux est un problème inverse non linéaire et mal posé. De nombreuses méthodes de décomposition de matériaux ont été développées. Certaines sont basées sur l'inversion d'un modèle physique prenant en compte la source, la réponse du détecteur et l'atténuation théorique des matériaux constituant l'objet. Toutefois, les algorithmes d'apprentissage profond peuvent également être utilisés pour résoudre des problèmes inverses. Des travaux récents en tomographie ont montré que ceux-ci peuvent améliorer les qualités des reconstructions et sont plus rapides que les algorithmes basés sur le modèle.

Le but de la thèse est de développer des algorithmes d'apprentissage profond pour la décomposition de matériaux et la reconstruction d'images mono-énergétiques en tomographie spectrale et de les évaluer par rapport à des méthodes basées sur des techniques d'optimisation. Nous nous sommes particulièrement focalisés sur l'application de la tomographie spectrale à l'arthrose du genou sans produit de contraste.

Des genoux humains excisés ont été scannés à l'aide d'un prototype de scanner spectral mais aussi par tomographie synchrotron mono-énergétique afin d'avoir une image de référence. Des fantômes de genoux ont été générés à partir de ces volumes synchrotron en segmentant les matériaux tels que les tissus mous, l'os et potentiellement le cartilage. Puis, les projections spectrales sont simulées en utilisant des modèles de scanners spectraux. Nous avons ensuite développé un algorithme d'apprentissage profond basé sur U-net afin d'effectuer la décomposition de matériaux (tissus mous et os) dans le domaine des projections. Les résultats sont comparés à un algorithme de décomposition itératif basé sur le modèle, Gauss-Newton régularisé. Nous reconstruisons également les images mono-énergétiques soit grâce à la décomposition de matériaux, soit directement grâce à une adaptation de la méthode d'apprentissage. Les méthodes sont comparées en calculant l'erreur quadratique moyenne et l'indice de similarité. Finalement, les méthodes basées sur le modèle et d'apprentissage sont appliquées sur des données réelles issues du prototype clinique de scanner spectral.

Mots clefs : Tomographie spectrale, Apprentissage profond, Décomposition de matériaux, Images mono-énergétiques, Arthrose du genou

Abstract

Spectral computed tomography is an emerging modality of X ray imaging that allows to acquire energy resolved data thanks to photon counting detectors. They are able to sort photons according to their energy. This extra energy dimension allows to decompose the objet into its material constituents or to reconstruct virtual mono-energetic images. In the medical field, material decomposition is often based on anatomic materials like soft tissue and bone, combined with contrast agents. From a mathematical point of view, material decomposition is an ill posed non linear inverse problem. Several material decomposition methods have been developped. Ones are based on the inversion of a physical model taking into account the source spectrum, detector reponse function and theoretical material constituents' attenuation. However, a lot of deep learning algorithms can be used to solve inverse problems. Recent studies in tomography show that they can improve reconstructions accuracy and are faster than the model-based one.

The goal of this thesis is developping deep learning algorithms for material decomposition and mono-energetic images reconstruction and evaluating them with respect to methods based on optimization. In particular, we focus on the application of spectral computed tomography to knee osteoarthritis without contrast agent. Excised human knees are scanned with a spectral computed tomography scanner but also with synchrotron mono-energetic computed tomography to have a gold standard. Phantoms are generated from this synchrotron computed tomography volumes, segmenting materials as soft tissue, bone and potentially cartilage. Then, spectral projections are simulated using spectral computed tomography scanners models. After, we developped a learning algorithm based on U-net to decompose materials (soft tissue and bone) in the projection domain. Results are compared with a model-based iterative decomposition algorithm, regularized Gauss-Newton. We also reconstruct mono-energetic images thanks to decomposed material maps, or directly with an adaptation of the learning method. We compare the methods computing the mean squared error and the structural similarity matrix. Finally, model-based and learning-based methods are applied to real data from the spectral computed tomography clinic prototype.

Keywords : Spectral computed tomography, Deep learning, Material decomposition, Mono-energetic images, Knee osteoarthritis

Remerciements

Je voudrais tout d'abord remercier mes encadrants, Françoise Peyrin et Nicolas Ducros avec l'appui de Juan Abascal pour m'avoir permis de faire le stage de fin d'étude et la thèse à leur côté et qui m'ont aidé pendant ces trois ans et demi pour mener à bien ce travail. Ils m'ont permis de me dépasser et de découvrir une nouvelle facette de moi-même.

Tout d'abord, je voudrais remercier Françoise qui a vu mon potentiel pour me proposer un stage et cette thèse. Elle a su me donner sa passion pour l'imagerie ostéo-articulaire. De plus, elle a joué un très bon rôle de garde fou pour conserver la cohérence de la thèse et m'aider à évoluer. Elle a été une référence pour moi avec son excellence et sa bienveillance.

Je voulais remercier Nicolas pour m'avoir permis de rester sur le droit chemin. Il m'a aidé avec les problèmes mathématiques. Il m'a poussé lors des coups de mou et m'a permise de prendre du recul sur mon travail.

Je voulais remercier Juan pour m'avoir permis de mener à bien mon stage et entamer ma thèse dans de bonnes conditions. Je le remercie pour m'avoir permis de travailler sur l'apprentissage profond. Il m'a été d'une grande aide pour comprendre les problèmes inverses et leurs enjeux.

Je tiens à remercier Olivier Beuf pour m'avoir accueillie au sein du laboratoire CREATIS dans un cadre de travail bienveillant et épanouissant.

Je voudrais aussi remercier les membres qui ont accepté de participer à mon jury et spécialement Thomas Rodet et Dimitris Visviks pour leurs retours bienveillants et constructifs. Ces derniers m'ont permis de me poser des questions sur certains éléments du manuscrit pour l'améliorer.

Je voudrais remercier tous les membres du laboratoire avec qui j'ai pu échanger sur des sujets de travail ou des sujets plus divers. Je les remercie pour leur joie, leur bienveillance et les conversations toujours intéressantes.

Je voudrais tout d'abord remercier Tom qui m'a formée pendant le stage, m'a montré le monde de la thèse et à ouvert le chemin pour mes travaux avec sa thèse.

Je ne pourrais pas faire ces remerciements sans citer Louise Friot- Giroux qui est devenue une amie au cours de la thèse. Elle a pris du temps pour m'aider à rédiger le premier chapitre et pour revoir l'esthétisme de ma présentation.

Je voudrais remercier mes co-bureau de stage : Eloïse, Maxime, Valeriya, Sixiang et Antonio, pour la bonne ambiance durant six mois et des merveilleux échanges avec Eloïse. Antonio a aussi apporté sa contribution tout au long de la thèse pour m'aider à appréhender l'apprentissage profond et m'aider à optimiser mon réseau pour gagner du temps d'apprentissage et avoir des meilleurs résultats.

Je remercie également mes co-bureaux de thèse : Audrey, Daria, Sébastien, Sofiane, Antoine, Kannara, Ruifen et Hao, pour l'environnement calme, l'entraide et des très bonnes discussions ; et mes co-bureau d'ingénieur : Baptiste, Martin et encore Maxime que j'ai pu voir évoluer après les trois ans de thèse.

Je voudrais remercier Frank pour sa bonne humeur et toujours son mot pour rire ; et Anne-Lise pour les moments de rigolade et le soutien.

Je tiens à remercier également ma famille et mes amis. Tout d'abord, je voudrais envoyer mille mercis à ma mère qui a toujours été là pour moi et qui m'a soutenue depuis toujours et encore plus durant ces trois ans de thèse. Je voudrais remercier Mélody qui m'a soutenue, m'a permis de me changer les idées et toujours être au top. Le spam, ils se reconnaîtront, ils m'ont permise de garder le sourire, de prendre des bols d'air pendant toute la thèse et qui sont toujours là pour des bons moments. Marianne, Isabelle et Thomas ont aussi été là, même de loin et ont toujours été un soutien solide. Lucien m'a aussi accompagnée et je suis contente d'avoir fini ma thèse dans de bonnes conditions grâce à lui. Je voudrais aussi remercier mes minettes, Glue et Cléo, pour m'avoir accompagnée de très près lors de ma rédaction, rajouté quelques mots incompréhensibles en marchant sur le clavier et annoncé qu'il était l'heure d'aller dormir.

Je tiens également à remercier l'école doctorale EEA pour avoir financé ma thèse et FEDORA pour l'accompagnement de la thèse.

Table des matières

Résumé	iii
Abstract	v
Remerciements	vii
Introduction	1
I Etat de l'art	3
1 Tomographie à rayons X	5
1.1 Introduction	5
1.2 Rayons X	5
1.2.1 Génération des rayons X	6
1.2.2 Interactions lumière / matière	6
1.3 Formation de l'image	10
1.3.1 Détecteurs	10
1.3.2 Acquisition tomographique	11
1.4 Reconstruction tomographique	13
1.4.1 Transformée de Radon	13
1.4.2 Géométrie parallèle	14
1.4.3 Géométrie en éventail	16
1.4.4 Géométrie conique 3D	16
1.5 Applications médicales	16
1.5.1 Applications diverses	16
1.5.2 Cas particulier de l'arthrose	17
1.6 Conclusion	18
2 Tomographie X spectrale	19
2.1 Introduction	19
2.2 Formation de l'image	19
2.2.1 Tomographie à double énergie	19
2.2.2 Tomographie spectrale à comptage photonique	21
2.2.3 Modèle direct de la tomographie spectrale	21
2.3 Reconstruction	24
2.3.1 Reconstruction tomographique pour la tomographie spectrale	24
2.3.2 Décomposition	24
2.3.3 Stratégies pour la décomposition de matériaux	28
2.4 Applications médicales	30
2.4.1 Application aux lithiases rénales	30
2.4.2 Application à la goutte	30
2.4.3 Applications cardiovasculaires	30

2.4.4	Application à l'arthrose	31
2.5	Conclusion	31
3	Apprentissage profond pour la résolution de problèmes inverses	33
3.1	Problèmes inverses	33
3.1.1	Problème direct et inverse	33
3.1.2	Résolution des problèmes inverses	34
3.1.3	Introduction à l'apprentissage profond pour la résolution de problèmes inverses	36
3.2	Introduction à l'apprentissage profond	36
3.2.1	Cas des réseaux de neurones convolutifs	37
3.2.2	Aspect techniques de la phase d'apprentissage	39
3.3	Apprentissage pour la résolution de problèmes inverses	46
3.3.1	Apprentissage profond et reconstruction tomographique	46
3.3.2	Cas particulier de la tomographie spectrale	48
3.4	Conclusion	50
II	Contributions	53
4	Données	55
4.1	Introduction	55
4.2	Échantillons et acquisitions	55
4.2.1	Tomographie synchrotron	56
4.2.2	Prototype de scanner spectral	58
4.3	Génération des fantômes de genoux	59
4.3.1	Segmentation des images	59
4.3.2	Changement d'unité	60
4.3.3	Augmentation géométrique des fantômes	61
4.4	Simulation des données spectrales	61
4.4.1	Paramètres du modèle de Schlomka	62
4.4.2	Paramètres du modèle du prototype Philips	63
4.5	Simulation des image mono-énergétiques	63
4.6	Bases de données pour l'apprentissage	63
4.6.1	Base de données réduite - BASE2LR	65
4.6.2	Bases de données étendues - BASE8	65
4.6.3	Données réelles	66
4.7	Conclusion	66
5	Méthodes de décomposition proposées	67
5.1	Introduction	67
5.2	Reconstruction tomographique	70
5.2.1	Projections simulées	70
5.2.2	Projections expérimentales	70
5.3	Décomposition de matériaux	71
5.3.1	Méthode basée sur le modèle	71
5.3.2	Méthodes par apprentissage	71
5.4	Reconstruction des images mono-énergétiques	72
5.4.1	A partir des décomposition en matériaux	72
5.4.2	Méthodes directes par apprentissage	73
5.5	Phase d'apprentissage	74

5.6	Métriques d'évaluation	74
5.6.1	Erreur quadratique moyenne	74
5.6.2	Indice de similarité structurel	75
5.6.3	Évaluation du bruit	75
5.7	Conclusion	76
6	Apprentissage de la décomposition dans le domaine des projections : étude sur des simulations	77
6.1	Introduction	77
6.2	Fantômes basse résolution	77
6.2.1	BASE2LR	77
6.2.2	BASE8LR	80
6.3	Fantômes à la résolution du scanner Philips	86
6.3.1	BASE8HR	87
6.4	Conclusion	93
7	Apprentissage de la décomposition dans le domaine image : étude sur des simulations	97
7.1	Introduction	97
7.2	Fantômes à la résolution du scanner Philips - BASE8HR	97
7.2.1	Décomposition de matériaux	98
7.2.2	Reconstruction des images mono-énergétiques	100
7.3	Conclusion	103
8	Application des méthodes d'apprentissage sur les données réelles	107
8.1	Introduction	107
8.2	Acquisitions	107
8.3	Domaine des projections	108
8.3.1	Décomposition de matériaux	108
8.3.2	Décompositions mono-énergétiques	110
8.4	Domaine image	111
8.4.1	Décomposition de matériaux	112
8.4.2	Reconstruction des images mono-énergétiques	113
8.5	Conclusion	114
	Conclusion	117
	Bibliographie	123

Table des figures

1.1	Cliché de la main de la femme de Wilhelm Röntgen, une des premières images avec des rayons X (source : Wikipédia).	6
1.2	Schéma d'un tube à rayons X avec C la cathode, A l'anode, W l'eau pour le refroidissement du dispositif, U_h le courant de chauffage de la cathode et U_a le courant d'accélération (source : Wikipédia).	7
1.3	Spectre des rayons X à la sortie du tube (en bleu) et après filtration avec 2.5mm d'épaisseur équivalente d'aluminium (en rouge); généré avec SpekCalc (Poludniowski et al., 2009).	7
1.4	Schéma de l'effet photoélectrique.	8
1.5	Schéma de la diffusion Compton.	9
1.6	Prépondérance des 3 interactions, présentes dans la gamme d'énergie médicale, en fonction de l'énergie.	10
1.7	Différence entre la conversion indirecte et la conversion directe	11
1.8	Différents types de géométrie. (A) Première génération, translation-rotation en faisceaux parallèles; (B) Deuxième génération, translation-rotation en faisceau en éventail; (C) Troisième génération, rotation en faisceau en éventail.	12
1.9	Projection à l'aide de la transformée de Radon dans le cas d'une géométrie parallèle.	13
1.10	Exemple, sur le fantôme de Shepp Logan, avec une image 200×200 , projetée avec 180 angles sur 180° , de la rétroprojection simple face à la rétroprojection filtrée avec le filtre rampe.	15
1.11	Comparaison de la reconstruction tomographique avec un nombre limité de projections, 45, sur 180° et sur 45° , et la reconstruction avec 180 projections sur 180°	15
1.12	Anatomie d'un genou atteint d'arthrose (à droite) et sain (à gauche) (source : <i>Issues in Osteoarthritis Care 2018</i>).	17
2.1	Différents types d'acquisitions en tomographie spectrale, avec (A) deux sources à différentes énergies, (B) un changement d'énergie, (C) un détecteur en couches et (D) le filtrage du spectre.	20
2.2	Schéma du détecteur à comptage photonique avec un capteur lumineux soit de type CdTe, soit CZT, un circuit électronique spécialisé pour une application (ASIC), un amplificateur, des comparateurs et des compteurs, pour un détecteur comptant sur I bins.	22
2.3	Différentes interactions agissant dans un détecteur à comptage de photons avec (A) l'empilement, (B) le partage de charge, (C) la mauvaise détection et (D) l'effet photoélectrique.	22
2.4	Différentes stratégies pour obtenir les cartes de matériaux reconstruites à partir des projections acquises. Ici, les projections sont acquises sur 4 bins et décomposées en 2 matériaux.	25

2.5	Différentes stratégies pour obtenir les images mono-énergétiques à partir des projections acquises. Ici, les projections sont acquises sur 4 bins et on récupère des images mono-énergétiques à 70 keV.	26
2.6	Coefficients d'atténuation massique des tissus mous, de l'os, de l'iode et du gadolinium entre 0 et 150 keV.	28
3.1	Schéma de l'architecture des cellules composant la rétine (source : (Purves et al., 2005)).	37
3.2	Schéma d'un neurone avec x_i les entrées, ω_i les poids, f la fonction d'activation non linéaire, β le biais et y la sortie. Le neurone apprend les poids ω_i et le biais β pour produire la sortie y compte tenu des données d'apprentissage x_i	38
3.3	Perceptron multicouche avec 4 entrées, 1 sortie et 1 couche cachée avec 5 neurones (source : (Hassan et al., 2015)).	38
3.4	Sortie de la première couche de convolution avec 96 filtres de taille $11 \times 11 \times 3$ pour un problème de classification sur la base ImageNet (source : (Krizhevsky, Sutskever et Hinton, 2017)).	39
3.5	Schéma d'un réseau avec une couche de convolution annoté avec le filtre de convolution en orange (source : (Rumelhart, Hinton et Williams, 1986)).	40
3.6	Architecture VGG16 de (Simonyan et Zisserman, 2015) (source : (<i>VGG16 - Convolutional Network for Classification and Detection</i> 2018)). Les deux premières dimensions représentent la taille de l'image et la troisième est le nombre de filtres.	41
3.7	Architecture du U-net de (Ronneberger, Fischer et Brox, 2015). Les deux dimensions sont la taille de l'image. Le nombre au dessus représente le nombre de filtres.	41
3.8	Augmentations courantes avec (A) l'image originale, (B) l'image retournée verticalement, (C) l'image avec une rotation de 30° , (D) l'image translaturée vers la gauche, (E) l'image avec un facteur d'échelle de 2 et (F) un patch de l'image	44
3.9	Différentes stratégies pour l'utilisation de l'apprentissage profond pour la reconstruction tomographique (source : (Ravishankar, Ye et Fessler, 2020)).	46
3.10	Schéma de la décomposition de matériaux dans deux domaines. (A) Décomposition dans le domaine des projections. (B) Décomposition dans le domaine image.	48
3.11	Architecture inspirée de U-net développée par (Clark, Holbrook et Badea, 2018), avec seulement deux changements de résolution, pour la décomposition de matériaux dans le domaine image.	51
4.1	Photo d'un échantillon de genou, juste sorti de la glace et emballé sous vide, sur la table, prêt à être scanné par le scanner spectral.	56
4.2	Tranche d'un genou scanné en imagerie synchrotron à trois énergies 55, 75 et 150 keV.	58
4.3	Rendu 3D des os des genoux #422 (A) et #438 (B), où l'on voit le fémur, la patella, le tibia et la tête de la fibula.	59
4.4	Volume original (A) et sa segmentation (B) avec l'os en rouge, les tissus mous en jaune et le cartilage en blanc.	60

4.5	Visualisation des cartes de matériaux avec du cartilage, des tissus mous et l'os avec (A) l'image du synchrotron, (B) la carte de segmentation des tissus mous, (C) de l'os et (D) du cartilage.	60
4.6	Projection d'un volume augmenté. (A) Projection originale. (B) Rotation de -10° autour de (x, y) . (C) Rotation de 10° autour de (x, y) . (D) Changement d'échelle avec un facteur de 0.55. (E) Changement d'échelle avec un facteur de 1.45.	61
4.7	Spectre de la source à rayons X avec $N_0 = 10^7$ photons et fonctions de réponse du détecteur à 4 bins d'énergie (source : (Schlomka et al., 2008)).	62
4.8	Projections spectrales à 90° pour 3 niveaux de bruit $N_0 = 10^7$ (A), 10^6 (B) et 10^5 (C). De haut en bas : les 4 bins d'énergie.	62
4.9	Spectre de la source à rayons X et fonctions de réponse du détecteur du prototype clinique de Philips (Si-Mohamed et al., 2017).	63
4.10	Projections spectrales à 0° avec le modèle de Philips sans normalisation (A) et normalisation avec la mesure de plein flux (B). Chaque ligne est un bin.	64
4.11	Reconstruction tomographique des projections spectrales normalisées. Visualisée sur une coupe avec les bins de gauche à droite.	64
4.12	Images mono-énergétiques du genou de 40 à 100 keV reconstruites à partir des données spectrales.	65
5.1	Différentes stratégies de décomposition. Décomposition de matériaux dans le domaine des projections avec la méthode (A) de Gauss-Newton et (B) d'apprentissage. Décomposition de matériaux dans le domaine image avec la méthode d'apprentissage (C).	68
5.2	Décomposition des projections mono-énergétiques avec la méthode d'apprentissage (A). Reconstruction des images mono-énergétiques avec la méthode d'apprentissage (B).	69
5.3	Stratégies pour la reconstruction des images mono-énergétiques; à partir de la décomposition de matériaux dans le domaine de projections (A) ou le domaine image (A)-partie droite, ou à partir de projections mono-énergétiques (B)	70
5.4	Architecture U-net utilisée pour la décomposition de matériaux.	71
5.5	Architecture U-net utilisée pour la reconstruction de projections ou d'images mono-énergétiques.	73
6.1	Fonctions de coût d'apprentissage (train-bleu) et de validation (valid-orange) sur 500 époques pour une simulation de la BASE2LR avec 10^7 photons.	78
6.2	Résultats de la décomposition de matériaux dans le domaine des projections avec U-net sur la base d'apprentissage de la base BASE2LR, exposés pour la projection à 90° . (A) Vérité terrain pour la décomposition. (B) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B). (C) Matériaux décomposés avec U-net. (E) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$. (F) Erreur entre les matériaux décomposés avec U-net et la vérité terrain.	79

- 6.3 Reconstruction tomographique des résultats de la décomposition de matériaux dans le domaine des projections avec U-net sur la base d'apprentissage de la BASE2LR. (A) Vérité terrain pour la décomposition. (B) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B). **(C) Matériaux décomposés avec U-net.** (E) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisé et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$. **(F) Erreur entre les matériaux décomposés avec U-net et la vérité terrain.** 81
- 6.4 Images mono-énergétiques à 70 keV obtenues à partir de la reconstruction tomographique des projections décomposées avec U-net sur la base d'apprentissage de la BASE2LR. (A) Vérité terrain. (B) Images mono-énergétiques issues des matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B). **(C) Images mono-énergétiques issues des matériaux décomposés avec U-net.** (E) Erreurs entre les images mono-énergétiques issues des matériaux décomposés avec la méthode de Gauss-Newton régularisé et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$. **(F) Erreur entre les images mono-énergétiques issues des matériaux décomposés avec U-net et la vérité terrain.** 82
- 6.5 Fonctions de coût d'apprentissage (train-bleu) et de validation (valid-orange) sur 200 époques pour une simulation de la BASE8LR avec 10^5 photons. 83
- 6.6 Résultats de la décomposition de matériaux dans le domaine des projections avec U-net sur la base de test de la BASE8LR, visualisés pour les projections décomposées pour un angle de 90° . (A) Vérité terrain pour la décomposition. (B,C) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B) et $\alpha = 3$ (C). **(D) Matériaux décomposés avec U-net.** (F,G) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$ (F) et $\alpha = 3$ (G). **(H) Erreur entre les matériaux décomposés avec U-net et la vérité terrain.** 84
- 6.7 Reconstruction tomographique des résultats de la décomposition de matériaux avec U-net dans le domaine des projections sur la base de test de la BASE8LR. (A) Vérité terrain pour la décomposition. (B,C) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B) et $\alpha = 3$ (C). **(D) Matériaux décomposés avec U-net.** (F,G) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$ (F) et $\alpha = 3$ (G). **(H) Erreur entre les matériaux décomposés avec U-net et la vérité terrain.** 85

- 6.8 Images mono-énergétiques à 70 keV obtenues à partir de la reconstruction tomographique des projections décomposées avec U-net sur la base de test de la BASE8LR. (A) Vérité terrain. (B) Images mono-énergétiques issues des matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B). **(C) Images mono-énergétiques issues des matériaux décomposés avec U-net.** (E) Erreurs entre les images mono-énergétiques issues des matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$. **(F) Erreur entre les images mono-énergétiques issues des matériaux décomposés avec U-net et la vérité terrain.** 86
- 6.9 Influence du taux d'apprentissage sur les fonctions de coût d'entraînement et de validation pour 100 époques pour la décomposition de matériaux avec U-net dans le domaine des projections. 88
- 6.10 Reconstruction tomographique des résultats de la décomposition de matériaux avec U-net sur la base de test de la BASE8HR dans le domaine des projections. (A) Vérité terrain pour la décomposition. (C) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$. (D) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.3$. **(E,G,I) Matériaux décomposés avec U-net avec un taux d'apprentissage initial de 10^{-3} (E), 10^{-2} (G) et 10^{-1} (I).** (F,H,J) **Erreur entre les matériaux décomposés avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (F), 10^{-2} (H) et 10^{-1} (J).** 89
- 6.11 Influence du taux d'apprentissage sur les fonctions de coût d'entraînement et de validation pour 100 époques pour la reconstruction directe des images mono-énergétiques avec U-net. 91
- 6.12 Reconstruction tomographique des projections mono-énergétiques à 70 keV reconstruites avec U-net sur la base de test de la BASE8HR à partir de la décomposition de matériaux ou directement. (A) Vérité terrain pour l'image mono-énergétique. (C) Image mono-énergétique à partir de la décomposition de matériaux avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$. (D) Erreurs entre l'image mono-énergétique à partir de la décomposition de matériaux avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ et la vérité terrain. **(E,I,M) Image mono-énergétique à partie de la décomposition de matériaux avec U-net avec un taux d'apprentissage initial de 10^{-3} (E), 10^{-2} (I) et 10^{-1} (M).** (F,J,N) **Erreur entre les image-mono-énergétiques à partir de la décomposition de matériaux avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (F), 10^{-2} (J) et 10^{-1} (N).** (G,K,O) **Image mono-énergétique directement reconstruites avec U-net avec un taux d'apprentissage initial de 10^{-3} (G), 10^{-2} (K) et 10^{-1} (O).** (H,L,P) **Erreur entre les image-mono-énergétiques directement reconstruites avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (H), 10^{-2} (L) et 10^{-1} (P).** 92

- 6.13 Zoom sur la zone du cartilage sur la reconstruction tomographique des projections mono-énergétiques reconstruites avec U-net sur la base de test de la BASE8HR. (A) et (B) Vérité terrain avec l'emplacement et le zoom de la zone du cartilage. (C) Image mono-énergétique à partir de la reconstruction tomographique des matériaux décomposés avec Gauss-Newton. **(D) Reconstruction tomographique des projections mono-énergétiques reconstruites avec U-net avec un taux d'apprentissage initial à 10^{-1} .** (E,F) Images mono-énergétiques à partir des décomposition de matériaux dans le domaine des projections avec U-net avec un taux d'apprentissage initial à 10^{-3} (E) et 10^{-2} (F). 93
- 7.1 Influence du taux d'apprentissage sur les fonctions de coût d'entraînement et de validation pour la décomposition de matériaux avec U-net dans le domaine image sur 100 époques. 99
- 7.2 Résultats de la décomposition de matériaux avec U-net sur la base de test de la BASE8HR dans le domaine image. (A) Vérité terrain pour la décomposition. (C) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$. (D) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.3$. **(E,G,I) Matériaux décomposés avec U-net avec un taux d'apprentissage initial de 10^{-3} (E), 10^{-2} (G) et 10^{-1} (I).** **(E,H,J) Erreur entre les matériaux décomposés avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (F), 10^{-2} (H) et 10^{-1} (J).** 100
- 7.3 Influence du taux d'apprentissage sur les fonctions de coût d'entraînement et de validation pour la reconstruction directe des images mono-énergétiques avec U-net sur 100 époques. 102
- 7.4 Résultat de la reconstruction des images mono-énergétiques à 70 keV avec U-net sur la base de test de la BASE8HR, à partir des décompositions et directement. (A) Vérité terrain pour l'image mono-énergétique. (C) Image mono-énergétique à partir de la décomposition de matériaux avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$. (D) Erreurs entre l'image mono-énergétique à partir de la décomposition de matériaux avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ et la vérité terrain. **(E,I,M) Image mono-énergétique à partir de la décomposition de matériaux avec U-net avec un taux d'apprentissage initial de 10^{-3} (E), 10^{-2} (I) et 10^{-1} (M).** **(F,J,N) Erreur entre les image-mono-énergétiques à partir de la décomposition de matériaux avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (F), 10^{-2} (J) et 10^{-1} (N).** **(G,K,O) Image mono-énergétique directement reconstruites avec U-net avec un taux d'apprentissage initial de 10^{-3} (G), 10^{-2} (K) et 10^{-1} (O).** **(H,L,P) Erreur entre les image mono-énergétiques directement reconstruites avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (H), 10^{-2} (L) et 10^{-1} (P).** 103

- 7.5 Zoom sur la zone du cartilage sur la reconstruction tomographique des projections mono-énergétiques reconstruites avec U-net sur la base de test de la BASE8HR. (A) et (B) Vérité terrain avec l'emplacement et le zoom de la zone du cartilage. (C) Image mono-énergétique à partir de la reconstruction tomographique des matériaux décomposés avec Gauss-Newton. **(D) Images mono-énergétiques à partir des décompositions de matériaux dans le domaine image avec U-net avec un taux d'apprentissage initial à 10^{-2} .** (E) **Images mono-énergétiques reconstruites avec U-net avec un taux d'apprentissage initial à 10^{-1} .** . . . 105
- 8.1 Reconstruction tomographique des projections spectrales des données réelles dans les 5 bins d'énergie. 108
- 8.2 Reconstruction tomographique des résultats de la décomposition de matériaux avec U-net dans le domaine des projections. Matériaux décomposés la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), **avec U-net avec un taux d'apprentissage à 10^{-3} (B) et 10^{-2} (C).** 109
- 8.3 Reconstruction tomographique des images mono-énergétiques à 70 keV avec U-net. Images mono-énergétiques à partir des projections décomposées la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), **avec U-net avec un taux d'apprentissage initial à 10^{-3} (C) et 10^{-2} (D); ou directement reconstruites avec U-net avec taux d'apprentissage initial à 10^{-1} (B).** 110
- 8.4 Reconstruction tomographique du résultat de la reconstruction des projections mono-énergétiques à 70 keV avec U-net. Images mono-énergétiques à partir des projections décomposées la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), **avec U-net avec un taux d'apprentissage initial à 10^{-3} (C) et 10^{-2} (D); ou directement reconstruites avec U-net avec taux d'apprentissage initial à 10^{-1} (B).** 111
- 8.5 Résultats de la décomposition de matériaux avec U-net dans le domaine image. Reconstruction tomographique des projections décomposées avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A) et **matériaux décomposés dans le domaine image avec U-net avec un taux d'apprentissage initial à 10^{-2} .** 112
- 8.6 Résultats de la reconstruction des images mono-énergétiques à 70 keV avec U-net dans le domaine image. Reconstruction tomographique des projections mono-énergétiques créées à partir des projections décomposées la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), **image mono-énergétique à partir de la décomposition de matériaux dans le domaine image avec U-net avec un taux d'apprentissage initial à 10^{-2} (B) et image mono-énergétique à partir de la décomposition directe avec U-net (3).** . . . 113
- 8.7 Résultats de la reconstruction des images mono-énergétiques à 70 keV avec U-net dans le domaine image. Reconstruction tomographique des projections mono-énergétiques créées à partir des projections de matériaux décomposées la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), **image mono-énergétique à partir de la décomposition de matériaux dans le domaine image avec U-net avec un taux d'apprentissage initial à 10^{-2} (B) et image mono-énergétique à partir de la décomposition directe avec U-net (3).** 114

Liste des tableaux

4.1	Description des échantillons acquis et modalités d'acquisition, les lignes en gras sont les genoux présent dans la BASE8, L'identifiant correspond à un identifiant unique du patient alloué lors de l'anonymisation des échantillon, l'âge et le sexe correspondent à l'âge du patient dont le genou est issu au moment de sa mort et son sexe, <i>OA</i> correspond au niveau d'arthrose, <i>SRCT</i> à l'acquisition synchrotron avec l'énergie annoncée en keV, <i>SCT</i> à l'acquisition avec le scanner spectral, <i>Date</i> est l'année d'acquisition.	57
6.1	Résultats quantitatifs pour la reconstruction des images mono-énergétiques à partir de la décomposition de matériaux, à 70 keV pour une coupe, avec U-net et comparaison à la méthode de Gauss-Newton.	80
6.2	Résultats quantitatifs pour la décomposition de matériaux dans le domaine des projections avec U-net sur la base de test et comparé à la méthode de Gauss-Newton. Calculé sur la projection à 90° et après reconstruction pour la coupe 50.	84
6.3	Résultats quantitatifs pour la décomposition de matériaux de la BASE8HR avec U-net pour la base de test et comparaison avec la méthode de Gauss-Newton, visualisés sur la projection à 90° et après reconstruction pour la coupe 50.	88
6.4	Résultats quantitatifs pour la reconstruction des images mono-énergétiques à 70 keV à partir des matériaux décomposées avec U-net ou directement reconstruites avec U-net, et comparées avec la méthode de Gauss-Newton.	91
7.1	Résultats quantitatifs pour la décomposition de matériaux de la base 8HR avec U-net sur la base de test et comparaison avec la méthode de Gauss-Newton. Exposés pour la coupe 50.	98
7.2	Résultats quantitatifs des images mono-énergétiques à 70 keV avec U-net à partir de la décomposition et avec U-net directement apprise sur les images mono-énergétiques et comparé à la méthode de Gauss-Newton. Exposés pour une coupe.	100
8.1	Récapitulatif des résultats sur les données réelles pour la décomposition de matériaux et mono-énergétique dans le domaine des projections et le domaine image.	116
8.2	Temps de calcul pour une projection/coupe par méthode.	119

Notations

m	Indice du matériau	
E	Énergie	keV
μ	Coefficient d'atténuation linéaire	cm^{-1}
ν	Coefficient d'atténuation linéaire projeté	
τ	Coefficient d'atténuation massique	$\text{cm}^2 \cdot \text{g}^{-1}$
ρ	Masse volumique	$\text{g} \cdot \text{cm}^{-3}$
a	Masse volumique projetée	$\text{g} \cdot \text{cm}^{-2}$
N	Spectre atténué	nombre de photons
N_0	Spectre de la source	nombre de photons
s	Projection	
S	Projection reconstruite	
Rf	Transformée de Radon	
$F[f]$	Transformée de Fourier de f	
Bg	Rétroprojection simple de g	
G	Opérateur du problème direct de la tomographie spectrale	
X	Opérateur de la transformée à rayons X	
F	Mixage spectral	
ω	Poids du réseau	
β	Biais du réseau	
λ	Taux d'apprentissage	
κ	Nombre de conditionnement	
σ	valeurs singulières	
C	Fonction de coût	
D	Terme d'attache aux données	
R	Fonction de régularisation	
α	Coefficient de régularisation	
Γ	matrice de Tikhonov	

Introduction

La tomographie X spectrale est une nouvelle modalité d'imagerie à rayons X. Avec les systèmes les plus récents, les données sont acquises avec une dimension énergétique grâce à des détecteurs à comptage de photons qui trient les photons en fonction de leur énergie dans des intervalles d'énergie nommés "bins" d'énergie. Grâce à cette dimension énergétique, il est possible d'estimer les matériaux constituant l'objet ou reconstruire des images mono-énergétiques. Les images mono-énergétiques correspondent à la reconstruction du coefficient d'atténuation linéaire à une énergie donnée, alors que les images Hounsfield, habituellement obtenues en tomographie X standard, intègrent cette information sur une large gamme d'énergie. La réduction des intervalles d'énergie permet de limiter les artefacts. Dans le domaine médical, pour la décomposition de matériaux, on utilise souvent des matériaux anatomiques comme les tissus mous et l'os, combinés ou non avec des produits de contraste.

La décomposition de matériaux est un problème inverse non linéaire et mal posé. Des méthodes basées sur la modélisation du problème direct ont été proposées. Elles tentent d'inverser un modèle physique qui prend en compte le spectre de la source, la réponse du détecteur et l'atténuation théorique des matériaux à décomposer. Elles peuvent être basées sur de la calibration (Alvarez et Macovski, 1976; Long et Fessler, 2014) qui utilise par exemple la connaissance d'un matériau pour en décomposer d'autres. Des méthodes d'optimisation régularisées (Ducros et al., 2017) sont également utilisées et se basent sur la minimisation d'une fonction de coût pour décomposer les matériaux. Des algorithmes d'apprentissage ont été développés pour résoudre des problèmes inverses (McCann, Jin et Unser, 2017), par exemple avec l'architecture U-net (Liu et al., 2018) pour le débruitage, développée par (Ronneberger, Fischer et Brox, 2015). Nous souhaitons utiliser l'apprentissage profond pour la décomposition de matériaux. (Clark, Holbrook et Badea, 2018) proposa l'utilisation de U-net pour la décomposition de matériaux dans le domaine image. Le problème de décomposition de matériaux peut être réalisé dans le domaine image ou dans le domaine des projections. Les images mono-énergétiques sont d'usage reconstruites à partir des cartes de matériaux décomposés. L'apprentissage profond pourrait également permettre de reconstruire directement ces images.

Cette thèse portera sur la décomposition de matériaux et la reconstruction des images mono-énergétiques à l'aide de l'apprentissage profond pour une application à l'arthrose du genou. En effet, l'arthrose est une maladie dégénérative des articulations qui est diagnostiquée à l'aide de la radiographie quand l'espace de l'articulation est déjà réduit. Avec le projet ANR-17-CE19-0011 *SALTO* nous souhaitons visualiser précocement l'arthrose. La thèse a été financée par une bourse gouvernementale délivrée par l'école doctorale EEA de l'université de Lyon au sein du laboratoire CREATIS dans l'équipe "Imagerie Tomographique et thérapie par radiation".

Le document est structuré en 8 chapitres. Les trois premiers font l'état de l'art concernant la tomographie classique, la tomographie spectrale et l'apprentissage

profond pour la résolution de problèmes inverses. Les suivants décrivent les contributions telles que les bases de données créées pour l'occasion, le choix des méthodes et les applications dans le domaine des projections, le domaine image et pour une application à des données réelles.

Le Chapitre 1 introduit la tomographie X classique avec la génération des rayons X et l'interaction lumière/matière à l'origine de la formation d'image. Nous exposons aussi les méthodes de reconstruction tomographique pour retrouver l'objet à partir des mesures dans le domaine des projections. Nous terminons avec les applications médicales.

Le Chapitre 2 introduit la tomographie X spectrale en présentant la tomographie à double énergie et la tomographie spectrale à comptage de photons. Nous exposons ainsi la formation de l'image, dans ce contexte, avec la modélisation du problème direct. Nous présentons les différentes stratégies de reconstruction du problème inverse avec la décomposition de matériaux, la reconstruction des images mono-énergétiques et la reconstruction tomographique. Nous terminons avec les applications médicales qui ont pu être améliorées et développées grâce à cette modalité d'imagerie.

Le Chapitre 3 introduit l'apprentissage profond avec les réseaux de neurones convolutifs. Nous focalisons notre étude de l'état de l'art sur les algorithmes de résolution de problèmes inverses avec l'introduction des problèmes inverses, le cas de la reconstruction tomographique et les problèmes inverses de la tomographie spectrale. Nous explorons aussi la gestion des données et des hyperparamètres du réseau pour de meilleurs résultats.

Le Chapitre 4 présente la création des bases de données de genoux. Il explore la création des fantômes numériques et la simulation des données spectrales. Nous utilisons des modèles spectraux de la littérature ou d'un prototype de scanner spectral. Les travaux ont été présentés à la conférence ISBI 2019 (Bussod et al., 2019).

Le Chapitre 5 décrit les méthodes proposées pour la décomposition de matériaux et la reconstruction des images mono-énergétiques. Nous utilisons une architecture U-net entraînée pour la décomposition de matériaux ou la reconstruction des images mono-énergétiques, dans le domaine des projections ou dans le domaine image. Nous comparons les résultats de test de ces réseaux avec les résultats obtenus avec la méthode Gauss-Newton régularisée grâce à l'erreur quadratique moyenne et l'indice de similarité.

Le Chapitre 6 présente les résultats de la décomposition de matériaux et mono-énergétique dans le domaine des projections sur toutes les bases de données. Les travaux sur la base de données basse résolution ont été présentés à EUSIPCO 2020 (Bussod et al., 2021).

Le Chapitre 7 présente les résultats de la décomposition de matériaux et de la reconstruction des images mono-énergétiques dans le domaine image sur la base de données haute résolution.

Le Chapitre 8 présente les résultats de la décomposition de matériaux et de la reconstruction des images mono-énergétiques, dans le domaine des projections et dans le domaine image, des données réelles issues du prototype clinique de scanner spectral.

Première partie

Etat de l'art

Chapitre 1

Tomographie à rayons X

Dans ce chapitre nous introduirons les bases de la tomographie à rayons X. Tout d'abord nous présenterons les rayons X et leur interaction avec la matière. Nous expliquerons comment cette interaction permet de former l'image. Ensuite, nous introduirons la reconstruction tomographique. Finalement, nous décrirons quelques applications médicales de la tomographie à rayons X.

1.1 Introduction

L'imagerie à rayons X, communément appelée radiographie, image l'atténuation des rayons X après leur passage à travers un objet à scanner. Cette imagerie est 2D, très rapide et peu onéreuse mais ne permet pas de localiser des informations en profondeur contrairement à la tomographie assistée par ordinateur. Dans la tomographie assistée par ordinateur à rayons X, une multitude de projections 1D sont acquises en tournant autour de l'objet ou en faisant tourner l'objet. Dans le cas clinique, le tube à rayons X et le détecteur tournent autour du patient qui est fixe. Dans d'autre cas, le dispositif tube/détecteur reste fixe et l'échantillon tourne sur lui même à l'aide d'une plate-forme. Chaque signal 1D est appelé projection et est défini par l'angle d'acquisition. Les projections pour l'ensemble des angles constituent le sinogramme. A partir de ces projections, ce sinogramme, et en connaissant les angles d'acquisition, on peut reconstruire une image 2D à l'aide d'algorithmes de reconstruction tomographique (Peyrin et Engelke, 2012).

Nous introduirons les rayons X et l'interaction lumière / matière qui permet d'acquérir des images dont nous détaillerons le processus de formation d'image. Nous introduirons ensuite les méthodes permettant de reconstruire une coupe en 2D à partir de projections 1D. Nous décrirons également leur généralisation pour obtenir un volume 3D à partir de projections 2D. Nous finirons par les applications médicales générales de la tomographie à rayons X. Nous nous focaliserons sur l'imagerie de l'arthrose qui servira principalement à tester les algorithmes développés lors de la thèse.

1.2 Rayons X

Les rayons X sont les radiations avec une longueur d'onde comprise entre 10 pm et 10 nm, ce qui correspond à une fréquence entre $30 \cdot 10^{15}$ Hz et $30 \cdot 10^{18}$ Hz et à des énergies entre quelques eV et quelques MeV. Ils ont été découverts en 1895 par Wilhelm Conrad Röntgen et nommés de la sorte car ils représentaient des radiations inconnues. Il a observé que le rayonnement X traversait les tissus biologiques en prenant une image de la main de sa femme où on voit la forte atténuation des os et une encore plus forte de son alliance (Figure 1.1).



FIGURE 1.1 – Cliché de la main de la femme de Wilhelm Röntgen, une des premières images avec des rayons X (source : Wikipédia).

1.2.1 Génération des rayons X

Pour la génération des rayons X, on utilise un tube à rayons X qui est composé de trois parties plongées dans le vide :

1. une cathode avec un filament (en tungstène) chauffé avec un courant électrique qui émet des électrons ;
2. une anode en matériau cible (tungstène, molybdène ou cuivre) qui reçoit les électrons ;
3. un flux de courant électrique de l'ordre de quelques dizaines de kV.

Le principe du tube à rayons X est représenté sur la Figure 1.2. Le réchauffement de la cathode permet la libération d'électrons provenant du filament de tungstène par effet thermoionique. La différence de potentiel entre la cathode et l'anode accélère les électrons qui entrent en collision avec l'anode et interagissent avec les électrons et le noyau du matériau de la cathode. Il en résulte une conversion de l'énergie cinétique en radiation électromagnétique. Les raies caractéristiques du spectre sont dues aux interactions avec les couches électroniques de l'atome du matériau composant de l'anode. En effet, l'électron incident va heurter un électron de l'atome et l'éjecter. Le reste du spectre est dû aux interactions avec le noyau de cet atome. L'électron incident passera à côté du noyau, plus il sera près, plus le rayonnement de freinage (Bremsstrahlung en Allemand) sera important et produira un rayonnement de haute fréquence. Le spectre des rayons X émis dépend donc du matériau de l'anode et du courant d'accélération des électrons.

En radiographie diagnostique, il est nécessaire de filtrer le spectre du tube pour réduire la dose de radiation infligée au patient en réduisant les basses énergies qui ne peuvent pas traverser entièrement le patient et qui seront totalement absorbées par ses tissus (Webb et Kagadis, 2003). La filtration du spectre est illustrée Figure 1.3 avec 2.5mm d'épaisseur équivalente d'aluminium qui peut être réalisée avec la superposition de plaques de béryllium, d'aluminium et de cuivre.

1.2.2 Interactions lumière / matière

Le rayonnement X filtré est émis en direction du patient et le traverse. Les rayons X interagissent avec les matériaux composant le patient de plusieurs manières :

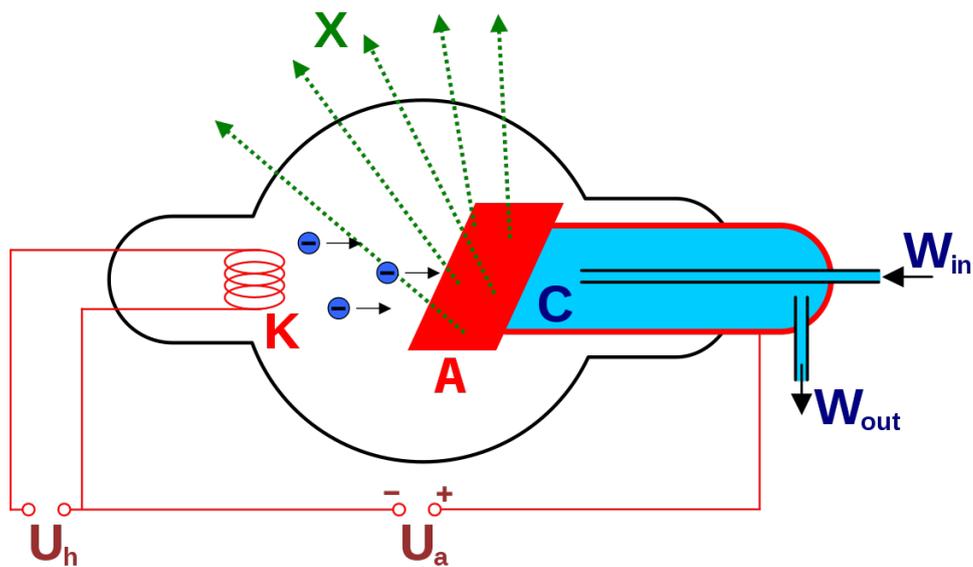


FIGURE 1.2 – Schéma d'un tube à rayons X avec C la cathode, A l'anode, W l'eau pour le refroidissement du dispositif, U_h le courant de chauffage de la cathode et U_a le courant d'accélération (source : Wikipédia).

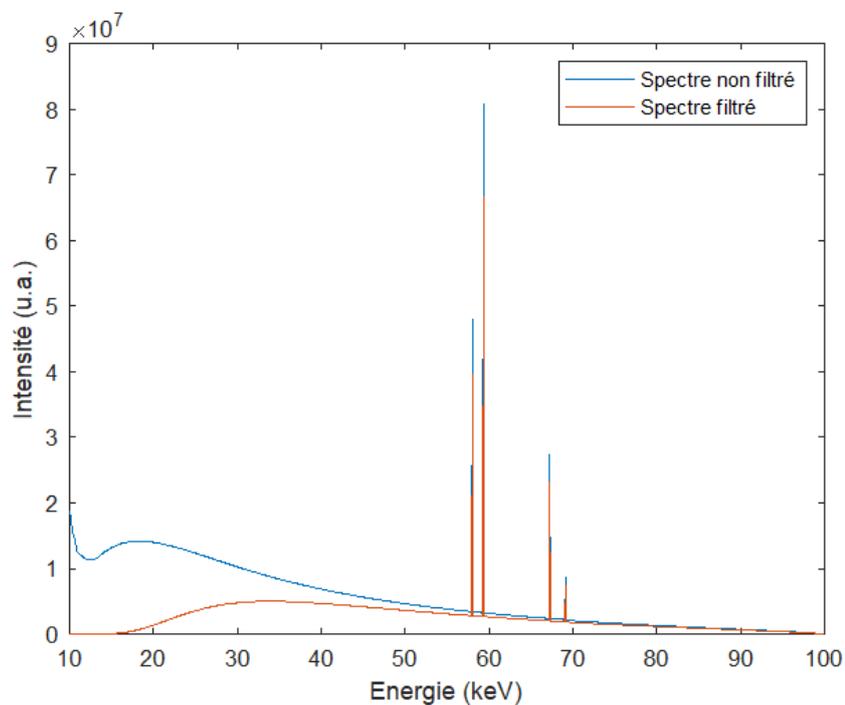


FIGURE 1.3 – Spectre des rayons X à la sortie du tube (en bleu) et après filtration avec 2.5mm d'épaisseur équivalente d'aluminium (en rouge); généré avec SpekCalc (Poludniowski et al., 2009).

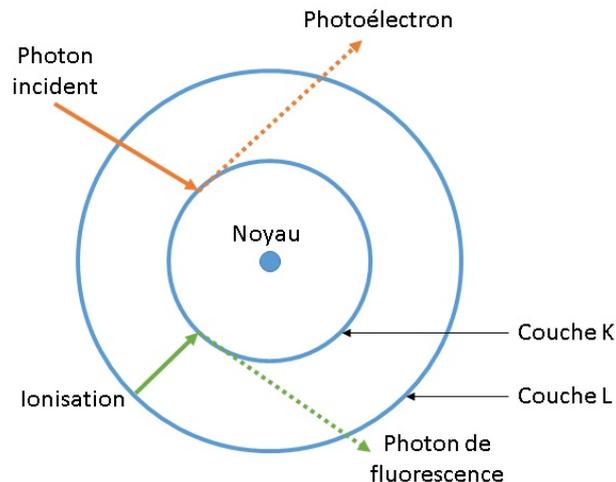


FIGURE 1.4 – Schéma de l'effet photoélectrique.

1. l'effet photoélectrique;
2. la diffusion Compton;
3. la diffusion cohérente.

Pour l'**effet photoélectrique** (photon/électron) (Figure 1.4), un photon incident donne toute son énergie à un électron d'une couche atomique. Si l'énergie est supérieure à l'énergie de liaison de l'électron, cela éjecte l'électron qui traverse ainsi la matière. L'énergie du photon incident est déposée dans la matière autour de l'endroit où a eu lieu l'interaction photoélectrique. L'interaction photoélectrique se fait donc en deux étapes, le transfert d'énergie et le dépôt d'énergie. La couche d'où a été arraché l'électron va se remplir grâce à un photon d'une couche supérieure et ainsi émettre un photon de fluorescence caractéristique de l'atome. L'énergie des photons de fluorescence est trop faible pour qu'ils puissent traverser le patient et être détectés. L'effet photoélectrique est la principale source d'atténuation des rayons X par la matière. Cette capacité d'atténuation dépend également du numéro atomique de l'atome avec lequel le photon interagit. En effet, le calcium des os à $Z=20$ atténue plus les rayons X que l'oxygène de l'eau présente dans les tissus mous avec $Z=8$. L'effet photoélectrique est prédominant pour les rayons X d'énergie inférieure à 30 keV.

Il y a **interaction Compton** (Figure 1.5) ou diffusion incohérente lorsqu'un photon incident rencontre un électron d'une couche avec une faible énergie de liaison. Comme pour l'effet photoélectrique, un électron est éjecté. En éjectant l'électron, une proportion de l'énergie du photon incident est absorbée, ce qui produit un photon de plus faible énergie. Ce nouveau photon quitte le lieu de l'interaction dans une direction différente du photon incident. Cette interaction prend en compte la conservation de l'énergie $h\nu = h\nu' + E$ avec $h\nu$ l'énergie du photon incident, $h\nu'$ l'énergie du photon après diffusion et E l'énergie du photon diffusé. Le fait que les photons soient diffusés dans toutes les directions crée du bruit dans les images car les photons n'arrivent pas sur le détecteur dans la même direction que la source. En effet, lors de la reconstruction tomographique, il est implicite que les photons arrivant sur le détecteur viennent de la direction de la source. Cet effet est prédominant au-dessus de 30 keV et ne dépend pas du numéro atomique de l'élément avec lequel le photon incident interagit.

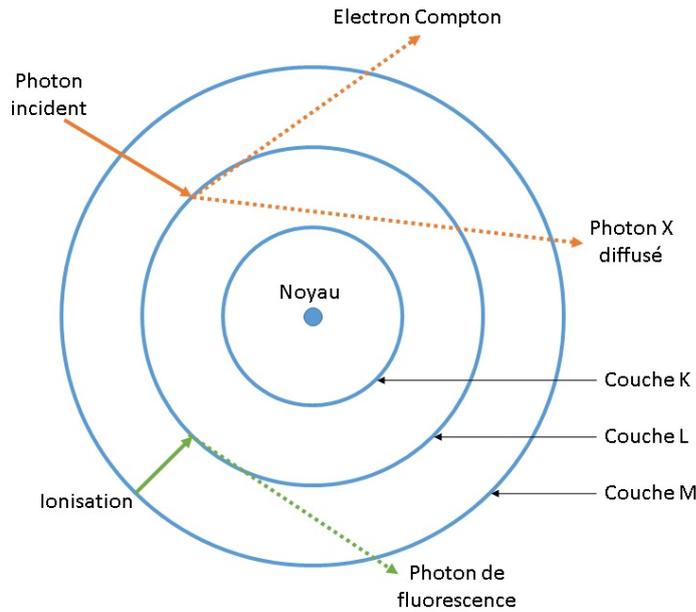


FIGURE 1.5 – Schéma de la diffusion Compton.

La **diffusion cohérente** ou diffusion de Rayleigh est uniquement le changement de direction du photon incident sans dépôt/perte d'énergie. Elle est définie sur tout le spectre des rayons X utilisé en imagerie médicale, soit 0 – 150 keV.

Ces interactions sont responsables de l'atténuation des rayons X qui traversent la matière. On nomme μ_m le vecteur d'atténuation linéaire du matériau traversé par un photon X en fonction de son énergie. Il est exprimé en cm^{-1} . Il dépend de l'atténuation massique du matériau τ_m qui dépend également de l'énergie et qui est exprimé en cm^2/g et de la masse volumique du matériau ρ_m , exprimée en g/cm^3 avec la fonction suivante

$$\mu_m = \tau_m \rho_m. \quad (1.1)$$

La figure 1.6 représente graphiquement les interactions lumière/matière dans la gamme d'énergie utilisée en imagerie médicale, soit de 0 à 150 keV, en montrant les interactions Compton, photoélectrique et de Rayleigh. En effet, dans cette figure nous montrons les coefficients d'atténuation massique qui correspondent à chacun des effets et qui pourraient être utilisés ultérieurement pour la décomposition des effets physiques en tomographie spectrale.

En considérant que les rayons X traversent un matériau homogène, leur atténuation sera modélisée par la loi de Beer-Lambert, tel que :

$$N = N_0 \exp(-\mu L) \quad (1.2)$$

avec N le nombre de photons qui atteignent le détecteur après avoir traversé l'objet, N_0 le nombre de photons émis par la source, μ l'atténuation linéaire du matériau traversé et L l'épaisseur de matériau traversée.

Un objet est généralement composé de différents matériaux donc on généralise la loi de Beer-Lambert à des objets homogènes par morceaux, tel que :

$$N = N_0 \exp\left(-\sum_{m=1}^M \mu_m L_m\right) \quad (1.3)$$

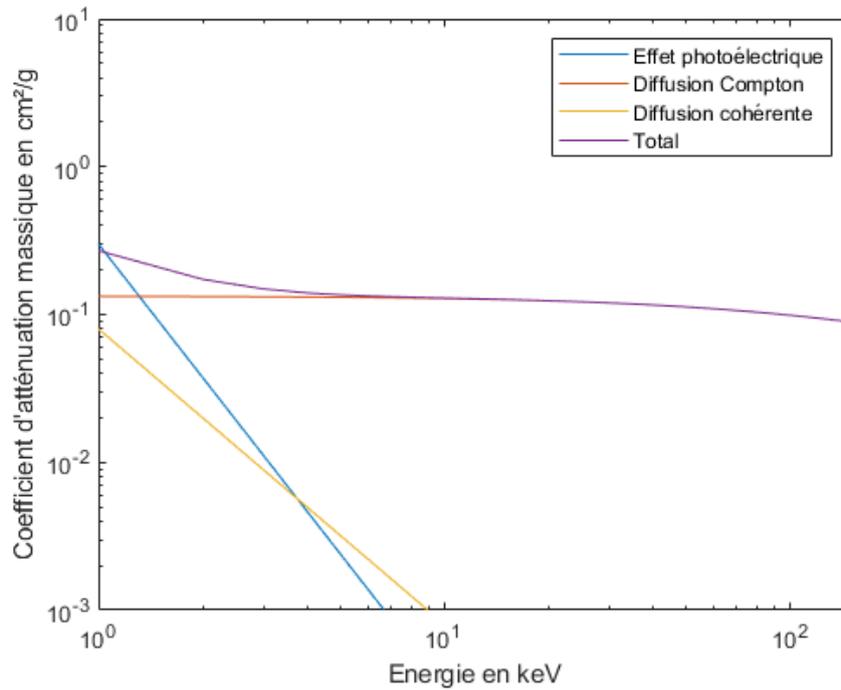


FIGURE 1.6 – Prépondérance des 3 interactions, présentes dans la gamme d'énergie médicale, en fonction de l'énergie.

avec μ_m l'atténuation linéaire du m -ième matériau et L_m l'épaisseur traversée dans ce matériau.

En considérant un μ continu alors, la loi de Beer Lambert devient

$$N = N_0 \exp \left(- \int_L \mu(l) dl \right). \quad (1.4)$$

Pour obtenir des mesures normalisées et linéarisées, on utilise la normalisation logarithmique en prenant en compte le nombre de photons émis par la source, la mesure de plein flux. Ainsi, le spectre atténué normalisé \hat{N} peut s'écrire de la forme suivante

$$\begin{aligned} \hat{N} &= \ln \frac{N_0}{N}, \\ &= \ln \frac{N_0}{N_0 \exp \left(- \int_{L_{\theta,p}} \mu(l) dl \right)}, \\ &= \int_L \mu(l) dl. \end{aligned} \quad (1.5)$$

\hat{N} correspond à la définition mathématique d'une projection.

1.3 Formation de l'image

1.3.1 Détecteurs

Historiquement, l'image radiographique était imprimée sur des films photosensibles, des films photographiques. Ces films étaient, en général, composés d'halogénures d'argent. Les radiographies étaient développées comme les photographies en plongeant le film dans une solution basique puis dans une solution acide. Les

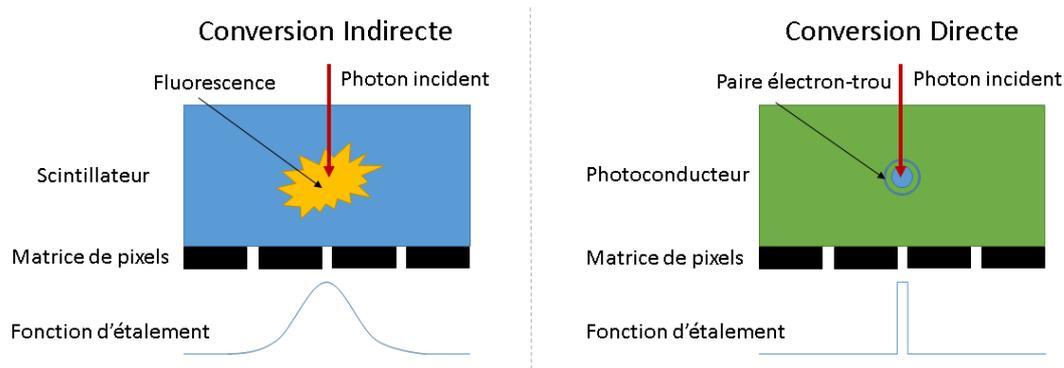


FIGURE 1.7 – Différence entre la conversion indirecte et la conversion directe

images peuvent ainsi être vues à l'aide d'un négatoscope car les zones non irradiées du film restent transparentes. C'est pour ça que les os ressortent blancs. Ensuite, les films ont été couplés avec des écrans renforçateurs, composés de sels fluorescents qui transforment les rayons X en photons lumineux. Ces rayons sont plus adaptés aux films photographiques. Le couple écran / film nécessite une quantité précise de photons pour obtenir un bon contraste. La radiographie passe au numérique avec l'apparition des capteurs plans dans les années 1990. Ces capteurs permettent de ne plus à avoir à développer de films radiographiques et permettent d'exploiter directement les images par traitement numérique. De plus, ils ont une meilleure efficacité de détection et donc permettent de diminuer la dose de rayons X émise pour une même qualité d'image. Les capteurs plans peuvent être de deux types, à conversion directe ou indirecte (Rowlands, 2002), et sont représentés Figure 1.7.

Les **capteurs à conversion indirecte** traduisent les photons X en photons lumineux pour pouvoir les quantifier. Dans ces capteurs, un scintillateur sous la forme d'une couche d'iodure de césium dopé au thallium convertit les photons X en photons lumineux. Cette conversion est faite, majoritairement, grâce à l'effet photoélectrique lors de l'interaction des photons X avec le thallium. Les photons de fluorescence sont ainsi émis dans le domaine du visible. Ces photons visibles sont transformés en signaux électriques par une matrice de photo-transistors. Ces capteurs sont moins précis que ceux à conversion directe car lors de la fluorescence, les rayons lumineux peuvent être projetés dans toutes les directions et donc ne pas être captés par la photo-diode alignée avec le photon X incident.

Les **capteurs à conversion directe** sont capables de quantifier directement les photons X. Les capteurs à conversion directe convertissent directement les photons X en signaux électriques. Ces détecteurs utilisent la photo-conductivité. Des photo-conducteurs, tels que le sélénium amorphe, capturent les photons X et créent des paires électron-trou par un effet photoélectrique. Les électrons sont attirés dans les trous en générant un courant proportionnel à l'intensité de l'irradiation.

Les capteurs plans sont plus efficaces tant sur leur sensibilité à des doses d'irradiation plus faibles que sur la rapidité d'acquisition en comparaison des méthodes avec film (*Digital Radiology 2013*).

1.3.2 Acquisition tomographique

Pour une acquisition tomographique, des images en deux dimensions sont acquises, à l'aide des capteurs présentés dans la partie précédente, pour des angles

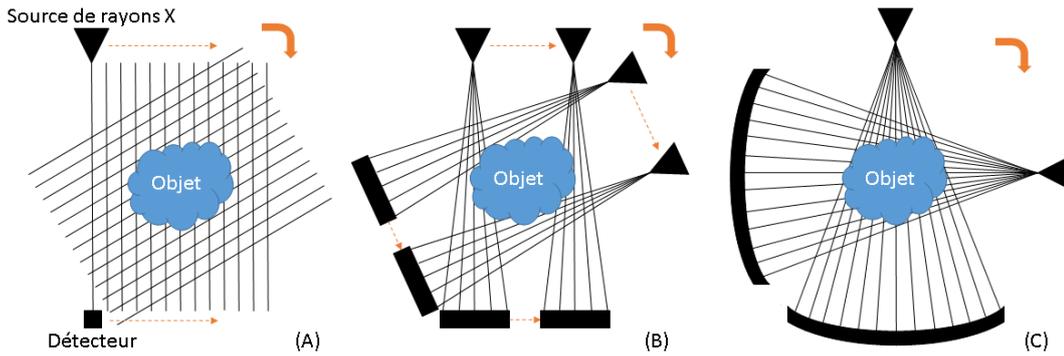


FIGURE 1.8 – Différents types de géométrie. (A) Première génération, translation-rotation en faisceaux parallèles ; (B) Deuxième génération, translation-rotation en faisceau en éventail ; (C) Troisième génération, rotation en faisceau en éventail.

de vue différents. Il a plusieurs types de géométrie présentés Figure 1.8. La première (Figure 1.8 (A)) représente la première génération de scanner qui a une source à faisceau parallèle et un détecteur avec un pixel unique qui se déplace latéralement pour scanner toute une ligne de l'objet et ensuite tournent autour de l'objet et ainsi de suite. Pour la deuxième génération (Figure 1.8 (B)), il a une source qui émet avec faisceau en éventail et un détecteur avec peu de pixels. Il est impossible de voir tout l'objet en une fois. Le détecteur et la source se déplacent latéralement pour réussir à scanner tout l'objet, ensuite ils tournent autour de l'objet, et ainsi de suite. La troisième génération (Figure 1.8 (C)) se compose d'une source en éventail avec un détecteur assez large pour scanner tout l'objet. Le détecteur peut être plan ou incurvé. Les projections représentent les rayons X qui, après avoir été atténués, arrivent sur le détecteur. Les projections sont définies par l'angle d'acquisition et le pixel sur le détecteur. L'ensemble des projections est appelé sinogramme. La transformée à rayons X est l'opérateur qui projette la coupe à tous les angles pour obtenir le sinogramme.

Ces principes se généralisent en 3D avec l'acquisition de projections 2D, comme avec le faisceau conique qui est l'équivalent du faisceau en éventail en deux dimensions. La projection sur le pixel p du détecteur pour l'angle de rotation θ est définie par :

$$s^\theta(p) = \int_E N^\theta(p, e) de \quad (1.6)$$

avec $s^\theta(p)$ la projection sur le pixel p du détecteur pour l'angle de rotation θ , $N^\theta(p, e)$ le spectre atténué arrivant au pixel p en suivant un vecteur orthogonal à celui porté par l'angle θ et à l'énergie e , et E l'intervalle d'énergie de la source. En reprenant l'équation (1.5) on obtient

$$s^\theta(p) = \int_{L_{\theta,p}} \int_E \mu(\{x, y, z\}, e) dedl \quad (1.7)$$

avec l l'abscisse curviligne de la droite L qui passe par p en réalisant un angle θ avec l'axe des ordonnées et passant par le pixel p et $\mu(\{x, y, z\}, e)$ l'atténuation linéaire des matériaux traversés pour l'énergie e .

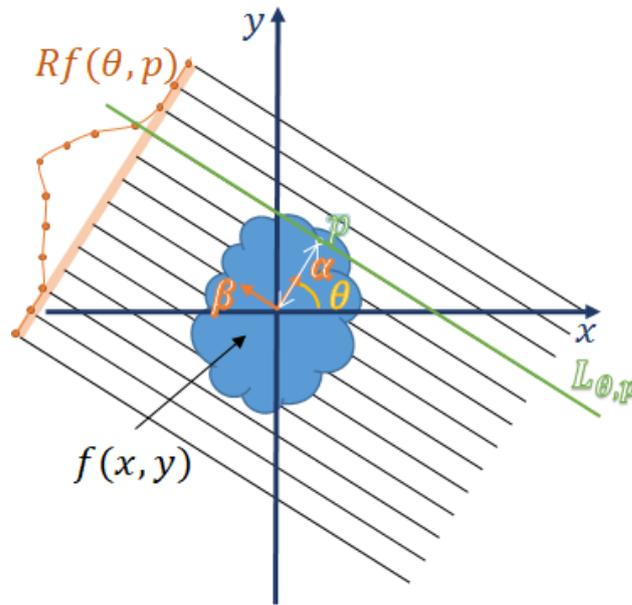


FIGURE 1.9 – Projection à l'aide de la transformée de Radon dans le cas d'une géométrie parallèle.

1.4 Reconstruction tomographique

La reconstruction tomographique permet de retrouver une coupe 2D à partir du sinogramme. Pour réaliser cela, il est nécessaire de connaître la géométrie de l'acquisition.

Pour la géométrie conique, on souhaite retrouver le volume 3D à partir des projections 2D pour les différents angles.

Lors d'une acquisition 1D, avec une géométrie parallèle et un faisceau monochromatique, on se trouve dans l'hypothèse de Radon (Radon, 1986).

1.4.1 Transformée de Radon

On considère une image 2D composée de pixels p définis par leur place dans l'espace (x, y) qui est appelé coupe. Pour obtenir la projection de cette coupe à l'angle θ on utilise la transformée de Radon en définissant le pixel p sur la droite L portée par β (Figure 1.9).

$$Rf(\theta, p) = \int_{L_{\theta, p}} f(x, y) dl \quad (1.8)$$

avec $f(x, y)$ le coefficient d'atténuation linéaire du pixel p intégré sur les énergies comme présenté dans l'équation (1.7), l l'abscisse curviligne le long de la droite L . Dans ce cas, la transformée à rayons X est équivalente à la transformée de Radon.

On peut définir $\alpha = \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix}$ un vecteur unitaire prenant origine au centre de la coupe et appliquant un angle θ avec l'axe des abscisses et $\beta = \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$ un vecteur orthogonal à α , la transformée de Radon sera alors

$$Rf(\theta, p) = \int_{\mathbb{R}} f(p\alpha + t\beta) dt. \quad (1.9)$$

1.4.2 Géométrie parallèle

Pour la géométrie parallèle, chaque projection est obtenue en appliquant la transformée de Radon à chaque angle de rotation comme le présente la Figure 1.9.

Théorème de la coupe centrale Le théorème de la coupe centrale (Bracewell, 1956) énonce que la transformée de Fourier d'une projection est égale à une ligne de la transformée de Fourier de l'image. Cette ligne passe par l'origine et fait un angle θ avec l'axe des abscisses. On écrit la transformée de Fourier d'une projection à θ fixé tel que :

$$\begin{aligned} F[Rf(\theta, \cdot)](\nu) &= \int_{\mathbb{R}} Rf(\theta, p) e^{-2i\pi\nu p} dp \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(p\alpha + t\beta) dt \right) e^{-2i\pi\nu p} dp. \end{aligned} \quad (1.10)$$

$x = p\cos\theta - t\sin\theta$ et $y = p\sin\theta + t\cos\theta$, ainsi $dx dy = dt dp$ et $p = x\cos\theta + y\sin\theta$. D'où,

$$F[Rf(\theta, \cdot)](\nu) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) e^{-2i\pi(x\cos\theta + y\sin\theta)\nu} dx dy \quad (1.11)$$

La transformée de Fourier 2D de f s'écrit

$$F[f](\nu\cos\theta, \nu\sin\theta) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x, y) e^{-2i\pi(x\cos\theta + y\sin\theta)\nu} dx dy \quad (1.12)$$

pour un θ fixé et un ν variable, le point de coordonnées $(x\cos\theta, y\sin\theta)$ sont sur la droite qui passe par le centre de la coupe et qui fait un angle θ avec l'axe des abscisses.

Donc

$$F[Rf](\nu) = F[f](\nu\cos\theta, \nu\sin\theta). \quad (1.13)$$

Le théorème de la coupe centrale nous dit qu'il est possible de reconstruire parfaitement une coupe en ayant une infinité de projections prises avec une infinité d'angles de vue. En effet, avec un nombre de projections fini, le domaine fréquentiel est incomplet surtout dans les hautes fréquences. Des interpolations devront donc être réalisées et pourront entraîner des erreurs. C'est pourquoi on introduit l'opérateur de rétroprojection

Rétroprojection simple Pour la rétroprojection simple, on considère toutes les contributions issues de toutes les projections g . Pour avoir toutes les projections, il nous suffit d'avoir les angles entre 0 et π . Ainsi la rétroprojection simple pour le pixel de coordonnées (x, y) est définie par

$$Bg(x, y) = \int_0^\pi g(\theta, x\cos\theta + y\sin\theta) d\theta. \quad (1.14)$$

Rétroprojection filtrée On utilise le théorème de la coupe centrale pour effectuer la reconstruction tomographique de f en ajoutant un filtre. Ainsi, on calcule la transformée de Fourier inverse, tel que :

$$f(x, y) = \int_{\mathbb{R}} \int_{\mathbb{R}} F[f](X, Y) e^{2i\pi(xX + yY)} dXdY \quad (1.15)$$

On passe en coordonnées polaires.

On pose $X = r\cos\theta$ et $Y = r\sin\theta$ alors $dXdY = |r|drd\theta$

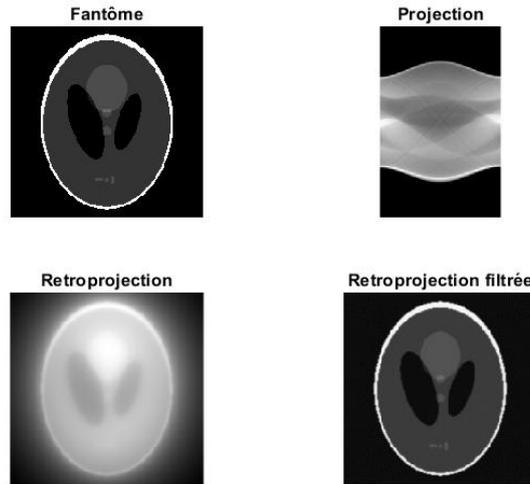


FIGURE 1.10 – Exemple, sur le fantôme de Shepp Logan, avec une image 200×200 , projetée avec 180 angles sur 180° , de la rétroprojection simple face à la rétroprojection filtrée avec le filtre rampe.

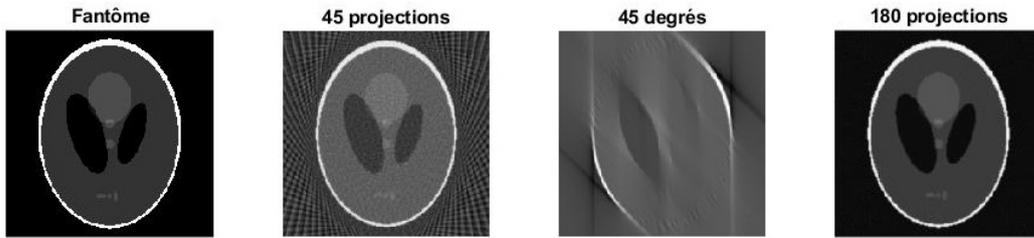


FIGURE 1.11 – Comparaison de la reconstruction tomographique avec un nombre limité de projections, 45, sur 180° et sur 45° , et la reconstruction avec 180 projections sur 180° .

$$\begin{aligned}
 f(x, y) &= \int_0^\pi \int_{\mathbb{R}} F[Rf(\theta, \cdot)](r) e^{2i\pi r(x\cos\theta + y\sin\theta)} |r| dr d\theta \\
 &= \int_0^\pi (g * K)(\theta, x\cos\theta + y\sin\theta) d\theta \\
 &= B(g * K)(x, y)
 \end{aligned} \tag{1.16}$$

où K correspond au filtre rampe d'équation $K(r) = |r|$ dans le domaine fréquentiel et Bg à la rétroprojection simple définie à l'équation (1.14).

Le filtre rampe est un filtre passe haut et permet de palier le problème du domaine fréquentiel incomplet, surtout dans les hautes fréquences, et ainsi améliorer la qualité de la reconstruction comme l'illustre la Figure 1.10. Ce filtre est nécessaire pour la rétroprojection filtrée. On peut également ajouter un filtre de Hann, de Butterworth, ou appliquer un filtre Gaussien dans le domaine image avant le passage en fréquentiel pour réaliser un débruitage.

Pour réduire au maximum le nombre d'interpolations, le nombre de projections ainsi que la part de cercle couverte est importante, comme présenté dans la Figure 1.11.

1.4.3 Géométrie en éventail

La géométrie en éventail est la plus répandue car elle permet d'accélérer l'acquisition par rapport à une géométrie parallèle. Pour l'acquisition, le détecteur peut être soit plat soit incurvé suivant si l'angle entre les faisceau est identique pour deux pixels consécutifs ou si la distance entre chaque pixel et le détecteur est identique. La rétroprojection filtrée peut se généraliser pour la géométrie en éventail. En effet, il est possible de réorganiser les pixels pour changer de base et se retrouver avec une géométrie parallèle (Zeng, 2010) ou prendre en compte la géométrie et les caractéristiques du détecteur (You, Zeng et Liang, 2005) dans les algorithmes de reconstruction.

1.4.4 Géométrie conique 3D

La transformée à rayons X est la projection du volume sur un axe perpendiculaire à l'axe de rotation du dispositif, soit le couple tube/détecteur, soit la plate-forme. Elle est caractérisée par la géométrie du faisceau émanant de la source et la forme du détecteur.

En 3D, l'algorithme de référence est le Feldkamp-Davis-Kress qui est la généralisation de la rétroprojection filtrée mais en 3D (Feldkamp, Davis et Kress, 1984). La généralisation ne permet pas une reconstruction exacte car elle est basée sur des approximations. Plus l'angle du cône est faible meilleure pourra être la reconstruction. La reconstruction FDK est rapide mais a besoin d'un grand nombre de projections pour être robuste. Avec un nombre plus réduit de projections, la reconstruction présente des artefacts. Ainsi, pour palier ces contraintes, les algorithmes itératifs sont beaucoup utilisés (Chlewicki et Badea, 2001 ; Kawata et Nalcioglu, 1985).

Pour la reconstruction tomographique de données acquises en hélicoïdal avec une acquisition divergente, qui se rapproche du conique 3D, on peut également utiliser l'algorithme FDK ou des algorithmes itératifs.

1.5 Applications médicales

Les applications médicales de la tomographie sont multiples et diverses grâce à la rapidité d'acquisition et au plus faible coût en comparaison de l'imagerie par résonance magnétique. Le premier scanner médical a été développé par (Hounsfield, 1973) et il a donné son nom à l'unité de l'image reconstruite en tomographie X. Elle est définie par :

$$HU = 1000 \frac{\mu - \mu_{eau}}{\mu_{eau} - \mu_{air}} \quad (1.17)$$

Ainsi, l'unité Hounsfield prend des valeurs de -1000 HU pour l'air à $+1000$ HU pour l'os.

1.5.1 Applications diverses

Originellement, la radiographie est l'imagerie des os et donc l'examen de première intention lors d'une suspicion de fracture. Plus généralement, la tomographie est l'imagerie des lésions traumatiques car elle a la possibilité d'imager rapidement tout le corps.

Grâce aux produits de contraste qui peuvent être injectés durant l'examen, la tomographie peut imager tout l'arbre vasculaire et ainsi caractériser la perfusion des organes tels que le cerveau, le foie ou encore des tumeurs.

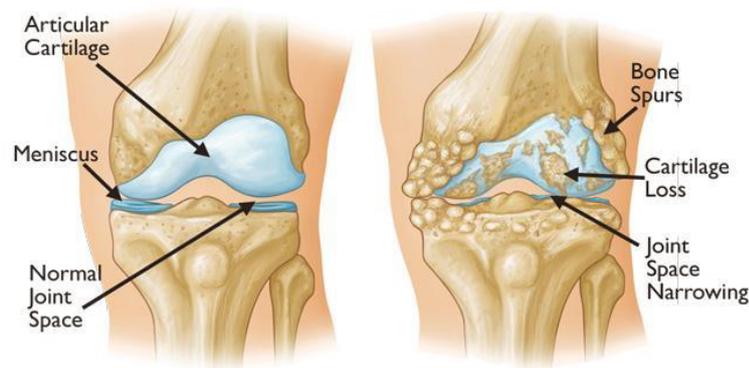


FIGURE 1.12 – Anatomie d'un genou atteint d'arthrose (à droite) et sain (à gauche) (source : (*Issues in Osteoarthritis Care 2018*)).

On peut planifier ou superviser un traitement ou une chirurgie. En effet, on utilise la tomographie pour superviser la pose d'un stent. Elle peut également servir à superviser des interventions telles que les biopsies. Elle est aussi utilisée pour planifier la dose et la localisation pour une radiothérapie (Buzug, 2011).

Les os étant très visibles avec la tomographie, elle est l'imagerie de référence pour les dentistes qui peuvent visualiser les dents et l'os de la mâchoire grâce à une acquisition avec une géométrie conique 3D. L'imagerie permet aussi de prévoir une pose d'implant par exemple.

1.5.2 Cas particulier de l'arthrose

L'arthrose est la plus commune des maladies dégénérative des articulations (Oei et al., 2014). 10 à 15% des plus de 60 ans sont symptomatiques avec une prévalence chez les femmes (*WHO | Chronic rheumatic conditions*). Cette maladie cause également des problèmes sociaux et de santé générale car les personnes souffrent et pour des stades avancés la personne ne peut plus bouger l'articulation correctement. L'arthrose est principalement due au vieillissement et empire avec l'obésité (Braun et Gold, 2012). Elle est caractérisée par une perte de cartilages et des changements pathologiques de l'os comme l'apparition d'excroissances osseuse, nommées ostéophytes, ou de géodes osseuses qui sont des zones moins denses à l'intérieur de l'os, ou de pincement de l'articulation par le rapprochement des os. Ces différentes modifications sont schématisées sur la Figure 1.12. On y voit la perte de cartilage, les ostéophytes, et le pincement de l'articulation du genou. Nous nous focalisons sur l'arthrose du genou car 40% des affections sont localisées sur le genou.

L'arthrose peut être décomposée en quatre stades. Le premier est un stade où les changements morphologiques ne sont pas évidents. Il est possible de voir un pincement douteux de l'articulation et le début de création d'ostéophytes. Le deuxième stade est défini par un possible pincement de l'articulation et des ostéophytes visibles. Le troisième stade est caractérisé par de multiples ostéophytes visibles, un pincement franc de l'articulation, de la sclérose et possiblement des déformations de l'articulation. Le quatrième stade est caractérisé par de larges ostéophytes, un pincement très marqué de l'articulation, de sévères scléroses et des déformations de l'articulation (Kellgren et Lawrence, 1957). Cette échelle a été créée à partir d'images

radiographiques. La radiographie et la tomographie conventionnelle sont les dispositifs d'imagerie diagnostic de l'arthrose grâce à leur rapidité et leur faible coût en comparaison de l'imagerie par résonance magnétique.

Malheureusement, la tomographie ne permet pas de détecter les stades peu avancés d'arthrose. Elle est performante pour imager l'os et donc pour voir les ostéophytes et les géodes osseuses. Le cartilage et les tissus mous, qui ont une atténuation très proche, ne peuvent pas être discriminés. Ainsi, on ne peut pas voir les changements morphologiques du cartilage. L'arthrose est donc détectée trop tard, quand l'espace de l'articulation a déjà commencé à se réduire.

1.6 Conclusion

Dans ce chapitre, nous avons décrit la formation des rayons X, leur émission et leur interaction avec la matière. Ce qui nous a intéressé dans son interaction avec la matière est comment le spectre à rayons X est atténué par les tissus pour obtenir les projections de l'objet, son sinogramme, qui est la base de l'imagerie tomographique. Divers outils de reconstruction tomographique peuvent être utilisés pour reconstruire l'objet en trois dimensions. La tomographie à rayons X est l'imagerie de référence dans de nombreux domaines tels que l'imagerie ostéo-articulaire.

La tomographie à rayons X présente certaines limitations, notamment, le phénomène de durcissement de faisceau. Celui-ci se traduit par l'absorption privilégiée des photons de basse énergie issus d'un faisceau polychromatique lors du passage à travers un objet. Il produit des artefacts comme des bandes très faiblement atténuées entre deux objets très atténuants ou une perte d'atténuation au centre d'un matériau homogène. Les limitations de la tomographie classique tendent à être surpassées par la tomographie spectrale qui fera l'objet du chapitre suivant. En effet, la tomographie spectrale réduit l'intervalle d'énergie intégré par le détecteur et donc les effets du durcissement de faisceau, par exemple.

Chapitre 2

Tomographie X spectrale

Dans ce chapitre, nous présenterons la tomographie X spectrale qui est une variante de la tomographie à rayons X. Sous le nom tomographie spectrale on trouve la tomographie à double énergie et la tomographie spectrale à comptage photonique. Nous présenterons la formation de l'image dans les deux cas. Ensuite, nous introduirons les différentes stratégies possibles en tomographie spectrale concernant la reconstruction tomographique et la décomposition de matériaux. Ceux-ci sont les deux sous problèmes de la tomographie spectrale. Enfin, nous présenterons les applications médicales spécifiques de la tomographie spectrale.

2.1 Introduction

La tomographie spectrale, aussi appelée tomographie couleur, est une nouvelle modalité d'imagerie à rayons X. Elle peut être séparée en deux types, la tomographie à double énergie et la tomographie à comptage photonique. La tomographie à double énergie acquiert deux jeux de données, un à basse énergie et un à haute énergie. La tomographie spectrale à comptage photonique acquiert deux jeux de données ou plus sur des intervalles d'énergie prédéfinis par des seuils. Cela rajoute une dimension énergétique aux images. Cette dimension supplémentaire dans les images permet de décomposer les matériaux composant un objet en connaissant leur signature énergétique (Schlomka et al., 2008 ; Taguchi et Iwanczyk, 2013). Il est possible de décomposer au plus autant de matériaux qu'il y a de jeux de données à différentes énergies. De plus, à partir de ces cartes de matériaux, il est possible de générer des images mono-énergétiques virtuelles. Chaque image représente la mesure de l'objet avec une source monochromatique à une énergie choisie. Il est aussi possible de générer des images Hounsfield en intégrant les bins d'énergie et considérant la contribution de l'eau.

2.2 Formation de l'image

Le modèle de formation de l'image en tomographie spectrale regroupe deux sous modèles : la projection à l'aide de la transformée à rayons X dépendant de la géométrie, comme en tomographie conventionnelle, et le mixage spectral qui permet d'obtenir la dimension énergétique. Le deuxième sous modèle diffère selon le type de tomographie spectrale, à double énergie ou à comptage photonique.

2.2.1 Tomographie à double énergie

Les systèmes de tomographie à double énergie peuvent suivre quatre principes présentés dans la Figure 2.1 :

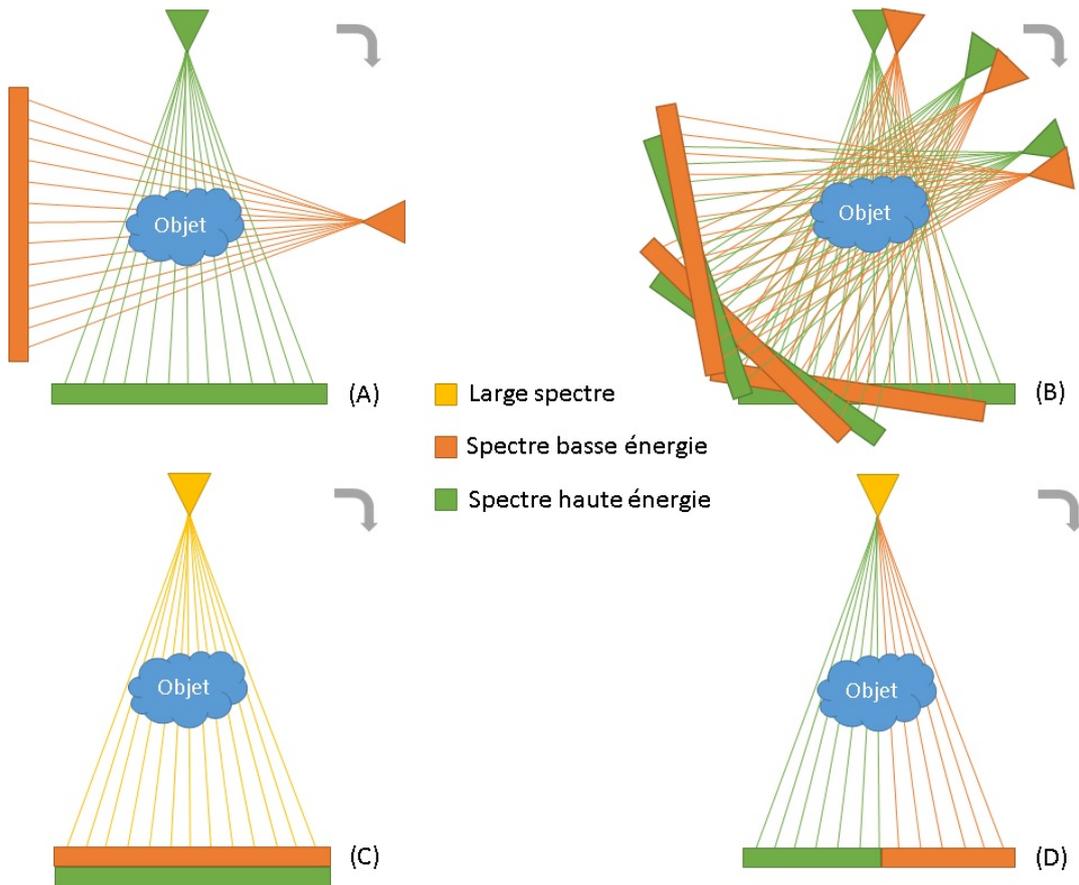


FIGURE 2.1 – Différents types d'acquisitions en tomographie spectrale, avec (A) deux sources à différentes énergies, (B) un changement d'énergie, (C) un détecteur en couches et (D) le filtrage du spectre.

1. Deux sources à différentes énergies et un détecteur intégrateur (Figure 2.1,(A));
2. Changement d'énergie d'une même source avec détecteur intégrateur (Figure 2.1,(B));
3. Une source avec un détecteur en couches (Figure 2.1,(C));
4. Une source avec un filtrage du spectre et un détecteur intégrateur (Figure 2.1,(D)).

Lorsqu'il y a deux sources (Figure 2.1,(A)), elles émettent simultanément et tournent autour de l'objet pour obtenir un sinogramme par source. Cela a été utilisé la première fois par (Genant et Boyd, 1977) pour une étude sur la minéralisation des os. Cette méthode permet d'augmenter la précision de la quantification des matériaux mais a deux désavantages majeurs, le doublement de la dose reçue par le patient et des artefacts de mouvement. En effet, les sources étant généralement positionnées à 90° , le patient a le temps de bouger entre le passage de la première et de la deuxième source. Malheureusement, ce grand angle entre les deux sources est nécessaire pour éviter que des photons diffusés de la première source ne se retrouvent sur le détecteur de la deuxième et inversement.

Pour diminuer les artefacts de mouvement, on peut utiliser le changement d'énergie de la source (Figure 2.1,(B)). L'énergie de la source est changée très rapidement au cours de la rotation (Kalender et al., 1986). Ainsi, une projection sur deux est haute énergie et inversement. L'acquisition est malheureusement plus lente que les doubles sources.

Pour accélérer l'acquisition en évitant les artefacts et le doublement de la dose, il est possible d'avoir recours aux détecteurs en couches ou détecteurs sandwich (Barnes et al., 1985) (Figure 2.1,(C)). Ce sont trois couches superposées, un matériau avec un numéro atomique modérément faible, avec un K-edge < 30 keV, couplé à une photodiode, une couche de cuivre, et un matériau avec grand numéro atomique, avec un K-edge > 50 keV, couplé à une photodiode. Le matériau à faible numéro atomique aura tendance à absorber les photons de basse énergie et à transmettre les photons de haute énergie. Le filtre en cuivre entre les deux couches de matériau augmente l'énergie effective des photons et donc permet une meilleure séparation entre basse et haute énergie.

Pour conserver la même source et le même détecteur qu'en tomographie conventionnelle et avoir des images à deux énergies, assez simplement, un scanner avec une séparation du spectre d'une même source grâce à des filtres a été développé (Euler et al., 2016) (Figure 2.1,(D)). Ces filtres permettent de créer artificiellement un spectre haute énergie, par exemple avec de l'or, et un spectre basse énergie, par exemple avec de l'étain. Le filtre est ainsi composé de deux parties de tailles égales, l'une avec de l'or et l'autre avec de l'étain. Comme cela la moitié de l'objet est scanné avec la haute énergie et l'autre moitié avec la basse énergie. Il est donc nécessaire de faire un deuxième scan avec le filtre inversé pour imager l'objet en entier.

2.2.2 Tomographie spectrale à comptage photonique

La tomographie spectrale à comptage photonique utilise une source similaire à celle utilisée en tomographie conventionnelle. Le détecteur diffère, c'est un détecteur à comptage de photons. Il est composé de photodiodes reliées à un circuit intégré propre à une application (ASIC) qui est capable de trier les photons incident en fonction de leur énergie et de les classer selon des seuils prédéfinis (Figure 2.2). Nous obtenons ainsi autant de projections par angle qu'il y a d'intervalles, appelés bins d'énergie (Taguchi et Iwanczyk, 2013). Ils sont inspirés des détecteurs à conversion directe introduits précédemment dans le 1.3.1. Malheureusement, ces détecteurs ne sont pas parfaits et l'ionisation du semiconducteur n'est pas toujours proportionnelle à l'énergie du photon incident. Il existe plusieurs phénomènes illustrés Figure 2.3 comme l'empilement des charges (Figure 2.3,(A)) où deux photons arrivant simultanément sur le même pixel du détecteur sont considérés comme un seul (Taguchi et al., 2010). Les autres phénomènes qui peuvent se produire sont le partage de charge (Figure 2.3,(B)) avec un photon qui est détecté sur deux pixels, la mauvaise détection (Figure 2.3,(C)) quand un photon qui arrive sur le bord d'un pixel est en partie détecté sur le pixel adjacent, et l'effet photoélectrique (Figure 2.3,(D)) qui est l'effet prépondérant dans les interactions photons / détecteur et qui donne une réponse du détecteur proportionnelle à l'ionisation dans la photodiode.

Pour la tomographie à double énergie, seulement deux projections par angle sont créées. Pour la tomographie spectrale à comptage photonique il a autant de projections par angle qu'il y a de bins d'énergie. (Schlomka et al., 2008) en utilise 4, le prototype de Philips (Si-Mohamed et al., 2017) en utilise 5 et le nombre de bins peut aller jusqu'à 32 (Xu et al., 2013).

2.2.3 Modèle direct de la tomographie spectrale

On considère un objet composé de M matériaux avec V voxels, imagé avec un capteur composé de P pixels, qui réalise des acquisitions sur I bins d'énergie et avec Θ angles de projection. On nomme s le vecteur de mesures qui correspond

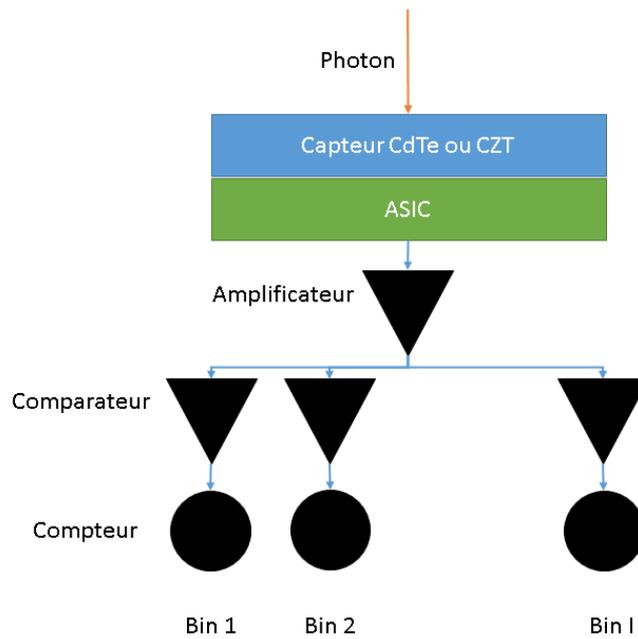


FIGURE 2.2 – Schéma du détecteur à comptage photonique avec un capteur lumineux soit de type CdTe, soit CZT, un circuit électronique spécialisé pour une application (ASIC), un amplificateur, des comparateurs et des compteurs, pour un détecteur comptant sur I bins.

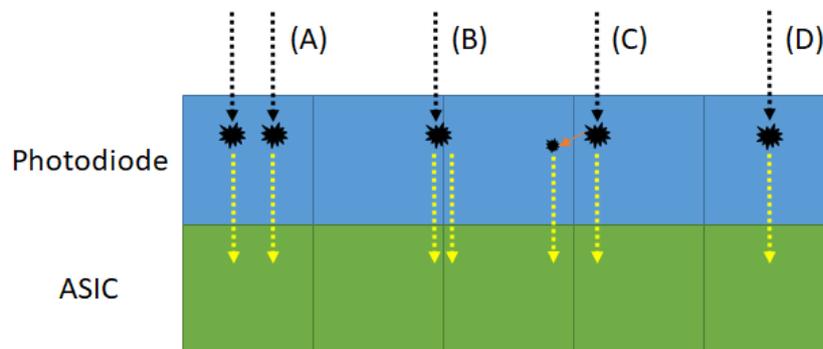


FIGURE 2.3 – Différentes interactions agissant dans un détecteur à comptage de photons avec (A) l'empilement, (B) le partage de charge, (C) la mauvaise détection et (D) l'effet photoélectrique.

au nombre de photons en prenant en compte les bins d'énergie et les angles de projection.

$$\mathbf{s} = (s_{1,1}^1, \dots, s_{i,p}^\theta, \dots, s_{I,P}^\Theta)^\top$$

avec $s_{i,p}^\theta$ les données mesurées sur le p -ième pixel pour la θ -ième projection et dans le i -ième bin d'énergie. \mathbf{s} est de taille $P \times \Theta$.

Le volume est défini par un vecteur de masse volumique inconnu $\boldsymbol{\rho}$ pour chaque matériau et voxel du volume.

$$\boldsymbol{\rho} = (\rho_{1,1}, \dots, \rho_{m,v}, \dots, \rho_{M,V})^\top$$

avec $\rho_{m,v}$ la masse volumique du m -ième matériau sur le v -ième voxel. $\boldsymbol{\rho}$ est de taille $V \times M$. Le modèle de matériaux est défini par l'atténuation linéaire μ telle que

$$\mu(v, E) = \sum_{m=1}^M \rho_m(v) \tau_m(E) \quad (2.1)$$

avec $\mu(v, E)$ le coefficient d'atténuation linéaire pour le voxel v à l'énergie E , $\rho_m(v)$ la masse volumique du matériau m au voxel v et $\tau_m(E)$ l'atténuation massique du matériau m à l'énergie E .

Le modèle direct G représente la création des projections spectrales à partir des masses volumiques tel que

$$\mathbf{s} = G(\boldsymbol{\rho}). \quad (2.2)$$

Ce modèle est la composition de X la transformée à rayons X et de F le mixage spectral tel que

$$G = F \circ X. \quad (2.3)$$

Le but de la tomographie spectrale est d'inverser (2.2).

Transformée à rayons X En tomographie spectrale, chaque matériau est indépendant car on considère des voxels purs dans le volume, chaque voxel ne possède qu'un seul matériau. Ainsi, on peut projeter chaque matériau indépendamment comme

$$\mathbf{a}_m^\theta = X(\boldsymbol{\rho}_m). \quad (2.4)$$

avec \mathbf{a} de la forme

$$\mathbf{a} = (a_{1,1}^1, \dots, a_{m,p}^\theta, \dots, a_{M,P}^\Theta)^\top$$

et $a_{m,p}^\theta$ est la masse volumique projetée pour le m -ième matériau au θ -ième angle et sur le p -ième pixel. \mathbf{a} est de taille $P \times M \times \Theta$

Mixage spectral Les détecteurs à comptage de photons permettent de récupérer le nombre de photons qui touchent le détecteur, dans un certain intervalle d'énergie. La quantité mesurée est le nombre de photons. La fonction de réponse du détecteur est la densité de probabilité d'un photon à une énergie donnée d'être caractérisé dans un certain bin d'énergie comme cela peut être représenté pour un pixel du détecteur dans la Figure 4.7 - droite, avec 4 bins d'énergies. Comme pour la tomographie conventionnelle, on utilise une source polychromatique dont le spectre est présenté Figure 4.7 - gauche.

Le mixage spectral permet de faire le passage entre les masses volumiques projetées pour chaque vue \mathbf{a}^θ aux nombres de photons pour chaque vue \mathbf{s}^θ . On calcule les projections spectrales dans différentes gammes d'énergie i comme

$$\mathbf{s}_i^\theta = F(\mathbf{a}^\theta). \quad (2.5)$$

On prend la formulation du mixage spectral générique décrit dans (Ducros et al., 2017).

$$s_{i,p}^\theta = N_0 \int_{\mathcal{E}} d_i(E) \exp \left[- \sum_{m=1}^M a_{m,p}^\theta \tau_m(E) \right] dE \quad (2.6)$$

avec \mathcal{E} l'intervalle d'énergie de la source, n_0 le spectre de la source, d_i la réponse du détecteur pour le i -ème bin d'énergie et τ_m le coefficient d'atténuation massique pour le m -ième matériau. Le nombre de photons de la source est défini par

$$N_0 = \int_{\mathcal{E}} n_0(E) dE \quad (2.7a)$$

quand la source irradie de manière homogène le détecteur. Quand la source irradie de manière inhomogène le détecteur on calcule

$$N_0(p) = \int_{\mathcal{E}} n_0(E, p) dE. \quad (2.7b)$$

2.3 Reconstruction

Comme le modèle de formation de l'image, le problème de reconstruction peut être divisé en deux sous problèmes : la reconstruction tomographique, comme pour la tomographie conventionnelle, et la décomposition de matériaux qui permet de récupérer les matériaux à partir des informations énergétiques. Leur résolution peut se faire l'un après l'autre ou les deux en même temps. Les différentes méthodes sont schématisées Figure 2.4. Pour la reconstruction des images mono-énergétiques, elle peut être réalisée à partir de la décomposition de matériaux ou directement, cela est schématisé Figure 2.5.

2.3.1 Reconstruction tomographique pour la tomographie spectrale

En tomographie conventionnelle, une seule projection est acquise pour chaque angle, en intégrant sur toutes les énergies de la source, alors qu'en tomographie spectrale plusieurs projections sont acquises pour le même angle mais sur différents intervalles d'énergie. La reconstruction tomographique classique (1.4) travaille avec une seule projection par angle. Ainsi pour reconstruire les projections dans les bins d'énergie, cela est réalisé pour chaque bin indépendamment. Pour reconstruire les cartes de matériaux projetées, cela est réalisé pour chaque matériau indépendamment.

2.3.2 Décomposition

Le décomposition a pour but de récupérer les matériaux ou les manifestations d'effets physiques à partir des données énergétiques. En effet, l'atténuation des matériaux a une dépendance énergétique. Le problème est mal posé s'il n'y a pas une infinité de bins d'énergie ou si un des matériaux est proche d'être une combinaison linéaire des autres.

La base utilisée pour décomposer peut être une base de matériaux anatomiques comme l'os, les tissus mous ou les tissus adipeux. Ils peuvent être couplés avec des agents de contraste tels que l'iode, le gadolinium ou encore l'or qui est peu utilisé en tomographie rayons X. Il est aussi possible de décomposer en effets physiques. Dans

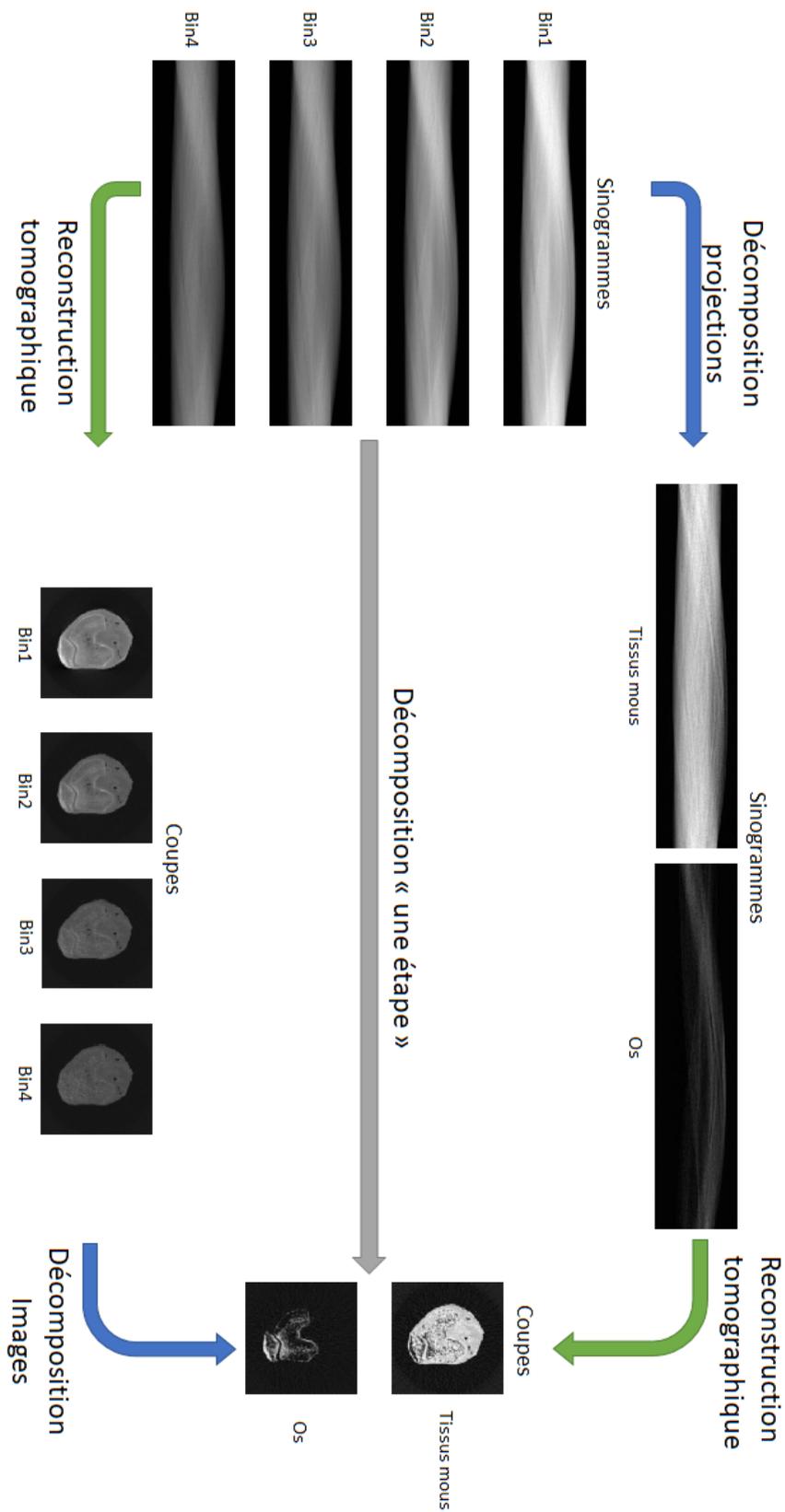


FIGURE 2.4 – Différentes stratégies pour obtenir les cartes de matériaux reconstruites à partir des projections acquises. Ici, les projections sont acquises sur 4 bins et décomposées en 2 matériaux.

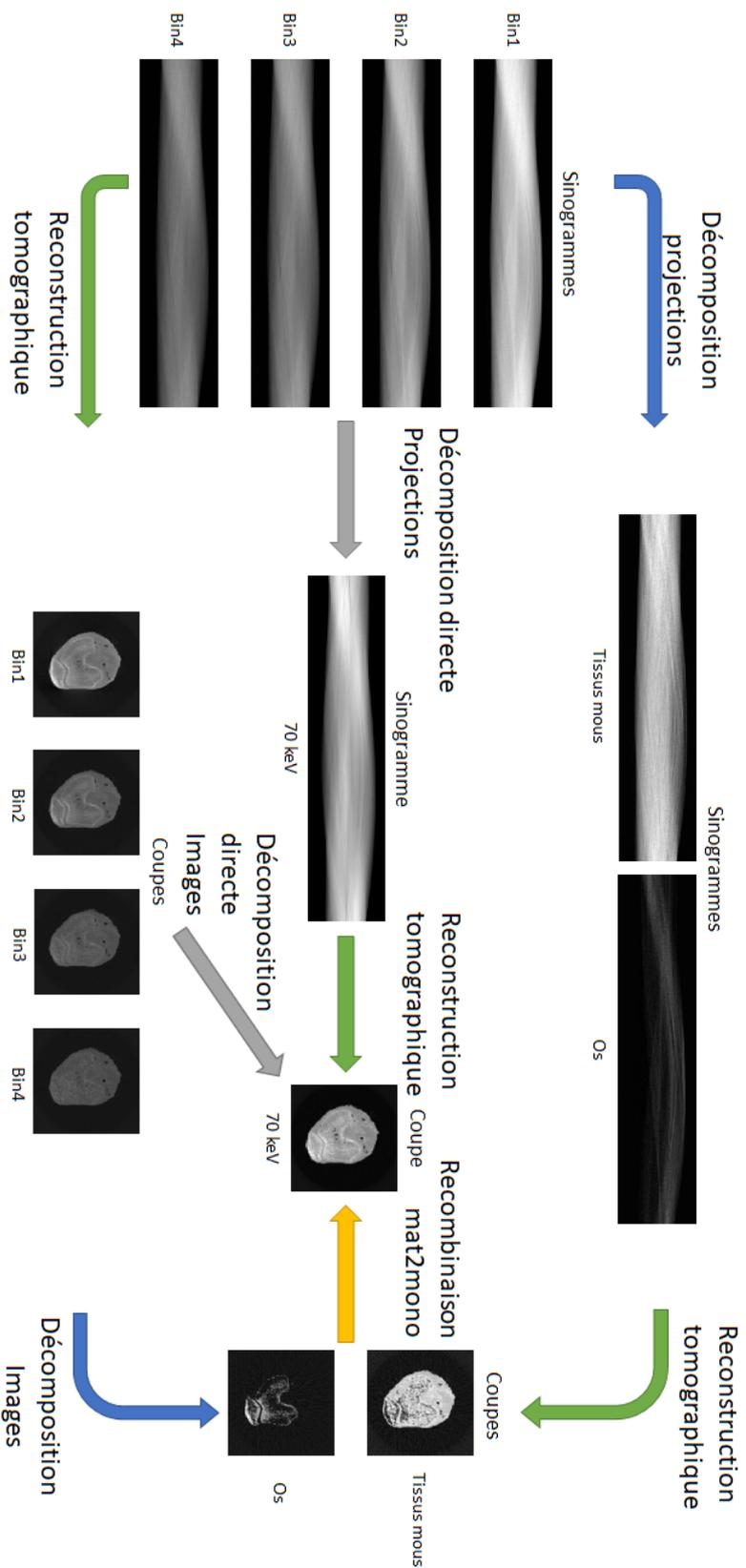


FIGURE 2.5 – Différentes stratégies pour obtenir les images mono-énergétiques à partir des projections acquises. Ici, les projections sont acquises sur 4 bins et on récupère des images mono-énergétiques à 70 keV.

ce cas, on décompose les données comme si elles avaient été acquises avec uniquement des photons réalisant une interaction Compton ou uniquement des photons réalisant une interaction photoélectrique. Pour cela, on utilise les coefficients d'atténuation massique de chaque matériau qui peuvent être dans la base de données du National Institute of Standards and Technology (Hubbell et Seltzer, 1995) qui référence les éléments chimiques et les tissus les plus communs. Des éléments ou tissus plus complexes peuvent être calculés à l'aide d'un petit logiciel, qui permet de calculer l'atténuation massique à partir de la combinaison linéaire des parts massiques de chaque élément composant le matériau (D'Errico, 2009).

La décomposition de matériaux ou d'effets physiques est opérée grâce à des algorithmes de décomposition. Dans leur papier, (Alvarez et Macovski, 1976) décrivent la décomposition dans une base et plus précisément les équations utilisées pour décomposer en effet Compton et effet photoélectrique. En effet, pour décrire une base, on utilise $\tau_m(E)$ l'atténuation massique d'un matériau m à l'énergie E comme décrite dans l'équation (2.1).

Décomposition Compton/Photoélectrique Il est possible de décomposer les images comme si elles avaient été uniquement produites avec l'effet Compton ou l'effet photoélectrique. Pour cela, on peut utiliser la modélisation de la dépendance en énergie de ces deux effets présentée dans (Alvarez et Macovski, 1976). Ainsi, il propose

$$\tau_{\text{photo}}(E) = \frac{1}{E^3} \quad (2.8a)$$

$$\tau_{\text{Compton}}(E) = \frac{1 + \alpha}{\alpha^2} \left[\frac{2(1 + \alpha)}{1 + 2\alpha} - \frac{1}{\alpha} \ln(1 + 2\alpha) \right] + \frac{1}{2\alpha} \ln(1 + 2\alpha) - \frac{1 + 3\alpha}{(1 + 2\alpha)^2} \quad (2.8b)$$

avec $\alpha = E/510.975$ keV. Il reprend la fonction de Klein-Nishina (Klein et Nishina, 1929) qui donne la dépendance en énergie d'une interaction Compton. Ces fonctions se retrouvent dans la Figure 1.6 à un facteur près.

(Roessl et Proksa, 2007) utilise les équations (2.8a) et (2.8b), et la log-vraisemblance pour décomposer en Compton, photoélectrique et agent de contraste, ici du gadolinium. (Cormode et al., 2010) décompose avec de l'or, en plus du gadolinium, qui a été utilisé comme agent de contraste. Les méthodes d'étalonnage sont aussi largement utilisées pour décomposer (Boussel et al., 2014; Ehn et al., 2016), elles utilisent des mesures d'étalonnage pour trouver la combinaison linéaire des matériaux et un terme quadratique prenant en compte les effets non-linéaires.

L'avantage de cette base est qu'elle contient un élément, l'effet photoélectrique, dépendant du numéro électronique donc du matériau et un élément, la diffusion Compton, qui dépend de la densité électronique et donc de la masse volumique.

Décomposition de matériaux Il est possible de décomposer les matériaux composant l'objet et non uniquement des effets physiques tels que les effets Compton et photo électrique. Les matériaux utilisés usuellement sont les tissus mous et l'os cortical, joints ou non à un ou plusieurs agents de contraste comme iode ou le gadolinium.

La Figure 2.6 représente les coefficient d'atténuation massique des matériaux les plus souvent utilisés, les tissus mous, l'os, l'iode et le gadolinium.

La décomposition d'un ou de plusieurs produits de contraste est souvent plus simple car ces éléments chimiques comportent une discontinuité dans leur coefficient d'atténuation massique dans la gamme d'énergie du détecteur, visible Figure 2.6. Cela permet de les décomposer très facilement en choisissant judicieusement

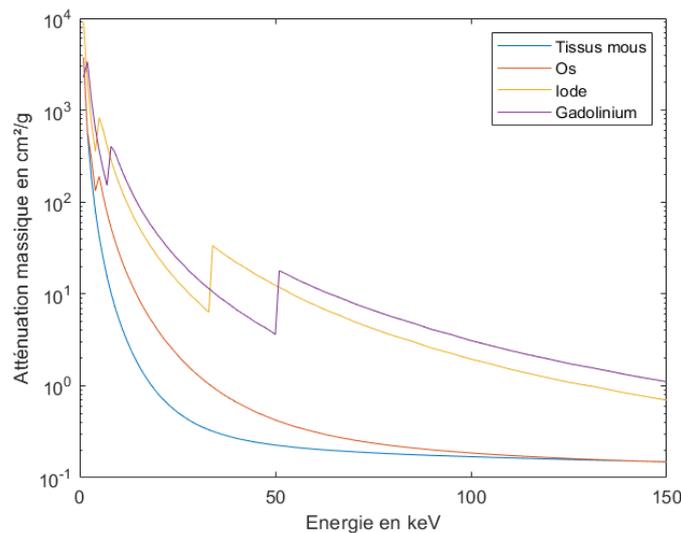


FIGURE 2.6 – Coefficients d'atténuation massique des tissus mous, de l'os, de l'iode et du gadolinium entre 0 et 150 keV.

des bins d'énergie avant et après la discontinuité. De plus, la concentration d'agent de contraste peut être quantifiée et ainsi permettre de classifier les lésions.

Décomposition mono-énergétiques Les images mono-énergétiques sont des images acquises à l'aide d'un faisceau monochromatique défini par son énergie E . La référence pour ces images est l'imagerie synchrotron. Il est possible de créer artificiellement ces images mono-énergétiques à partir de la décomposition de matériaux. En effet, en reprenant l'équation (2.1) qui définit le coefficient d'atténuation linéaire, on voit très facilement qu'à partir de l'atténuation massique d'un matériau et de sa masse volumique, on obtient le coefficient d'atténuation linéaire. Ce coefficient est la valeur des voxels d'un volume mono-énergétique. Il est également possible de décomposer directement les projections ou images mono-énergétiques à partir des projections spectrales ou des projections spectrales reconstruites. Les différentes possibilités pour la reconstruction des images mono-énergétiques sont illustrées sur la Figure 2.5.

L'avantage des images mono-énergétiques est que ce sont des images semblables à des images en unité Hounsfield ce qui les rend facilement interprétable par des médecins et elles permettent d'obtenir un meilleur contraste entre les tissus et une diminution du bruit (Hickethier et al., 2017). De plus, il semble que les énergies autour de 70 keV soient le meilleur compromis entre une bonne résolution spatiale et un faible niveau de bruit (Ozguner et al., 2018).

2.3.3 Stratégies pour la décomposition de matériaux

Comme représenté sur la Figure 2.4, il est possible de décomposer dans le domaine des projections, dans le domaine image ou de réaliser une décomposition "une étape".

Domaine des projections Il est possible de décomposer les matériaux dans le domaine des projections et ensuite reconstruire chaque matériau indépendamment. On

cherche donc à reconstruire la masse volumique projetée à partir des projections acquises. Le résultat est ensuite reconstruit en 2D ou 3D. Travailler dans le domaine des projections permet de garder un *a priori* sur la physique liée à la formation d'image. De plus, travailler sur chaque projection indépendamment permet de paralléliser le problème et gagner du temps au vu du grand nombre de projections acquises en général. Avant tout, le fait de travailler directement sur les données acquises diminue le risque d'erreurs et d'artefacts créés lors d'un traitement antérieur.

Différentes techniques ont été développées pour faire suite à (Alvarez et Macovski, 1976). (Long et Fessler, 2014) décomposent trois matériaux, tissus adipeux, sang et os, à partir d'images de tomographie à double énergie. Il utilise une méthode de vraisemblance pénalisée et contrainte. Ils arrivent ainsi à résoudre ce problème sous-dimensionné. D'autres, comme (Huy et Molloy, 2011), utilisent les caractéristiques des agents de contraste. Ils tentent de décomposer en 4 matériaux : iode, hydroxyapatite, tissus glandulaires et adipeux. Ils utilisent diverses versions de la méthode des moindres carrés qui par exemple nécessite un étalonnage en connaissant la concentration d'un matériau donné. Cela nécessite un fort *a priori* sur la composition de l'objet scanné. (Ducros et al., 2017) quant à eux, utilisent une méthode de Gauss-Newton pondérée et régularisée, pour modéliser des statistiques poissonniennes, et la comparent à la méthode du maximum de vraisemblance. La régularisation est spécifique à chaque matériau mais qui ne demande pas de connaître la concentration d'un matériau spécifique. Ils décomposent trois matériaux, tissus mous, os et gadolinium, à partir de quatre bins d'énergie. Ils recherchent également le meilleur paramètre de régularisation pour réduire au mieux le bruit sans flouter l'image.

Domaine image Il est aussi possible de réaliser la reconstruction tomographique de chaque bin d'énergie et ensuite décomposer en matériaux dans le domaine image. On cherche donc à reconstruire la masse volumique à chaque voxel du volume à partir de la reconstruction tomographique des projections acquises. C'est une décomposition assez simple mais les artefacts créés lors de la reconstruction peuvent altérer la décomposition. De plus si on décompose tout le volume en même temps, cela peut être très chronophage.

La première décomposition a été réalisée sur la tomographie à double énergie (Brooks, 1977) mais a ensuite été étendue à plus de plages d'énergies pour décomposer plus de matériaux. (Clark et Badea, 2014) décomposent trois matériaux, l'iode, l'or et le gadolinium en utilisant une méthode itérative qui mêle une régression avec un noyau constant par morceau et une méthode Bregman. Elle nécessite un étalonnage qui est réalisé avec des concentrations connues d'iode, d'or et de gadolinium. (Persson et al., 2014) utilisent un très grand nombre de plages d'énergie pour sur-dimensionner grandement son problème. Ils décomposent un fantôme, composé d'inserts d'eau, de graisse, de solution aqueuse de calcium, d'iode et de gadolinium avec des concentrations différentes, à l'aide d'une méthode des moindres carrés sous contrainte d'une conservation de la masse dans le volume. (Rajendran et al., 2016) décomposent en calcium, eau et iode pour l'application à l'arthrose avec un logiciel propriétaire en accompagnant la décomposition avec une procédure d'étalonnage.

Décomposition "une étape" Il est également possible de résoudre les problèmes de reconstruction tomographique et décomposition de matériaux en même temps. On appelle cela la décomposition "une étape" ou one-step. On cherche donc à reconstruire la masse volumique de chaque voxel d'un volume à partir des projections acquises. Ce type de méthode peut conduire à plus d'erreur considérant que l'on

résout deux problèmes en même temps. Le problème est algorithmiquement plus complexe et plus difficile à résoudre d'autant que la parallélisation des calculs est impossible. Ainsi (Mory et al., 2018) comparent la vitesse de convergence de cinq algorithmes de décomposition directe pour trois matériaux.

2.4 Applications médicales

L'amélioration de la qualité d'image et le bénéfice liés à la décomposition de matériaux et aux images mono-énergétiques, a permis d'améliorer des applications existantes pour la tomographie classique et d'en développer d'autre.

2.4.1 Application aux lithiases rénales

Les lithiases rénales ou calculs rénaux sont des cristaux durs qui se forment dans le rein. 80% d'entre eux sont éliminés sans douleur mais dans le reste des cas, les patients se retrouvent avec de fortes douleurs que l'on appelle colique néphrétique. Les coliques néphrétiques touchent 5 à 10% des individus au cours de leur vie.

Les lithiases peuvent être composées de calcium, de phosphate ammoniacal magnésien, d'acide urique ou de cystine. En utilisant des images mono-énergétiques virtuelles à deux énergies et connaissant la signature énergétique des matériaux composant les calculs, (Wisnbaugh et al., 2014) ont pu améliorer la caractérisation des lithiases rénales.

2.4.2 Application à la goutte

L'application à la goutte est proche de celle des lithiases. En effet, la goutte est caractérisée par des cristaux d'urate monosodique qui se fixent dans les articulations, souvent dans un contexte d'arthrite. Cela touche 4% de la population adulte américaine et a été décrite pour la première fois en Égypte antique. La tomographie classique ne peut visualiser que l'érosion sur l'os et des masses dans les tissus péri-articulaires. L'imagerie par résonance magnétique peut visualiser l'érosion de l'os avec des oedèmes de la moelle et les dépôts de cristaux. Grâce à la tomographie spectrale, il est possible de détecter ces cristaux et de les caractériser comme étant de la goutte. En effet, comme pour les lithiases rénales, on cherche l'acide urique dans les tissus autour de l'articulation (Chou, Chin et Peh, 2017).

2.4.3 Applications cardiovasculaires

Les maladies cardiovasculaires sont caractérisées par des problèmes au cœur et aux vaisseaux sanguins. Elles sont la première cause de décès dans le monde et provoquent en moyenne 31% des décès. Elles peuvent être provoquées par de l'hypertension, l'augmentation du taux de lipides et glucides sanguins ou le surpoids et l'obésité (*Cardiovascular diseases*).

Il est primordial d'arriver à détecter ces maladies lors de routines cliniques pour traiter et adapter au mieux le mode de vie du patient.

La tomographie spectrale peut être utilisée pour les applications cardiovasculaires telles que l'athérosclérose (Cormode et al., 2010). En effet, en injectant un produit de contraste dans le sang les vaisseaux sont visibles. De plus, la capacité de la tomographie spectrale à décomposer les matériaux nous permet de trouver les caractéristiques des vaisseaux comme leur taille ou la quantité d'agent de contraste dans le sang.

Il est ainsi possible d'injecter des micro-particules d'or pour marquer les plaques d'athérosclérose et de connaître le risque qu'elles se décrochent et qu'elles causent un bouchage du vaisseau. L'or présente une discontinuité dans son coefficient d'atténuation massique aux alentours de 81 keV. En sélectionnant correctement les bins d'énergie, il est possible d'avoir une image avec et une image sans or.

(Si-Mohamed et al., 2017) ont testé plusieurs applications cardiovasculaires cliniques. Il est possible de mieux visualiser l'architecture d'un stent avec la tomographie spectrale en comparaison de la tomographie classique. De plus, la visualisation de la perfusion des organes permet de pouvoir avoir une analyse quantitative des lésions ischémiques ou de tumeurs.

2.4.4 Application à l'arthrose

Comme présenté dans la partie 1.5.2, la radiographie et la tomographie conventionnelle sont les dispositifs d'imagerie diagnostic de l'arthrose. Malheureusement, il est impossible de diagnostiquer l'arthrose précoce avec ces dispositifs qui ne montrent pas de changement évidents.

La tomographie spectrale a déjà été utilisée pour détecter l'arthrose précoce (Rajendran et al., 2016) en marquant des protéines du cartilage grâce à de l'iode. Cet élément présente une discontinuité autour de 33 keV et permet de séparer l'iode et de pouvoir voir l'état microscopique du cartilage.

Les images mono-énergétiques donnant un meilleur contraste, elles semblent être un bon point d'entrée pour la détection de l'arthrose précoce. Nous nous intéresserons particulièrement à ce problème dans nos contributions en étudiant l'apport de la tomographie spectrale à la visibilité du cartilage dans les images de genoux.

2.5 Conclusion

Dans ce chapitre, nous avons présenté les caractéristiques de la tomographie spectrale. La tomographie à double énergie s'est améliorée pour arriver à des détecteur pouvant avoir de nombreux bins d'énergie grâce aux détecteurs à comptage photonique. Cette imagerie nous permet d'avoir une information sur la composition de l'objet grâce à la décomposition des tissus, des informations sur la perfusion des organes et caractériser des lésions grâce à la décomposition et la quantification de la concentration de produits de contraste. Ces qualités la rendent très utile dans de nombreux domaines comme l'imagerie cardiovasculaire ou l'imagerie des cristaux dans la goutte. En effet, elle résout certains problèmes rencontrés en tomographie conventionnelle. Notre objectif dans la suite sera d'explorer le potentiel des nouvelles méthodes de reconstruction spectrale en particulier pour l'imagerie de l'arthrose sans produit de contraste.

Chapitre 3

Apprentissage profond pour la résolution de problèmes inverses

Dans ce chapitre nous introduirons brièvement les problèmes inverses et les principales méthodes pour les résoudre. Nous présenterons ensuite l'apprentissage profond, les réseaux de neurones de convolutifs et comment ils sont utilisés pour la résolution de problèmes inverses. Nous nous focaliserons sur les problèmes rencontrés en tomographie spectrale, qui sont la reconstruction tomographique et la décomposition de matériaux.

3.1 Problèmes inverses

3.1.1 Problème direct et inverse

Les problèmes inverses sont le fait de remonter à une grandeur d'intérêt à partir de mesures. Le problème direct consiste à obtenir un modèle de la forme

$$s = F(x) + b \quad (3.1)$$

avec F l'opérateur, linéaire ou non, qui modélise les données acquises s à partir de l'image à reconstruire x avec b le bruit de mesure. En imagerie, l'opérateur F peut correspondre à la fonction d'étalement des points. En tomographie, les données acquises sont les projections d'un objet. Elles sont définies par l'équation (1.6).

Alors que le problème direct consiste à obtenir s à partir de x , le problème inverse consiste à obtenir x à partir de s . Un problème inverse est dit mal posé s'il ne satisfait pas au moins une des conditions suivantes selon (Hadamard, 1923) :

1. la solution existe ;
2. la solution est unique ;
3. les erreurs sur la solution dépendent continûment des perturbations sur les données.

Un problème peut être bien posé mais instable en présence de bruit. C'est notamment le cas des problèmes discrets pour lesquels l'opérateur F est une matrice de plein rang mal conditionnée. On définit le nombre de conditionnement d'une matrice F par :

$$\kappa(F) = \frac{\sigma_{\max}}{\sigma_{\min}} \quad (3.2)$$

où σ_{\max} et σ_{\min} sont les valeurs singulières maximale et minimale de la matrice F . Un problème avec un nombre de conditionnement faible sera considéré comme bien conditionné et un problème avec un grand nombre de conditionnement (de l'ordre de 10^5 ou supérieur) sera considéré comme mal conditionné. Dans ce cas, une faible

perturbation sur les mesures peut donner une forte perturbation sur la solution tel que

$$\frac{\|\Delta \mathbf{x}\|}{\|\mathbf{x}\|} \leq \kappa \frac{\|\Delta \mathbf{s}\|}{\|\mathbf{s}\|}. \quad (3.3)$$

Une des raisons pour laquelle un problème est mal conditionné est qu'une partie des données est fortement corrélée entre elle (Hansen, 1998).

3.1.2 Résolution des problèmes inverses

La résolution de problèmes inverses est généralement réalisée minimisant une fonction de coût C , soit

$$\hat{\mathbf{x}} = \underset{x}{\operatorname{argmin}} C(x). \quad (3.4)$$

Dans les cas simples, par exemple dans le cas des moindres carrés

$$C(\mathbf{x}) = \|\mathbf{s} - \mathbf{F}\mathbf{x}\|_2^2, \quad (3.5)$$

il est possible d'exprimer la solution de manière analytique, soit en inversant la matrice \mathbf{F} , et si \mathbf{F} n'est pas inversible, il est possible d'utiliser la pseudo-inverse de Moore-Penrose (Penrose, 1955) tel que

$$\hat{\mathbf{x}} = \mathbf{F}^\dagger \mathbf{s} \quad (3.6)$$

\mathbf{F}^\dagger la pseudo-inverse de la matrice de l'opérateur et \mathbf{s} les données. La pseudo-inverse dans les cas sur-déterminés s'écrit

$$\mathbf{F}^\dagger = (\mathbf{F}^\top \mathbf{F})^{-1} \mathbf{F}^\top \quad (3.7a)$$

avec \mathbf{F}^\top la transposée de \mathbf{F} , et

$$\mathbf{F}^\dagger = \mathbf{F}^\top (\mathbf{F}\mathbf{F}^\top)^{-1} \quad (3.7b)$$

pour les cas sous-déterminés.

Pour la tomographie \mathbf{F}^\dagger correspond à la rétroprojection filtrée.

La solution obtenue avec la pseudo-inverse est très sensible au bruit lorsque le problème est mal conditionné. Dans le cas sous-déterminé, où il existe un ensemble de solutions à (3.6), la solution de plus faible variance est choisie.

Pour palier ce problème, surtout en présence de bruit, il est possible de rajouter un terme de régularisation R dans la fonction de coût C qui prend la forme

$$C(\mathbf{x}) = \|\mathbf{s} - \mathbf{F}\mathbf{x}\|_2^2 + \alpha R(\mathbf{x}) \quad (3.8)$$

où R est un terme de régularisation qui permet de donner un *a priori* sur la solution. α est le paramètre de régularisation. Si $\alpha = 0$, la fonction de coût n'est pas régularisée. La régularisation de Tikhonov (Tikhonov, 1943) conduit à choisir

$$R(\mathbf{x}) = \|\Gamma \mathbf{x}\|_2^2 \quad (3.9)$$

où Γ est une matrice judicieusement choisie. Γ peut être le gradient ou le laplacien pour limiter les variations trop rapides de la solution.

En acquisition compressée, avec peu d'acquisitions en comparaison de la taille de la solution, le problème est sous-déterminé. Pour le résoudre, une minimisation des moindres carrés avec une régularisation L_1 sera souhaitable. La régularisation

L_1 , de la forme $\|\mathbf{x}\|_1 = \sum_i |x_i|$, favorise une solution parcimonieuse avec beaucoup de zéros. Cette solution pourra être décrite avec moins de paramètres.

La régularisation à l'aide de la variation totale (Jordan, 1881) est utilisée pour le débruitage (Rudin, Osher et Fatemi, 1992). Elle correspond à la norme L_1 de l'amplitude du gradient. La fonction coût de la variation totale anisotrope devient ainsi

$$C(\mathbf{x}) = \|\mathbf{s} - \mathbf{F}\mathbf{x}\|_2^2 + \alpha \|\nabla \mathbf{x}\|_1, \quad (3.10)$$

où ∇ est le gradient. Elle a l'avantage de débruiter assez efficacement. Elle favorise des résultats constants par morceaux ce qui peut être intéressant pour les images telles que des mires mais qui peut être plus problématique pour les images naturelles.

Certains problèmes inverses ne possèdent pas de solution analytique, comme les problèmes linéaires avec une fonction de coût non quadratique. Pour les résoudre, on peut utiliser des algorithmes qui minimisent la fonction de coût de manière itérative avec la mise à jour

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \lambda_k \mathbf{d}_k \quad (3.11)$$

où λ_k est le pas et \mathbf{d}_k est la direction de descente à l'itération k . Le pas optimal peut être calculé par recherche linéaire le long de la direction de descente tel que

$$\lambda_k = \underset{\lambda}{\operatorname{argmin}} C(\mathbf{x}_k + \lambda \mathbf{d}_k). \quad (3.12)$$

Pour la descente de gradient, la direction est calculée telle que

$$\mathbf{d}_k = -\mathbf{g}_k \quad (3.13)$$

où \mathbf{g}_k est le gradient à l'itération k .

La méthode des gradients conjugués (Hestenes et Stiefel, 1952) accélère la vitesse de convergence en comparaison de la simple descente de gradient en choisissant une direction de descente dépendant du gradient et de sa transposée tel que :

$$\beta_k = \frac{\mathbf{g}_{k+1}^\top \mathbf{g}_{k+1}}{\mathbf{g}_k^\top \mathbf{g}_k} \quad (3.14a)$$

$$\mathbf{d}_{k+1} = -\mathbf{g}_{k+1} + \beta_k \mathbf{d}_k. \quad (3.14b)$$

Quand $k = 0$, $\mathbf{d}_0 = -\mathbf{g}_0$. Cette méthode peut être utilisée, par exemple, pour de la reconstruction tomographique (Rit et al., 2014).

Pour l'algorithme de Gauss-Newton (Gauss, 1809), la direction est calculée à partir de la Hessienne et du gradient de la fonction de coût, en inversant le système

$$\mathbf{H}_k \mathbf{d}_k = -\mathbf{g}_k \quad (3.15)$$

où \mathbf{H}_k la Hessienne de la fonction de coût à l'itération k de la forme

$$\mathbf{g}_k = \mathbf{g}_{k,D} + \alpha \mathbf{g}_{k,R} \quad (3.16a)$$

$$\mathbf{H}_k = \mathbf{H}_{k,D} + \alpha \mathbf{H}_{k,R} \quad (3.16b)$$

où $\mathbf{g}_{k,D}$, $\mathbf{g}_{k,R}$, $\mathbf{H}_{k,D}$ et $\mathbf{H}_{k,R}$ sont les gradients et Hessiennes pour le terme d'attache aux données et pour le terme de régularisation.

Les problèmes inverses peuvent aussi être résolus grâce à des algorithmes d'apprentissage profond.

3.1.3 Introduction à l'apprentissage profond pour la résolution de problèmes inverses

Les architectures classiques, proposées pour des problèmes de segmentation ou de détection d'objet ont été, par la suite, appliquées à la résolution de problèmes inverses. Les réseaux de neurones convolutifs ont été développés dès 1988 pour la restauration d'images (Zhou et al., 1988). Il sont utilisés pour de nombreuses applications comme en imagerie médicale ou satellitaire (McCann, Jin et Unser, 2017) pour résoudre des problèmes comme le débruitage (Jain et Seung, 2009), la déconvolution (Schawinski et al., 2017) ou encore la reconstruction tomographique (Gupta et al., 2018).

L'apprentissage profond est de plus en plus utilisé pour la résolution de problèmes inverses en imagerie. Par exemple, pour le débruitage les réseaux sont capables d'apprendre les différents types de bruits tels que le bruit gaussien (Jain et Seung, 2009; Zhang et al., 2017; Lefkimmiatis, 2018). Dans ce cas, les entrées du réseau seront les images bruitées et les sorties, les images non bruitées. En posant x la solution débruitée et s la mesure bruitée alors pour un lot avec N paires d'entrées/sorties d'apprentissage $\{s^{(n)}, x^{(n)}\}$, $1 \leq n \leq N$, on minimise une fonction de coût pour chaque paire telle que

$$C^{(n)}(\omega) = \|h_{\omega}(s^{(n)}) - x^{(n)}\|_2^2 \quad (3.17)$$

avec ω les poids des neurones du réseau h . La fonction de coût de tout le lot sera

$$C(\omega) = \begin{cases} \sum_{n=1}^N C^{(n)}(\omega) & \text{si somme} \\ \frac{1}{N} \sum_{n=1}^N C^{(n)}(\omega) & \text{si moyenne} \end{cases} \quad (3.18)$$

Pour certains apprentissages, on donne déjà un *a priori* sur le bruit et dans d'autres cas on opère un débruitage aveugle sans connaître les caractéristiques du bruit. Les réseaux utilisés sont des réseaux spécifiques à l'application ou des réseaux bien connus comme les réseaux résiduels ou l'architecture U-net (Abascal et al., 2021a).

3.2 Introduction à l'apprentissage profond

L'apprentissage profond a été inspiré par la façon dont le cerveau humain résout des problèmes d'apparence simple mais très complexes à traiter mathématiquement. En prenant l'exemple de la vue, pour réussir à analyser le monde qui nous entoure, des centaines de millions de cellules photosensibles sont présentes sur la rétine. Plusieurs de ces cellules sont connectées à de plus grosses cellules et ainsi de suite comme schématisé sur la Figure 3.1. Chaque cellule est considérée comme un neurone et peut être modélisée mathématiquement comme sur la Figure 3.2. Ce modèle du *neurone formel* a été introduit par (McCulloch et Pitts, 1943). Un réseau de cellules est organisé de manière plutôt aléatoire à la naissance, les poids du réseau sont souvent aléatoires au départ. Grâce à de nombreux stimuli, la probabilité d'utiliser une liaison plutôt qu'une autre va se modifier pour réussir au mieux l'opération demandée. Cela correspond à l'adaptation des poids du réseau pendant l'apprentissage. L'adaptation des poids se fait en minimisant une erreur judicieusement choisie pour le problème à résoudre.

Le développement de l'apprentissage par ordinateur date des années 50 avec le perceptron (Rosenblatt, 1958). Ce modèle a été développé pour surpasser les limitations des modèles statistiques avec des associations linéaires (Le Cun et Fogelman-Soulié, 1987). En effet, dans les neurones 3.2, il y a une fonction d'activation qui est

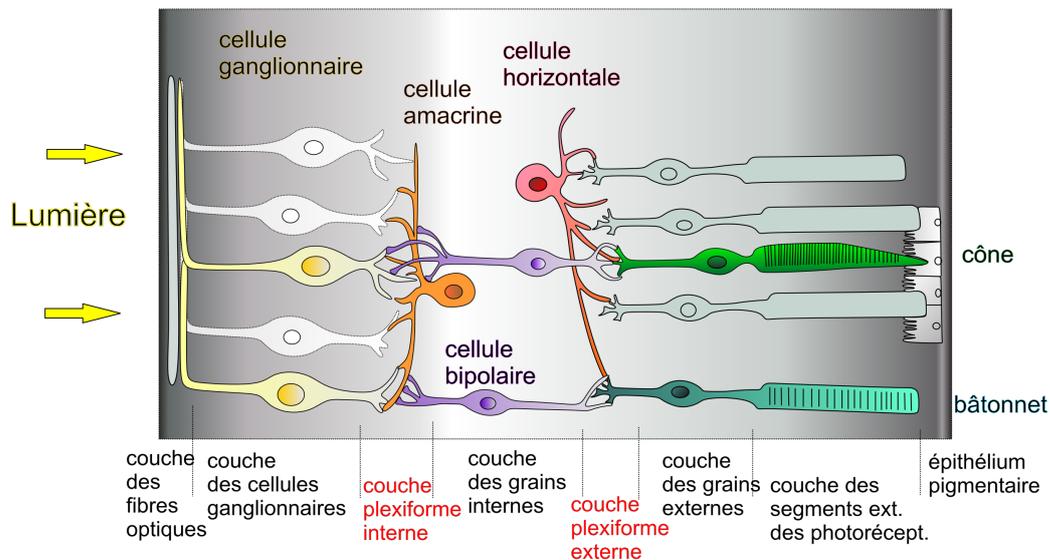


FIGURE 3.1 – Schéma de l'architecture des cellules composant la rétine (source : (Purves et al., 2005)).

non linéaire. Cela a permis d'apprendre le *où exclusif* qui est un opérateur très important en électronique et informatique. En associant plusieurs neurones, on aboutit à l'exemple le plus connu d'apprentissage profond, le perceptron multicouches (Figure 3.3). Il est caractérisé par différentes couches avec un nombre donné de neurones par couche. Les neurones des couches adjacentes sont tous connectés les uns aux autres. On les appelle couches pleinement connectées. Chaque couche permet d'apprendre une fonction potentiellement plus complexe que la précédente grâce à la fonction non-linéaire qui est appliquée après la combinaison linéaire des neurones reliés issus de la couche précédente. (Le Cun et Fogelman-Soulié, 1987) réalise une vue d'ensemble des progrès du classifieur linéaire au perceptron multicouches et souligne l'importance des données présentées à l'algorithme pour réaliser l'apprentissage. En effet, il est nécessaire de bien connaître les données et de les présenter dans le bon ordre.

Deux bibliothèques sont majoritairement utilisées pour coder les réseaux de neurones. Tensorflow (Abadi et al., 2015) et Pytorch (Paszke et al., 2017) se basent sur le langage Python (Van Rossum et Drake, 2009).

3.2.1 Cas des réseaux de neurones convolutifs

La première architecture de réseaux de neurones convolutifs pour la reconnaissance de motifs simples a été publiée par (Rumelhart, Hinton et Williams, 1986). Ceux-ci sont particulièrement bien adaptés au traitement d'image. Ils utilisent un noyau ou filtre de convolution pour traiter plusieurs pixels à la fois et limiter ainsi le nombre de paramètres par couche en comparaison des couches pleinement connectées. La Figure 3.5 présente un réseau avec une couche de convolution. La fonction d'activation reste toujours présente pour permettre de créer de la non-linéarité entre chaque couche de convolution. Pour obtenir la valeur du pixel en sortie y on calcule

$$y = f(x * \omega). \quad (3.19)$$

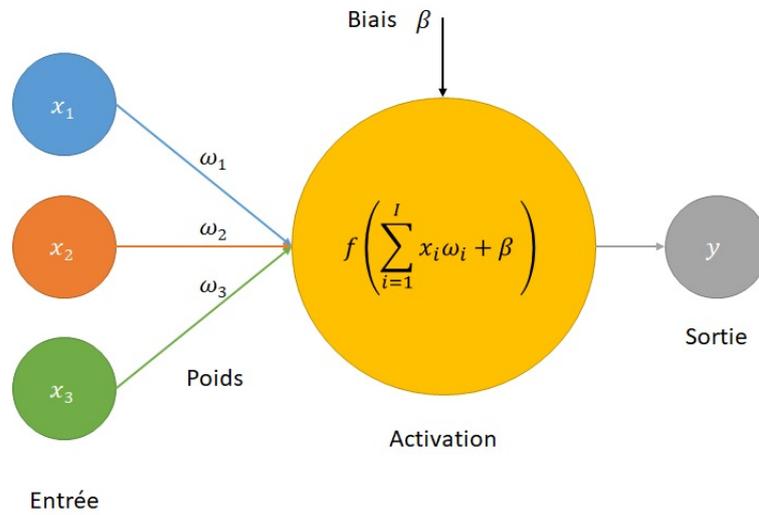


FIGURE 3.2 – Schéma d’un neurone avec x_i les entrées, ω_i les poids, f la fonction d’activation non linéaire, β le biais et y la sortie. Le neurone apprend les poids ω_i et le biais β pour produire la sortie y compte tenu des données d’apprentissage x_i .

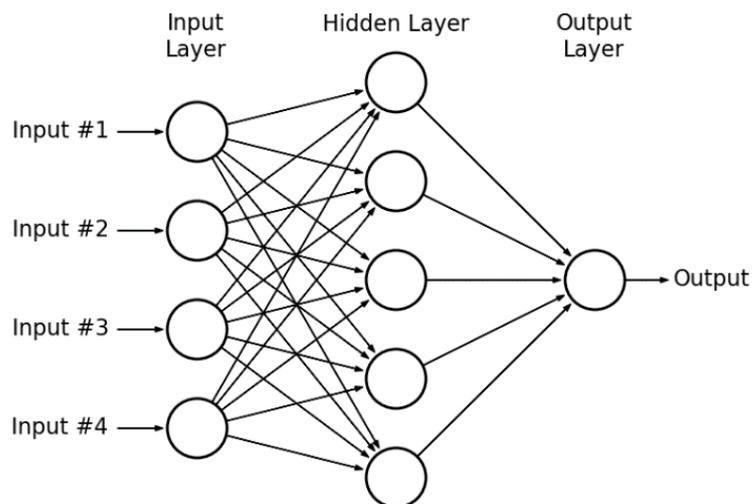


FIGURE 3.3 – Perceptron multicouche avec 4 entrées, 1 sortie et 1 couche cachée avec 5 neurones (source :(Hassan et al., 2015)).



FIGURE 3.4 – Sortie de la première couche de convolution avec 96 filtres de taille $11 \times 11 \times 3$ pour un problème de classification sur la base ImageNet (source : (Krizhevsky, Sutskever et Hinton, 2017)).

Les poids du réseau ω dépendent uniquement du noyau de convolution et non plus de chaque relation entre pixel de la couche inférieure et celle de la couche supérieure. Ainsi avec I la taille du filtre, on calcule la sortie avec les I pixels coïncidents avec le filtre. On réalise cette opération sur toute l'image avec les I ω_i fixes et les x_i qui dépendent de la localisation dans l'image. Avec n pixels sur la couche en amont et m pixels sur la couche en aval, et un noyau de convolution avec k coefficients, on passe donc de $n \times m$ poids pour une couche pleinement connectée à uniquement k poids pour une couche convolutive. Si on a une image en niveau de gris de taille 100×100 pixels avec une couche cachée de la même taille que l'image, il nous faudra 10'000 poids avec une couche pleinement connectée. En prenant la même image et un filtre de convolution 3×3 , nous n'en aurons que 9. La taille du noyau de convolution dépend du nombre de canaux couleur. Le nombre de filtre est la troisième composante de la taille de ce noyau.

Les réseaux de neurones convolutifs, décrivent l'image par des formes simples, des motifs, comme des gradients de couleur ou de niveau de gris (Figure 3.4). Les sorties des couches de convolution vont pouvoir décrire des structures géométriques de plus en plus complexes plus la couche est profonde grâce à la fonction d'activation non linéaire.

Le succès des architectures basées sur des couches de convolution a été accéléré avec l'apparition de défis tels que *ImageNet Large Scale Visual Recognition Challenge* à partir 2011 (Russakovsky et al., 2015) avec des tâches de détection, classification et segmentation. Des réseaux tels que VGG16 (Simonyan et Zisserman, 2015) ou U-net (Ronneberger, Fischer et Brox, 2015; Çiçek et al., 2016) font aussi leur apparition en 2015. VGG16, présenté Figure 3.6, est une architecture qui réduit la taille des filtres de convolution et augmente le nombre de couches avec 13 couches de convolution et 3 couches pleinement connectées pour la classification. L'augmentation du nombre de couches permet de multiplier les chances d'apprendre des fonctions complexes. La réduction de la taille du filtre de convolution permet d'apprendre des statistiques très locales et réduire le nombre de paramètres. L'architecture du réseau U-net, présenté Figure 3.7, permet de conserver une information avant le sous-échantillonnage. Cette information plus haute résolution peut être très utile pour la segmentation de petites structures.

3.2.2 Aspect techniques de la phase d'apprentissage

Après avoir choisi une des architecture suscitée, défini le problème à résoudre comme la segmentation ou le débruitage d'objet par exemple, on doit choisir une fonction de coût qui sera en rapport avec le problème. Pour évaluer la qualité d'une segmentation on aura tendance à utiliser la mesure de l'indice DICE (Dice, 1945). Pour évaluer la qualité d'un débruitage on peut utiliser l'erreur quadratique moyenne. Ensuite, on peut passer à la phase d'apprentissage puis de test.

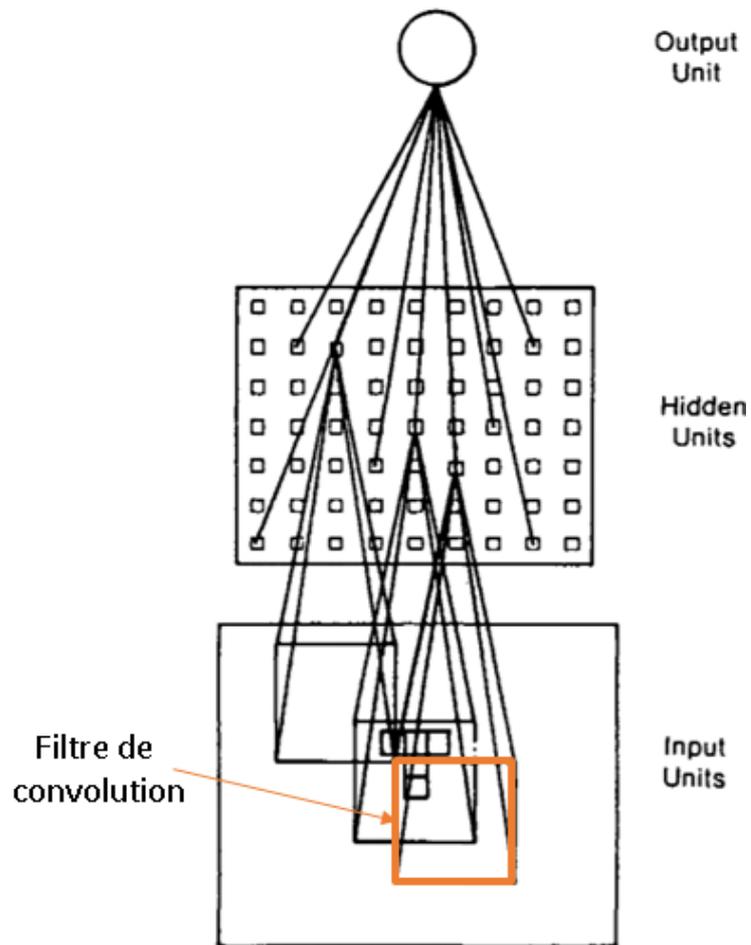


FIGURE 3.5 – Schéma d'un réseau avec une couche de convolution annoté avec le filtre de convolution en orange (source : (Rumelhart, Hinton et Williams, 1986)).

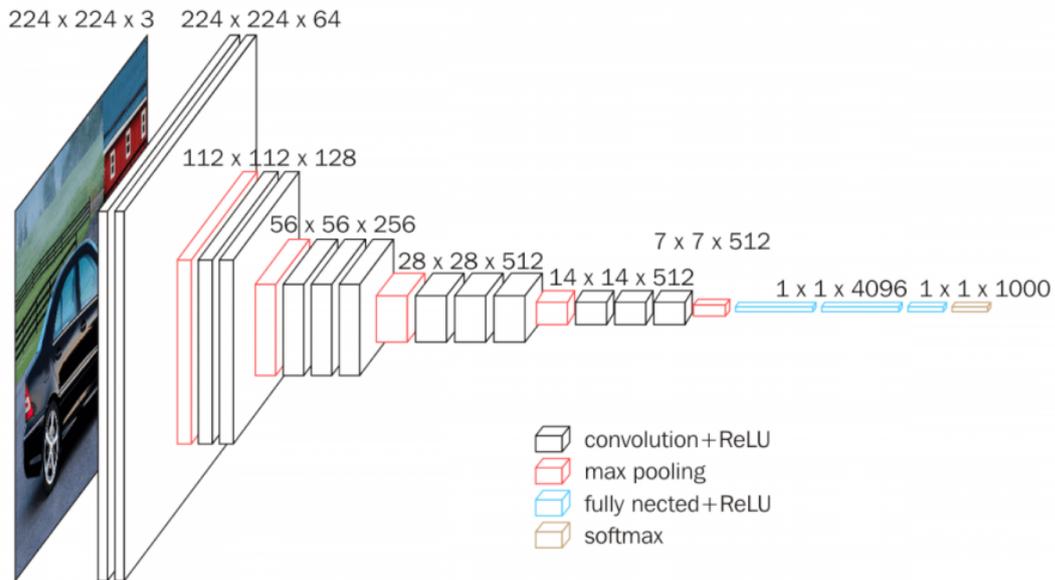


FIGURE 3.6 – Architecture VGG16 de (Simonyan et Zisserman, 2015) (source : (*VGG16 - Convolutional Network for Classification and Detection* 2018)). Les deux premières dimensions représentent la taille de l'image et la troisième est le nombre de filtres.

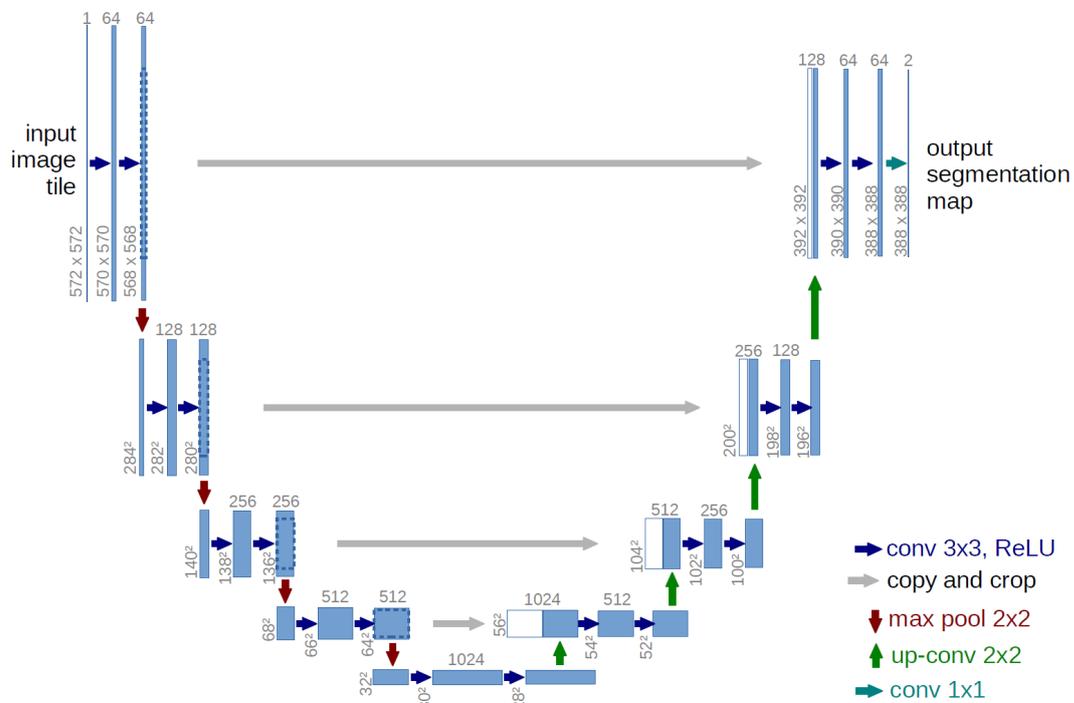


FIGURE 3.7 – Architecture du U-net de (Ronneberger, Fischer et Brox, 2015). Les deux dimensions sont la taille de l'image. Le nombre au dessus représente le nombre de filtres.

Apprentissage et test L'apprentissage profond se décompose en deux phases, l'apprentissage ou entraînement et le test ou l'inférence. On sépare les images disponibles en trois bases, la base d'entraînement qui se sera utilisée pour ajuster les poids du réseau lors de l'entraînement, la base de validation pour vérifier le bon apprentissage du réseau et la base de test qui permet de vérifier que le réseau peut généraliser à des données inconnues.

Les hyperparamètres du réseau sont la **fonction de coût**, qui mesure la similarité entre les entrées et les sorties, qui est minimisée grâce à un **optimiseur** avec un **taux d'apprentissage**.

Phase d'apprentissage Lors de la phase d'apprentissage, de nombreuses images de la base de données d'apprentissage sont présentés au réseau. Les images peuvent être présentées par lots avec la **taille de lot** égale à 1 si on ne présente qu'une seule image à la fois. Une fois la base terminée, on adapte les poids du réseau.

On peut calculer la fonction de coût C_l pour chaque lot indépendamment en suivant (3.17) et la fonction de coût totale C sera la moyenne de ces fonctions de coût par lot l de la forme

$$C(\omega) = \frac{1}{n} \sum_{l=1}^n C_l(\omega) \quad (3.20)$$

ω correspond au vecteur de poids du réseau. Ainsi, on va vouloir connaître le gradient de chacune de ces fonctions de coût en fonction du lot l en utilisant par exemple la méthode du gradient stochastique (Bottou, 1999). La mise à jour des poids est réalisée dans chaque lot telle que

$$\begin{aligned} \omega_{k+1} &= \omega_k - \lambda \nabla C_l(\omega_k) \\ \omega_{k+1} &= \omega_k - \frac{\lambda}{n} \sum_{l=1}^n \nabla C_l(\omega_k), \end{aligned} \quad (3.21)$$

où k est l'indice du poids du réseau et ne dépend pas du lot l , ∇ est l'opérateur du gradient et λ correspond au pas de l'optimiseur qui est nommé taux d'apprentissage ou d'entraînement en apprentissage profond. On utilise la plupart du temps la méthode de rétro-propagation du gradient. Pour une meilleure convergence, il est possible d'utiliser des taux d'apprentissage adaptatifs (Murphy, 2022).

Il est possible d'ajouter de la mémoire lors de l'apprentissage que l'on appelle moment. Ainsi, la méthode du moment est définie telle que

$$\Delta\omega_{k+1} = \lambda \nabla C_l(\omega_k) + \alpha \Delta\omega_k \quad (3.22a)$$

$$\omega_{k+2} = \omega_k - \Delta\omega_{k+1}. \quad (3.22b)$$

Un des algorithmes les plus connus utilisant la méthode du moment est ADAM pour *Adaptative Moment Estimation* (Kingma et Ba, 2015) qui se définit tel que

$$m_{k+1} = \beta_1 m_{\omega,k} + (1 - \beta_1) \nabla C(\omega_k) \quad (3.23a)$$

$$v_{k+1} = \beta_2 v_{\omega,k} + (1 - \beta_2) (\nabla C(\omega_k))^2 \quad (3.23b)$$

$$\hat{m}_{k+1} = \frac{m_{\omega,k+1}}{1 - \beta_1} \quad (3.23c)$$

$$\hat{v}_{k+1} = \frac{v_{\omega,k+1}}{1 - \beta_1} \quad (3.23d)$$

$$\omega_{k+1} = w_k - \lambda \frac{\hat{m}_{k+1}}{\sqrt{\hat{v}_{k+1} + \epsilon}} \quad (3.23e)$$

où m est le moment du gradient avec un facteur d'oubli β_1 et v est le second moment du gradient avec un facteur d'oubli β_2 , ϵ est un petit nombre pour éviter la division par 0. Les valeurs de base dans PyTorch sont $\beta_1 = 0.9$, $\beta_2 = 0.999$ et $\epsilon = 10^{-8}$.

Le taux d'apprentissage peut aussi être modifié au cours de l'apprentissage. En effet, à un moment de l'apprentissage, il est possible que la fonction de coût ne puisse plus décroître car le pas de l'optimiseur est trop grand. Ainsi il est possible de le diminuer, par exemple avec une division par un facteur toutes les n époques.

La phase d'apprentissage se termine lorsqu'il y a convergence. On choisit de terminer l'entraînement lorsque la fonction de coût ne varie plus assez. On peut aussi arrêter l'entraînement grâce à une base de validation qui va être une base inconnue du réseau et qui permet de se rendre compte de la capacité du réseau à généraliser, mais durant l'entraînement. Quand la fonction de coût de la base de validation ne varie plus ou réaugmente, on arrête l'entraînement. Cette phase est la plus lente et peut prendre de quelques minutes à quelques mois suivant la capacité de calcul, la taille du modèle et le nombre de paramètres à optimiser et la taille de la base utilisée pour l'entraînement.

Il y a ensuite une phase de test. Pour celle-ci, on utilise une base inconnue du réseau, la base de test, on passe les images dans le réseau avec les poids déjà adaptés. Les poids ne seront pas modifiés dans cette phase. Cette phase est donc très rapide et permet de surpasser la plupart des algorithmes itératifs.

Il est d'usage de présenter les images issues de la base de test car elles reflètent à la fois la qualité d'apprentissage mais aussi de généralisation du réseau à des données inconnues.

Augmentation de données Pour apprendre un modèle, les algorithmes d'apprentissage profond ont besoin de suffisamment de données. Ces données doivent être représentatives du problème à apprendre, en prenant en compte toute sa variabilité. Il y a de nombreuses banques d'images naturelles mais peu de banques d'images médicales même si celles-ci se démocratisent. On peut retrouver des banques d'images lors de concours comme KITS19 (Heller et al., 2020). Les banques d'images médicales ont besoin d'être annotées par un ou plusieurs médecins radiologues pour étiqueter les zones, pour des problèmes de segmentation ou de classification, ce qui prend énormément de temps. De plus, pour des applications très spécifiques, il est d'usage que les groupes de chercheurs fassent leurs propres acquisitions. Dans ce cas, ils auront peu de données mais devront quand même apprendre à partir de celles-ci. Ils chercheront donc à augmenter artificiellement le jeu de données.

Avec AlexNet, (Krizhevsky, Sutskever et Hinton, 2017) augmentent leurs données en réalisant des rognages aléatoires puis en les retournant horizontalement et en changeant la couleur dans le domaine RGB avec l'augmentation de couleur par analyse en composante principale. Depuis, l'augmentation de données se développe. On peut augmenter les données en manipulant les images ou en modifiant des éléments d'apprentissage profond (Shorten et Khoshgoftaar, 2019).

Modifications géométriques Pour augmenter les données, on peut appliquer des transformations simples comme retourner l'image (Figure 3.8,(B)), faire une rotation de l'image (Figure 3.8,(C)), traduire l'objet d'intérêt de l'image (Figure 3.8,(D)), ajouter un facteur d'échelle (Figure 3.8,(E)). La rotation de l'image permet à l'algorithme d'être invariant à la rotation et de simuler une prise de vue différente par



FIGURE 3.8 – Augmentations courantes avec (A) l’image originale, (B) l’image retournée verticalement, (C) l’image avec une rotation de 30° , (D) l’image translaturée vers la gauche, (E) l’image avec un facteur d’échelle de 2 et (F) un patch de l’image

exemple un scan de haut en bas par rapport à un scan de bas en haut quand on retourne l’image de 180° . La rotation peut être à des angles plus fins qui ne pourront pas conserver les dimensions de départ et les données dans les angles pourront être perdues si on conserve la même taille d’image avant et après la transformation. La translation de l’objet change uniquement l’objet de place au sein de l’image. Ainsi le réseau apprendra que l’objet peut se trouver partout dans l’image et ne pas supposer que l’objet se trouve uniquement à un seul endroit de l’image, souvent le centre. Ajouter un facteur d’échelle permet de voir l’objet plus grand ou plus petit dans l’image. Il peut imiter un organe plus petit ou plus grand. Dans le cas d’une base de données très stéréotypée avec toujours la même morphologie cela peut permettre de diversifier artificiellement les morphologies et la quantité de tissus traversée. Conventionnellement la taille de l’image est conservée. Il est aussi possible d’ajouter du bruit aux images ou de changer l’espace ou la dynamique couleur pour simuler des conditions d’acquisition différentes.

Augmentation des données dans le réseau On différencie des augmentations réalisées dans le réseau à celles réalisées en amont. Il y a de nombreuses manières d’augmenter les données dans le réseau, on ne citera que quelques exemples.

Pour réaliser de l’augmentation de données à l’intérieur du réseau, il est possible de modifier la taille de l’espace d’apprentissage. Cet espace correspond à la partie du réseau où les données sont représentées avec le moins de paramètres. Il est possible d’améliorer les résultats en sortie du réseau de manière significative, lors de la phase de test, en modifiant des couches après entraînement (Konno et Iwazume, 2018). Cela peut être vu comme de l’adaptation de domaine.

Les réseaux antagonistes génératifs permettent également d’augmenter les données en créant des données réalistes à partir de bruit ou en n’altérant pas les images en injectant du bruit (Bowles et al., 2018). Ce que l’on peut appeler le générateur de visage de NVidia (Karras et al., 2018) est plutôt impressionnant quant à sa capacité

à créer des visages très réalistes de personnes qui n'existent pas. Ainsi, l'entrée du réseau n'est pas une entrée qui a un sens pour l'oeil humain ou un sens physique mais des éléments ayant un sens sont créés à partir de bruit puis exploités.

Approche par patch Pour l'approche par patch, on ne sélectionne qu'une partie l'image (Figure 3.8,(F)). Ensuite, soit l'image est remise à sa taille d'origine soit on garde la taille de la découpe qu'on pourra appeler ultérieurement patch. L'image peut être coupée aléatoirement selon une taille donnée. Cela permet de porter toute l'attention du réseau sur des petites zones, indépendamment, sans que celui-ci ne sache de quel endroit il s'agit. De plus, conserver la taille de la zone découpée permet de traiter des plus petites images et de mieux gérer la mémoire.

Cette approche a été développée en conséquence du grand nombre de données par rapport à la mémoire disponible pour l'apprentissage. (Louppe et Geurts, 2012) utilise des patchs choisis aléatoirement pour apprendre des estimateurs et ensuite choisir le meilleur. Il démontre qu'il est possible d'apprendre avec une partie de la base de données sans pour autant perdre en précision. Il est aussi possible de choisir les patchs et non plus en prendre des aléatoires. Ils peuvent être choisis avec uniquement des informations pertinentes (Roy et al., 2019) ou en ajoutant des informations non-pertinentes dans le patch (Kao et al., 2020).

En PyTorch, il est possible de prendre des patchs aléatoires lors du chargement des données et de les changer à chaque époque. Les patchs seront ainsi choisis aléatoirement et différents au sein du lot et entre les lots.

Normalisation La normalisation permet d'avoir moins de variabilité dans la base de données et dans les poids qui pourraient empêcher le bon apprentissage du réseau. La normalisation peut être réalisée sur les données où au sein du réseau.

Normalisation de lot La normalisation de lot (batch normalization en Anglais) normalise les sorties d'un neurone avant ou après la fonction d'activation. Cela peut augmenter la vitesse d'apprentissage en utilisant un taux d'apprentissage plus grand, en réduisant la régularisation, accélérer la décroissance du taux d'apprentissage et permettre de mélanger des échantillons très différents (Ioffe et Szegedy, 2015).

Normalisation des données Il est possible de normaliser les données avant de les injecter dans le réseau. Il est possible de centrer et réduire les données telles que :

$$\hat{X}[:, i] = \frac{X[:, i] - m_i}{\sigma_i} \quad (3.24)$$

avec $X[k, i]$ la valeur du pixel k pour le canal i , $m_i = \frac{1}{N} \times \sum_{k=1}^N (X[k, i])$ et $\sigma_i = \sqrt{\frac{1}{N-1} \times \sum_{k=1}^N (X[k, i] - m_i)^2}$.

On peut également normaliser les données entre 1 et -1

$$\hat{X}[:, i] = \left(\frac{X[:, i] - \min(X[:, i])}{\max(X[:, i]) - \min(X[:, i])} \right) \times 2 - 1 \quad (3.25)$$

La normalisation peut améliorer la stabilité du modèle et la vitesse de convergence (Bishop, 1995).

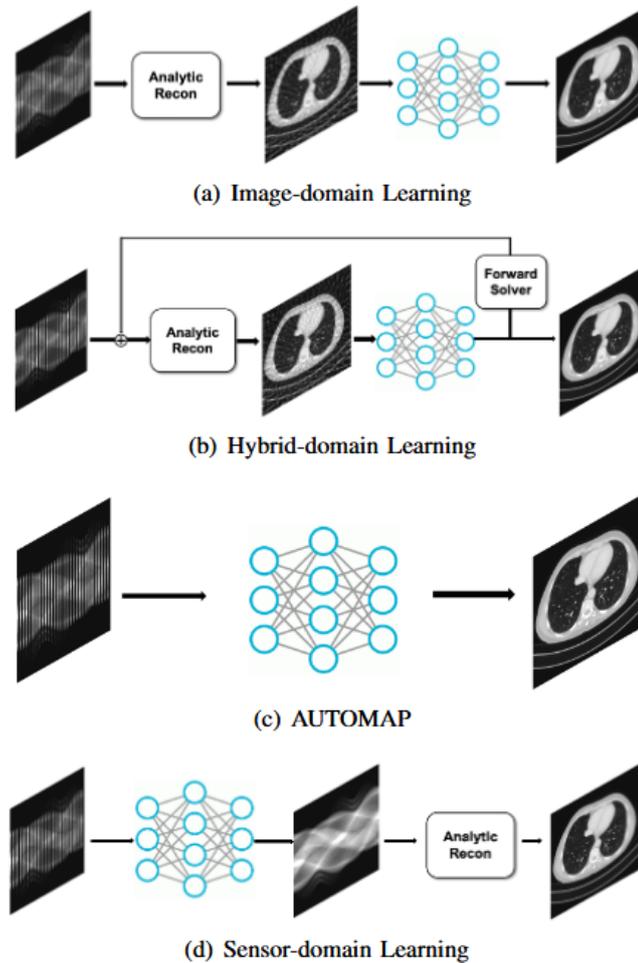


FIGURE 3.9 – Différentes stratégies pour l'utilisation de l'apprentissage profond pour la reconstruction tomographique (source : (Ravishankar, Ye et Fessler, 2020)).

Pour les problèmes dont le modèle direct est exponentiel, comme l'atténuation des rayons X, La normalisation logarithmique permet de linéariser le problème tel que

$$\hat{X}[j, i] = \ln \frac{X[j, i]}{\tilde{X}[j, i]}. \quad (3.26)$$

où $\tilde{X}[j, i]$ est correspond à valeur maximum possible sur ce pixel. En imagerie X, elle correspond à la mesure de plein flux. On peut retrouver cette normalisation dans l'équation (1.5).

3.3 Apprentissage pour la résolution de problèmes inverses

3.3.1 Apprentissage profond et reconstruction tomographique

La reconstruction tomographique est un sujet brûlant dans le domaine de l'apprentissage (Wang, Ye et De Man, 2020). En effet, les calculs peuvent durer plusieurs heures voire jours et nécessitent d'être parallélisés. Comme présenté sur la Figure 3.9, il y a grossièrement 4 stratégies pour l'utilisation de l'apprentissage profond pour la reconstruction tomographique.

La première méthode (Figure 3.9,(c)), la transformée automatique par approximation de variété, apprend à modéliser les résultats dans le domaine image directement à partir des mesures dans le domaine des projections (Zhu et al., 2018). Ainsi, en nommant s le sinogramme et x la reconstruction, on minimise une fonction de coût de la forme

$$C(\omega) = \|h_\omega(s) - x\|_2^2. \quad (3.27)$$

Des couches pleinement connectées et des couches de convolution sont nécessaires, ainsi que beaucoup de mémoire car les couches pleinement connectées ont beaucoup de paramètres. Ces contraintes rendent ces méthodes assez coûteuses en temps et mémoire.

La deuxième méthode (Figure 3.9,(a)), l'apprentissage dans le domaine image apprend à réduire les artefacts présents sur des reconstructions issues de méthodes analytiques (Kang, Min et Ye, 2017; Chen et al., 2017; Wolterink et al., 2017; Shan et al., 2019). Ainsi en nommant \tilde{X} la fonction de reconstruction analytique, telle que $\tilde{x} = \tilde{X}(s)$, et x la solution débruitée, on minimise une fonction de coût de la forme

$$C(\omega) = \|h_\omega(\tilde{x}) - x\|_2^2. \quad (3.28)$$

L'avantage est que l'on apprend uniquement un débruitage dans le domaine image, cela peu coûteux et efficace. Le précurseur pour la reconstruction tomographique avec l'apprentissage profond est le FBPConvNet qui combine la rétroprojection filtrée et un réseau de neurones convolutif de type U-net (Jin et al., 2016), l'initialisation est faite avec une reconstruction grossière. Les algorithmes itératifs d'apprentissage montrent aussi des bons résultats pour la reconstruction tomographique ou de la régression (Adler et Öktem, 2017; Mozumder et al., 2021). Ceux-ci changent l'entrée du réseau en fonction de la sortie trouvée et de manière itérative. L'apprentissage profond permettra d'encoder un *a priori* sur les données (Gupta et al., 2018). Dans ces deux cas, une première reconstruction sans apprentissage est nécessaire.

La troisième (Figure 3.9,(b)), l'apprentissage dans un domaine hybride impose que sortie du réseau soit proche de la reconstruction à partir de l'inversion de l'opérateur du modèle direct et réalise un apprentissage comme la première méthode (Kang et al., 2018; Adler et Öktem, 2018; Chen et al., 2018; Wu et al., 2017). On minimise une fonction de coût de la forme

$$C(\omega) = \|h_\omega(\tilde{x}) - x\|_2^2. \quad (3.29)$$

Ajouter la connaissance d'une reconstruction grossière permet d'améliorer encore les reconstructions qui maintiendront toujours un sens physique.

La quatrième méthode (Figure 3.9,(d)), l'apprentissage dans le domaine du capteur, apprend à améliorer les images dans le domaine des projections avec des algorithmes qui peuvent apprendre l'interpolation entre les angles et le débruitage (Lee et al., 2019; Würfl et al., 2016; Ghani et Karl, 2019). Dans ce cas on nomme \tilde{s} le sinogramme sous échantillonné et bruité alors on opérera une minimisation d'une fonction de coût de la forme

$$C(\omega) = \|h_\omega(\tilde{s}) - s\|_2^2. \quad (3.30)$$

L'avantage de traiter les données dans le domaine du capteur est que l'on travaille directement sur des données ayant un sens physique et qui n'ont pas pu être altérées au préalable. De plus, les traitements ne sont pas plus coûteux que pour la première méthode. De nouveaux réseaux permettent de reconstruire directement depuis le sinogramme (Yang et al., 2020). La reconstruction tomographique est un défi avec la

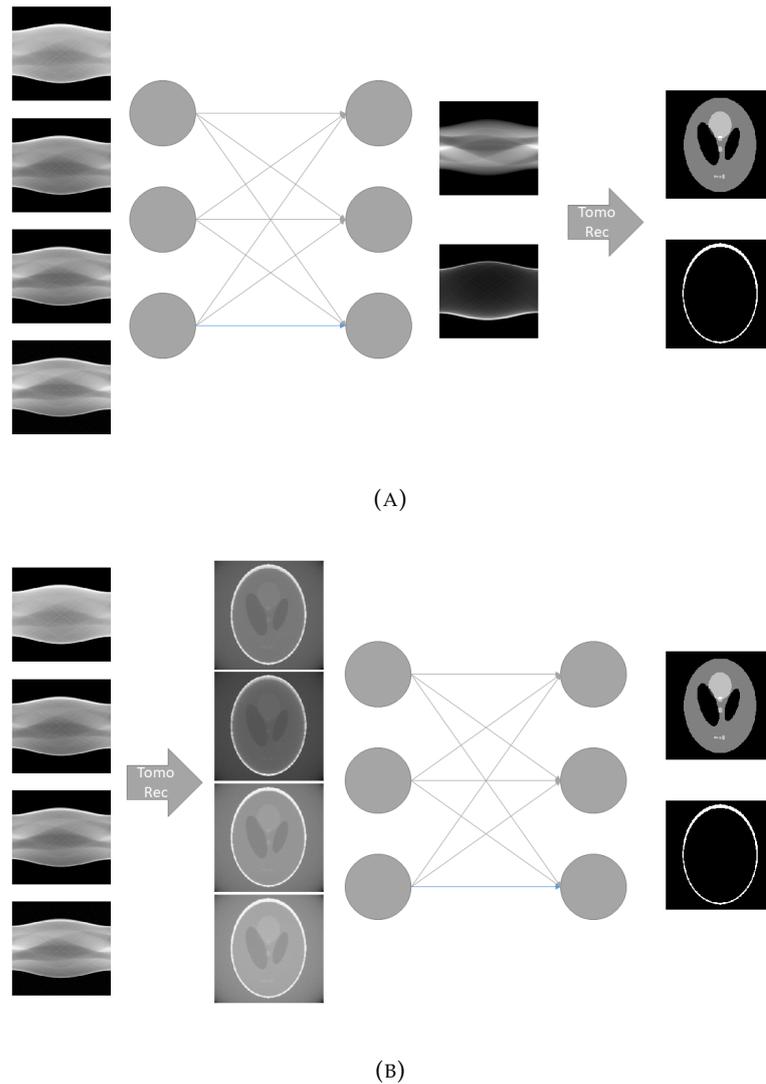


FIGURE 3.10 – Schéma de la décomposition de matériaux dans deux domaines. (A) Décomposition dans le domaine des projections. (B) Décomposition dans le domaine image.

tomographie à faible dose. En effet, le bruit dans les acquisition peut créer des artefacts dans les images reconstruites. Des algorithmes d'apprentissage profond ont été développés pour palier ce problème. Ils peuvent être basés sur des algorithmes de reconstruction itératifs (Zheng et al., 2018) ou des algorithmes de réduction d'artefacts.

3.3.2 Cas particulier de la tomographie spectrale

Comme présenté dans (2.3) le modèle direct de la tomographie spectrale peut être divisé en deux sous-modèles : la projection et le mixage spectral. Il se pose ainsi un nouveau problème inverse en plus de la reconstruction tomographique, qui peut être résolu à l'aide de l'apprentissage profond. On considère ainsi le problème de la décomposition de matériaux qui peut être appréhendé dans le domaine des projections ou dans le domaine image.

Dans le domaine de projections comme schématisé Figure 3.10,(A), on cherche à apprendre le réseau en minimisant la fonction de coût

$$C(\omega) = \|h_\omega(\hat{s}) - a\|_2^2. \quad (3.31)$$

avec h_ω le réseau de paramètres ω et \hat{s} les mesures spectrales normalisées. Cela correspond aux projections calculées grâce à un modèle comme décrit par l'équation (2.6) et normalisées avec la mesure en plein flux.

Dans le domaine image comme schématisé Figure 3.10,(B), on cherche à apprendre le réseau en minimisant la fonction de coût

$$C(\omega) = \|h_\omega(\hat{S}) - \rho\|_2^2. \quad (3.32)$$

avec \hat{S} la reconstruction tomographique des mesures spectrales normalisées.

Domaine image Pour résoudre ce problème dans le domaine image, il a été proposé de nombreux réseaux.

(Chen et Li, 2017) propose une version simplifiée de l'architecture VGG16 pour décomposer 5 matériaux à partir de 8 bins d'énergie sur des fantômes. Ce réseau a été choisi car il montre de bonnes performances pour la classification. De plus, l'avantage d'utiliser un réseau de neurones convolutif permet de s'émanciper des problèmes spatiaux. Les entrées du réseaux sont les projections reconstruites dans les bins d'énergie. Les sorties du réseau sont les matériaux décomposés. La fonction de coût est une erreur quadratique moyenne. Cela donne des meilleurs résultats que l'inversion simple du problème qui correspond à la résolution du système linéaire du modèle de matériaux (2.1) à plusieurs énergies.

(Clark, Holbrook et Badea, 2018) propose un réseau U-net, moins profond que l'originale, visible sur la Figure 3.11. Il décompose en 4 matériaux à partir de 5 bins d'énergie. Ainsi les entrées du réseau sont des images représentant l'atténuation linéaire de différents matériaux combinés de manière géométrique dans les bins d'énergie. Les sorties du réseau sont les cartes de matériaux dans le domaine image. La fonction de coût est une erreur quadratique moyenne. La décomposition est robuste, même sur des données réelles. L'avantage de U-net par rapport à VGG16, par exemple, est la conservation de la résolution dans le réseau grâce aux connexions résiduelles à chaque étage. La décomposition avec ce type de réseau conduit à une réduction du bruit en apprenant sur des cartes de matériaux qui ne présentent pas de bruit.

(Zhang et al., 2019) propose le Butterfly-Net pour décomposer en 2 matériaux à partir de mesures à double énergie. Les entrées du réseau sont séparées en deux, l'entrée basse énergie et l'entrée haute énergie. Les sorties du réseau sont les cartes de matériaux. La fonction de coût est la distance Euclidienne dans chaque paire du réseau. En effet ce réseau avec la connexion transversale permet de réaliser deux minimisations. La décomposition est meilleure qu'avec l'inversion directe ou avec une méthode itérative sans apprentissage, de même que la gestion du bruit. L'architecture est particulièrement travaillée pour la décomposition de deux énergies à deux matériaux. De plus, comme pour les réseaux présentés ultérieurement, l'utilisation des réseaux de neurones aident à réduire le bruit.

(Li, Chun et Long, 2020) propose BCD-net, pour *Block coordinate descent*, qui est un réseau de neurones itératif pour la décomposition en tissus mous et os, ou tissu mou, os et produit de contraste à partir de mesures à double énergie. Le réseau utilise plusieurs blocs et plusieurs fonctions de coût. La fonction de coût qui permet

de reconstruire les cartes de matériaux est une erreur quadratique moyenne pondérée. Celle-ci est combinée avec un *a priori* sur les statistiques du bruit dans les images. Le bloc de réduction du bruit utilise une fonction de régularisation. Le réseau est capable de retirer les artefacts des images lors de la décomposition et donne de meilleurs résultats, en terme d'erreur quadratique moyenne, que l'inversion directe ou des algorithmes de décomposition usuels en tomographie double énergie.

(Gong et al., 2020) utilise un réseau, avec des blocs *Inception* (Szegedy et al., 2014). Il décompose en 4 matériaux à partir de 3 bins d'énergie. Pour ce réseau, les entrées sont les masses volumiques dans les bins d'énergie. Les sorties sont les masses volumiques des matériaux. La fonction de coût est une erreur quadratique moyenne avec une fonction de régularisation qui préserve les bords. Il surpasse les réseaux cités précédemment avec une meilleure robustesse au bruit et moins de sensibilité au changement de dose, pour une meilleure précision. En effet, les blocs *Inception* permettent à la fois la robustesse de la décomposition et la suppression du bruit en préservant les bords des objets.

Domaine des projections Dans le domaine des projections, des algorithmes pour la décomposition dans une base avec agent de contraste ont également été développés. (Abascal et al., 2021b) propose une architecture de type U-net avec une connexion résiduelle pour décomposer 2 matériaux à partir de 5 bins d'énergie. Les entrées du réseau sont les mesures spectrales dans les bins d'énergie normalisées à l'aide de la mesure de plein flux. Les sorties sont les projections des matériaux normalisées. La fonction de coût minimisée est une erreur quadratique moyenne. La décomposition avec ce réseau surpasse celle avec la méthode Gauss-Newton régularisée présentée dans la sous-section 3.1.2. En effet, il est possible de réduire le bruit dans les décompositions sans altérer la qualité des images. La quantification des concentrations des agents de contraste est similaire à la méthode de l'état de l'art. La décomposition a aussi été réalisée dans le domaine image et comparé au domaine des projections. La décomposition dans le domaine image donne de meilleurs résultats que ceux dans le domaine des projections. La connexion résiduelle permet de garder un *a priori* sur les données d'entrées et donne de meilleurs résultats que sans celle-ci. (Zimmerman et al., 2020) propose un perceptron multicouches avec adaptation de domaine pour décomposer en 3 matériaux à partir de 4 bins d'énergie. Les entrées du réseau sont les mesures spectrales normalisées de manière logarithmique. La fonction de coût est spécialement conçue pour des petites bases de données. L'adaptation de domaine améliore la qualité des images mais ne permet pas d'améliorer la quantification. L'adaptation de domaine est réalisée à partir d'un entraînement sur des images étalonnées et d'un entraînement sur des simulations.

3.4 Conclusion

Dans ce chapitre nous avons introduit les problèmes inverses et leur résolution grâce à des méthodes analytiques ou itératives. Ces problèmes sont très présents en imagerie et peuvent être résolus grâce à l'apprentissage profond, que nous avons présenté, et plus précisément, les réseaux de neurones convolutifs. Il sont utilisés résoudre des problèmes de débruitage mais aussi de tomographie. Pour la tomographie spectrale, des travaux récents sur les algorithmes d'apprentissage profond pour la décomposition de matériaux ont été présentés dans le domaine image et dans le domaine des projections.

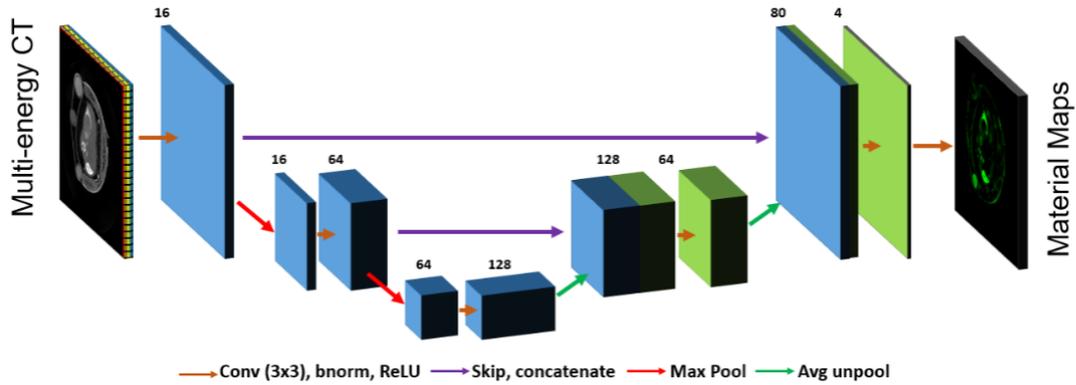


FIGURE 3.11 – Architecture inspirée de U-net développée par (Clark, Holbrook et Badea, 2018), avec seulement deux changements de résolution, pour la décomposition de matériaux dans le domaine image.

Pour apprendre correctement les problèmes, il faut également se focaliser les données, savoir s'il y en a assez pour résoudre le problème. Dans le cas contraire, il est possible d'augmenter les données. De plus, pour un meilleur apprentissage, il est également possible de normaliser les données. Les paramètres du réseaux doivent également être ajustés pour correspondre au mieux au problème à résoudre.

Deuxième partie

Contributions

Chapitre 4

Données

Dans ce travail, nous nous sommes focalisés sur l'application des méthodes d'apprentissage profond en tomographie spectrale dans le cadre de la détection précoce de l'arthrose. Dans ce but, nous avons pu acquérir différents jeux de données qui ont été utilisés tout au long de la thèse. Nous présenterons les échantillons qui ont été acquis avec plusieurs méthodes d'imagerie à rayons X. Nous introduirons ensuite les méthodes utilisées pour la création de fantômes numériques et réalistes du genou humain. Ces fantômes sont simulés avec différents paramétrages du modèle direct de la tomographie spectrale pour obtenir des données spectrales et des images mono-énergétiques. Nous finirons avec les différentes bases de données qui nous ont permis d'entraîner et de tester nos algorithmes d'apprentissage.

4.1 Introduction

Nous souhaiterions être capable de caractériser d'arthrose précoce à l'aide de la tomographie spectrale. L'arthrose se plaçant très souvent sur le genou, nous focalisons notre étude sur l'arthrose du genou. Il nous a donc fallu des données acquises en tomographie spectrale. De plus, comme énoncé au chapitre 2, l'imagerie synchrotron est la référence pour les images mono-énergétiques. Nous avons des données de genoux que nous avons acquises sur différents appareils, un scanner synchrotron à l'European Synchrotron Radiation Facility à Grenoble et un scanner spectral au CERMEP à Lyon.

4.2 Échantillons et acquisitions

Nous avons utilisé des pièces de genoux excisées d'une taille d'environ 15cm de long où se trouvent la rotule, la tête du fémur et la tête du tibia. Les tendons et ligaments ont été retirés. Certains ont la peau retirée et d'autre l'ont toujours. Ils sont conservés dans de la glace carbonique après avoir passé 20 jours dans une solution de conservation au formol. Les genoux sont préparés ainsi par l'Institut d'Anatomie Paris Descartes et fournis par le laboratoire B3OA (Biologie, Bioingénierie et Bioimagerie ostéo-articulaire) CNRS 7052, Université Paris Diderot. L'étude a été approuvée par le comité d'éthique de l'Université Paris Descartes. La Figure 4.1 montre un échantillon de genou sur la table du scanner spectral.

Sur 3 ans nous avons eu à disposition une vingtaine de genoux. Chaque échantillon est référencé par l'âge du donneur et son sexe et a été scanné en tomographie haute résolution (HR-pQCT) par le B3OA pour déterminer préalablement son niveau d'arthrose. La table 4.1 récapitule les échantillons disponibles et comment ils ont été acquis. Les échantillons présentent une grande variabilité. En effet, en visualisant cette table, les genoux présentent différents niveaux d'arthrose depuis des



FIGURE 4.1 – Photo d'un échantillon de genou, juste sorti de la glace et emballé sous vide, sur la table, prêt à être scanné par le scanner spectral.

sujets sains à des sujets avec une arthrose très avancée. De plus, la morphologie des genoux n'est pas non plus la même, il y a des personnes de forte stature et des personnes très minces. La taille et la minéralisation des os changent aussi d'un sujet à l'autre.

Les acquisitions des échantillons ont donc été réalisées en imagerie synchrotron et tomographie spectrale. 7 genoux ont été scannés en 2017 et 19 ont été scannés en 2018 ou 2019. Certains échantillons ont été imagés avec les deux appareils alors que d'autres ont été scannés avec uniquement l'un des deux. Dans le cas d'une acquisition avec les deux appareils, on préfère imager d'abord avec le scanner spectral et ensuite avec le scanner synchrotron car l'ionisation importante de celui-ci détériore les échantillons.

Tous les genoux n'ont pas été scannés à l'aide d'un scanner haute résolution au préalable. C'est pour cela que certains n'ont pas de valeur pour le niveau d'arthrose.

4.2.1 Tomographie synchrotron

La tomographie synchrotron a été réalisée à l'European Synchrotron Radiation Facility (ESRF) à Grenoble sur la ligne de lumière ID17 qui est aussi appelée ligne de lumière biomédicale. Le rayonnement synchrotron est un rayonnement émis quand une particule qui se déplace à vitesse quasi relativiste est déviée par un champ magnétique. A l'ESRF, les particules sont accélérées tout d'abord par un accélérateur linéaire, puis circulaire jusqu'à une énergie de 6 GeV puis injectées dans un anneau de stockage de 844 m de circonférence. Elles passent à travers différents éléments d'insertion et émettent alors un rayonnement synchrotron qui est redirigé vers des lignes de lumière. Sur la ligne de lumière, le faisceau synchrotron est alors mis en forme pour être utilisé comme une source de rayons X pour une acquisition tomographique. En particulier, on peut obtenir un faisceau monochromatique sur une large gamme d'énergie, de très haut flux comparé à une source standard de rayons X, et cohérent. La tomographie synchrotron a donc l'avantage de pouvoir fournir des images monochromatiques, à fort rapport signal sur bruit, jusqu'à des résolutions spatiales élevées (Salomé et al., 1999). Différentes lignes peuvent être utilisées pour faire de la tomographie synchrotron, notamment la ligne ID17 qui permet de

Id	Age	Sexe	OA	SRCT				SCT		Date
				55	75	100	Disp	Scan	Disp	
#62	91	F	/	X	0	X	0	0	0	2017
#124	84	F	/	0	0	0	0	X	0	2019
#196	94	M	0	X	X	X	X	0	0	2017
#274	99	M	/	0	0	0	0	X	0	2019
#234	90	M	2	X	X	X	X	0	0	2017
#296	82	F	/	0	0	0	0	0	0	2019
#313	85	F	4	X	X	X	X	X	0	2017
#314	90	M	0	X	X	X	X	X	X	2017
#386	56	F	/	0	0	0	0	X	0	2019
#391	96	F	/	0	0	0	0	X	0	2019
#405	88	M	/	X	0	0	X	0	0	2018
#410	89	F	/	0	0	0	0	X	0	2019
#415	104	F	4	0	0	0	0	X	0	2019
#422	58	M	1	X	X	X	X	X	0	2017
#433	91	F	/	X	0	X	X	X	0	2017
#434	93	F	1	X	0	0	0	X	0	2018
#435	87	F	/	0	0	0	0	X	0	2019
#437	77	M	/	0	0	0	0	X	X	2019
#438	89	M	3	X	X	X	X	X	0	2017
#440	92	F	0	X	0	X	X	0	0	2017
#448	90	M	1	X	X	X	X	0	0	2017
#450	94	F	1	X	X	X	X	X	0	2017
#505	91	F	/	X	0	X	X	X	0	2017
#514	76	F	/	X	0	0	X	X	0	2019
#581	63	M	2	X	X	X	X	X	0	2017

TABLE 4.1 – Description des échantillons acquis et modalités d’acquisition, les lignes en gras sont les genoux présent dans la BASE8, L’identifiant correspond à un identifiant unique du patient alloué lors de l’anonymisation des échantillon, l’âge et le sexe correspondent à l’âge du patient dont le genou est issu au moment de sa mort et son sexe, OA correspond au niveau d’arthrose, SRCT à l’acquisition synchrotron avec l’énergie annoncée en keV, SCT à l’acquisition avec le scanner spectral, Date est l’année d’acquisition.



FIGURE 4.2 – Tranche d’un genou scanné en imagerie synchrotron à trois énergies 55, 75 et 150 keV.

travailler sur une gamme d’énergie d’environ 30 à 130 keV et sur laquelle nous avons réalisé nos acquisitions (expériences *MD1045* - 2017 ; *MD1142* - 2019).

Les genoux ont été scannés à différentes énergies 55, 75 et 100 keV, avec une taille de pixel sur le détecteur de $46 \mu\text{m}$. Pour pouvoir avoir un champ de vue de 16 cm de diamètre, nous avons utilisé le mode « demi- acquisition » qui consiste à décaler le centre du détecteur. De plus, la hauteur du faisceau n’étant que de l’ordre de quelques millimètres (entre 7 et 10 mm), il a fallu réaliser typiquement 15 scans se chevauchant légèrement pour couvrir la région d’intérêt du genou (environ 8 cm scannés), incluant la tête du fémur, la rotule, la tête du tibia et la tête de la fibula. Pour chaque scan, nous avons donc acquis 2995 projections sur 360° d’angle de vue et un temps de pose de 0.15 s par vue. Le temps d’acquisition par scan était de 18 minutes avec un temps total d’acquisition par échantillon et par énergie, d’environ 4 heures.

La reconstruction tomographique a été réalisée avec les outils dédiés accessibles à l’ESRF, basé sur un algorithme de rétroprojection filtré. Il a ainsi été nécessaire de reconstruire chaque scan correspondant à un sous-volume, puis de recombinaison ceux-ci et éliminer les zones de chevauchement afin d’obtenir le volume entier. Les volumes reconstruits avec tous les sous volumes recombinaison sont stockés sur 32 bits pour 70 Go de mémoire par volume. Au total, nous avons pu acquérir 16 genoux avec l’imagerie synchrotron et seulement 8 ont été exploités car ce sont ceux qui étaient disponibles au début de la thèse. La Figure 4.2 illustre une tranche d’un genou scanné à 55 keV - Le noir dans l’image correspond à de l’air qui est entré dans l’échantillon pendant la décongélation lors de l’acquisition. La Figure 4.3 montre une visualisation 3D des os de deux patients (#422 et #438), un patient avec un très faible niveau d’arthrose (#422) et un avec un niveau plus avancé (#438).

4.2.2 Prototype de scanner spectral

Les genoux ont également été scannés avec le scanner spectral disponible au CERMEP, Lyon. Les 7 échantillons de 2017 ont été scannés avec la première version du prototype de scanner spectral de Philips et les 12 de 2019 ont été scannés avec le prototype clinique de Philips. Les deux prototypes fournissent des images 3D avec une taille cubique de voxel de $250 \mu\text{m}$.

Le prototype de scanner spectral utilisé en 2017 fournit un champ de vue de 500 mm. Il utilise un tube à rayons X de tension 80, 100 ou 120 kVp et d’intensité comprise entre 10 et 500 mA. La plage d’énergie du spectre émis est incluse entre 25 et 120 keV. Le détecteur à comptage de photons est composé de 2 mm de tellurure de zinc-cadmium, avec une distance entre les pixels de $270 \mu\text{m}$ dans les deux directions,

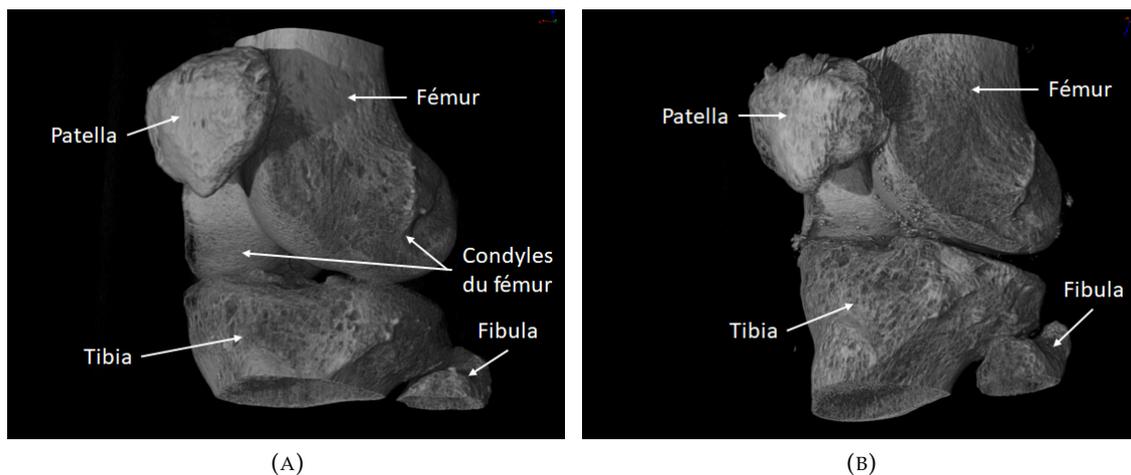


FIGURE 4.3 – Rendu 3D des os des genoux #422 (A) et #438 (B), où l'on voit le fémur, la patella, le tibia et la tête de la fibula.

et d'un circuit intégré spécialisé ChromAIX2. Les échantillons de 2017 ont été scannés en acquisition axiale et ceux de 2019 en acquisition hélicoïdale. Une rotation en acquisition axiale représente 2400 projections sur 360° et couvre 17.5 mm (Pourmorteza et al., 2016). Le décalage effectué lors d'un tour en géométrie hélicoïdale est de 18.1 mm, il y a le même nombre de projections par tour que l'acquisition axiale.

Les genoux ont été scannés avec le tube à rayons X à 120 kVp et 330 mA, et avec une acquisition axiale, en 2017, et à 120 kVp, 226 mA, avec une géométrie hélicoïdale, en 2019. Les données brutes prétraitées de seulement 1 genou de 2017 et 1 de 2019 ont pu être exportées. Il a fallu des démarches administratives avec Philips pour avoir accès aux données.

4.3 Génération des fantômes de genoux

Dans un premier temps, nous avons généré des fantômes numériques 3D de genoux à partir des acquisitions synchrotron que nous avons préalablement segmenté. Ces images 3D seront ensuite utilisées pour simuler leurs projections spectrales à l'aide de la modélisation du problème direct de la tomographie spectrale. Elles seront utilisées pour l'entraînement, la validation et le test de nos réseaux de neurones. Elles serviront aussi pour le test de l'algorithme de l'état de l'art.

Les données synchrotrons étant acquises à très haute résolution, on les sous-échantillonne d'un facteur 5 pour avoir la même taille de pixel qu'avec le scanner spectral. Le sous-échantillonnage est réalisé après un filtrage gaussien. Il résulte un effet de volume partiel que nous ne prenons pas en compte pour simplifier les fantômes.

4.3.1 Segmentation des images

Deux images synchrotron 3D de genoux de 2017 ont été segmentées grossièrement par seuillage pour séparer les tissus mous, l'os et le fond. Ensuite, le cartilage a été segmenté manuellement à partir de cette image des tissus mous à l'aide du logiciel (*Avizo Software | Avizo2D Software - FR*). La segmentation est représentée Figure 4.4. Pour les autres genoux, acquis en 2018 et 2019, seuls les tissus mous, l'os et le fond sont segmentés à l'aide de méthode de classification des K-moyennes (Lloyd,

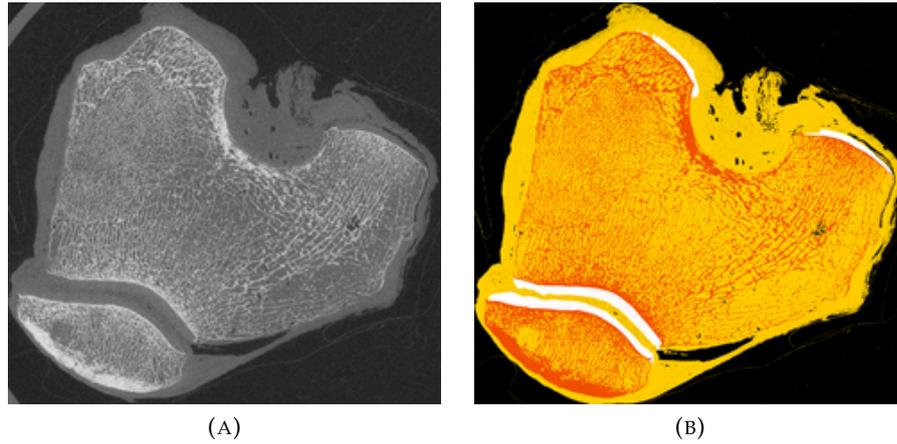


FIGURE 4.4 – Volume original (A) et sa segmentation (B) avec l'os en rouge, les tissus mous en jaune et le cartilage en blanc.

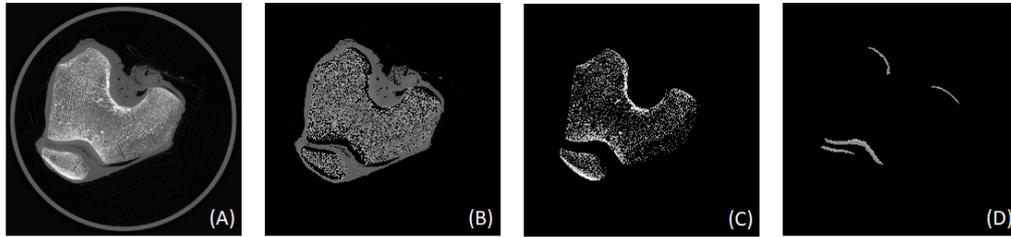


FIGURE 4.5 – Visualisation des cartes de matériaux avec du cartilage, des tissus mous et l'os avec (A) l'image du synchrotron, (B) la carte de segmentation des tissus mous, (C) de l'os et (D) du cartilage.

1982) avec trois graines. La Figure 4.5 représente les cartes de matériaux segmentés à l'aide de la méthode des K-moyennes.

Pour la simulation à basse résolution les volumes sont encore sous échantillonnés par 4 pour obtenir une taille de pixel de 1mm.

4.3.2 Changement d'unité

Les volumes synchrotron représentent l'atténuation linéaire à une énergie donnée $\mu(v, E)$ et on considère que l'on connaît le coefficient d'atténuation massique de chaque matériau $\tau_m(E)$. On suppose que les voxels ne sont constitués que d'un seul matériaux. En connaissant la localisation de chaque matériau grâce aux masques de segmentation, on peut définir la masse volumique par

$$\rho_m(v) = \frac{\mu_m(v, E)}{\tau_m(E)} \quad (4.1)$$

avec v le voxel sur lequel on se place, $\rho_m(v)$ la masse volumique dans le masque de segmentation du matériau m au voxel v et $\mu_m(v, E)$ l'atténuation linéaire dans le masque de segmentation correspondant au matériau m au voxel v du volume acquis à l'énergie E . Ainsi on peut retrouver masse volumique dans chaque masque de segmentation.

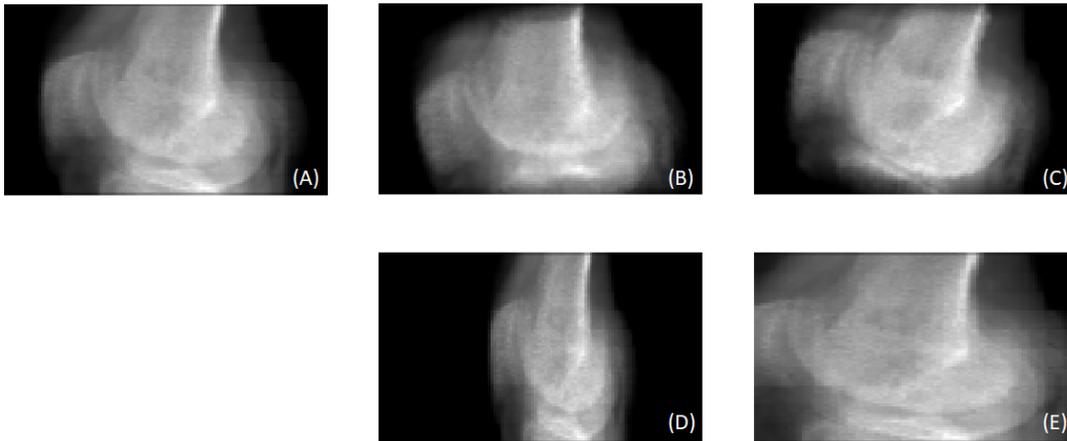


FIGURE 4.6 – Projection d’un volume augmenté. (A) Projection originale. (B) Rotation de -10° autour de (x, y) . (C) Rotation de 10° autour de (x, y) . (D) Changement d’échelle avec un facteur de 0.55. (E) Changement d’échelle avec un facteur de 1.45.

4.3.3 Augmentation géométrique des fantômes

Une partie des fantômes servira pour apprendre nos réseaux de neurones, les autres serviront pour la validation et d’autres pour le test de ces derniers.

Dans le but d’avoir une base de données assez grande et variée pour réaliser l’apprentissage du réseau et sachant que nous n’avons que 8 fantômes disponibles, on augmente chaque fantôme. Les modifications apportées font partie des modifications décrites dans le 3.2.2. Nous effectuons des rotations autour de l’axe formé par (x, y) tous les degrés de -10° à 10° . En effet, il est inutile de réaliser des rotations autour de l’axe z car c’est l’axe de rotation du scanner. Nous effectuons un changement d’échelle dans le plan (x, y) avec un facteur allant de 0.55 à 1.45 tous les 0.05, pour simuler des genoux plus ou moins gros. Il est inutile de réaliser le changement d’échelle dans le plan (x, z) ou (y, z) car cela ne ferait que faire sortir le genou du champ de vue en hauteur ou rajouter du noir en dessus ou en dessous. Ces changements d’échelles ne prennent pas en compte la réalité anatomique des genoux mais nous donne simplement de nouveaux fantômes. Nous réalisons ces modifications sur le volume 3D avant projection et mixage spectral. Une illustration d’une projection de ces volumes est fournie Figure 4.6.

4.4 Simulation des données spectrales

L’objectif est de simuler les mesures brutes acquises par un scanner X spectral décrit dans la sous-section 2.2.3 à partir des fantômes numériques de genoux. On simule les données spectrales du fantôme à partir du modèle direct du scanner spectral défini dans le 2.2.3 en choisissant deux paramétrages de modèle, celui de Schlomka (Schlomka et al., 2008) et celui de Philips (Si-Mohamed et al., 2017). Celui de Philips est fourni par le scanner avec les données brutes. Dans ces modèle le rayonnement diffusé n’est pas pris en compte bien qu’il n’y ait pas de dispositif anti-diffusion dans le prototype du scanner de Philips. Pour la projection tomographique on choisi de projeter suivant une géométrie parallèle sur 720 angles de 0.25° à 180° tous les 0.25° .

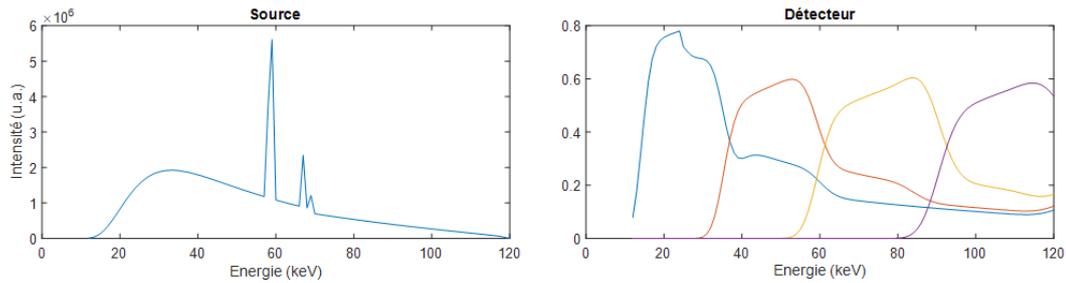


FIGURE 4.7 – Spectre de la source à rayons X avec $N_0 = 10^7$ photons et fonctions de réponse du détecteur à 4 bins d'énergie (source : (Schlomka et al., 2008)).

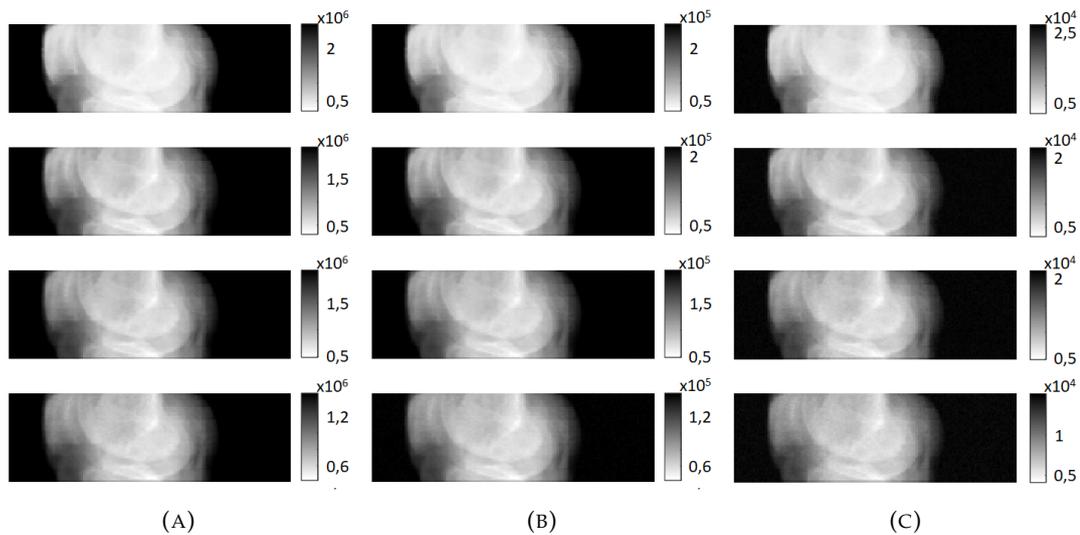


FIGURE 4.8 – Projections spectrales à 90° pour 3 niveaux de bruit $N_0 = 10^7$ (A), 10^6 (B) et 10^5 (C). De haut en bas : les 4 bins d'énergie.

4.4.1 Paramètres du modèle de Schlomka

Le modèle de Schlomka est défini dans (Schlomka et al., 2008) par la réponse du détecteur avec 4 bins d'énergie. Nous utilisons une source simulée avec logiciel présenté dans (Poludniowski et al., 2009) avec une tension du tube de 120 kVp, une anode orientée à 12° et un filtre en aluminium de 1.2 mm d'épaisseur. Le spectre de source et les fonctions de réponse du détecteur sont visibles sur la Figure 4.7.

On remplacera donc dans l'équation (2.6) $d_i(E)$ par les valeurs correspondant à ce détecteur et N_0 par le nombre de photons de la source. C'est un modèle simple car on considère que la source irradie de la même manière tous les pixels. Les données sont simulées pour trois niveaux de bruit. Les niveaux de bruit sont réalisés avec un changement du nombre de photons émis par la source. La Figure 4.8 montre les projections spectrales pour $N_0 = 10^7$, 10^6 et 10^5 photons produits par la sources. On voit que plus on diminue la quantité de photons émis par la source, plus il y a de bruit dans les projections. La vue à 90° donne une vision latérale du genou avec à gauche la rotule, les deux demi cercles vers le bas sont les condyles du fémur qui est l'os en haut et le plateau tibial en bas.

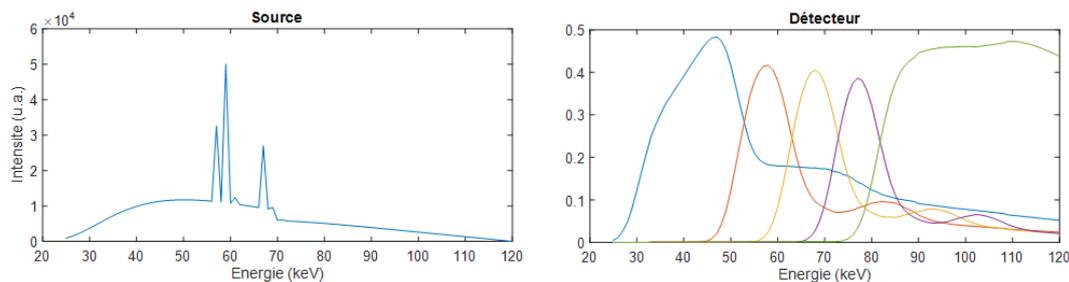


FIGURE 4.9 – Spectre de la source à rayons X et fonctions de réponse du détecteur du prototype clinique de Philips (Si-Mohamed et al., 2017).

4.4.2 Paramètres du modèle du prototype Philips

Le modèle de Philips présenté dans (Si-Mohamed et al., 2017) est caractérisé par une source uniforme dans la direction z mais pas le plan (x, y) et le détecteur est avec 5 bins d'énergie. La Figure 4.9 représente l'intensité de l'irradiation du détecteur sans objet au niveau du pixel central et la fonction de réponse du détecteur.

La Figure 4.10,(A) représente les données spectrales simulées avec le modèle de Philips. La Figure 4.10,(B) représente les données spectrales normalisées avec la mesure de plein flux (3.26).

Pour le travail dans le domaine image, il est nécessaire de reconstruire les bins d'énergie. On voit sur la Figure 4.11 les mesures spectrales normalisées et reconstruites pour les différents bins d'énergie.

4.5 Simulation des image mono-énergétiques

Pour la simulation des images mono-énergétiques à partir des cartes de matériaux décomposées et reconstruites, on reprend l'équation (2.1) et les valeurs de masse volumique qui ont été calculées pour chaque matériau dans 4.3.2. Ainsi, pour créer les images mono-énergétiques à partir des cartes de tissus mous et d'os, on utilise

$$\mu(v, E) = \rho_{\text{tissus mous}}(v)\tau_{\text{tissus mous}}(E) + \rho_{\text{os}}(v)\tau_{\text{os}}(E). \quad (4.2)$$

Nous choisissons de nous focaliser sur des images à 70 keV en s'appuyant sur les travaux de (Ozguner et al., 2018) qui ont montré que cette énergie était un bon compromis entre résolution spatiale et bruit. Ces résultats ont été confirmés par (Garcelon et al., 2020; Chappard et al., 2020) sur les données de genou. Dans la Figure 4.12, on montre aussi que la qualité d'image pour notre base concorde avec ces résultats. En effet, à 70 keV, il est possible de voir à la fois l'os et un bon contraste entre les tissus mous, en particulier du cartilage.

4.6 Bases de données pour l'apprentissage

En vue d'entraîner nos réseaux, nous devons créer des couples entrées/sorties. Nous considérons 4 réseaux qui seront détaillés dans le Chapitre 5, le premier réalise la décomposition de matériaux dans le domaine des projections et prend en entrée les mesures spectrales dans les bins d'énergie dans le domaine des projections et

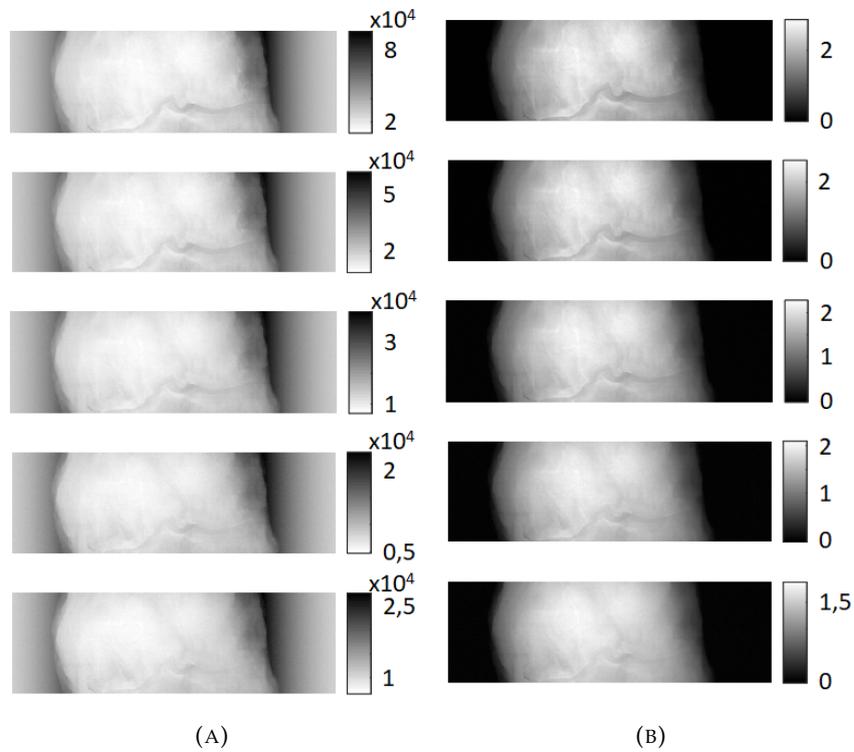


FIGURE 4.10 – Projections spectrales à 0° avec le modèle de Philips sans normalisation (A) et normalisation avec la mesure de plein flux (B). Chaque ligne est un bin.

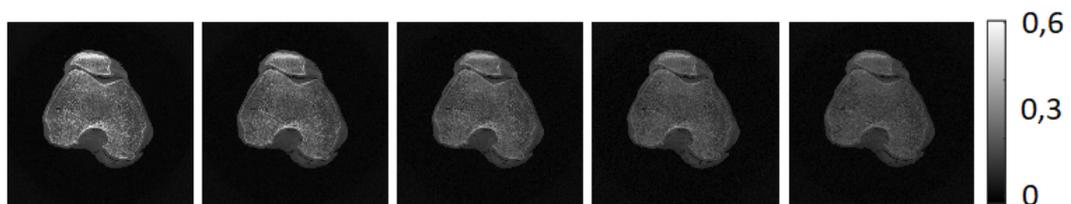


FIGURE 4.11 – Reconstruction tomographique des projections spectrales normalisées. Visualisée sur une coupe avec les bins de gauche à droite.

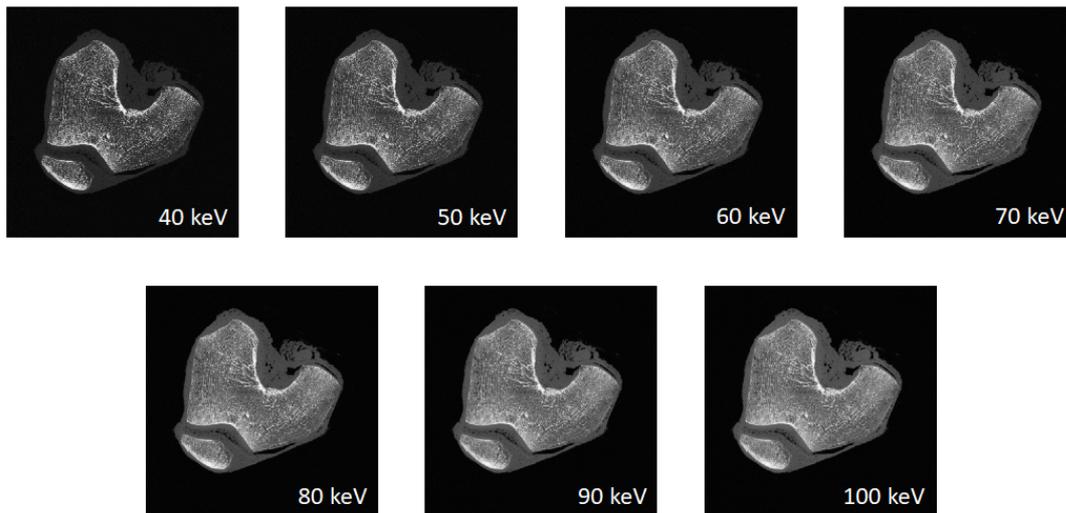


FIGURE 4.12 – Images mono-énergétiques du genou de 40 à 100 keV reconstruites à partir des données spectrales.

en sortie les cartes de matériaux dans le domaine des projections. Le deuxième décompose des projections mono-énergétiques et prend les mêmes entrées que le premier mais les sorties sont les projections mono-énergétiques. Le troisième réalise la décomposition de matériaux dans le domaine image. Il prend en entrée la reconstruction tomographique des mesures spectrales dans les bins d'énergie et en sortie la reconstruction tomographique des cartes de matériaux. Le quatrième réalise la reconstruction des images mono-énergétiques prend les mêmes entrées que le troisième mais les sorties sont les images mono-énergétiques.

4.6.1 Base de données réduite - BASE2LR

La base de données réduite est composée de deux genoux, ceux de 2017. Un genou augmenté avec les rotations de -10° à 10° et les changements d'échelle avec un facteur de 0.75 à 0.95, décrit dans le 4.3.3, représente la base d'entraînement, le même genou dans sa version originale représente la base de validation et un deuxième genou dans sa version originale, la base de test. Le genou pour la base de test est de validation est un genou avec un fort niveau d'arthrose. Le genou pour la base de test est un genou contrôle, c'est-à-dire sans arthrose. Pour des raisons de temps de calcul ; cette base de données est à basse résolution. Cela signifie que les fantômes avec une taille de voxel de $250 \mu\text{m}$ sont sous-échantillonnés par 4 pour avoir une taille de voxel de 1 mm. Les volumes sont de taille $161 \times 161 \times 51$ pixels. De plus, la version augmentée présente 25 fantômes. La base d'entraînement représente 25 fantômes, la base de validation 1 et la base de test 1.

4.6.2 Bases de données étendues - BASE8

Pour les bases de données complètes, on utilise les 8 genoux dont les fantômes ont été créés. Nous avons choisi d'utiliser 4 genoux pour l'apprentissage, 1 pour la validation et 3 pour le test. En effet, nous avons besoin du plus grand nombre d'échantillons pour apprendre, très peu pour contrôler si l'apprentissage se passe bien et un peu plus pour vérifier que le réseau pourra généraliser à de nouvelles données. Les stades d'arthroses dans les bases sont aléatoirement répartis.

Base basse résolution - BASE8LR Il existe une base basse résolution avec ces 8 genoux. Les 4 genoux pour l'entraînement sont dans leur version augmentée incluant le fantôme original. Ces augmentations sont réalisées avec un facteur d'échelle allant de 0.55 à 1.45 tous les 0.05 et des rotations de -10° à 10° tous les 1° . La base d'entraînement présente donc 164 fantômes. La base de validation présente 1 fantôme original et la base de test présente 3 fantômes originaux. A basse résolution, les fantômes sont constitués $161 \times 161 \times 51$ voxels de 1mm de coté.

Base haute résolution - BASE8HR Pour la base haute résolution, la taille des voxel est de $250 \mu\text{m}$ de coté. La taille du volume est de $635 \times 635 \times 202$ voxels. La base d'entraînement n'est plus augmentée.

4.6.3 Données réelles

On dispose également d'un genou scanné en 2018 avec le prototype clinique de scanner spectral de Philips. Les projections spectrales ont une taille de pixel de $250 \mu\text{m}$ de coté. C'est le genou #437 de la Table 4.1. Il est acquis avec une géométrie cône 3D en suivant une trajectoire hélicoïdale avec 2400 projections par tour.

4.7 Conclusion

Dans ce chapitre nous avons présenté les données utilisées dans la thèse pour mettre en oeuvre les méthodes d'apprentissage. Nous avons décrit les échantillons, les pièces de genoux humains excisés, que nous avons scanné à l'aide de l'imagerie synchrotron et d'un prototype de scanner spectral. Les fantômes sont ensuite générés à partir des données acquises en imagerie synchrotron. Les données spectrales sont ensuite simulées à l'aide de modèles prédéfinis. Les fantômes sont ensuite répartis dans des bases de données pour l'apprentissage des réseaux de neurones. Il existe aussi un jeu de données réelles pour tester les réseaux.

Chapitre 5

Méthodes de décomposition proposées

Dans ce chapitre, nous détaillerons les méthodes que nous proposons pour la reconstruction tomographique des données de tomographie spectrale. Nous chercherons soit à obtenir des cartes de matériaux, soit des images mono-énergétiques. Pour cela, nous proposons de développer des méthodes de décomposition par apprentissage profond, soit en travaillant dans le domaine des projections, soit dans le domaine image. Nous introduirons également les métriques qui nous permettront d'évaluer nos méthodes en comparaison à l'état de l'art, une méthode de décomposition des projections en matériaux.

5.1 Introduction

Nous réalisons la décomposition de matériaux et mono-énergétiques à l'aide d'algorithmes d'apprentissage. Les résultats de la décomposition de matériaux seront comparés avec les résultats obtenus avec la méthode de Gauss-Newton illustré sur la Figure 5.1,(A). Dans le cas de la reconstruction d'images mono-énergétiques, celles-ci seront reconstruites grâce au modèle de matériaux (2.1) représenté sur la Figure 5.3,(A) - droite. Nous travaillons dans différents domaines, le domaine des projections et le domaine image comme décrit sur la Figure 3.10 avec la reconstruction tomographique préalable des mesures spectrales par bin d'énergie pour le travail dans le domaine image.

Ainsi, nous considérons les 4 réseaux suivants, deux pour la décomposition de matériaux et deux pour la décomposition mono-énergétique :

- La décomposition de matériaux dans le domaine des projections (Figure 5.1,(B));
- La décomposition de matériaux dans le domaine image (Figure 5.1,(C));
- La décomposition des projections mono-énergétiques (Figure 5.2,(A));
- La reconstruction des images mono-énergétiques (Figure 5.2,(B)).

Nous apprenons donc avec 4 réseaux qui sont codés sous PyTorch. Les métriques sont calculées dans le domaine image, ainsi lors du travail dans le domaine des projection, on applique une reconstruction tomographique sur les résultats comme présenté sur la Figure 5.3,(A) - gauche et (B). Nous choisissons de réaliser la reconstruction des images mono-énergétiques, à partir des décompositions de matériaux, dans le domaine image car c'est une combinaison linéaire (équation (4.2)) et elle est commutative avec la reconstruction tomographique.

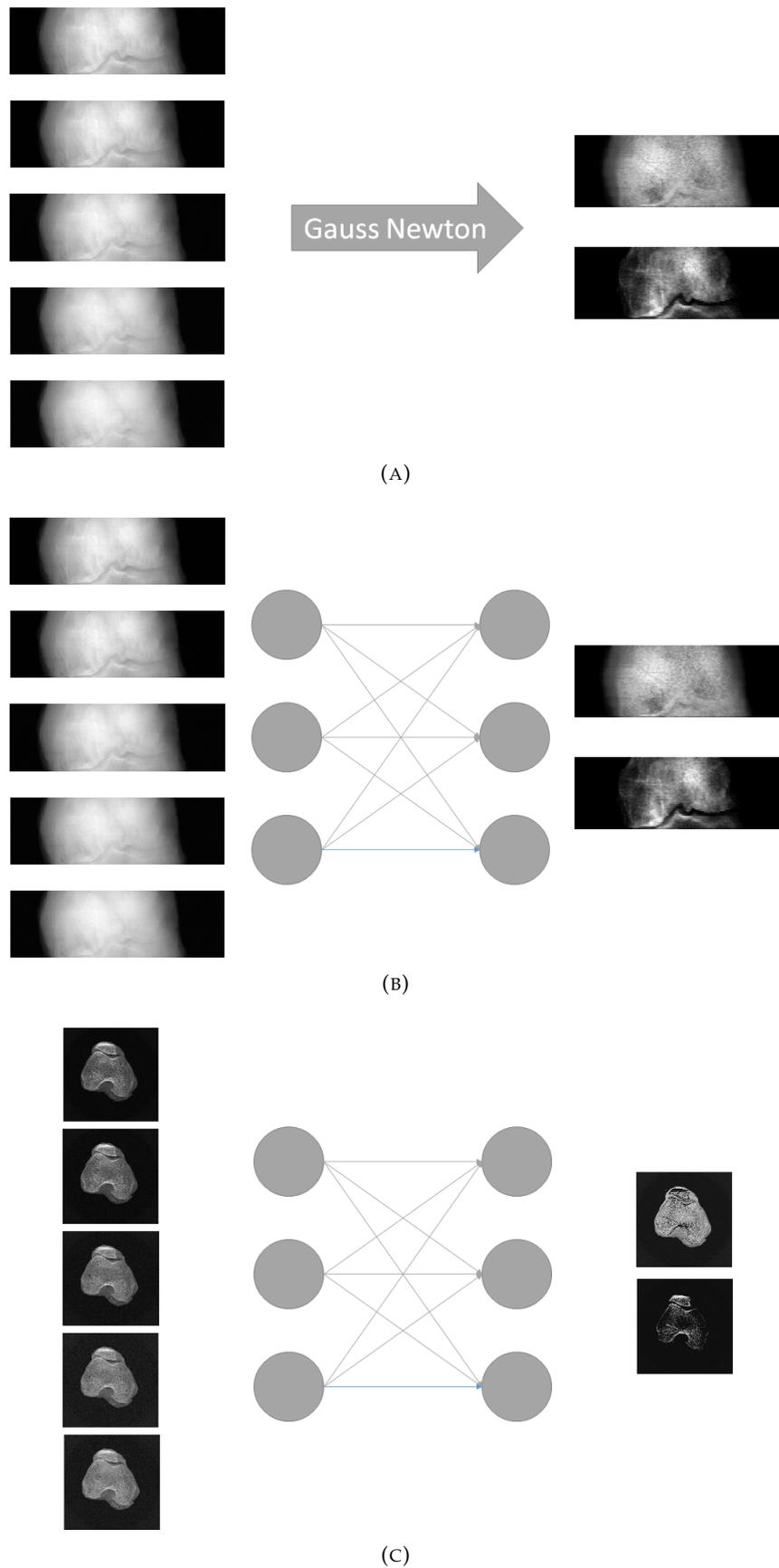
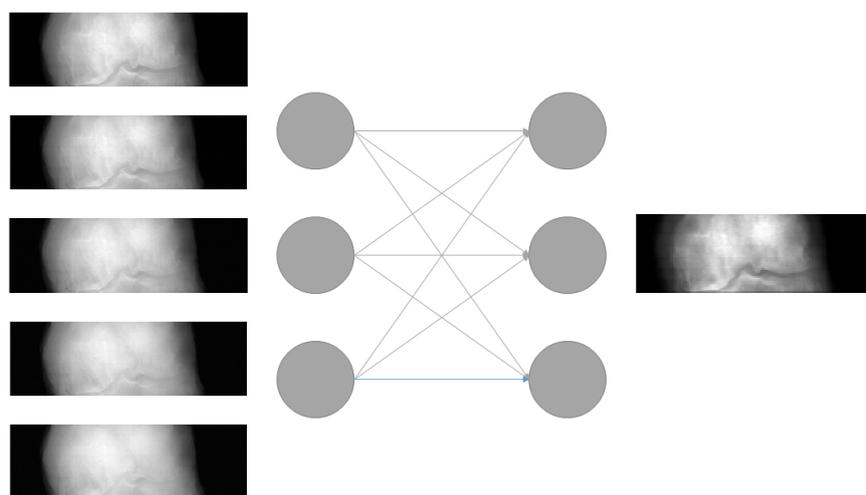
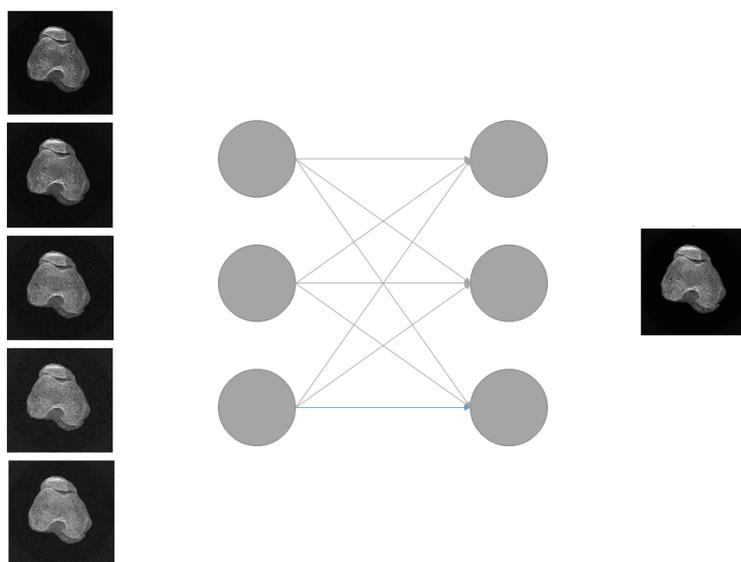


FIGURE 5.1 – Différentes stratégies de décomposition. Décomposition de matériaux dans le domaine des projections avec la méthode (A) de Gauss-Newton et (B) d'apprentissage. Décomposition de matériaux dans le domaine image avec la méthode d'apprentissage (C).



(A)



(B)

FIGURE 5.2 – Décomposition des projections mono-énergétiques avec la méthode d'apprentissage (A). Reconstruction des images mono-énergétiques avec la méthode d'apprentissage (B).

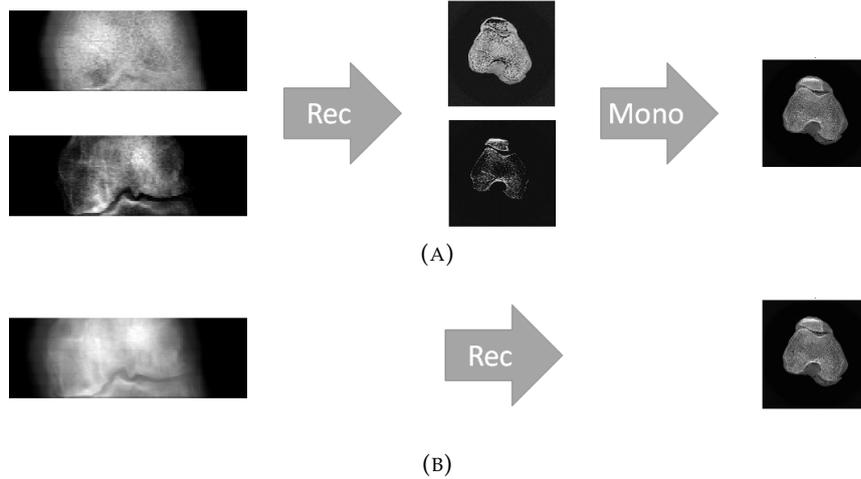


FIGURE 5.3 – Stratégies pour la reconstruction des images mono-énergétiques; à partir de la décomposition de matériaux dans le domaine de projections (A) ou le domaine image (A)-partie droite, ou à partir de projections mono-énergétiques (B)

5.2 Reconstruction tomographique

Pour traiter les données dans le domaine image ou pour reconstruire les données traitées dans le domaine des projections, nous avons besoin de méthodes de projection et de reconstruction tomographique.

5.2.1 Projections simulées

Pour simuler les projections, nous considérons une géométrie parallèle 2D. Pour cela, nous utiliserons la fonction *radon* de Matlab. Celle-ci implémente la transformée de Radon, qui est appliquée pour chaque coupe indépendamment. Ainsi pour reconstruire les données simulées, on utilisera la fonction *iradon* de Matlab avec un filtre rampe. Cette fonction implémente la rétroprojection filtrée avec le filtre rampe. Elle est appliquée pour chaque ligne de la projection 2D indépendamment pour reconstruire chaque coupe.

5.2.2 Projections expérimentales

Les données réelles sont acquises avec une géométrie conique 3D en suivant une trajectoire hélicoïdale avec un détecteur incurvé comme décrit dans le 4.2.2. Pour reconstruire ces données on utilise le logiciel *RTK* (Rit et al., 2014) et la méthode du gradient conjugué qui minimise la fonction de coût telle que :

$$\Phi(\rho_m) = \|X(\rho_m) - a_m\|_2^2 + \gamma \|\nabla \rho_m\|_2^2, \quad (5.1)$$

où γ est le paramètre de régularisation. On peut adapter cette formule pour la reconstruction tomographique des projections spectrales dans les bins d'énergie, ou pour la reconstruction tomographique des projections mono-énergétiques. Dans le cas de nos données expérimentales, nous avons choisi un paramètre de régularisation de 100 et 25 itérations comme dans (Abascal et al., 2021b). Nous gardons les mêmes valeurs de paramètres dans tous les cas.

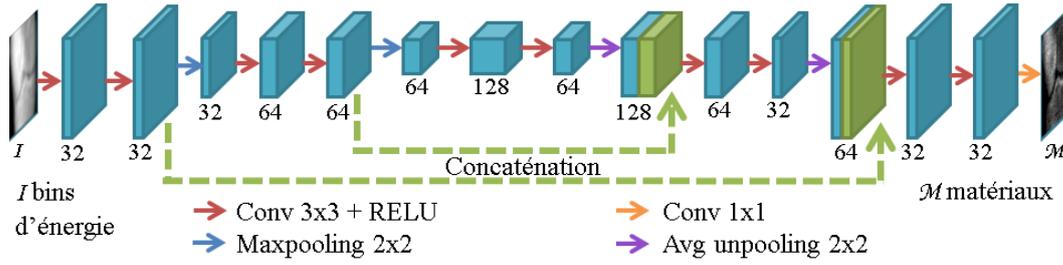


FIGURE 5.4 – Architecture U-net utilisée pour la décomposition de matériaux.

5.3 Décomposition de matériaux

La décomposition de matériaux cherche à inverser l'équation (2.5). La structure de la méthode d'apprentissage proposée est un réseau de neurones convolutifs basé sur une modification de l'architecture U-net. Nous précisons tout d'abord la méthode basée sur un modèle utilisée comme référence, puis décrivons les réseaux utilisés.

5.3.1 Méthode basée sur le modèle

La méthode basée sur le modèle que l'on utilise comme référence est la méthode Gauss-Newton régularisée (Ducros et al., 2017). Cette méthode est utilisée dans le domaine des projections. Elle utilise la fonction de coût suivante pour tenter d'inverser (2.5).

$$C(\mathbf{a}) = \frac{1}{2} \|\mathbf{s} - F(\mathbf{a})\|_W^2 + \alpha R(\mathbf{a}). \quad (5.2)$$

avec $F(\mathbf{a})$ défini par l'équation (2.6), W une matrice de pondération de la forme $W = \text{diag}(1/s)$ pour que le terme d'attache aux données soit adapté à des données corrompues avec un bruit poissonnien.

Cette fonction de coût est minimisée itérativement avec la méthode de Gauss-Newton à partir d'une estimation initiale. Nous choisissons les décompositions initiales comme étant deux images avec tous les pixels à 0.

La régularisation est réalisée spécifiquement pour chaque matériau en choisissant une régularisation Tikhonov d'ordre 1 pour l'os de la forme $\|\Delta \mathbf{a}_{\text{os}}\|_2^2$ et Tikhonov d'ordre 2 pour les tissus mous de la forme $\|\nabla \mathbf{a}_{\text{tissus}}\|_2^2$. Cette régularisation permet de débruiter. Nous choisissons l'ordre 1 pour l'os qui a tendance à avoir des changements brutaux et l'ordre 2 pour les tissus mous qui présentent des changements plus doux. En fonction du niveau de bruit, α est choisi à la valeur 0.03 pour décomposer avec une très faible régularisation, 0.3 pour décomposer avec une faible régularisation et 3 pour décomposer avec une forte régularisation.

5.3.2 Méthodes par apprentissage

Nous utilisons un réseau de neurones basé sur l'architecture U-net (Ronneberger, Fischer et Brox, 2015), en reprenant celle développée par (Clark, Holbrook et Badae, 2018) représentée Figure 5.4. Ce réseau a montré de bons résultats pour la décomposition de matériaux dans le domaine image avec du produit de contraste (Abascal et al., 2021b). On utilise la même architecture que l'on applique et dans le domaine des projections et le domaine image.

Domaine des projections Dans le domaine des projections (Bussod et al., 2021), on apprend à décomposer les deux cartes de matériaux projetées à partir des bins des acquisitions spectrales normalisées en suivant

$$h_{\omega} : \hat{s}^{\theta} \mapsto \mathbf{a}^{\theta} \quad (5.3)$$

où \hat{s} sont les acquisitions normalisées telles que $\hat{s}^{\theta} = \ln(\bar{s}^{\theta}/s^{\theta})$ où \bar{s}^{θ} est l'acquisition sans objet soit $\bar{s}^{\theta} = F(0)$ et ω sont les paramètres du réseau h .

En donnant N paires d'entrées/sorties d'apprentissage $\{\hat{s}^{(n)}, \mathbf{a}^{(n)}\}$, $1 \leq n \leq N$, on minimise une fonction de coût des moindres carrés définie par

$$L(\omega) = \frac{1}{N} \sum_{n=1}^N \frac{1}{2} \sum_{m=1}^2 \|h_{\omega}(\hat{s}_m^{(n)}) - \gamma_m \mathbf{a}_m^{(n)}\|^2. \quad (5.4)$$

Dans ce cas, les entrées sont les projections spectrales normalisées et les sorties sont les cartes de matériaux projetées. γ_m est une valeur de normalisation de la sortie pour le matériau m . Quand il est à la valeur 1 le matériau en sortie n'est pas normalisé.

Domaine image Dans le domaine image, on apprend à décomposer directement les cartes de matériaux à partir des reconstructions des images des différents bins d'énergie normalisées, \hat{S} , en suivant

$$h_{\omega} : \hat{S} \mapsto \rho \quad (5.5)$$

où ω sont les poids du réseau h .

En donnant N paires d'entrées/sorties d'apprentissage $\{\hat{S}^{(n)}, \rho^{(n)}\}$, $1 \leq n \leq N$, on minimise une fonction de coût des moindres carrés telle que

$$L(\omega) = \frac{1}{N} \sum_{n=1}^N \frac{1}{2} \sum_{m=1}^2 \|h_{\omega}(\hat{S}_m^{(n)}) - \gamma_m \rho_m^{(n)}\|^2. \quad (5.6)$$

Dans ce cas, les entrées sont les images reconstruites des différents bins d'énergie normalisées et les sorties sont les cartes de matériaux reconstruites.

5.4 Reconstruction des images mono-énergétiques

On cherche également à reconstruire les images mono-énergétiques, soit directement avec une méthode d'apprentissage, soit à partir des décompositions de matériaux obtenues avec la méthode de Gauss-Newton ou la méthode d'apprentissage. Comme précisé dans le Chapitre 4, nous avons choisi de créer un seul volume mono-énergétique à 70 keV mais la méthode peut, bien entendu, être adaptée à n'importe quelle valeur d'énergie.

5.4.1 A partir des décomposition en matériaux

En reprenant le modèle de matériaux décrit dans l'équation (4.2), il est possible de reconstruire l'image mono-énergétique à partir de la décomposition en tissus mous et os, qu'elle soit réalisée avec la méthode de Gauss-Newton ou la méthode d'apprentissage. Il est possible de combiner les matériaux avec les cartes de matériaux dans le domaine des projections ou dans le domaine image. Nous préférons le réaliser dans le domaine image car la combinaison linéaire du modèle de matériaux est commutative avec la reconstruction tomographique.

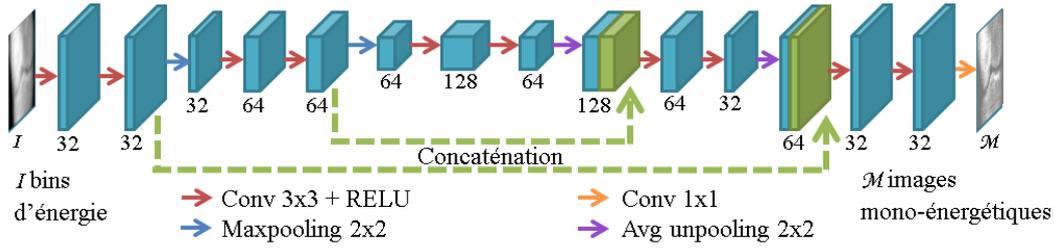


FIGURE 5.5 – Architecture U-net utilisée pour la reconstruction de projections ou d'images mono-énergétiques.

5.4.2 Méthodes directes par apprentissage

Nous proposons également d'obtenir les images mono-énergétiques par apprentissage. Nous avons gardé la même architecture que précédemment, avec des sorties différentes et une phase d'apprentissage adaptée. Comme pour la décomposition de matériaux, il est possible de réaliser la reconstruction des images mono-énergétiques dans le domaine des projections et dans le domaine image.

Domaine des projections Pour la décomposition des projections mono-énergétiques avec la méthode d'apprentissage, on suit l'équation :

$$h_{\omega} : \hat{s}^{\theta} \mapsto \mathbf{v}^{\theta} \quad (5.7)$$

où \hat{s} sont les acquisitions normalisées et ω sont les paramètres du réseau h . \mathbf{v} représente les projections des images mono-énergétiques à 70 keV.

En donnant N paires d'entrées/sorties d'apprentissage $\{\hat{s}^{(n)}, \mathbf{v}^{(n)}\}$, $1 \leq n \leq N$, on minimise une fonction de coût des moindres carrés régularisée telle que

$$L(\omega) = \frac{1}{N} \sum_{n=1}^N \|\mathbf{h}_{\omega}(\hat{s}^{(n)}) - \mathbf{v}^{(n)}\|^2. \quad (5.8)$$

Dans ce cas, les entrées sont les projections spectrales normalisées et les sorties sont les projections mono-énergétiques à 70 keV.

Domaine image Dans le domaine image, on suit la formule

$$h_{\omega} : \hat{S} \mapsto \boldsymbol{\mu} \quad (5.9)$$

où \hat{S} sont les reconstructions des bins d'énergie normalisées, ω sont les paramètres du réseau h et $\boldsymbol{\mu}$ est l'image mono-énergétique à 70 keV.

En donnant N paires d'entrées/sorties d'apprentissage $\{\hat{S}^{(n)}, \boldsymbol{\mu}^{(n)}\}$, $1 \leq n \leq N$, on minimise une fonction de coût des moindres carrés telle que

$$L(\omega) = \frac{1}{N} \sum_{n=1}^N \|\mathbf{h}_{\omega}(\hat{S}^{(n)}) - \boldsymbol{\mu}^{(n)}\|^2. \quad (5.10)$$

Dans ce cas, les entrées sont les images spectrales normalisées et les sorties sont les images mono-énergétiques.

La Figure 5.5 représente l'architecture pour la reconstruction des projections ou des images mono-énergétiques.

5.5 Phase d'apprentissage

Lors de la phase d'apprentissage, les paramètres du réseau seront estimés en minimisant les fonctions de coût décrites dans les deux sections précédentes. Celles-ci représentent l'erreur quadratique moyenne avec ou non une pondération sur la sortie pour la décomposition de matériaux et sans pondération pour la reconstruction des projections / images mono-énergétiques. La pondération est utilisée uniquement pour l'apprentissage sur les bases basse résolution.

Pour minimiser les fonctions de coût, on utilisera l'optimiseur ADAM présenté au 3.2.2 avec les paramètres par défaut de Pytorch. On utilise un taux d'apprentissage fixe ou qui décroît au fil du temps. Quand le taux d'apprentissage décroît, on le multiplie par 0.3 toutes les époques multiples de 30. Le taux d'apprentissage ne décroît que pour l'étude sur la base haute résolution.

Dans chaque partie des résultats, on précisera la taille de lot qui est de 15 pour la basse résolution et 45 pour la haute résolution. Les données sont mélangées automatiquement avant d'être mises en lot grâce aux fonctionnalités de l'opérateur de chargement des données (`torch.utils.data.DataLoader`).

Pour l'étude avec la base haute résolution, on utilise une approche par patch en utilisant la fonction `RandomCrop` de la bibliothèque TorchVision qui rend les coordonnées de départ de la découpe, sa hauteur et sa largeur. On considère la hauteur et la largeur fixées car on veut des patches 64×64 . On découpe les images dans chaque canal d'énergie et de matériau indépendamment avec ces coordonnées. Chaque projection ou coupe est découpée différemment, même au sein du même lot. A chaque époque, la découpe change également.

Les apprentissages en basse résolution ont été réalisés en local avec une NVidia RTX 2080 Ti et ceux en haute résolution sur le cluster du laboratoire CREATIS avec une NVidia Tesla V100.

Pour les trois chapitres suivants, les résultats présentés, sauf les fonctions de coût, sont le test des réseaux de neurones avec la base de test inconnue du réseau.

5.6 Métriques d'évaluation

Pour évaluer la qualité des reconstructions en matériaux ou mono-énergétiques on utilise l'erreur quadratique moyenne, l'indice de similarité et parfois une mesure du bruit.

5.6.1 Erreur quadratique moyenne

On nomme n le nombre de pixels dans chaque image A et B alors l'erreur quadratique moyenne est calculée telle que

$$MSE(A, B) = \frac{1}{n} \sum_{i=1}^n (A_i - B_i)^2 \quad (5.11)$$

avec A_i la valeur du pixel i de l'image A , de même pour B . Nous calculons l'erreur quadratique moyenne à l'aide la fonction `IMMSE` de Matlab.

Domaine des projections Dans le domaine des projections, l'erreur quadratique moyenne est calculée pour chaque matériau indépendamment ou pour chaque énergie indépendamment et chaque projection.

Pour la décomposition de matériaux, on remplace A et B de l'équation (5.11) par $a_{th_m}^\theta$ la masse volumique projetée de la vérité terrain pour le matériau m et à l'angle θ et $a_{th_m}^\theta$ la masse volumique projetée obtenue pour le matériau m et à l'angle θ . n est le nombre de pixels d'une projection, soit $161 \times 51 = 8'211$ en basse résolution et $693 \times 202 = 139'986$ pour la haute résolution.

Pour la décomposition de projections mono-énergétiques, on remplace A et B de l'équation (5.11) par v_{th}^θ la projection mono-énergétique pour l'angle θ de la vérité terrain et v^θ la projection mono-énergétique obtenue pour l'angle θ .

Domaine image Dans le domaine image, l'erreur quadratique moyenne est calculée pour chaque matériau indépendamment ou pour chaque énergie indépendamment entre chaque coupe.

Pour la décomposition de matériaux, on remplace A et B de l'équation (5.11) par $\rho_t h_m^z$ la masse volumique de la coupe z pour le matériau m de la vérité terrain et ρ_m^z la masse volumique obtenue pour la coupe z et le matériau m .

Pour la reconstruction d'images mono-énergétiques, on remplace A et B de l'équation (5.11) par μ_{th}^z l'atténuation linéaire pour la coupe z de la vérité terrain et μ^z l'atténuation linéaire obtenue pour la coupe z .

5.6.2 Indice de similarité structurel

L'indice de similarité peut être calculé pour deux images A et B tel que

$$SSIM(A, B) = \frac{(2m_A m_B + c_1)(2\sigma_{AB} + c_2)}{(m_A^2 + m_B^2 + c_1)(\sigma_A^2 + \sigma_B^2 + c_2)} \quad (5.12)$$

avec m_A la moyenne des valeurs de pixel de A , de même pour B , σ_A l'écart type des valeurs de pixels de A , de même pour B et σ_{AB} la covariance entre A et B . Nous calculons l'indice de similarité à l'aide de la fonction SSIM de Matlab (Wang et al., 2004) avec les valeurs par défaut des constantes qui sont $c_1 = 10^{-4}$ et $c_2 = 9 \cdot 10^{-4}$. Grâce à cette fonction il est possible de le calculer sur un volume en 3 dimensions. Ainsi, dans le domaine des projections, on le calcule sur toutes les projections. Dans le domaine image, on le calcule sur toutes les coupes, tout le volume.

5.6.3 Évaluation du bruit

L'évaluation du bruit est réalisée sur les images reconstruites, soit après reconstruction tomographique, soit obtenues directement dans le domaine image. Pour évaluer le bruit, on sélectionne une zone homogène dans les tissus mous et on calcule l'écart-type à l'intérieur de cette zone. Cela est réalisé dans 3 zones, de forme carrée, différentes sur chaque tranche pour plus de précision. Ainsi le calcul du bruit est calculé tel que

$$b = \sum_{i=1}^3 \frac{1}{3} \sigma_{ROI_i} \quad (5.13)$$

avec σ_{ROI_i} l'écart-type de la région d'intérêt i .

5.7 Conclusion

Dans ce chapitre nous avons introduit les méthodes de décomposition proposées dans la thèse. Au préalable, nous avons introduit les méthodes de reconstruction tomographique pour les données simulées et expérimentales qui ont une géométrie plus complexe. Ces méthodes de reconstruction nous permettent de reconstruire les résultats après avoir travaillé dans le domaine des projections ou directement travailler dans le domaine image. Nous avons présenté la méthode de Gauss-Newton régularisée qui nous sert de référence pour la décomposition de matériaux dans le domaine des projections. Nous avons présenté ensuite les méthodes d'apprentissage proposées basées sur l'architecture U-net pour apprendre la décomposition de matériaux dans le domaine des projections mais aussi dans le domaine image. Si c'est la reconstruction tomographique qui prend le plus de temps, alors si l'on doit effectuer beaucoup de tests, le domaine image sera plus avantageux pour le temps de calcul. Si l'on a qu'un seul test, ça sera le domaine des projections car reconstruire 2 canaux est plus rapide qu'en reconstruire 5. Pour les simulations, la reconstruction prend environ 46s par canal mais pour les données réelles cela prend environ 5h30 donc travailler dans le domaine image est avantageux. Nous avons décrit ensuite la reconstruction des images mono-énergétiques à partir de la combinaison linéaire des décompositions et avec la méthode d'apprentissage. Finalement, nous avons introduit les métriques que nous utilisons pour comparer les méthodes d'apprentissage à la méthode de Gauss-Newton.

Chapitre 6

Apprentissage de la décomposition dans le domaine des projections : étude sur des simulations

Dans ce chapitre nous présenterons la décomposition de matériaux ou mono-énergétiques dans le domaine des projections. Tout d'abord, nous travaillons en basse résolution. Nous opérerons la décomposition de matériaux sur une base très réduite pour tester la capacité de décomposition du réseau. Puis nous diversifions la base pour montrer que augmenter le nombre de données en entraînement augmente la précision du réseau. Enfin, nous passons en haute résolution pour palier le manque de détails dans les images basse résolution et nous introduisons ainsi l'approche par patch.

6.1 Introduction

Nous entraînons et testons les réseaux présentés dans le chapitre précédent. Pour cela, nous utilisons les bases présentées au Chapitre 4. Tous les résultats présentés, sauf les fonctions de coût, sont sur la base de test, sauf pour la BASE2LR.

Pour les bases basse résolution, on calcule les métriques dans le domaine des projections et dans le domaine image. Pour la base haute résolution, on calcule les métrique sur les reconstructions tomographiques des résultats.

6.2 Fantômes basse résolution

Les projections des fantômes basse résolution ont une taille de pixel de 1 mm de coté et chaque projection a une taille de 161×51 pixels.

6.2.1 BASE2LR

Cette base de données, décrite dans le 4.6.1, est une base de données construite à partir de seulement deux genoux mais augmentée pour obtenir 25 fantômes. Les projections spectrales ont été simulées, à partir des projections des matériaux du fantôme, uniquement avec les paramètres du modèle de Schlomka (4.7). Comme énoncé dans 4.4.1, les projections spectrales sont simulées pour 3 niveaux de bruit (Figure 4.8). Pour vérifier la performance de nos algorithmes, on choisit dans un premier temps des simulations avec très peu de bruit, avec $N_0 = 10^7$ photons. Ce nombre de photon nous permettra de nous placer dans un cas très simple.

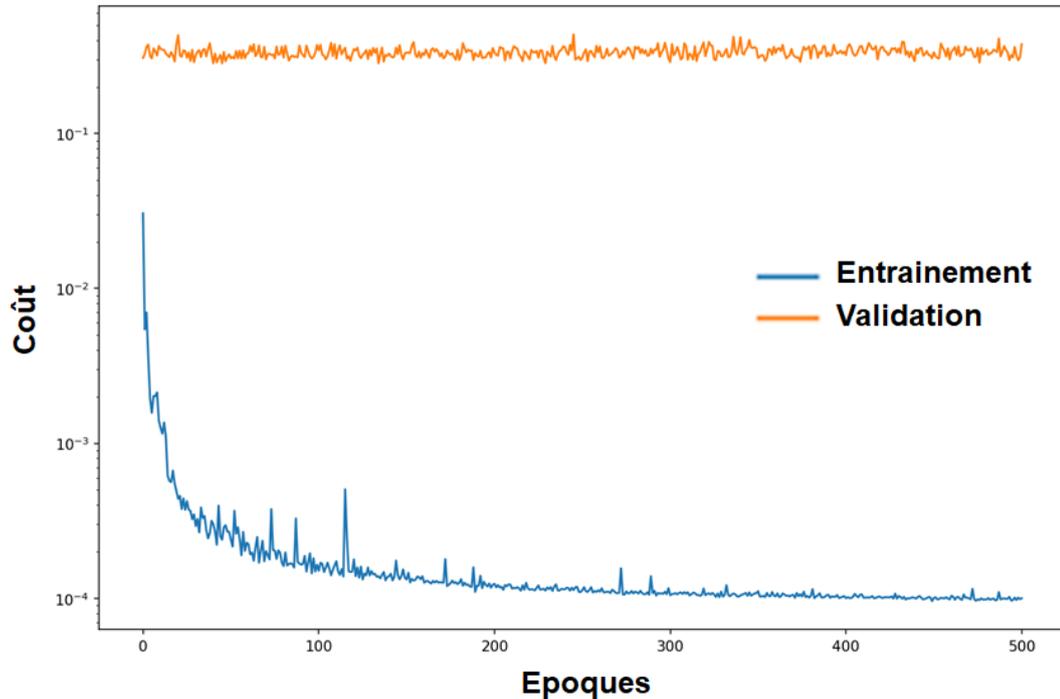


FIGURE 6.1 – Fonctions de coût d'apprentissage (train-bleu) et de validation (valid-orange) sur 500 époques pour une simulation de la BASE2LR avec 10^7 photons.

Décomposition de matériaux La taille de lot est de 15 projections de 4 bins pour les entrées et 2 matériaux, tissus mous et os, pour les sorties, 1200 itérations sont donc nécessaires pour réaliser une époque. En effet, on a $25 \times 720 = 18'000$ projections en considérant les 25 fantômes. Le taux d'apprentissage est de 10^{-3} et est fixe, la fonction de coût est l'erreur quadratique moyenne et l'optimiseur utilisé est ADAM (3.2.2). Nous apprenons sur 500 époques, bien qu'un comportement se détache dès les 100 premières époques. En effet, en regardant sur la Figure 6.1 l'évolution des fonctions de coût d'entraînement et de validation, on remarque que le réseau semble apprendre correctement sur la base d'entraînement mais ne peut pas généraliser à la base de validation. La fonction de coût d'apprentissage décroît tout au long des 500 époques mais celle de validation reste coincée à la valeur de départ. Plusieurs hypothèses sont considérées. La base d'entraînement étant basée sur un seul genou, elle n'est sûrement pas assez variée pour réussir à généraliser. La base de validation étant également sur un seul genou, elle est sûrement trop éloignée de la base d'apprentissage. Cette hypothèse semble peu probable car la base d'entraînement est une version augmentée de la base de validation. Pour vérifier le bon déroulement de l'entraînement, on peut visualiser la décomposition sur le genou de la base d'entraînement Figure 6.2. On compare les résultats obtenus avec U-net avec ceux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ qui correspond au meilleur paramètre de régularisation. On peut voir que les résultats de l'apprentissage sont similaires à ceux obtenus avec la méthode de Gauss-Newton.

La Figure 6.3 montre les images reconstruites à partir des projections décomposées pour un fantôme de la base d'entraînement. Comme pour les projections, les images issues des projections décomposées avec U-net sont de qualité similaire à celles issues des projections décomposées avec la méthode de Gauss-Newton. Dans

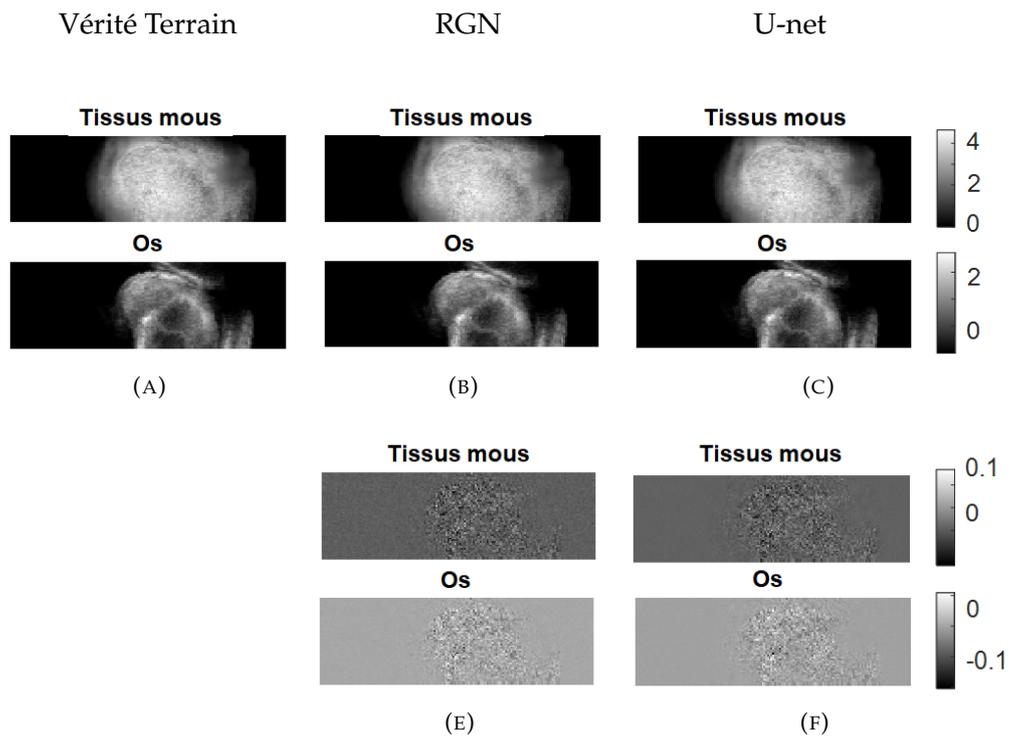


FIGURE 6.2 – Résultats de la décomposition de matériaux dans le domaine des projections avec U-net sur la base d'apprentissage de la base BASE2LR, exposés pour la projection à 90° . (A) Vérité terrain pour la décomposition. (B) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B). (C) **Matériaux décomposés avec U-net.** (E) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$. (F) **Erreur entre les matériaux décomposés avec U-net et la vérité terrain.**

TABLE 6.1 – Résultats quantitatifs pour la reconstruction des images mono-énergétiques à partir de la décomposition de matériaux, à 70 keV pour une coupe, avec U-net et comparaison à la méthode de Gauss-Newton.

	MSE	SSIM
RGN $\alpha = 0.03$	0.0411	0.9999
U-net	0.0650	0.9999

le domaine image, on se rend compte de l’influence de la taille de pixel sur la qualité des images car celles-ci sont très grossières.

Reconstruction des images mono-énergétiques grâce au modèle de matériaux À ce stade du développement de notre algorithme d’apprentissage profond, nous ne reconstruisons pas directement les images mono-énergétiques avec le réseau de neurones mais à partir de la décomposition de matériaux. On utilise ainsi le modèle de matériaux présenté avec l’équation (2.1). En regardant les résultats quantitatifs dans la Table 6.1, on remarque que les résultats sur les données d’apprentissage sont toujours moins bon que les résultats avec la méthode de Gauss-Newton en terme d’erreur quadratique moyenne mais sont similaires en terme d’indice de similarité. Sur la Figure 6.4 on peut bien visualiser que la qualité d’image est similaire pour la méthode de Gauss-Newton et U-net mais les erreurs sont plus fortes pour U-net et surtout au niveau du bord du genou.

6.2.2 BASE8LR

Pour palier le problème de manque de données rencontré avec la BASE2LR, on choisit de prendre plus de fantômes. Ainsi, la base de données choisie est une base de données avec 8 genoux dans leur forme augmentée. La description de cette base est réalisée dans le 4.6.2. L’entraînement est réalisé avec la base de données augmentée de 164 fantômes. On simule cette base avec les paramètres du modèle de Schlomka (4.7) pour un nombre de photons de la source égal à $N_0 = 10^5$. Nous réduisons le nombre de photons par rapport à l’étude précédente pour se rapprocher au plus du modèle d’un scanner réel. Les résultats suivants ont été présentés à *EUSIPCO 2020* (Bussod et al., 2021).

Décomposition de matériaux La taille de lot est de 15 projections de 4 bins pour les entrées et 2 matériaux pour les sorties, 7872 itérations sont nécessaires pour réaliser une époque. Le taux d’apprentissage est à 10^{-3} qui est fixe, la fonction de coût est l’erreur quadratique moyenne pondérée. On pondère la sortie en plaçant la valeur γ_m de l’équation (5.4) à $1/8$ pour les tissus mous et $1/3$ pour l’os. Les valeurs 8 et 3 correspondent à la valeur maximale moyenne dans les projections de tissus mous et d’os. L’optimiseur est ADAM (3.2.2). Nous apprenons sur 200 époques car un certain comportement sur la fonction de coût d’entraînement semble se détacher avec ce nombre d’époques.

En regardant sur la Figure 6.5 la fonction de coût d’entraînement et de validation, on se rend compte que le réseau est capable d’apprendre mais on remarque que nous avons des difficultés à généraliser sur la base de validation. En effet, la fonction de coût de validation a un comportement assez anarchique. On teste le réseau sur la base de test, soit des genoux inconnus du réseau. On compare ensuite les résultats à

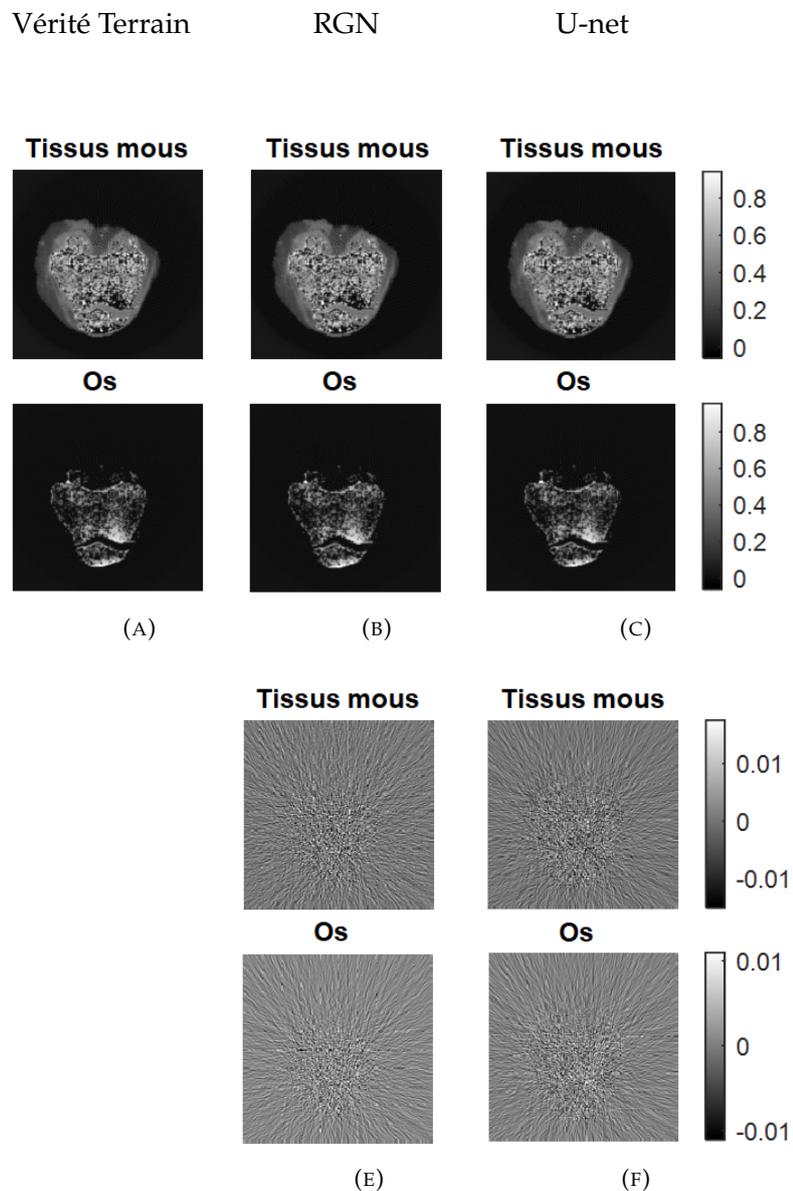


FIGURE 6.3 – Reconstruction tomographique des résultats de la décomposition de matériaux dans le domaine des projections avec U-net sur la base d'apprentissage de la BASE2LR. (A) Vérité terrain pour la décomposition. (B) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B). (C) **Matériaux décomposés avec U-net.** (E) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisé et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$. (F) **Erreur entre les matériaux décomposés avec U-net et la vérité terrain.**

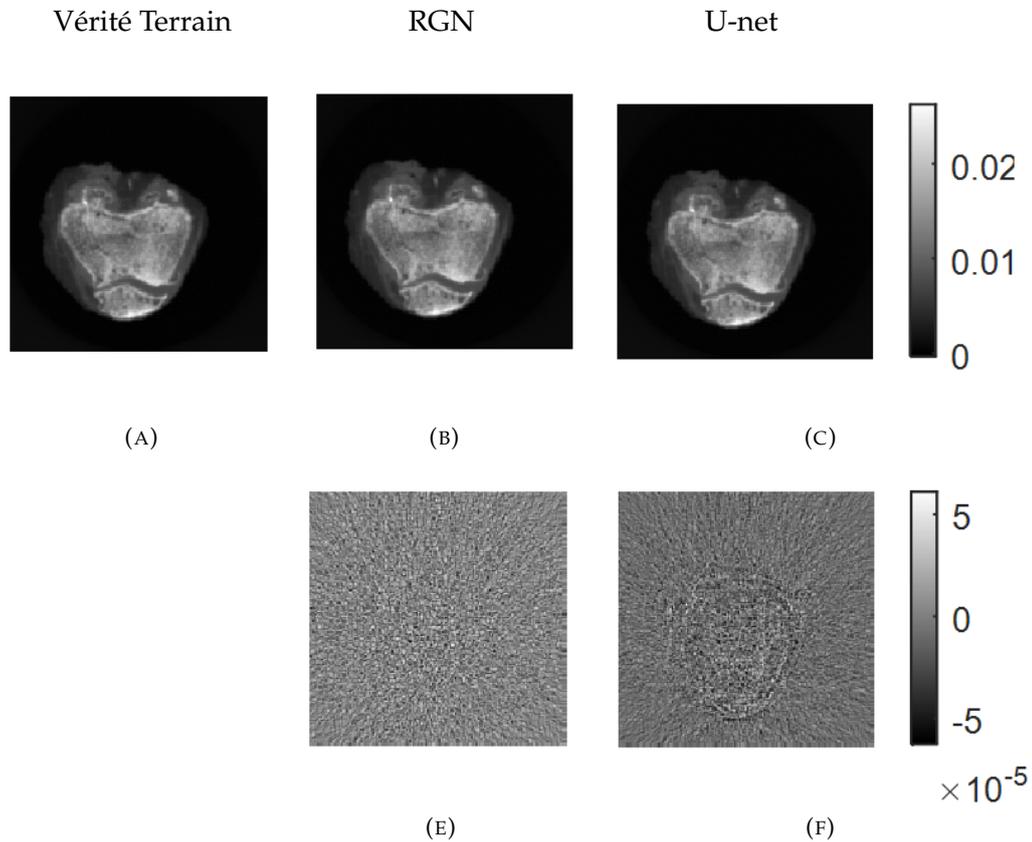


FIGURE 6.4 – Images mono-énergétiques à 70 keV obtenues à partir de la reconstruction tomographique des projections décomposées avec U-net sur la base d’apprentissage de la BASE2LR. (A) Vérité terrain. (B) Images mono-énergétiques issues des matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B). (C) **Images mono-énergétiques issues des matériaux décomposés avec U-net.** (E) Erreurs entre les images mono-énergétiques issues des matériaux décomposés avec la méthode de Gauss-Newton régularisé et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$. (F) **Erreur entre les images mono-énergétiques issues des matériaux décomposés avec U-net et la vérité terrain.**

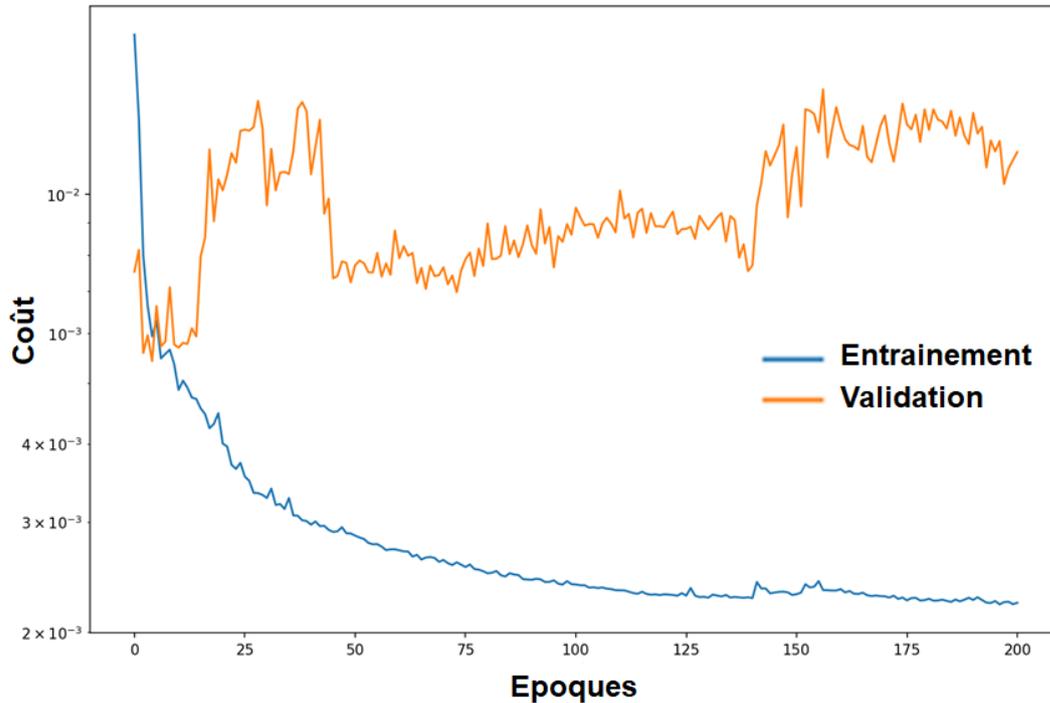


FIGURE 6.5 – Fonctions de coût d'apprentissage (train-bleu) et de validation (valid-orange) sur 200 époques pour une simulation de la BASE8LR avec 10^5 photons.

l'état de l'art, la méthode de Gauss-Newton régularisée avec deux paramètres de régularisation : une faible régularisation avec $\alpha = 0.03$ et une régularisation moyenne avec $\alpha = 3$.

La Table 6.2, donne l'erreur quadratique moyenne, l'indice de similarité et la valeur de bruit pour une projection, de la base de test, à 90° , car les valeurs sont uniformes sur tous les angles de projection, et pour les décompositions reconstruites. La méthode de Gauss-Newton peu régularisée donne les meilleures valeurs d'erreur quadratique moyenne et d'indice de similarité mais présente le plus fort bruit. La méthode de Gauss-Newton moyennement régularisée lutte le mieux contre le bruit mais au dépend des autres métriques. Quant à elle, la méthode avec U-net donne des décompositions avec un bon compromis entre un faible bruit et, une faible erreur quadratique moyenne et un bon indice de similarité.

La Figure 6.6 présente les résultats de la décomposition de matériaux dans le domaine des projections et la Figure 6.7, la reconstruction tomographique de ces résultats.

Ces figures montrent que la méthode de Gauss-Newton faiblement régularisée donne des résultats plus bruités que ceux de U-net et la méthode de Gauss-Newton moyennement régularisée rend les résultats plus flous. Visuellement, U-net donne des résultats avec à la fois peu de bruit et une bonne conservation des détails.

Malheureusement, le fait de travailler sur une base basse résolution ne nous permet pas de visualiser totalement la capacité du réseau à conserver des détails plus fins. De plus U-net nous permet également d'accélérer le calcul de la décomposition. En effet, il nécessite 40 secondes pour décomposer 720 projections alors que la méthode de Gauss-Newton régularisée nécessite 1 heure, sur le processeur central. En utilisant une carte graphique NVidia RTX 2080 Ti, la décomposition avec U-net ne nécessite que 2 secondes, sur la base de test, après apprentissage.

TABLE 6.2 – Résultats quantitatifs pour la décomposition de matériaux dans le domaine des projections avec U-net sur la base de test et comparé à la méthode de Gauss-Newton. Calculé sur la projection à 90° et après reconstruction pour la coupe 50.

		Tissus mous projetés	Os projetés	Tissus mous	Os
MSE	RGN $\alpha = 0.03$	0.020	0.009	0.0013	0.0006
	RGN $\alpha = 3$	0.021	0.010	0.0038	0.0018
	U-net	0.038	0.015	0.0032	0.0015
SSIM	RGN $\alpha = 0.03$	0.89	0.84	0.96	0.96
	RGN $\alpha = 3$	0.91	0.67	0.51	0.52
	U-net	0.91	0.72	0.69	0.69
Bruit	RGN $\alpha = 0.03$	0.76	0.43	0.32	0.26
	RGN $\alpha = 3$	0.74	0.39	0.18	0.18
	U-net	0.71	0.46	0.21	0.19

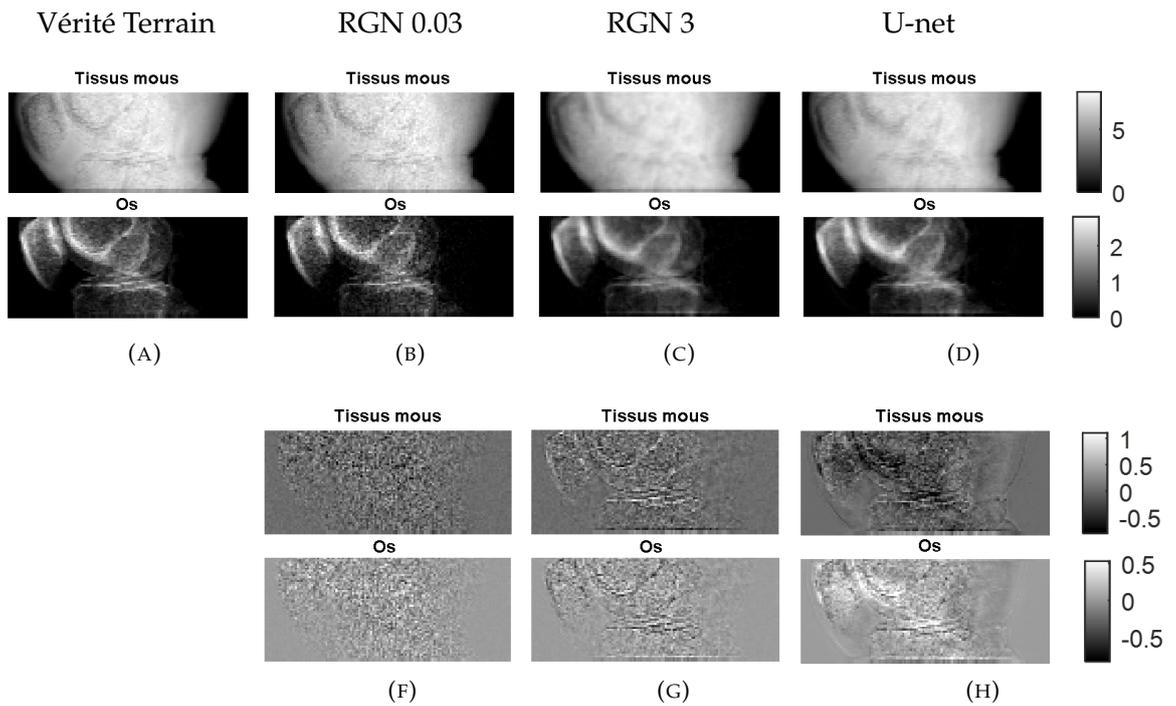


FIGURE 6.6 – Résultats de la décomposition de matériaux dans le domaine des projections avec U-net sur la base de test de la BASE8LR, visualisés pour les projections décomposées pour un angle de 90° . (A) Vérité terrain pour la décomposition. (B,C) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B) et $\alpha = 3$ (C). (D) **Matériaux décomposés avec U-net.** (F,G) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$ (F) et $\alpha = 3$ (G). (H) **Erreur entre les matériaux décomposés avec U-net et la vérité terrain.**

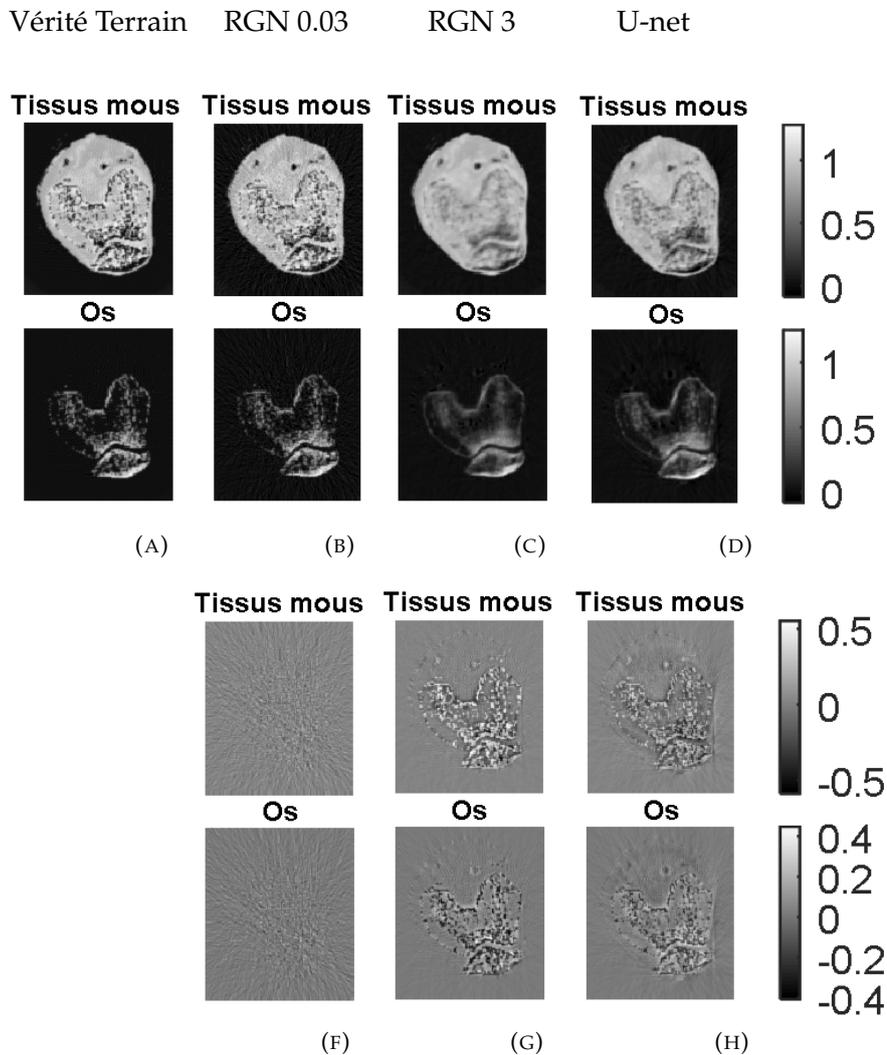


FIGURE 6.7 – Reconstruction tomographique des résultats de la décomposition de matériaux avec U-net dans le domaine des projections sur la base de test de la BASE8LR. (A) Vérité terrain pour la décomposition. (B,C) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B) et $\alpha = 3$ (C). (D) **Matériaux décomposés avec U-net.** (F,G) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$ (F) et $\alpha = 3$ (G). (H) **Erreur entre les matériaux décomposés avec U-net et la vérité terrain.**

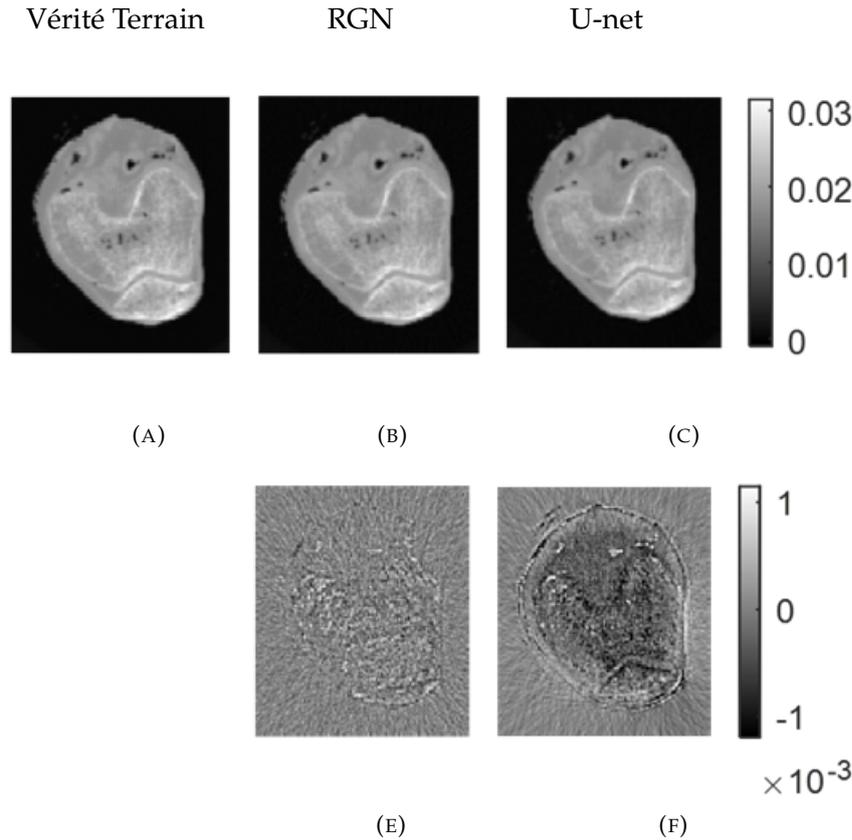


FIGURE 6.8 – Images mono-énergétiques à 70 keV obtenues à partir de la reconstruction tomographique des projections décomposées avec U-net sur la base de test de la BASE8LR. (A) Vérité terrain. (B) Images mono-énergétiques issues des matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.03$ (B). (C) **Images mono-énergétiques issues des matériaux décomposés avec U-net.** (E) Erreurs entre les images mono-énergétiques issues des matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.03$. (F) **Erreur entre les images mono-énergétiques issues des matériaux décomposés avec U-net et la vérité terrain.**

Reconstruction des images mono-énergétiques Comme précédemment, les images mono-énergétiques sont obtenues à partir des décompositions en utilisant le modèle de matériaux. On peut voir sur la Figure 6.8 les images mono-énergétiques à 70 keV obtenues à partir des différentes méthodes. Comme pour la décomposition de matériaux, la qualité d’image est similaire avec la méthode de Gauss-Newton et la méthode d’apprentissage. De même, les erreurs sur les images mono-énergétiques avec U-net sont surtout focalisées sur les bords.

6.3 Fantômes à la résolution du scanner Philips

Les fantômes à la résolution du scanner ont une taille de pixel de $250 \mu\text{m}$ de côté et chaque projection a une taille de 643×202 pixels. Pour l’apprentissage, on utilise une approche par patch. En effet, pour une meilleure gestion de la mémoire,

à chaque itération, on sélectionne un patch aléatoirement dans la projection, pour l'étude dans le domaine des projections, ou l'image, pour l'étude dans le domaine image, de 64×64 pixels.

6.3.1 BASE8HR

Cette base de données prend en compte les 8 genoux disponibles. La description de cette base est réalisée dans le 4.6.2. Les données sont simulées avec le modèle de Philips. On utilise une approche par patch pour une meilleure gestion de la mémoire. A chaque itération, pour chaque lot, le réseau prend un patch aléatoire et différent, de taille 64×64

Décomposition de matériaux Pour l'apprentissage, on utilise une taille de lot de 45 projections. Il a été possible d'augmenter la taille du lot car nous travaillons sur des patchs 64×64 pixels qui sont plus petits que les projections de taille 161×51 pixels. Ils prennent donc moins d'espace mémoire. Avec cette taille de lot, il faut 64 itérations pour réaliser une époque. La fonction de coût est l'erreur quadratique moyenne et l'optimiseur est ADAM. On réalise des essais avec des taux d'apprentissage initiaux de 10^{-3} , 10^{-2} et 10^{-1} . Ils décroissent au fur et à mesure de l'apprentissage en étant multipliés par 0.3 toutes les 30 époques. Le but de décroître le taux d'apprentissage est de pouvoir atteindre plus précisément le minimum.

On voit sur la Figure 6.9 les fonctions de coût d'entraînement et de validation pour plusieurs taux d'apprentissage. Lorsque l'on augmente le taux d'apprentissage, la fonction de coût d'entraînement est de plus en plus faible mais la fonction de coût de validation est de plus en plus instable, elle se stabilise sur les dernières époques. La fonction de coût de validation suit la fonction de coût d'entraînement, cela prouve que l'on arrive à apprendre et généraliser.

En testant sur la base de test, on montre sur la Figure 6.10 les reconstructions tomographiques des décompositions de matériaux dans le domaine de projections. On voit que l'on arrive à décomposer un genou totalement inconnu du réseau. On compare ces décomposition pour plusieurs taux d'apprentissage avec celle réalisée avec la méthode de Gauss-Newton faiblement régularisée avec $\alpha = 0.3$. Nous choisissons un taux de régularisation légèrement plus fort pour l'étude en basse résolution car nous n'en utilisons qu'un. En effet, une très faible régularisation avec $\alpha = 0.03$ nous donnait beaucoup de bruit.

Comme pour la base basse résolution, U-net donne le meilleur compromis entre la réduction du bruit et la conservation des détails. En effet, on visualise que le bruit du fond est très bien géré et que les structures comme l'architecture trabéculaire de l'os sont bien conservées. La Table 6.3 expose la comparaison des résultats quantitatifs pour l'erreur quadratique moyenne et l'indice de similarité. On voit que U-net surpasse la méthode de Gauss-Newton pour toutes les métriques. De plus, les meilleurs résultats pour U-net sont avec des taux d'apprentissage initiaux de 10^{-3} et 10^{-2} .

Décomposition des projections mono-énergétiques par apprentissage Les projections mono-énergétiques sont reconstruites à partir des décompositions de matériaux opérées avec le réseau de neurones. Les projections mono-énergétiques peuvent être directement calculées avec le réseau de neurones U-net entraîné à cet effet. Les paramètres du réseau sont les même que pour la décomposition de matériaux. Ainsi, la taille de lot est de 45 projections, la fonction de coût est l'erreur quadratique moyenne, l'optimiseur est ADAM et le taux d'apprentissage démarre à 10^{-3} , 10^{-2}

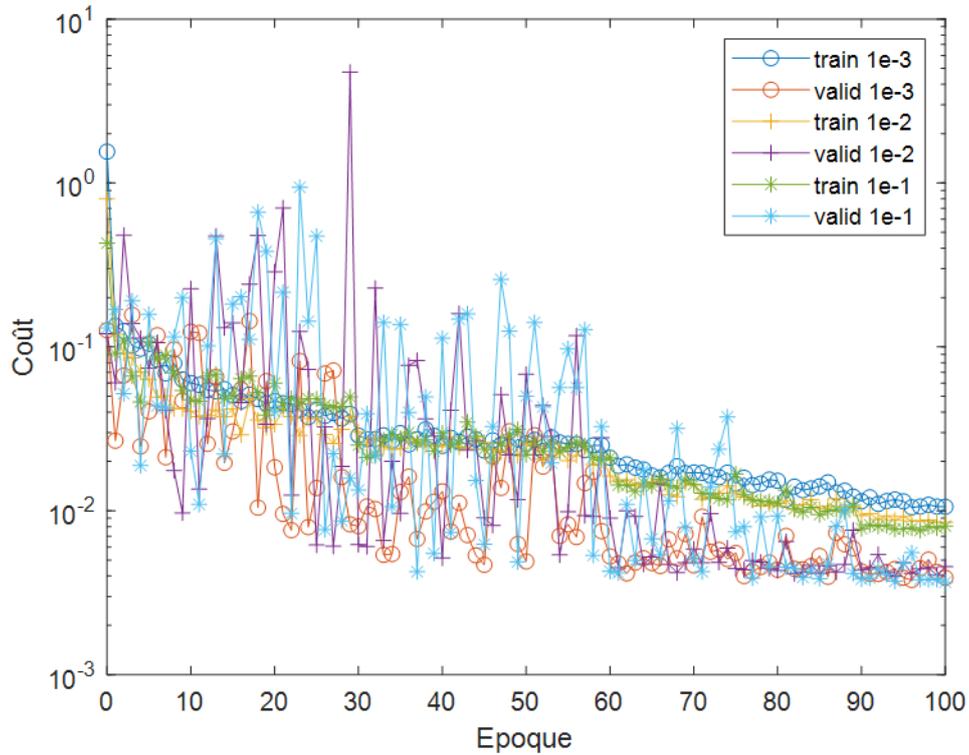


FIGURE 6.9 – Influence du taux d'apprentissage sur les fonctions de coût d'entraînement et de validation pour 100 époques pour la décomposition de matériaux avec U-net dans le domaine des projections.

TABLE 6.3 – Résultats quantitatifs pour la décomposition de matériaux de la BASE8HR avec U-net pour la base de test et comparaison avec la méthode de Gauss-Newton, visualisés sur la projection à 90° et après reconstruction pour la coupe 50.

		Tissus mous projetés	Os projetés	Tissus mous	Os
MSE	RGN $\alpha = 0.3$	0.0083	0.0034	0.0066	0.0028
	U-net $\lambda = 10^{-3}$	0.0076	0.0044	0.0055	0.0031
	U-net $\lambda = 10^{-2}$	0.0073	0.0039	0.0066	0.0031
	U-net $\lambda = 10^{-1}$	0.0095	0.0033	0.0055	0.0033
SSIM	RGN $\alpha = 0.3$	0.4787	0.5407	0.2844	0.3721
	U-net $\lambda = 10^{-3}$	0.7014	0.5198	0.6304	0.7032
	U-net $\lambda = 10^{-2}$	0.6932	0.8063	0.5890	0.7375
	U-net $\lambda = 10^{-1}$	0.5633	0.5566	0.4260	0.5541

ou 10^{-1} , et est multiplié par 0.3 toutes les 30 époques. On entraîne sur 100 époques car les fonctions de coût se stabilisent vers ce nombre d'époques.

La Figure 6.11 montre les fonctions de coût d'entraînement et de validation pour différents taux d'apprentissage. On voit la convergence des fonctions de coût d'entraînement et de validation pour la reconstruction directe des projections mono-énergétiques avec U-net. De plus, comme pour les fonctions de coût pour la décomposition de matériaux, on améliore le coût en augmentant le taux d'apprentissage mais la fonction de coût de validation est instable sur les petites époques et se stabilise sur la fin.

La Figure 6.12 compare les projections mono-énergétiques, obtenues à partir de la décomposition de matériaux réalisée par le réseau U-net ou issues directement de l'apprentissage, avec celles issues de la décomposition de matériaux avec la méthode de Gauss-Newton. Les images issues de l'apprentissage, qu'il soit direct ou pour la décompositions de matériaux, sont moins bruitées et permettent une meilleure visualisation des détails.

La Table 6.4 montre les résultats quantitatifs de la décomposition mono-énergétique. On voit que les images mono-énergétiques reconstruites à partir des décomposition de matériaux ou directement décomposés donnent des résultats équivalents ou qui surpassent la méthode de Gauss-Newton sur toutes les métriques.

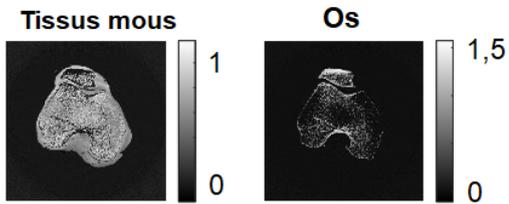
On peut voir que les images mono-énergétiques issues des décompositions sont très légèrement meilleures. On peut expliquer cela, car s'il y a un mélange entre les matériaux dans la décomposition, il sera pallié par la recombinaison des matériaux. Si une erreur est réalisée en apprenant la directement la décomposition mono-énergétique dans le réseau, elle ne pourra pas être compensée. Néanmoins, apprendre directement à reconstruire les projections mono-énergétique nous permet de nous émanciper du modèle de matériaux et de gagner une étape lors du post-traitement. Contrairement à la décomposition de matériaux, les meilleures métriques sont pour le taux d'apprentissage initial à 10^{-1} .

On cherche à visualiser le cartilage sur les images mono-énergétiques. Malheureusement, sur la coupe que nous avons utilisé et qui montre le mieux les partie du genou avec une bonne portion de la rotule et du fémur, nous avons de l'air le long du cartilage. Ainsi il est difficile de voir la différence entre le cartilage et les tissus environnants quand nous avons du noir entre les deux. Nous choisissons donc une autre tranche et nous allons choisir de regarder le cartilage à l'arrière d'un condyle.

Sur la Figure 6.13, on choisit d'évaluer la qualité de la visualisation du cartilage sur les images mono-énergétiques à partir de la reconstruction tomographique des décomposition de matériaux avec U-net dans le domaine des projections avec un taux d'apprentissage à 10^{-3} et 10^{-2} et à partir de la reconstruction tomographique de

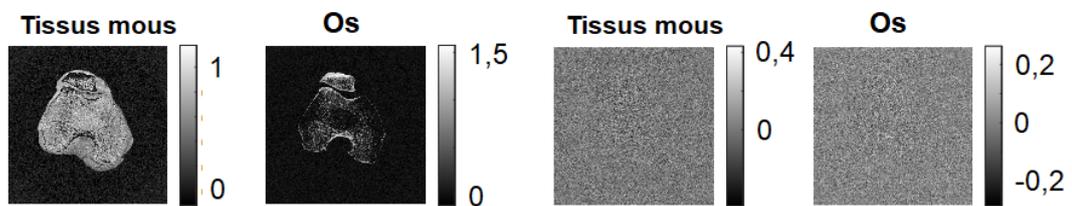
FIGURE 6.10 (page suivante) – Reconstruction tomographique des résultats de la décomposition de matériaux avec U-net sur la base de test de la BASE8HR dans le domaine des projections. (A) Vérité terrain pour la décomposition. (C) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$. (D) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.3$. **(E,G,I) Matériaux décomposés avec U-net avec un taux d'apprentissage initial de 10^{-3} (E), 10^{-2} (G) et 10^{-1} (I).** (F,H,J) Erreur entre les matériaux décomposés avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (F), 10^{-2} (H) et 10^{-1} (J).

Vérité terrain



(A)

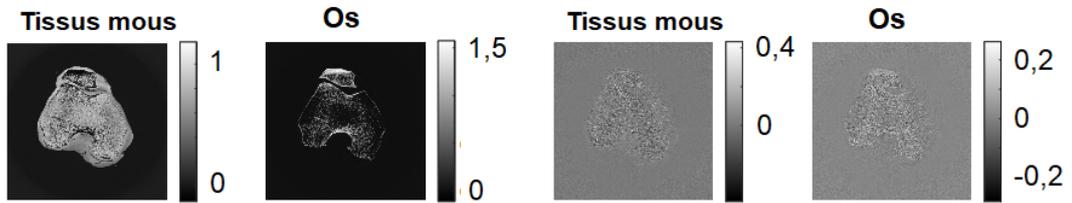
RGN



(C)

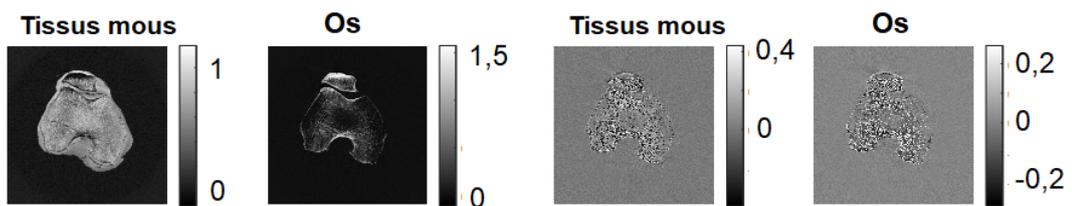
(D)

U-net



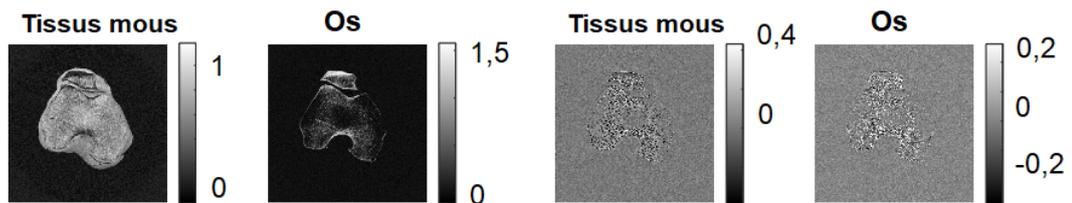
(E)

(F)



(G)

(H)



(I)

(J)

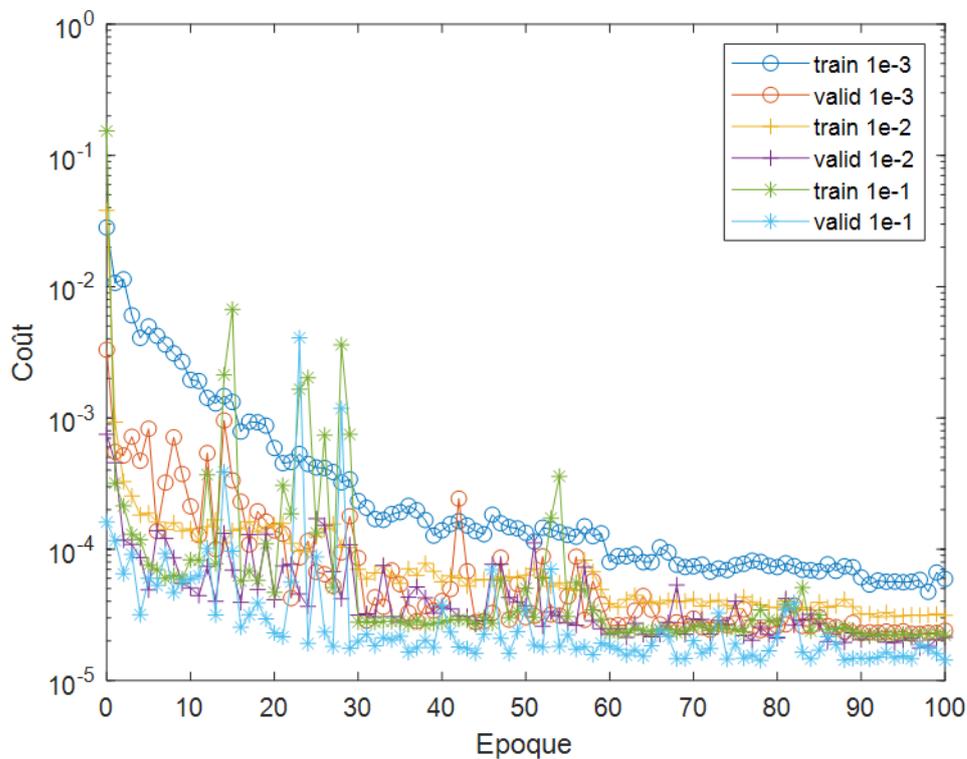


FIGURE 6.11 – Influence du taux d'apprentissage sur les fonctions de coût d'entraînement et de validation pour 100 époques pour la reconstruction directe des images mono-énergétiques avec U-net.

TABLE 6.4 – Résultats quantitatifs pour la reconstruction des images mono-énergétiques à 70 keV à partir des matériaux décomposées avec U-net ou directement reconstruites avec U-net, et comparées avec la méthode de Gauss-Newton.

	MSE	SSIM
RGN $\alpha = 0.3$	$1.16 \cdot 10^{-5}$	0.9883
U-net dec $\lambda = 10^{-3}$	$3.74 \cdot 10^{-5}$	0.9786
U-net dec $\lambda = 10^{-2}$	$1.24 \cdot 10^{-5}$	0.9901
U-net dec $\lambda = 10^{-1}$	$1.15 \cdot 10^{-5}$	0.9903
U-net dir $\lambda = 10^{-3}$	$2.08 \cdot 10^{-5}$	0.9826
U-net dir $\lambda = 10^{-2}$	$1.76 \cdot 10^{-5}$	0.9856
U-net dir $\lambda = 10^{-1}$	$1.37 \cdot 10^{-5}$	0.9876

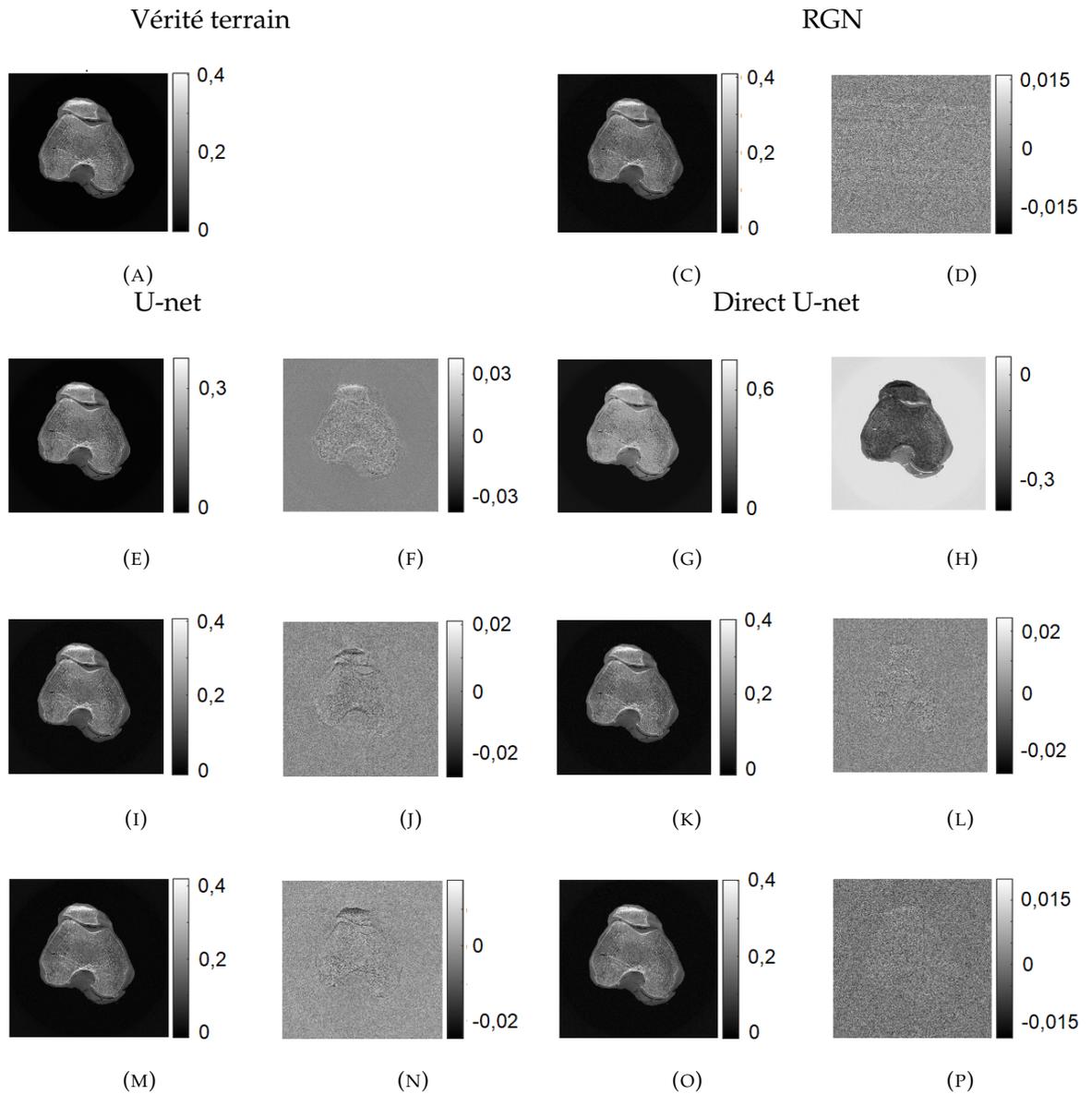


FIGURE 6.12 – Reconstruction tomographique des projections mono-énergétiques à 70 keV reconstruites avec U-net sur la base de test de la BASE8HR à partir de la décomposition de matériaux ou directement. (A) Vérité terrain pour l’image mono-énergétique. (C) Image mono-énergétique à partir de la décomposition de matériaux avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$. (D) Erreurs entre l’image mono-énergétique à partir de la décomposition de matériaux avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ et la vérité terrain. (E,I,M) Image mono-énergétique à partir de la décomposition de matériaux avec U-net avec un taux d’apprentissage initial de 10^{-3} (E), 10^{-2} (I) et 10^{-1} (M). (F,J,N) Erreur entre les image-mono-énergétiques à partir de la décomposition de matériaux avec U-net et la vérité terrain pour un taux d’apprentissage initial de 10^{-3} (F), 10^{-2} (J) et 10^{-1} (N). (G,K,O) Image mono-énergétique directement reconstruites avec U-net avec un taux d’apprentissage initial de 10^{-3} (G), 10^{-2} (K) et 10^{-1} (O). (H,L,P) Erreur entre les image-mono-énergétiques directement reconstruites avec U-net et la vérité terrain pour un taux d’apprentissage initial de 10^{-3} (H), 10^{-2} (L) et 10^{-1} (P).

la décomposition des projections mono-énergétiques créées avec U-net avec un taux d'apprentissage à 10^{-1} . On compare ces résultats avec les images mono-énergétiques issues de la reconstruction tomographique des projections de matériaux décomposées avec la méthode de Gauss-Newton régularisée. On peut voir que le cartilage est visible à l'œil nu sur toutes les images. De plus, il est possible de voir une partie plus foncée au milieu du cartilage qui pourrait correspondre avec une lésion induite par l'arthrose. Nous pouvons également visualiser précisément l'architecture de l'os qui est très nette pour l'image mono-énergétique issue de la reconstruction tomographique des projections mono-énergétique et à partir de la décomposition avec U-net avec un taux d'apprentissage à 10^{-2} et Gauss-Newton. L'image mono-énergétique issue de la décomposition de matériaux avec U-net avec un taux d'apprentissage à 10^{-3} est assez floue et ne permet pas une bonne visualisation de l'os.

6.4 Conclusion

Nous avons présenté dans ce chapitre les résultats obtenus par nos méthodes d'apprentissage pour la décomposition dans le domaine des projections de matériaux ou mono-énergétiques à partir des données spectrales.

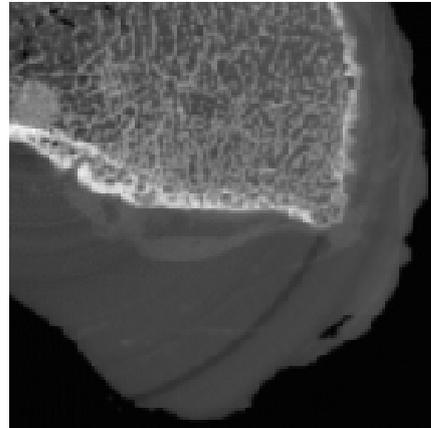
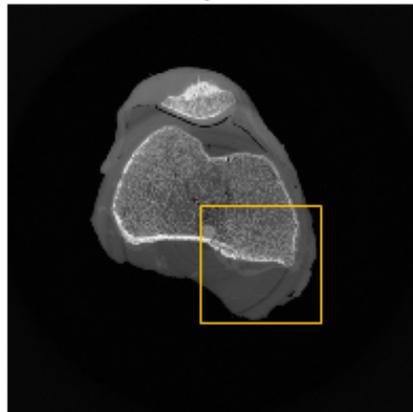
On a ainsi vu que la base de donnée réduite est trop faible pour une décomposition de matériaux généralisable à des nouvelles données. En augmentant la base, on obtient de meilleurs résultats qui sont comparables avec l'état de l'art. En effet, cela donne un meilleur compromis entre le bruit et la reconstruction des détails en s'émancipant du choix d'un paramètre de régularisation qui est indispensable avec la méthode de Gauss-Newton régularisée. Malheureusement travailler en basse résolution ne permet pas de montrer l'étendue des capacités du réseau de neurones car il y a trop peu de détails dans les images.

En passant en haute résolution on a davantage de pixels et donc plus de données d'apprentissage par rapport au nombre de paramètres du réseau. En utilisant cette base, on surpasse la méthode de Gauss-Newton régularisée en décomposant les matériaux et les projections mono-énergétiques avec un plus faible niveau de bruit et une meilleure visualisation des détails. Nous avons choisi de réaliser plusieurs tests avec plusieurs taux d'apprentissage pour trouver les meilleurs résultats et pouvoir choisir ultérieurement en fonction des métriques primordiales pour le problème.

Pour la reconstruction des projections mono-énergétiques, on se rend compte que les images mono-énergétiques issues des reconstructions tomographiques des matériaux décomposés, que ce soit avec U-net ou la méthode de Gauss-Newton, sont légèrement meilleures que celles issues de la reconstruction tomographique des projections mono-énergétiques directement apprises avec U-net. La visualisation du cartilage est possible peu importe la méthode.

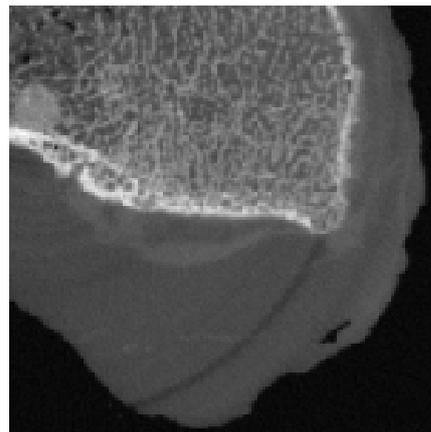
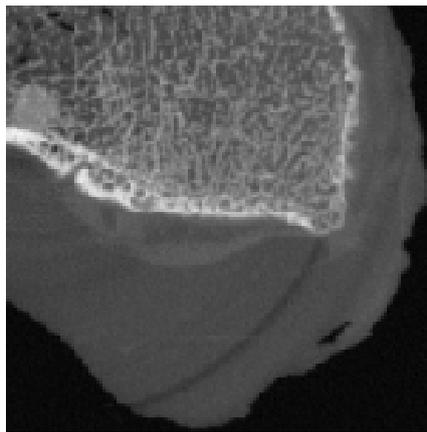
FIGURE 6.13 (page suivante) – Zoom sur la zone du cartilage sur la reconstruction tomographique des projections mono-énergétiques reconstruites avec U-net sur la base de test de la BASE8HR. (A) et (B) Vérité terrain avec l'emplacement et le zoom de la zone du cartilage. (C) Image mono-énergétique à partir de la reconstruction tomographique des matériaux décomposés avec Gauss-Newton. **(D) Reconstruction tomographique des projections mono-énergétiques reconstruites avec U-net avec un taux d'apprentissage initial à 10^{-1} .** **(E,F) Images mono-énergétiques à partir des décomposition de matériaux dans le domaine des projections avec U-net avec un taux d'apprentissage initial à 10^{-3} (E) et 10^{-2} (F).**

Vérité terrain



RGN (A)

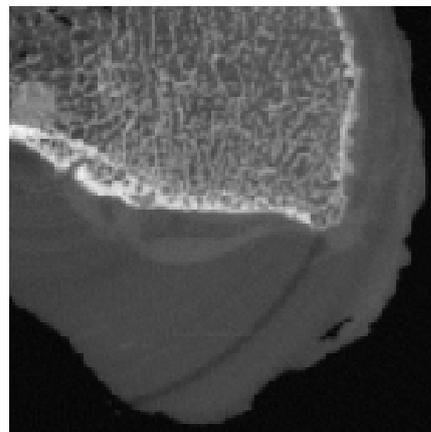
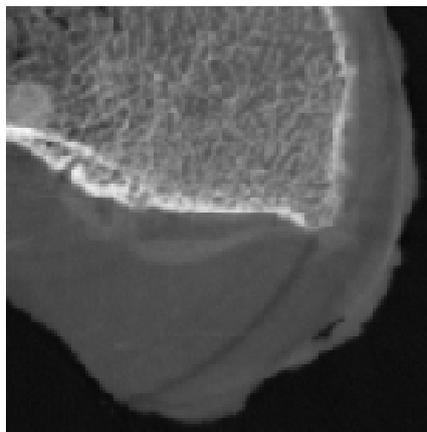
(B)
Direct U-net



(C)

(D)

U-net



(E)

(F)

De plus, l'utilisation des réseaux de neurones nous permet de gagner beaucoup de temps. En effet l'apprentissage prend environ 6h; mais pour tester ensuite les données, nous n'avons besoin que 3 minutes, chargement des données compris, pour décomposer les matériaux ou reconstruire les projections mono-énergétiques sur tout le volume, soit 0.25 seconde par projection. Pour la méthode de Gauss-Newton, il faudra 2h par volume.

Chapitre 7

Apprentissage de la décomposition dans le domaine image : étude sur des simulations

Dans ce chapitre nous étendrons l'étude précédente sur les méthodes d'apprentissage réalisées dans le domaine des projections au domaine image. Nous travaillerons avec la base de données de 8 genoux avec la même résolution que le prototype de scanner spectral simulée avec le modèle du prototype de scanner. Nous utiliserons le réseau U-net pour apprendre la reconstruction de matériaux et mono-énergétiques. Nous comparerons les résultats de ces algorithmes avec la reconstruction tomographique de ceux obtenus avec la méthode de Gauss-Newton régularisée qui n'est opérationnelle que dans la domaine des projections.

7.1 Introduction

Travailler dans le domaine image peut permettre un traitement plus rapide. En effet, ici, la partie qui prend le plus de temps est la reconstruction tomographique. Avec l'étude dans le domaine image nous la faisons une fois pour toute sur les projections spectrales. Réaliser la reconstruction tomographique après la décomposition de matériaux peut rajouter des artefacts sur les résultats des décompositions.

Pour obtenir la base de données dans le domaine image, on doit reconstruire, au préalable les données spectrales acquises par le scanner et la vérité terrain des cartes de matériaux projetées. Comme expliqué dans le 5.3.2, on reconstruit les données spectrales préalablement normalisées pour conserver une échelle de niveaux de gris raisonnable pour l'apprentissage. On compare les résultats de U-net avec les résultats reconstruits de la méthode de Gauss-Newton faiblement régularisée avec $\alpha = 0.3$.

7.2 Fantômes à la résolution du scanner Philips - BASE8HR

On utilise la base de données à la résolution du prototype de scanner spectral avec les 8 fantômes de genoux disponibles décrite dans le 4.6.2. Les fantômes à la résolution du scanner ont une taille de voxel de $250 \mu\text{m}$ de côté et chaque coupe a une taille de 635×635 pixels. Cette base compile les projections spectrales, dans les bins d'énergie, normalisées puis reconstruites, les cartes de matériaux reconstruites et les images mono-énergétiques reconstruites. Comme pour l'étude dans le domaine des projections on utilise une approche par patch en choisissant des patches de taille 64×64 pixels. Dans le domaine des projections on apprendait projection par projection,

TABLE 7.1 – Résultats quantitatifs pour la décomposition de matériaux de la base 8HR avec U-net sur la base de test et comparaison avec la méthode de Gauss-Newton. Exposés pour la coupe 50.

		Tissus mous	Os
MSE	RGN $\alpha = 0.3$	0.0066	0.0028
	U-net $\lambda = 10^{-3}$	0.0013	0.0007
	U-net $\lambda = 10^{-2}$	0.0012	0.0006
	U-net $\lambda = 10^{-1}$	0.0016	0.0007
SSIM	RGN $\alpha = 0.3$	0.2844	0.3721
	U-net $\lambda = 10^{-3}$	0.8237	0.8305
	U-net $\lambda = 10^{-2}$	0.8527	0.8677
	U-net $\lambda = 10^{-1}$	0.7441	0.8008

dans le domaine image, on apprend coupe par coupe pour un nombre total de 202 coupes par volume.

7.2.1 Décomposition de matériaux

On garde la même architecture de réseau que pour l’analyse dans le domaine des projections, on change la base de données et on réentraîne le réseau. On choisit une taille de lot de 45 coupes de 5 bins d’énergie pour les entrées et 2 matériaux pour les sorties. 18 itérations sont nécessaires pour réaliser une époque. On choisit l’erreur quadratique moyenne comme fonction de coût, l’optimiseur ADAM 3.2.2 et un taux d’apprentissage initial à 10^{-3} , 10^{-2} et 10^{-1} qui décroît en étant multiplié par 0.3 toutes les 30 époques. L’apprentissage est limité à 100 époques car les fonctions de coût convergent aux alentours ce nombre d’époques.

On visualise sur la Figure 7.1, la convergence de la fonction de coût d’entraînement et de validation. On peut ainsi voir que le réseau est capable d’apprendre sur une base et généraliser sur une nouvelle base. De plus, comme pour l’étude dans le domaine des projections, la valeur des fonctions de coûts est plus faible avec des taux d’apprentissage initiaux plus grands et les fonctions de validation avec de forts taux d’apprentissage initiaux sont instables sur les petites époques et se stabilisent sur la fin.

Sur la Figure 7.2, on compare les reconstructions tomographiques des matériaux décomposés dans le domaine des projections avec la méthode de Gauss-Newton. Les résultats sont calculés sur la base de test, totalement inconnue du réseau. Comme dans le domaine des projections, la décomposition de matériaux avec U-net donne un meilleur compromis entre un faible bruit et une bonne reconstruction des détails. La Table 7.1 nous montre que quantitativement, les décompositions dans le domaine image avec U-net ont une plus faible erreur quadratique moyenne et un meilleur indice de similarité que les reconstructions tomographiques des matériaux décomposés avec la méthode de Gauss-Newton régularisée. Les meilleures métriques sont pour un taux d’apprentissage initial à 10^{-2} et sont également bien meilleures que les meilleures métriques de U-net dans le domaine des projections.

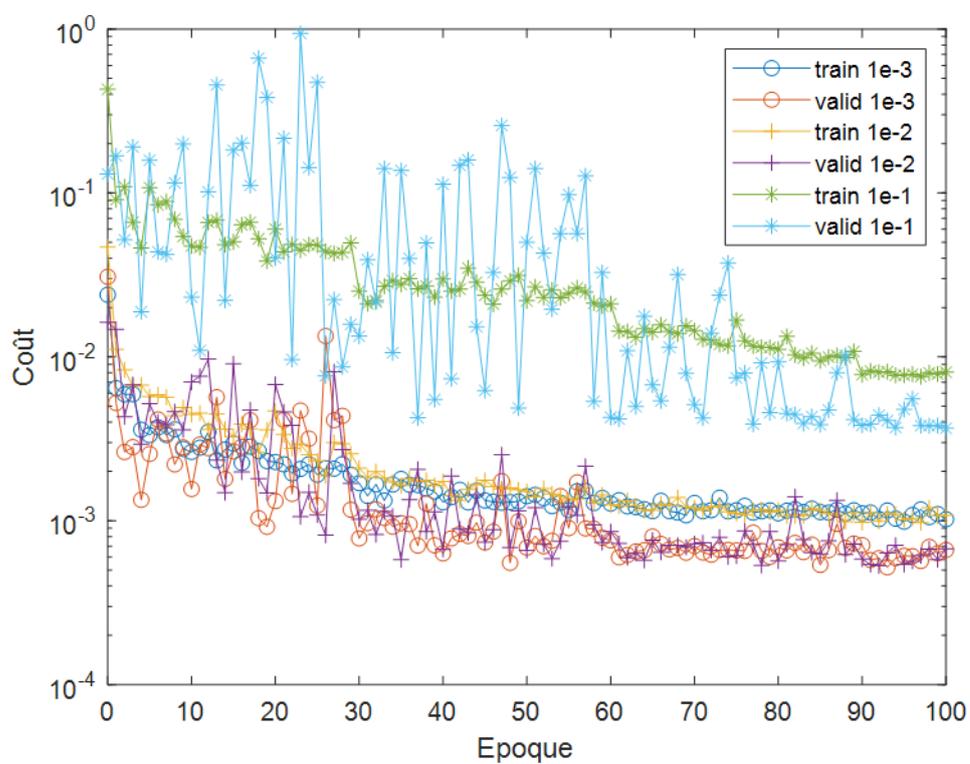


FIGURE 7.1 – Influence du taux d'apprentissage sur les fonctions de coût d'entraînement et de validation pour la décomposition de matériaux avec U-net dans le domaine image sur 100 époques.

TABLE 7.2 – Résultats quantitatifs des images mono-énergétiques à 70 keV avec U-net à partir de la décomposition et avec U-net directement apprise sur les images mono-énergétiques et comparé à la méthode de Gauss-Newton. Exposés pour une coupe.

	MSE	SSIM
RGN $\alpha = 0.3$	$1.18 \cdot 10^{-5}$	0.9883
U-net dec $\lambda = 10^{-3}$	$1.84 \cdot 10^{-5}$	0.9774
U-net dec $\lambda = 10^{-2}$	$0.81 \cdot 10^{-5}$	0.9930
U-net dec $\lambda = 10^{-1}$	$0.77 \cdot 10^{-5}$	0.9936
U-net dir $\lambda = 10^{-3}$	$1.37 \cdot 10^{-5}$	0.9914
U-net dir $\lambda = 10^{-2}$	$1.38 \cdot 10^{-5}$	0.9855
U-net dir $\lambda = 10^{-1}$	$1.14 \cdot 10^{-5}$	0.9914

7.2.2 Reconstruction des images mono-énergétiques

Pour l'apprentissage direct des images mono-énergétiques, on utilise le réseau de neurones avec une taille de lot de 45 coupes, l'erreur quadratique moyenne comme fonction de coût, ADAM 3.2.2 comme optimiseur et un taux d'entraînement initial de 10^{-3} , 10^{-2} et 10^{-1} .

La Figure 7.3 montre la convergence des fonction de coût d'entraînement et de validation pour la reconstruction directe des images mono-énergétiques avec U-net dans le domaine image

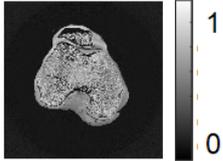
La Table 7.2 montre que la reconstruction des images mono-énergétiques à partir de la décomposition de matériaux est légèrement meilleure sur les métriques observées que la reconstruction directe, sauf pour un taux d'apprentissage initial à 10^{-3} pour la décomposition de matériaux. Le taux d'apprentissage initial à 10^{-1} rend les meilleurs résultats pour les images mono-énergétiques mais on choisit le taux initial à 10^{-2} qui est le meilleur pour la décomposition et très bon pour la reconstruction mono-énergétique à partir de la décomposition de matériaux. Le taux d'apprentissage à 10^{-1} donne les meilleurs résultats pour la reconstruction directe des images mono-énergétiques. Sur la Figure 7.4, on visualise les images mono-énergétiques reconstruites à partir des matériaux et directement avec le réseau de neurones.

Comme pour l'étude dans le domaine des projections, nous décidons d'utiliser une autre coupe que celle présentée précédemment pour une meilleure visualisation du cartilage. Sur la Figure 7.5, on montre la partie du cartilage à l'arrière d'un condyle pour les images mono-énergétiques issues de la décomposition de matériaux avec U-net dans le domaine image avec un taux d'apprentissage initial à 10^{-2}

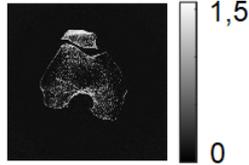
FIGURE 7.2 (page suivante) – Résultats de la décomposition de matériaux avec U-net sur la base de test de la BASE8HR dans le domaine image. (A) Vérité terrain pour la décomposition. (C) Matériaux décomposés avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$. (D) Erreurs entre les matériaux décomposés avec la méthode de Gauss-Newton régularisée et la vérité terrain avec un paramètre de régularisation $\alpha = 0.3$. (E,G,I) Matériaux décomposés avec U-net avec un taux d'apprentissage initial de 10^{-3} (E), 10^{-2} (G) et 10^{-1} (I). (F,H,J) Erreur entre les matériaux décomposés avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (F), 10^{-2} (H) et 10^{-1} (J).

Vérité terrain

Tissus mous



Os

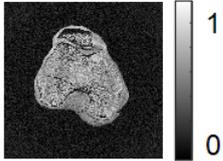


(A)

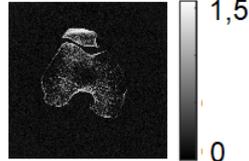
(B)

RGN

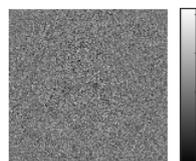
Tissus mous



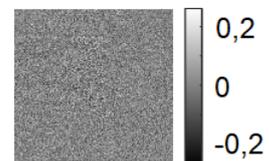
Os



Tissus mous



Os

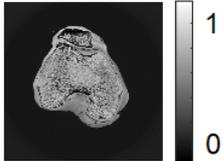


(C)

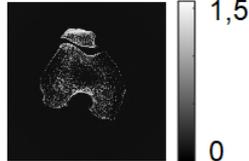
(D)

U-net

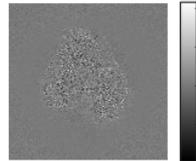
Tissus mous



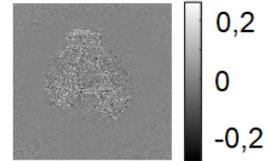
Os



Tissus mous



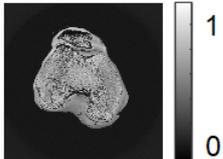
Os



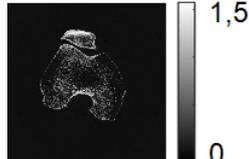
(E)

(F)

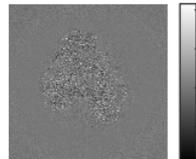
Tissus mous



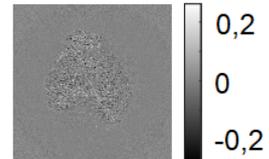
Os



Tissus mous



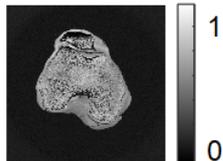
Os



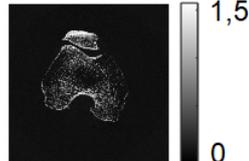
(G)

(H)

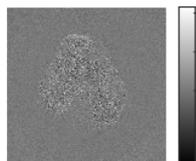
Tissus mous



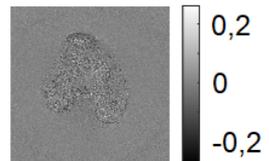
Os



Tissus mous



Os



(I)

(J)

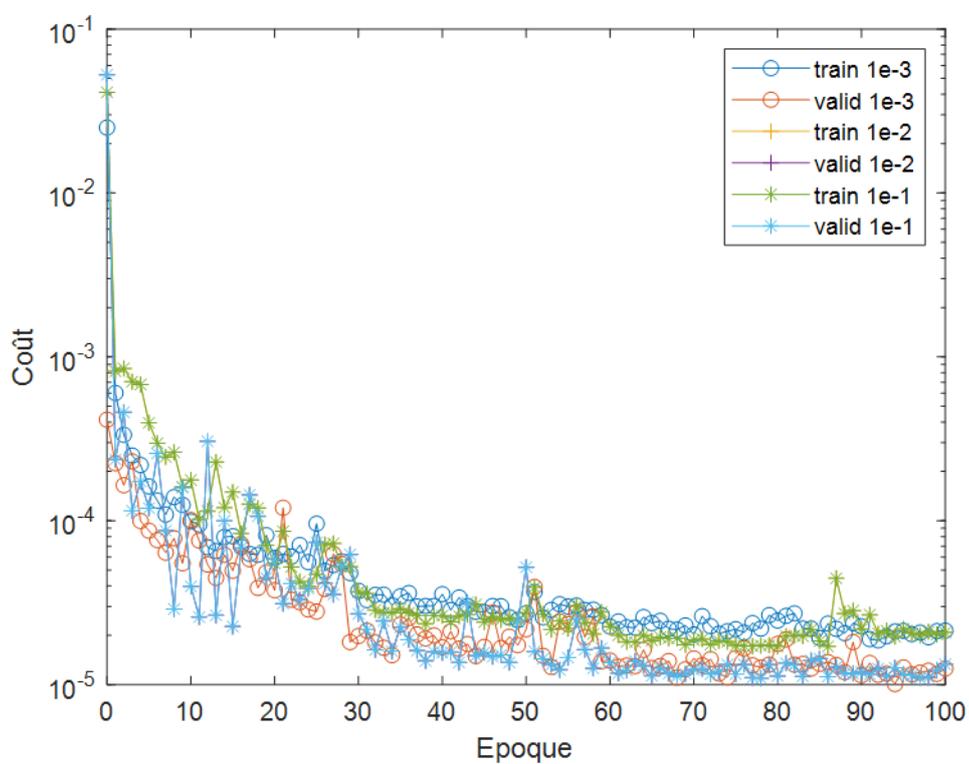


FIGURE 7.3 – Influence du taux d'apprentissage sur les fonctions de coût d'entraînement et de validation pour la reconstruction directe des images mono-énergétiques avec U-net sur 100 époques.

et directement reconstruites avec U-net avec un taux d'apprentissage initial à 10^{-1} . En effet, ces valeurs de taux d'apprentissage donnent des meilleurs résultats pour chaque méthode indépendamment. On voit que le cartilage est visible à l'oeil nu peu importe la méthode. On voit les défauts dans le cartilage qui peuvent être causés par de l'arthrose. Une chose étrange se produit avec les images mono-énergétiques issues de la décomposition de matériaux avec U-net. En effet, il y a comme du cartilage supplémentaire à l'endroit où celui-ci est abîmé. Ainsi, pour la détection des lésions du cartilage, la reconstruction directe des images mono-énergétiques semble plus pertinente. Contrairement au domaine des projections, nos résultats ne présentent pas d'altération des détails à certains taux d'apprentissage.

7.3 Conclusion

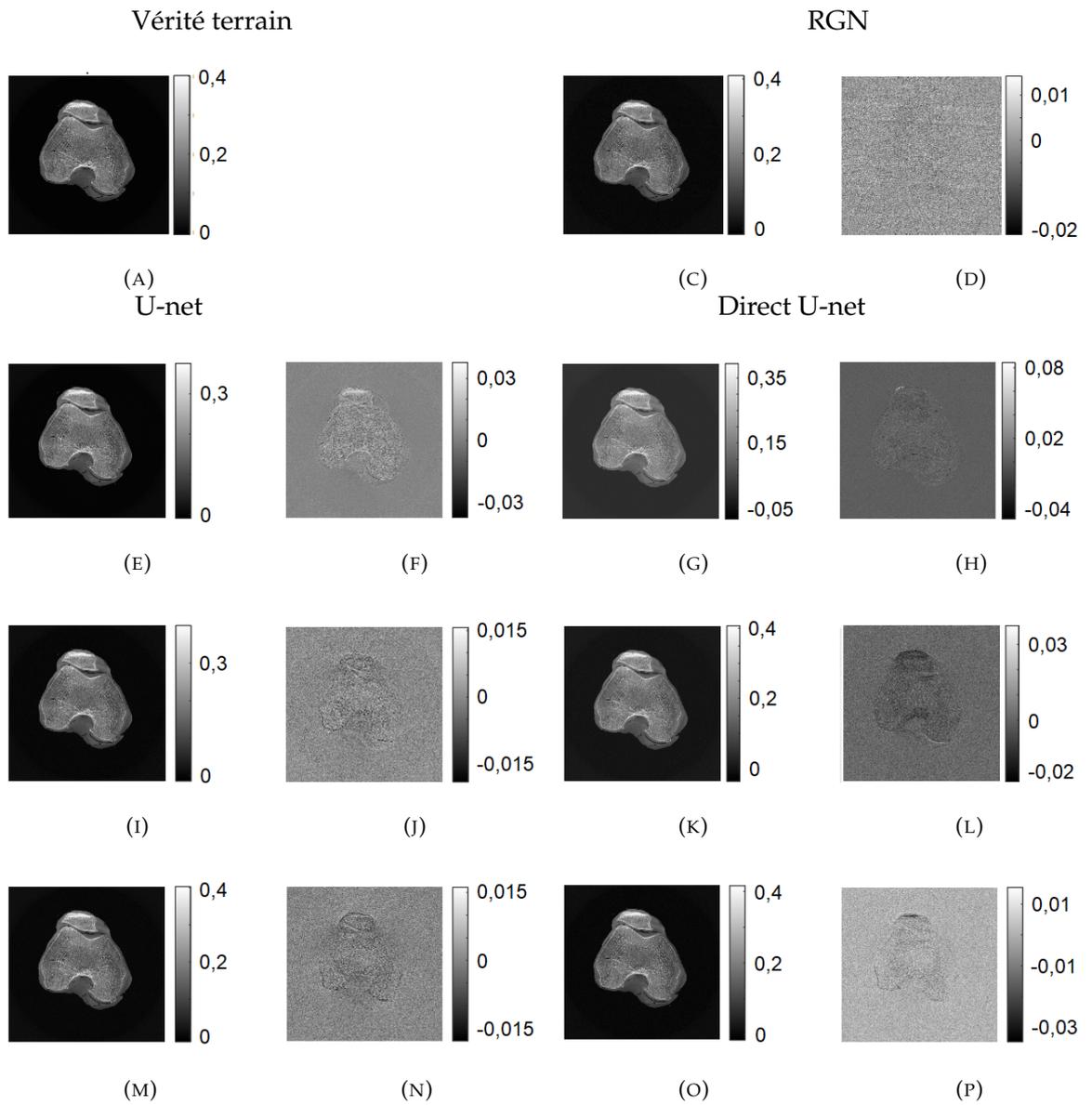
Dans ce chapitre, nous avons montré qu'il était possible de faire un apprentissage des cartes de matériaux ou des images mono-énergétiques directement dans le domaine image. Les décompositions donnent de meilleurs résultats en terme d'erreur quadratique moyenne et d'indice de similarité que la méthode de l'état de l'art, Gauss-Newton faiblement régularisée.

L'étude dans le domaine image donne de meilleurs résultats que la reconstruction des décompositions avec U-net dans le domaine des projections. Les images mono-énergétiques issues des décompositions de matériaux donnent également des meilleurs résultats avec U-net qu'avec la méthode de Gauss-Newton régularisée.

La reconstruction directe des images mono-énergétiques dans le domaine image donne de moins bon résultats qu'à partir de la décomposition et donne des résultats équivalents à la décomposition directe des projections mono-énergétiques. Malgré cela, elle permet d'obtenir des résultats satisfaisants pour reconstruire directement les images mono-énergétiques sans avoir besoin de passer par la décomposition de matériaux.

Le cartilage est visible à l'oeil nu sur les images mono-énergétiques peu importe la méthode même si les détails au niveau des lésions du cartilage semblent être mieux reconstruites sur les images mono-énergétiques directement reconstruites avec U-net.

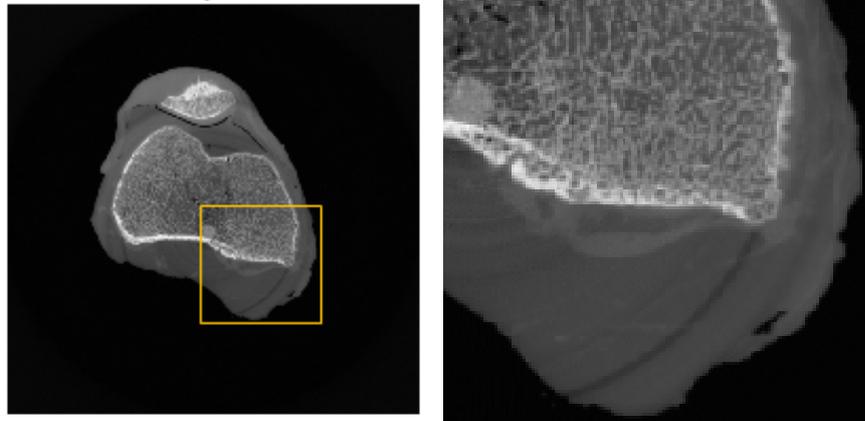
FIGURE 7.4 (page suivante) – Résultat de la reconstruction des images mono-énergétiques à 70 keV avec U-net sur la base de test de la BASE8HR, à partir des décompositions et directement. (A) Vérité terrain pour l'image mono-énergétique. (C) Image mono-énergétique à partir de la décomposition de matériaux avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$. (D) Erreurs entre l'image mono-énergétique à partir de la décomposition de matériaux avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ et la vérité terrain. **(E,I,M) Image mono-énergétique à partir de la décomposition de matériaux avec U-net avec un taux d'apprentissage initial de 10^{-3} (E), 10^{-2} (I) et 10^{-1} (M).** **(F,J,N) Erreur entre les image-mono-énergétiques à partir de la décomposition de matériaux avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (F), 10^{-2} (J) et 10^{-1} (N).** **(G,K,O) Image mono-énergétique directement reconstruites avec U-net avec un taux d'apprentissage initial de 10^{-3} (G), 10^{-2} (K) et 10^{-1} (O).** **(H,L,P) Erreur entre les image mono-énergétiques directement reconstruites avec U-net et la vérité terrain pour un taux d'apprentissage initial de 10^{-3} (H), 10^{-2} (L) et 10^{-1} (P).**



Comme pour le domaine des projections, après 4h d'apprentissage, les données peuvent être testées en 1 minutes par volume, soit 0.3 seconde par coupe, alors qu'il faut 2h avec la méthode de Gauss-Newton pour opérer la décomposition de matériaux.

FIGURE 7.5 (page suivante) – Zoom sur la zone du cartilage sur la reconstruction tomographique des projections mono-énergétiques reconstruites avec U-net sur la base de test de la BASE8HR. (A) et (B) Vérité terrain avec l'emplacement et le zoom de la zone du cartilage. (C) Image mono-énergétique à partir de la reconstruction tomographique des matériaux décomposés avec Gauss-Newton. **(D) Images mono-énergétiques à partir des décompositions de matériaux dans le domaine image avec U-net avec un taux d'apprentissage initial à 10^{-2} .** **(E) Images mono-énergétiques reconstruites avec U-net avec un taux d'apprentissage initial à 10^{-1} .**

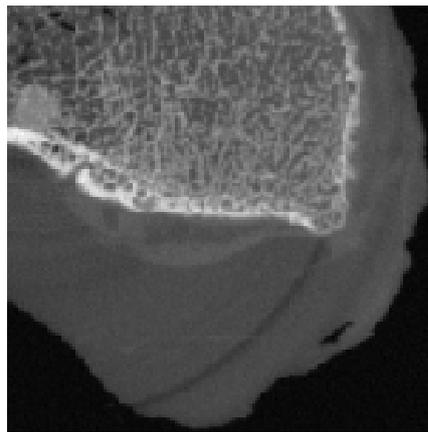
Vérité terrain



(A)

(B)

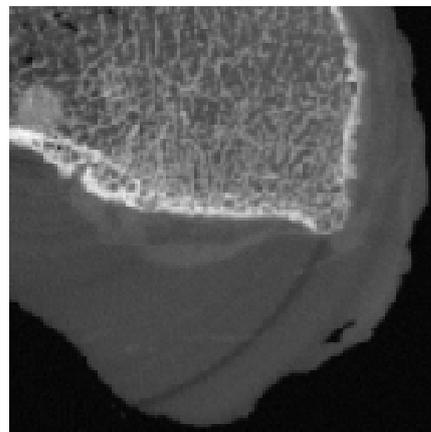
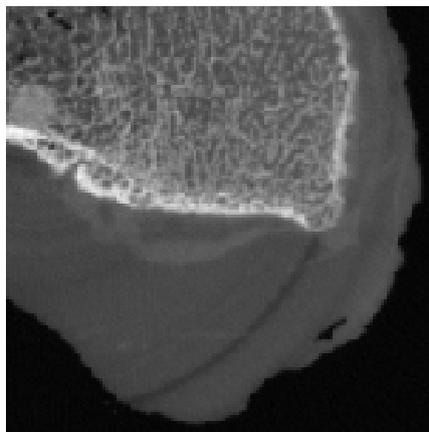
RGN



(C)

U-net

Direct U-net



(D)

(E)

Chapitre 8

Application des méthodes d'apprentissage sur les données réelles

Dans ce chapitre, nous utiliserons les réseaux développés dans les chapitres 6 et 7, pré-entraînés pour la décomposition de matériaux et la reconstruction d'images mono-énergétiques, dans le domaine des projections et dans le domaine image, et nous les testerons avec des données issues d'un prototype clinique de scanner spectral. Nous présenterons les données réelles acquises avec le prototype clinique de scanner spectral. Dans ce chapitre nous étudierons donc la généralisation aux données réelles d'un réseau pré-entraîné pour la décomposition de matériaux dans le domaine des projections, d'un réseau pré-entraîné pour la décomposition directe des projections mono-énergétiques, d'un réseau pré-entraîné pour la décomposition de matériaux dans le domaine image et d'un réseau pré-entraîné pour la reconstruction directe des images mono-énergétiques.

8.1 Introduction

La généralisation aux données réelles permet de tester la viabilité de notre méthode pour une application dans d'autres scanners spectraux. Les données réelles n'ayant pas de vérité terrain pour réaliser un apprentissage avec des données appareillées, il est plus compliqué d'apprendre leur distribution avec des réseaux simples comme U-net. Ainsi, il est plus simple de tester si nos apprentissages à partir de simulations sont généralisables à ces données.

8.2 Acquisitions

Nous utiliserons l'acquisition de scanner spectral de 2018 qui est marquée comme disponible dans la Table 4.1. Cette acquisition est réalisée avec une trajectoire hélicoïdale et en acquisition conique 3D. Nous ne pouvons donc pas utiliser la rétro-projection filtrée comme nous le faisons pour les données simulées en géométrie parallèle. Nous utilisons la méthode du gradient conjugué comme décrite dans le 5.2.2 implémentée dans RTK (Rit et al., 2014) pour reconstruire ces genoux. La géométrie de l'acquisition avec les matrices de projection pour chaque position a été intégrée dans le programme. Nous mettons environ 5h30 à reconstruire chaque bin ou matériau.

La Figure 8.1 montre la reconstruction tomographique des 5 bins d'énergie, nous avons les mêmes échelles de valeurs que pour les simulations visibles Figure 4.11. Contrairement au Chapitre 4, il est difficile de visualiser les données spectrales dans

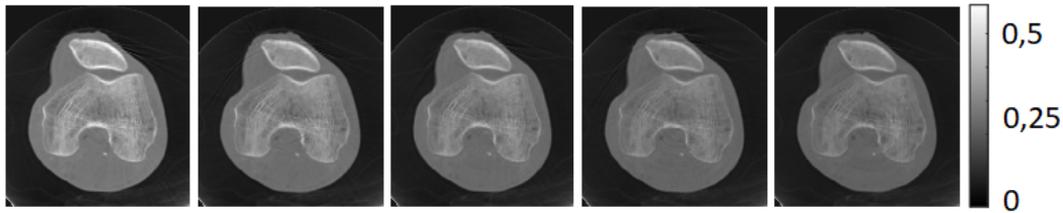


FIGURE 8.1 – Reconstruction tomographique des projections spectrales des données réelles dans les 5 bins d'énergie.

le domaine des projections car chaque projection a très peu de pixels selon l'axe z . Sur ces images, on voit déjà quelques artefacts qui peuvent résulter du durcissement de faisceaux et de problèmes de calibration du détecteur.

8.3 Domaine des projections

Les données que nous injectons dans les réseaux choisis ci-dessous sont la normalisation logarithmique de chaque bin d'énergie comme nous l'avons fait pour les simulations. En effet, pour chaque acquisition, les mesures de plein flux sont disponibles. Ainsi, nous pouvons réaliser la normalisation de la même manière sur les données réelles que sur les données simulées. Une fois les projections de matériaux ou les projections mono-énergétiques décomposées, il faut opérer la reconstruction tomographique qui prendra 11h pour les deux matériaux et 5h30 pour une image mono-énergétique. Cela est un désavantage car à la durée de chaque test, environ 3 minutes, on ajoute les 11h ou 5h30 de reconstruction.

8.3.1 Décomposition de matériaux

Nous avons vu dans le Chapitre 6 que les taux d'apprentissage initiaux les plus performants pour la décomposition de matériaux étaient 10^{-3} et 10^{-2} . Nous allons donc opérer la décomposition de matériaux pour les données réelles dans le domaine des projections avec les réseaux entraînés avec ces deux taux d'apprentissage.

La Figure 8.2 montre la reconstruction tomographique des projections décomposées avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ et la méthode d'apprentissage avec U-net pour des taux d'apprentissage initiaux à 10^{-3} et 10^{-2} . On peut voir que l'on arrive à décomposer les matériaux avec les trois méthodes. On s'attendrait à avoir un maximum des valeurs aux alentours de 1 pour les tissus mous et 1.6 pour l'os.

On trouve des mélanges de matériaux surtout là où l'os est le plus dense comme sur la pointe de la patella. Les méthodes de Gauss-Newton et U-net avec un taux d'apprentissage initial à 10^{-3} présentent moins de mélange de matériaux que U-net avec un taux d'apprentissage initial à 10^{-2} . Ainsi, compte tenu des valeurs attendues et des mélanges de matériaux, la méthode de Gauss-Newton semble donner de meilleurs résultats car le maximum des tissus mous est aux alentours de 1.5 et représente l'os mal décomposé dans l'image de tissus mous. Il semble que la décomposition avec U-net avec un taux d'apprentissage initial à 10^{-2} est plus nette que les deux autres. Sur toutes les images, nous pouvons voir des artefacts en bande sur les zones très dense mais aussi des artefacts circulaires qui doivent être le résultat de pixels mal calibrés sur le détecteur.

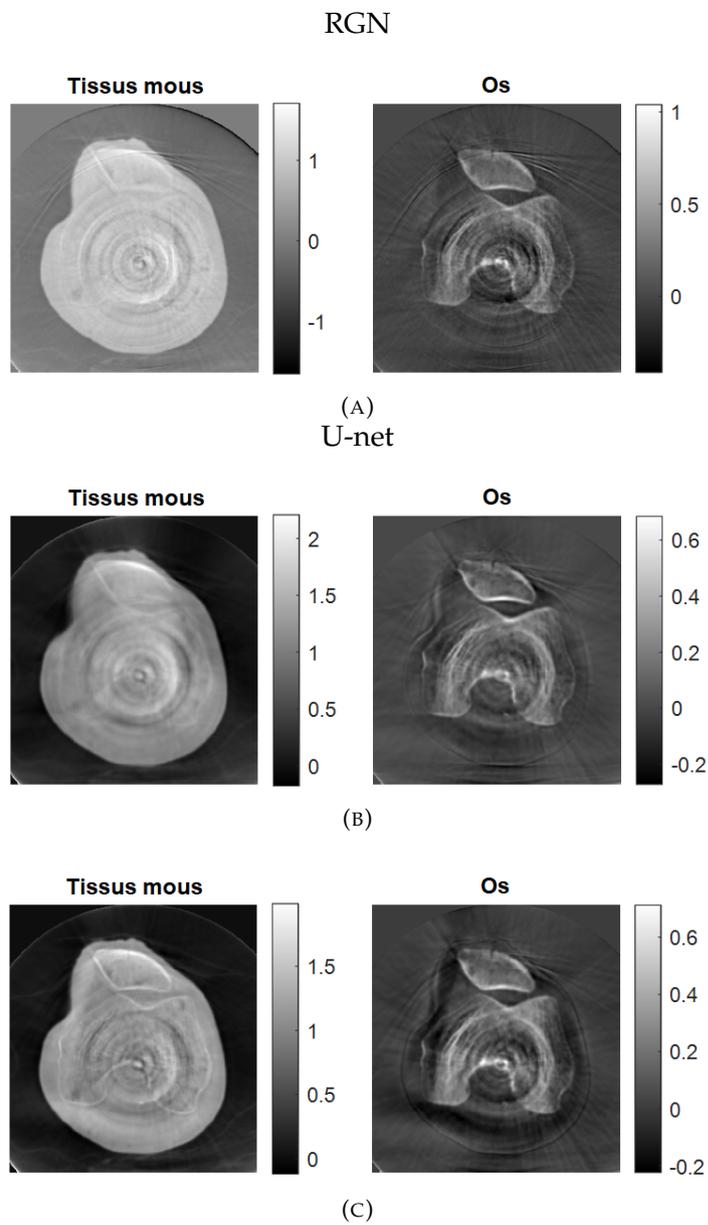


FIGURE 8.2 – Reconstruction tomographique des résultats de la décomposition de matériaux avec U-net dans le domaine des projections. Matériaux décomposés la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), avec U-net avec un taux d'apprentissage à 10^{-3} (B) et 10^{-2} (C).

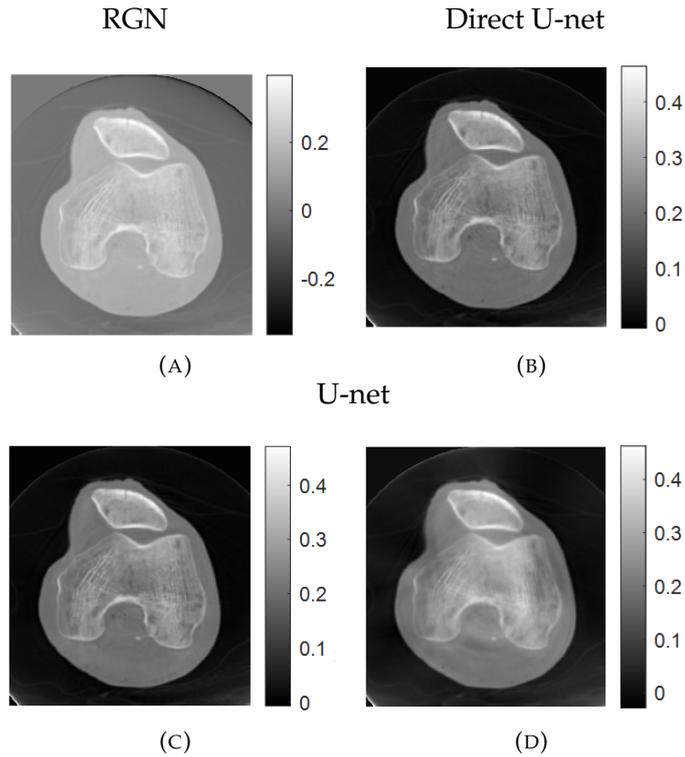


FIGURE 8.3 – Reconstruction tomographique des images mono-énergétiques à 70 keV avec U-net. Images mono-énergétiques à partir des projections décomposées la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), avec U-net avec un taux d'apprentissage initial à 10^{-3} (C) et 10^{-2} (D); ou directement reconstruites avec U-net avec taux d'apprentissage initial à 10^{-1} (B).

8.3.2 Décompositions mono-énergétiques

Nous avons vu dans le Chapitre 6 que le taux d'apprentissage initial le plus performant pour la décomposition des projections mono-énergétique était 10^{-1} , nous allons donc utiliser le réseau entraîné avec ce taux d'apprentissage pour la décomposition des projections mono-énergétiques des données réelles.

Dans la Figure 8.3, on compare les images mono-énergétiques reconstruites à partir de la reconstruction tomographique de la décomposition de matériaux avec la méthode de Gauss-Newton régularisée, avec la méthode d'apprentissage pour les taux d'apprentissage initiaux à 10^{-3} et 10^{-2} et la reconstruction tomographique des projections mono-énergétiques décomposées directement grâce à U-net avec un taux d'apprentissage initial à 10^{-1} .

On remarque que les artefacts en bande et circulaire ont disparu en comparaison des images de la décomposition de matériaux. De la même manière, contrairement à la décomposition de matériaux, ce n'est plus la méthode de Gauss-Newton mais U-net qui donne les valeurs d'atténuation linéaire les plus réalistes. On peut également voir que U-net avec un taux d'apprentissage à 10^{-3} donne une image assez floue où il est difficile de voir la structure de l'os. Au contraire, il est facile d'observer les spécificités de l'os pour la méthode de Gauss-Newton, U-net avec un taux d'apprentissage initial à 10^{-2} et U-net pour décomposer directement les images mono-énergétiques. Ces trois images montrent une qualité similaire à de la tomographie conventionnelle.

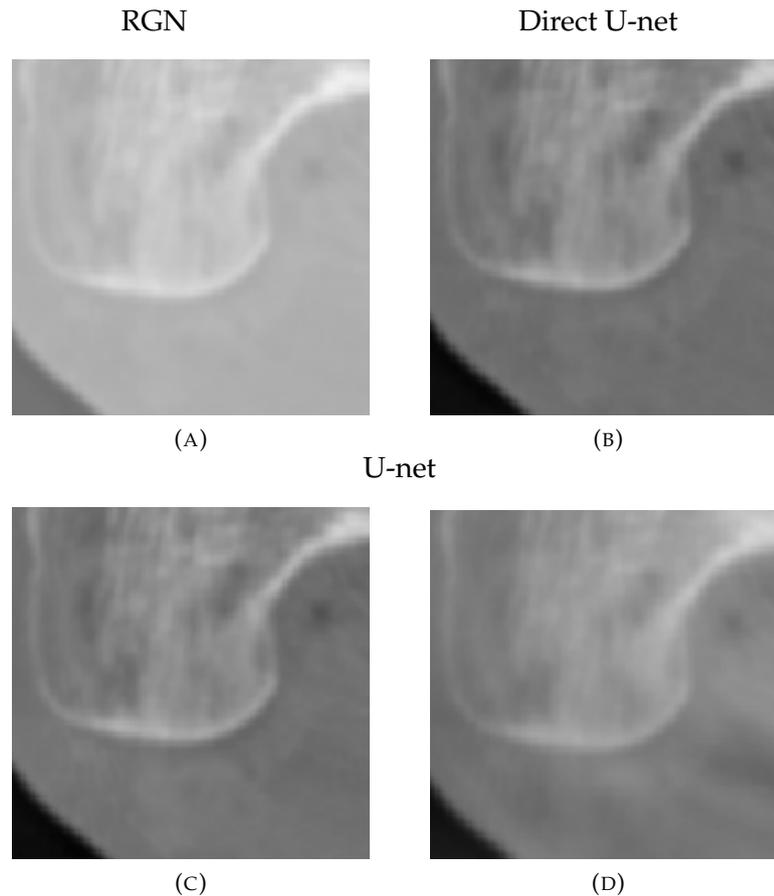


FIGURE 8.4 – Reconstruction tomographique du résultat de la reconstruction des projections mono-énergétiques à 70 keV avec U-net. Images mono-énergétiques à partir des projections décomposées la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), avec U-net avec un taux d'apprentissage initial à 10^{-3} (C) et 10^{-2} (D); ou directement reconstruites avec U-net avec taux d'apprentissage initial à 10^{-1} (B).

Ainsi les images mono-énergétiques, obtenues à partir de la décomposition de matériaux avec des hyperparamètres adéquats et directement obtenues, sont de bonne qualité et montre la capacité de U-net à décomposer les projections mono-énergétiques.

Comme pour les simulations, on cherche à visualiser le cartilage, on va donc zoomer à l'arrière d'un condyle. La Figure 8.4 montre un zoom sur l'arrière d'un condyle des images de la Figure 8.3. On voit que le cartilage est visible à l'oeil nu, bien qu'il soit bien moins contrasté que sur les simulation. Comme pour les images entières, le taux d'apprentissage initial à 10^{-3} pour les images mono-énergétiques à partir de de la décomposition de matériaux, est trop flou pour visualiser correctement ce détail de l'image. De plus, les niveaux de gris pour les résultats avec U-net étant meilleurs, le cartilage est plus simple à visualiser que pour les résultats avec Gauss-Newton.

8.4 Domaine image

Pour le domaine image, nous appliquons la reconstruction tomographique de chaque bin d'énergie individuellement et nous les injectons ensuite dans les réseaux

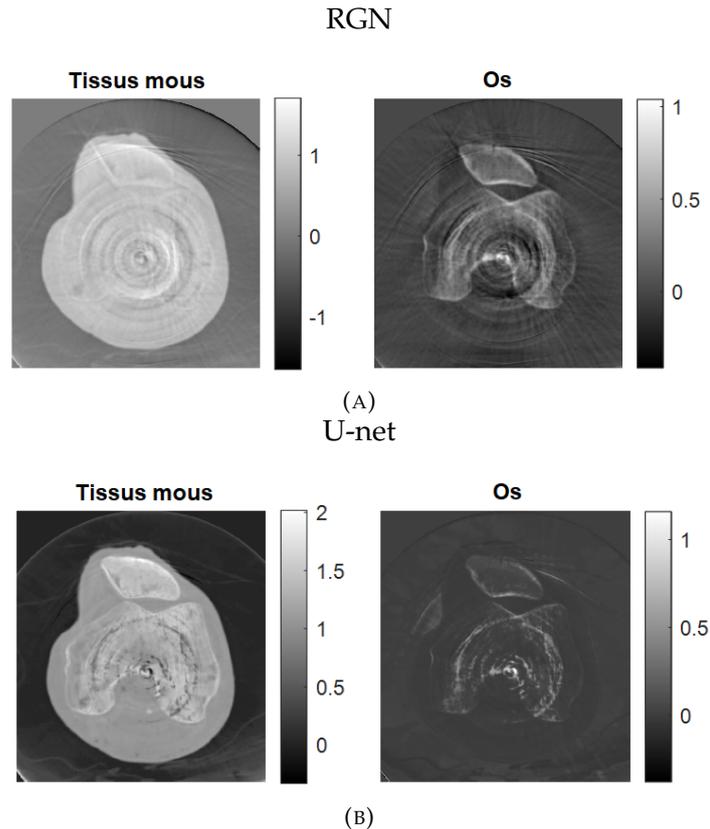


FIGURE 8.5 – Résultats de la décomposition de matériaux avec U-net dans le domaine image. Reconstruction tomographique des projections décomposées avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A) et **matériaux décomposés dans le domaine image avec U-net avec un taux d'apprentissage initial à 10^{-2} .**

pré-entraînés décrits dans chaque partie. Les bins reconstruits ont été préalablement normalisés, comme pour les simulations. L'avantage du domaine image est que nous passons 27h30 pour reconstruire les 5 bins d'énergie, ensuite les calculs prennent 1 minute. On peut donc faire énormément de tests sans avoir la contrainte de temps de la reconstruction.

8.4.1 Décomposition de matériaux

Nous avons vu dans le chapitre précédent que le taux d'apprentissage le plus performant pour la décomposition de matériaux dans le domaine image était 10^{-2} , nous allons donc uniquement tester les données réelles avec le réseau appris avec ce taux d'apprentissage.

La Figure 8.5 compare les matériaux dans le domaine image avec U-net avec un taux d'apprentissage initial à 10^{-2} et la reconstruction tomographique des projections décomposées avec la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$. On peut voir que les valeurs de masse volumique sont toujours meilleures pour la méthode de Gauss-Newton en comparaison de U-net.

De plus, les résultats avec la méthode d'apprentissage ont énormément de mélange de matériaux. Quasiment tout l'os se retrouve dans l'image de tissus mous. Les artefacts de reconstruction circulaires portent aussi préjudice à la décomposition de

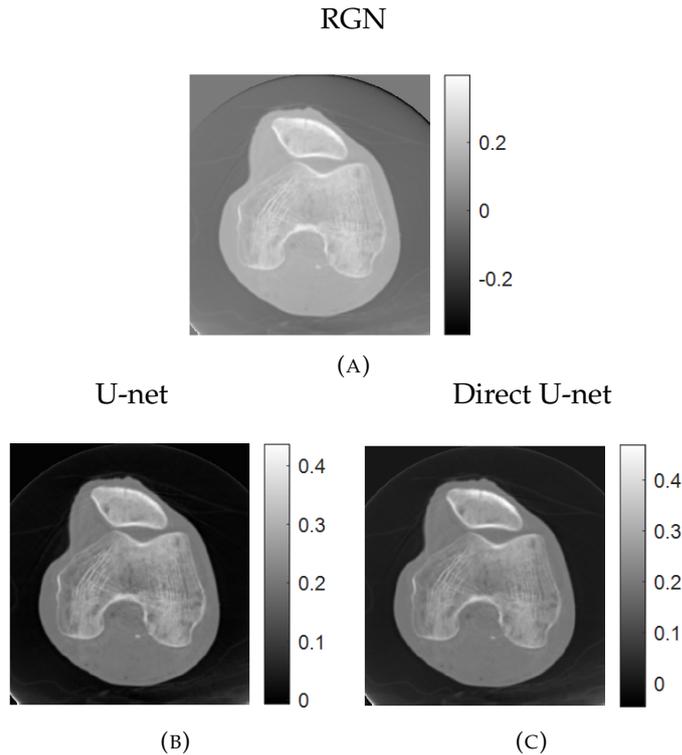


FIGURE 8.6 – Résultats de la reconstruction des images mono-énergétiques à 70 keV avec U-net dans le domaine image. Reconstruction tomographique des projections mono-énergétiques créées à partir des projections décomposées la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), **image mono-énergétique à partir de la décomposition de matériaux dans le domaine image avec U-net avec un taux d'apprentissage initial à 10^{-2}** (B) et **image mono-énergétique à partir de la décomposition directe avec U-net (3)**.

matériaux dans le domaine image, on le voit avec le cercle au centre de l'image qui semble avoir aussi été décomposé.

8.4.2 Reconstruction des images mono-énergétiques

Nous avons pu observer dans le chapitre précédent que la reconstruction directe des images mono-énergétiques était la meilleure avec le taux d'apprentissage initial à 10^{-1} et que les images mono-énergétiques issues de la décompositions étaient très bonnes avec un taux d'apprentissage initial à 10^{-2} . La Figure 8.6 montre la comparaison entre les images mono-énergétiques créées à partir des différentes méthodes : U-net direct avec un taux d'apprentissage initial à 10^{-1} et U-net décomposition de matériaux avec un taux d'apprentissage initial à 10^{-2} .

Comme dans le domaine des projections, les artefacts disparaissent lors de la reconstruction des images mono-énergétiques. De même, les valeurs d'atténuation linéaire sont meilleures pour les images mono-énergétiques créées avec U-net qu'avec la méthode de Gauss-Newton. On peut visualiser sur les deux images les caractéristiques morphologiques de l'os. La reconstruction des images mono-énergétiques de manière directe ou à partir de la décomposition de matériaux ne semble pas influencer beaucoup le résultat.

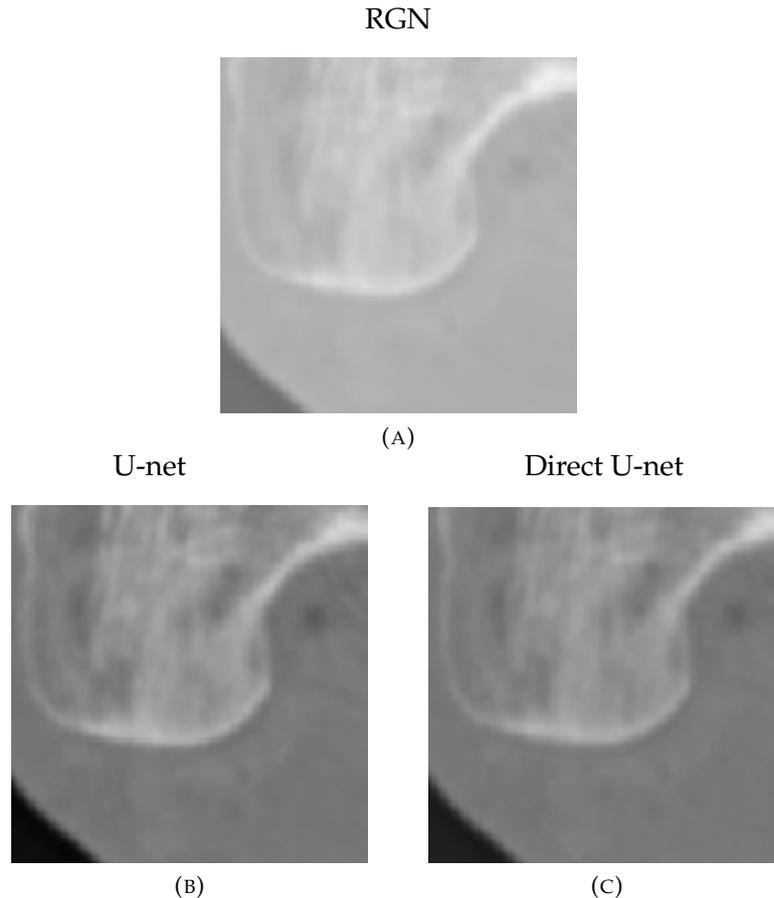


FIGURE 8.7 – Résultats de la reconstruction des images mono-énergétiques à 70 keV avec U-net dans le domaine image. Reconstruction tomographique des projections mono-énergétiques créées à partir des projections de matériaux décomposées la méthode de Gauss-Newton régularisée avec un paramètre de régularisation $\alpha = 0.3$ (A), **image mono-énergétique à partir de la décomposition de matériaux dans le domaine image avec U-net avec un taux d'apprentissage initial à 10^{-2}** (B) et **image mono-énergétique à partir de la décomposition directe avec U-net** (3).

Comme pour l'étude des données réelles dans le domaine des projections, nous allons effectuer un zoom au niveau de l'arrière d'un condyle pour visualiser une partie avec du cartilage. La Figure 8.7 montre un zoom sur l'arrière d'un condyle des images de la Figure 8.6. Comme pour le domaine des projections, les résultats issus de traitements dans le domaine image permettent une visualisation du cartilage à l'oeil nu. Il est également plus visible sur les résultats issus de l'apprentissage que de la méthode de Gauss-Newton.

8.5 Conclusion

On a montré que l'on pouvait opérer des décomposition de matériaux et la reconstruction des images mono-énergétiques sur les données réelles. L'utilisation de données acquises en géométrie conique 3D avec une acquisition hélicoïdale rend les reconstructions tomographiques plus longues et moins précises que sur les simulations en géométrie parallèle. De plus, nous avons pris des paramètres arbitraires

pour la reconstruction. Ainsi celles-ci pourraient être améliorées en réglant au mieux les paramètres pour notre application. Le rendu flou laisse croire que nous avons utilisé un paramètre de régularisation trop important pour la méthode du gradient conjugué.

A cause de la longue durée de reconstruction tomographique, les traitements dans le domaine image sont plus rapides. Pourtant, ce sont les décompositions de matériaux dans le domaine des projections qui sont les meilleures malgré des mélanges de matériaux. Dans les deux domaines, U-net donne de meilleurs résultats que la méthode de Gauss-Newton, pour la reconstruction des images mono-énergétiques. En prenant en compte la qualité de celles-ci, on peut dire que nos algorithmes d'apprentissage, pré-entraînés sur les données simulées, sont généralisables aux données réelles. La Table 8.1 fait le récapitulatif des points forts et faibles des résultats présentés dans ce chapitre.

Il est possible de visualiser le cartilage à l'oeil nu sur les images mono-énergétiques et la reconstruction tomographique des projections mono-énergétiques, qu'elles soit apprises directement ou reconstruites à partir de la décomposition de matériaux. La méthode de Gauss-Newton donne des résultats moins contrastés, le cartilage est donc légèrement moins visible.

Cependant, il faudrait améliorer la qualité des reconstructions pour palier les problèmes dans la décomposition de matériaux. De plus, en ayant plus de données réelles, nous pourrions faire de l'adaptation de domaine pour obtenir de meilleurs résultats encore.

Domaine des projections	Matériaux	GN $\alpha = 0.3$	Bons niveaux de gris et peu de mélange de matériaux ; artefacts circulaires
		U-net $\lambda = 10^{-3}$	Niveaux de gris sur estimés, flou et peu de mélange de matériaux ; artefacts circulaires
		U-net $\lambda = 10^{-2}$	Niveaux de gris sur estimés, moins flou et plus de mélange de matériaux ; artefacts circulaires
	Mono-énergétique	GN $\alpha = 0.3$	Niveaux de gris très sous estimés et détails visibles mais peu de contraste ; peu d'artefacts
		U-net dir $\lambda = 10^{-1}$	Bons niveaux de gris et bonne visualisation des détails ; peu d'artefacts
		U-net dec $\lambda = 10^{-3}$ U-net dec $\lambda = 10^{-2}$	Bons niveaux de gris et détails visibles ; peu d'artefacts Bons niveaux de gris mais flou ; peu d'artefacts
Domaine image	Matériaux	U-net $\lambda = 10^{-2}$	Beaucoup de mélange de matériaux ; beaucoup d'artefacts de reconstruction qui pénalisent la décomposition
	Mono-énergétique	U-net dir $\lambda = 10^{-1}$	Bons niveaux de gris, détails visibles mais légèrement flou ; peu d'artefacts
		U-net dec $\lambda = 10^{-2}$	Niveaux de gris légèrement sous estimés, très bon contraste et détails visibles ; peu d'artefacts

TABLE 8.1 – Récapitulatif des résultats sur les données réelles pour la décomposition de matériaux et mono-énergétique dans le domaine des projections et le domaine image.

Conclusion

La tomographie spectrale une nouvelle modalité d'imagerie à rayons X qui donne une nouvelle dimension énergétique aux données. De celle-ci, il est possible de décomposer les matériaux constitutifs de l'objet et de reconstruire des images mono-énergétiques.

L'objectif de la thèse était de développer des méthodes d'apprentissage profond pour la décomposition de matériaux et la reconstruction des images mono-énergétiques ; et d'utiliser ces méthodes pour une application à l'arthrose du genou.

Dans le Chapitre 4 nous avons présenté la création de fantômes de genoux et la simulation des données spectrales pour créer les bases de données à différentes résolutions et avec un nombre différent de genoux.

Dans le Chapitre 5 nous avons présenté les méthodes d'apprentissage profond développées pour la décomposition de matériaux et mono-énergétique. Le réseau a été choisi pour avoir déjà montré sa performance pour la décomposition de matériaux dans le domaine image. Nous avons aussi présenté la méthode de Gauss-Newton régularisée, qui est une méthode basée sur le modèle pour la décomposition de matériaux dans le domaine des projections.

Dans les Chapitres 6 et 7 nous avons présenté les résultats de la décomposition de matériaux et de la reconstruction des images mono-énergétiques dans le domaine des projections et dans le domaine image à partir des simulations des données spectrales des fantômes présentés au Chapitre 4. Les images mono-énergétiques sont soit reconstruites à partir de la décomposition de matériaux, soit directement avec U-net.

Nous montrons que les décompositions avec les méthodes d'apprentissage dans le domaine des projections ou image donnent de meilleurs résultats que la méthode de l'état de l'art basée sur le modèle. De plus, la décomposition dans le domaine image donne de meilleurs résultats que dans le domaine des projections. De même, les images mono-énergétiques reconstruites à partir de la décomposition de matériaux avec U-net dans le domaine des projections ou dans le domaine image donnent de meilleurs résultats que celles reconstruites à partir de la décomposition de matériaux réalisée avec la méthode de Gauss-Newton.

La décomposition directe des projections mono-énergétiques donne des résultats légèrement moins bons que les autres sus-citées, mais en restant dans les mêmes ordres de grandeur. Cette décomposition a quand même l'avantage de s'émanciper du modèle de matériaux. Dans le domaine image, la reconstruction directe des images mono-énergétiques donne des résultats similaires à ceux dans le domaine des projections et est toujours surpassée par la reconstruction des images mono-énergétiques à partir des décompositions de matériaux. Sur les images mono-énergétiques, il est possible de visualiser le cartilage à l'oeil nu, ce qui est impossible en tomographie classique. Cela offre de belles perspectives pour l'étude de l'arthrose à l'aide de la tomographie X spectrale.

Les méthodes par apprentissage ont montré que dans la plupart des cas, elles donnaient de meilleurs résultats que la méthode basée sur le modèle. De plus, ces méthodes ont l'avantage de ne nécessiter qu'un seul apprentissage par application qui dure environ 6h pour l'entraînement dans le domaine des projections et 4h dans le domaine image. Ensuite, chaque test ne demande que 1 min pour 202 coupes ou 3 min pour 720 projections alors que la méthode de Gauss-Newton demande environ 2 heures par test. Ces données sont référencées dans la Table 8.2.

Dans le Chapitre 8 nous avons appliqué les méthodes d'apprentissage, pré-entraînées sur les simulations, aux données réelles. Nous avons choisi les réseaux entraînés avec les hyperparamètres ayant donné les meilleurs résultats pour les simulations. Nous avons testés ces réseaux sur des données réelles de genoux issues d'un prototype clinique de scanner spectral. Les calculs dans ce chapitre ont pris énormément de temps car chaque canal nécessite 5h30 de reconstruction tomographique alors que celui-ci est uniquement de 46s pour les données simulées. Ainsi, les calculs dans le domaine image sont plus rentables en terme de temps car on reconstruit les bins une fois pour toute alors que dans le domaine des projections, nous devons reconstruire chaque résultat.

Pourtant, la décomposition dans le domaine des projections rend de meilleurs résultats en comparaison du domaine image. Les images mono-énergétiques décomposées dans le domaine des projections ou dans le domaine image donnent des images comparables à la tomographie classique et sont plus lisibles avec les méthodes d'apprentissage qu'avec la méthode basée sur le modèle. Il est également possible de voir le cartilage à l'oeil nu sur les images mono-énergétiques même si le contraste est moins fort que pour les simulations.

Résultats soulignés Les 4 réseaux présentés au Chapitre 5, la décomposition de matériaux et de projections mono-énergétiques dans le domaine des projections et la décomposition de matériaux et d'images mono-énergétiques dans le domaine image, donnent de bons résultats. La décomposition de matériaux fonctionne mieux dans le domaine image que dans le domaine des projections pour les simulations. La décompositions des images / projections mono-énergétiques est similaire dans les deux domaines. Il est possible d'obtenir des résultats satisfaisants dans le domaine image avec ces 4 méthodes. Que ce soit sur les simulations ou les données réelles, dans le domaine image ou des projections, il est possible de visualiser le cartilage à l'oeil nu sur les reconstructions mono-énergétiques.

Comme présenté dans la Table 8.2, les tests sont 33 à 40 fois plus rapides avec U-net qu'avec la méthode de Gauss-Newton. Pour la méthode de Gauss-Newton et U-net dans le domaine des projections, on a pris les temps de calcul pour une projection de taille 643×202 (129'886 pixels); et pour U-net dans le domaine image, une coupe de taille 635×635 (403'225 pixels). Le changement de nombre de pixels entre le domaine des projections et le domaine image peut expliquer que le domaine image est légèrement plus lent. On y gagne quand même au bout de plusieurs tests car le temps pris par la reconstruction tomographique est ajouté une fois pour toute pour le domaine image et multiplié par le nombre de test dans le domaine des projections.

Perspectives Nous avons peu de données en considérant la nécessité d'un grand nombre de données pour l'apprentissage profond. La première perspective serait la création de nouveaux fantômes et la récupération de plus de données réelles pour que nos réseaux puissent apprendre une plus grande variabilité de données.

	Entraînement	Test
Gauss-Newton		10s
U-net Projection	30s	0.25s
U-net Image	71s	0.30s

TABLE 8.2 – Temps de calcul pour une projection/coupe par méthode.

Nous nous sommes cantonnés à une seule architecture de réseaux de neurones, or (Gong et al., 2020) a montré que d’autres réseaux pouvaient surpasser U-net pour la décomposition de matériaux. Ainsi, il pourrait être intéressant de tester de nouvelles architectures de réseaux. Par exemple, les réseaux antagonistes génératifs (Goodfellow et al., 2014) pourraient nous permettre d’ajouter les données réelles dans l’apprentissage en utilisant des données non-appareillées.

Bien que changer d’architecture pourrait améliorer nos résultats, un réglage approfondi des hyperparamètres de l’architecture que nous utilisons pourrait très aisément donner de bien meilleurs résultats. Il est possible d’utiliser une version de U-net, avec des paramètres et des hyperparamètres optimisés, nommée nnU-net (Isensee et al., 2018).

Le travail sur les données réelles est pénalisé par la qualité des reconstructions tomographiques. Ainsi, le réglage des paramètres de l’algorithme de reconstruction pourrait permettre d’améliorer les images reconstruites et donc les décompositions de matériaux et d’images mono-énergétiques.

Toujours pour les données réelles, les réseaux sont pré-entraînés sur des simulations qui ne prennent certainement pas en compte toutes les particularités des données réelles. Ainsi, faire de l’adaptation de domaine en incluant des données réelles lors de l’apprentissage pourrait améliorer les résultats sur les données réelles.

On peut également réaliser du post-traitement sur les images pour, par exemple, segmenter le cartilage pour pouvoir quantifier l’épaisseur de cartilage. Cette quantification pourra grandement nous aider pour la détection anticipée de l’arthrose. Nous pouvons aussi faire du débruitage comme a pu le faire (Abascal et al., 2021a) sur les données réelles de genoux.

Travaux personnels

- Abascal, J., S. Bussod, N. Ducros, S. Si-Mohamed, P. Douek, C. Chappard et F. Peyrin (jan. 2021a). « A residual U-Net network with image prior for 3D image denoising ». In : *2020 28th European Signal Processing Conference (EUSIPCO)*. DOI : [10.23919/Eusipco47968.2020.9287607](https://doi.org/10.23919/Eusipco47968.2020.9287607).
- Abascal, J., N. Ducros, V. Pronina, S. Rit, P.-A. Rodesch, T. Broussaud, S. Bussod, P. Douek, A. Hauptmann, S. Arridge et F. Peyrin (2021b). « Material Decomposition in Spectral CT Using Deep Learning : A Sim2Real Transfer Approach ». In : *IEEE Access* 9, p. 25632-25647. ISSN : 2169-3536. DOI : [10.1109/ACCESS.2021.3056150](https://doi.org/10.1109/ACCESS.2021.3056150).
- Bussod, S., J. Abascal, S. Arridge, A. Hauptmann, C. Chappard, N. Ducros et F. Peyrin (jan. 2021). « Convolutional Neural Network for Material Decomposition in Spectral CT Scans ». In : *2020 28th European Signal Processing Conference (EUSIPCO)*, p. 1259-1263. DOI : [10.23919/Eusipco47968.2020.9287781](https://doi.org/10.23919/Eusipco47968.2020.9287781).
- Bussod, S., J. Abascal, N. Ducros, C. Olivier, S. Si-Mohamed, P. Douek, C. Chappard et F. Peyrin (avr. 2019). « Human Knee Phantom for Spectral CT : Validation of a Material Decomposition Algorithm ». In : *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, p. 1170-1173. DOI : [10.1109/ISBI.2019.8759192](https://doi.org/10.1109/ISBI.2019.8759192).
- Chappard, C., J. Abascal, S. Bussod, S. Uk, S. Si-Mohamed, P. Douek et F. Peyrin (avr. 2020). « Feasibility of spectral computed tomography to assess knee cartilage ». English. In : *Osteoarthritis and Cartilage* 28, S277-S278. ISSN : 1063-4584, 1522-9653. DOI : [10.1016/j.joca.2020.02.442](https://doi.org/10.1016/j.joca.2020.02.442).

Bibliographie

- Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu et X. Zheng (2015). *TensorFlow : Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org.
- Abascal, J., S. Bussod, N. Ducros, S. Si-Mohamed, P. Douek, C. Chappard et F. Peyrin (jan. 2021a). « A residual U-Net network with image prior for 3D image denoising ». In : *2020 28th European Signal Processing Conference (EUSIPCO)*. DOI : [10.23919/Eusipco47968.2020.9287607](https://doi.org/10.23919/Eusipco47968.2020.9287607).
- Abascal, J., N. Ducros, V. Pronina, S. Rit, P.-A. Rodesch, T. Broussaud, S. Bussod, P. Douek, A. Hauptmann, S. Arridge et F. Peyrin (2021b). « Material Decomposition in Spectral CT Using Deep Learning : A Sim2Real Transfer Approach ». In : *IEEE Access* 9, p. 25632-25647. ISSN : 2169-3536. DOI : [10.1109/ACCESS.2021.3056150](https://doi.org/10.1109/ACCESS.2021.3056150).
- Adler, J. et O. Öktem (nov. 2017). « Solving ill-posed inverse problems using iterative deep neural networks ». en. In : *Inverse Problems* 33.12, p. 124007. ISSN : 0266-5611. DOI : [10.1088/1361-6420/aa9581](https://doi.org/10.1088/1361-6420/aa9581).
- (juin 2018). « Learned Primal-Dual Reconstruction ». In : *IEEE Transactions on Medical Imaging* 37.6, p. 1322-1332. ISSN : 1558-254X. DOI : [10.1109/TMI.2018.2799231](https://doi.org/10.1109/TMI.2018.2799231).
- Alvarez, R. et A. Macovski (1976). « Energy-selective reconstructions in X-ray computerised tomography ». en. In : *Physics in Medicine & Biology* 21.5, p. 733. ISSN : 0031-9155. DOI : [10.1088/0031-9155/21/5/002](https://doi.org/10.1088/0031-9155/21/5/002).
- Avizo Software | Avizo2D Software - FR. fr.
- Barnes, G., R. Sones, M. Tesic, D. Morgan et J. Sanders (août 1985). « Detector for dual-energy digital radiography. » In : *Radiology* 156.2, p. 537-540. ISSN : 0033-8419. DOI : [10.1148/radiology.156.2.4011921](https://doi.org/10.1148/radiology.156.2.4011921).
- Bishop, C. (1995). *Neural Networks for Pattern Recognition*. USA : Oxford University Press, Inc. ISBN : 978-0-19-853864-6.
- Bottou, L. (jan. 1999). « On-line Learning and Stochastic Approximations ». en. In : *On-Line Learning in Neural Networks*. Sous la dir. de D. Saad. 1^{re} éd. Cambridge University Press, p. 9-42. ISBN : 978-0-521-65263-6 978-0-521-11791-3 978-0-511-56992-0. DOI : [10.1017/CB09780511569920.003](https://doi.org/10.1017/CB09780511569920.003).
- Boussel, L., P. Coulon, A. Thran, E. Roessl, G. Martens, M. Sigovan et P. Douek (juil. 2014). « Photon counting spectral CT component analysis of coronary artery atherosclerotic plaque samples ». EN. In : *The British Journal of Radiology*. DOI : [10.1259/bjr.20130798](https://doi.org/10.1259/bjr.20130798).
- Bowles, C., L. Chen, R. Guerrero, P. Bentley, R. Gunn, A. Hammers, D. Dickie, M. Hernández, J. Wardlaw et D. Rueckert (oct. 2018). « GAN Augmentation : Augmenting Training Data using Generative Adversarial Networks ». In : *arXiv :1810.10863 [cs]*.

- Bracewell, R. N. (1956). « Strip Integration in Radio Astronomy ». en. In : *Australian Journal of Physics* 9.2, p. 198-217. ISSN : 1446-5582. DOI : [10.1071/ph560198](https://doi.org/10.1071/ph560198).
- Braun, H. et G. Gold (août 2012). « Diagnosis of Osteoarthritis : Imaging ». In : *Bone* 51.2, p. 278-288. ISSN : 8756-3282. DOI : [10.1016/j.bone.2011.11.019](https://doi.org/10.1016/j.bone.2011.11.019).
- Brooks, R. A. (oct. 1977). « A quantitative theory of the Hounsfield unit and its application to dual energy scanning ». eng. In : *Journal of Computer Assisted Tomography* 1.4, p. 487-493. ISSN : 0363-8715. DOI : [10.1097/00004728-197710000-00016](https://doi.org/10.1097/00004728-197710000-00016).
- Bussod, S., J. Abascal, S. Arridge, A. Hauptmann, C. Chappard, N. Ducros et F. Peyrin (jan. 2021). « Convolutional Neural Network for Material Decomposition in Spectral CT Scans ». In : *2020 28th European Signal Processing Conference (EUSIPCO)*, p. 1259-1263. DOI : [10.23919/Eusipco47968.2020.9287781](https://doi.org/10.23919/Eusipco47968.2020.9287781).
- Bussod, S., J. Abascal, N. Ducros, C. Olivier, S. Si-Mohamed, P. Douek, C. Chappard et F. Peyrin (avr. 2019). « Human Knee Phantom for Spectral CT : Validation of a Material Decomposition Algorithm ». In : *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, p. 1170-1173. DOI : [10.1109/ISBI.2019.8759192](https://doi.org/10.1109/ISBI.2019.8759192).
- Buzug, T. (2011). « Computed Tomography ». en. In : *Springer Handbook of Medical Technology*. Sous la dir. de R. Kramme, K.-P. Hoffmann et R. S. Pozos. Springer Handbooks. Berlin, Heidelberg : Springer, p. 311-342. ISBN : 978-3-540-74658-4. DOI : [10.1007/978-3-540-74658-4_16](https://doi.org/10.1007/978-3-540-74658-4_16).
- Cardiovascular diseases*. en.
- Chappard, C., J. Abascal, S. Bussod, S. Uk, S. Si-Mohamed, P. Douek et F. Peyrin (avr. 2020). « Feasibility of spectral computed tomography to assess knee cartilage ». English. In : *Osteoarthritis and Cartilage* 28, S277-S278. ISSN : 1063-4584, 1522-9653. DOI : [10.1016/j.joca.2020.02.442](https://doi.org/10.1016/j.joca.2020.02.442).
- Chen, H., Y. Zhang, Y. Chen, J. Zhang, W. Zhang, H. Sun, Y. Lv, P. Liao, J. Zhou et G. Wang (juin 2018). « LEARN : Learned Experts' Assessment-Based Reconstruction Network for Sparse-Data CT ». In : *IEEE Transactions on Medical Imaging* 37.6, p. 1333-1347. ISSN : 1558-254X. DOI : [10.1109/TMI.2018.2805692](https://doi.org/10.1109/TMI.2018.2805692).
- Chen, H., Y. Zhang, W. Zhang, P. Liao, K. Li, J. Zhou et G. Wang (fév. 2017). « Low-dose CT via convolutional neural network ». EN. In : *Biomedical Optics Express* 8.2, p. 679-694. ISSN : 2156-7085. DOI : [10.1364/BOE.8.000679](https://doi.org/10.1364/BOE.8.000679).
- Chen, Z. et L. Li (2017). « Preliminary Research on Multi-Material Decomposition of Spectral CT Using Deep Learning ». In : *The 14th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, p. 52-526.
- Chlewicki, W. et C. Badaea (2001). « Cone based 3D Reconstruction : a FDK - SART Comparison for Limited Number of Projections ». en. In : p. 3.
- Chou, H., T. Chin et W. Peh (mar. 2017). « Dual-energy CT in gout - A review of current concepts and applications ». eng. In : *Journal of Medical Radiation Sciences* 64.1, p. 41-51. ISSN : 2051-3909. DOI : [10.1002/jmrs.223](https://doi.org/10.1002/jmrs.223).
- Clark, D., M. Holbrook et C. Badaea (2018). « Multi-energy CT decomposition using convolutional neural networks ». In : *Medical Imaging 2018 : Physics of Medical Imaging*. T. 10573. SPIE, p. 59.
- Clark, D. et C. Badaea (oct. 2014). « Spectral diffusion : an algorithm for robust material decomposition of spectral CT data ». en. In : *Physics in Medicine and Biology* 59.21, p. 6445-6466. ISSN : 0031-9155. DOI : [10.1088/0031-9155/59/21/6445](https://doi.org/10.1088/0031-9155/59/21/6445).
- Cormode, D., E. Roessl, A. Thran, T. Skajaa, R. Gordon, J.-P. Schlomka, V. Fuster, E. Fisher, W. Mulder, R. Proksa et Z. Fayad (sept. 2010). « Atherosclerotic plaque composition : analysis with multicolor CT and targeted gold nanoparticles ». eng. In : *Radiology* 256.3, p. 774-782. ISSN : 1527-1315. DOI : [10.1148/radiol.10092473](https://doi.org/10.1148/radiol.10092473).

- D'Errico, J. (2009). *A Matlab function to compute the attenuation coefficient*.
- Dice, L. (1945). « Measures of the Amount of Ecologic Association Between Species ». en. In : *Ecology* 26.3, p. 297-302. ISSN : 1939-9170. DOI : [10.2307/1932409](https://doi.org/10.2307/1932409).
- Digital Radiology (fév. 2013). *Digital Radiology : Global Transition of the X-ray Image Capture Process*. en.
- Ducros, N., J. Abascal, B. Sixou, S. Rit et F. Peyrin (2017). « Regularization of Non-linear Decomposition of Spectral X-ray Projection Images ». In : *Medical Physics* 44.9, e174-e187.
- Ehn, S., T. SELLERER, K. MECHLEM, A. FEHRINGER, M. EPPEL, J. HERZEN, F. PFEIFFER et P. NOËL (déc. 2016). « Basis material decomposition in spectral CT using a semi-empirical, polychromatic adaption of the Beer-Lambert model ». en. In : *Physics in Medicine and Biology* 62.1, N1-N17. ISSN : 0031-9155. DOI : [10.1088/1361-6560/aa4e5c](https://doi.org/10.1088/1361-6560/aa4e5c).
- Euler, A., A. Parakh, A. Falkowski, S. Manneck, D. Dashti, B. Krauss, Z. Szucs-Farkas et S. Schindera (août 2016). « Initial Results of a Single-Source Dual-Energy Computed Tomography Technique Using a Split-Filter : Assessment of Image Quality, Radiation Dose, and Accuracy of Dual-Energy Applications in an In Vitro and In Vivo Study ». eng. In : *Investigative Radiology* 51.8, p. 491-498. ISSN : 1536-0210. DOI : [10.1097/RLI.000000000000257](https://doi.org/10.1097/RLI.000000000000257).
- Feldkamp, L., L. Davis et J. Kress (juin 1984). « Practical cone-beam algorithm ». EN. In : *JOSA A* 1.6, p. 612-619. ISSN : 1520-8532. DOI : [10.1364/JOSAA.1.000612](https://doi.org/10.1364/JOSAA.1.000612).
- Garcelon, C., J. Abascal, C. Olivier, S. Si-Mohamed, S. Uk, H. K. Ea, L. Bousset, P. Douek, F. Peyrin et C. Chappard (déc. 2020). « Analyse quantitative morphologique du cartilage à partir du scanner spectral à comptage de photons ». fr. In : *Revue du Rhumatisme*. 33e Congrès Français de Rhumatologie 87, A261. ISSN : 1169-8330. DOI : [10.1016/j.rhum.2020.10.470](https://doi.org/10.1016/j.rhum.2020.10.470).
- Gauss, C. F. (1809). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium auctore Carolo Friderico Gauss*. la. sumtibus Frid. Perthes et I.H. Besser.
- Genant, H. et D. Boyd (déc. 1977). « Quantitative bone mineral analysis using dual energy computed tomography ». eng. In : *Investigative Radiology* 12.6, p. 545-551. ISSN : 0020-9996. DOI : [10.1097/00004424-197711000-00015](https://doi.org/10.1097/00004424-197711000-00015).
- Ghani, M. et W. Karl (août 2019). « Fast Enhanced CT Metal Artifact Reduction using Data Domain Deep Learning ». en. In : *arXiv :1904.04691 [cs]*.
- Gong, H., S. Tao, K. Rajendran, W. Zhou, C. McCollough et S. Leng (déc. 2020). « Deep-learning-based direct inversion for material decomposition ». In : *Medical physics* 47.12, p. 6294-6309. ISSN : 0094-2405. DOI : [10.1002/mp.14523](https://doi.org/10.1002/mp.14523).
- Goodfellow, I., J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville et Y. Bengio (juin 2014). « Generative Adversarial Networks ». In : *arXiv :1406.2661 [cs, stat]*.
- Gupta, H., K. Jin, H. Nguyen, M. McCann et M. Unser (juin 2018). « CNN-Based Projected Gradient Descent for Consistent CT Image Reconstruction ». In : *IEEE Transactions on Medical Imaging* 37.6, p. 1440-1453. ISSN : 1558-254X. DOI : [10.1109/TMI.2018.2832656](https://doi.org/10.1109/TMI.2018.2832656).
- Hadamard, J. (1923). *Lectures on Cauchy's Problem in Linear Partial Differential Equations, by Jacques Hadamard ...* en. Yale University Press.
- Hansen, P. (jan. 1998). *Rank-Deficient and Discrete Ill-Posed Problems*. Mathematical Modeling and Computation. Society for Industrial et Applied Mathematics. ISBN : 978-0-89871-403-6. DOI : [10.1137/1.9780898719697](https://doi.org/10.1137/1.9780898719697).
- Hassan, H., A. Negm, M. Zahran et O. Saavedra (déc. 2015). « Assessment of Artificial Neural Network for Bathymetry Estimation using High Resolution Satellite

- Imagery in Shallow Lakes : case study El Burullus Lake. » In : *International Water Technology Journal* 5.
- Heller, N., N. Sathianathan, A. Kalapara, E. Walczak, K. Moore, H. Kaluzniak, J. Rosenberg, P. Blake, Z. Rengel, M. Oestreich, J. Dean, M. Tradewell, A. Shah, R. Tejpaul, Z. Edgerton, M. Peterson, S. Raza, S. Regmi, N. Papanikolopoulos et C. Weight (mar. 2020). « The KiTS19 Challenge Data : 300 Kidney Tumor Cases with Clinical Context, CT Semantic Segmentations, and Surgical Outcomes ». In : *arXiv :1904.00445 [cs, q-bio, stat]*.
- Hestenes, M. et E. Stiefel (1952). « Methods of Conjugate Gradients for Solving Linear Systems ». en. In : p. 28.
- Hickethier, T., B. Baeßler, J. Kroeger, J. Doerner, G. Pahn, D. Maintz, G. Michels et A. Bunck (fév. 2017). « Monoenergetic reconstructions for imaging of coronary artery stents using spectral detector CT : In-vitro experience and comparison to conventional images ». eng. In : *Journal of Cardiovascular Computed Tomography* 11.1, p. 33-39. ISSN : 1876-861X. DOI : [10.1016/j.jcct.2016.12.005](https://doi.org/10.1016/j.jcct.2016.12.005).
- Hounsfield, G. (déc. 1973). « Computerized transverse axial scanning (tomography) : Part 1. Description of system ». In : *The British Journal of Radiology* 46.552, p. 1016-1022. ISSN : 0007-1285. DOI : [10.1259/0007-1285-46-552-1016](https://doi.org/10.1259/0007-1285-46-552-1016).
- Hubbell, J. et S. Seltzer (1995). *Tables of x-ray mass attenuation coefficients and mass energy-absorption coefficients 1 keV to 20 MeV for elements Z = 1 to 92 and 48 additional substances of dosimetry interest*. en. Rapp. tech. NIST IR 5632. Gaithersburg, MD : National Institute of Standards et Technology, NIST IR 5632. DOI : [10.6028/NIST.IR.5632](https://doi.org/10.6028/NIST.IR.5632).
- Huy, Q. et S. Molloy (jan. 2011). « Least squares parameter estimation methods for material decomposition with energy discriminating detectors ». eng. In : *Medical Physics* 38.1, p. 245-255. ISSN : 0094-2405. DOI : [10.1118/1.3525840](https://doi.org/10.1118/1.3525840).
- Ioffe, S. et C. Szegedy (mar. 2015). « Batch Normalization : Accelerating Deep Network Training by Reducing Internal Covariate Shift ». In : *arXiv :1502.03167 [cs]*.
- Isensee, F., J. Petersen, A. Klein, D. Zimmerer, P. Jaeger, S. Kohl, J. Wasserthal, G. Koehler, T. Norajitra, S. Wirkert et K. Maier-Hein (sept. 2018). « nnU-Net : Self-adapting Framework for U-Net-Based Medical Image Segmentation ». In : *arXiv :1809.10486 [cs]*. arXiv : 1809.10486.
- Issues in Osteoarthritis Care* (juin 2018). *Issues in Osteoarthritis Care : Concepts and Controversies*.
- Jain, V. et S. Seung (2009). « Natural Image Denoising with Convolutional Networks ». In : *Advances in Neural Information Processing Systems*. Sous la dir. de D. Koller, D. Schuurmans, Y. Bengio et L. Bottou. T. 21. Curran Associates, Inc.
- Jin, K., M. McCann, E. Froustey et M. Unser (nov. 2016). « Deep Convolutional Neural Network for Inverse Problems in Imaging ». In : *arXiv :1611.03679 [cs]*. DOI : [10.1109/TIP.2017.2713099](https://doi.org/10.1109/TIP.2017.2713099).
- Jordan, C. (1881). « Sur la série de Fourier ». French. In : *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences, Paris* 92. ISSN : 0001-4036.
- Kalender, W., W. Perman, J. Vetter et E. Klotz (juin 1986). « Evaluation of a prototype dual-energy computed tomographic apparatus. I. Phantom studies ». eng. In : *Medical Physics* 13.3, p. 334-339. ISSN : 0094-2405. DOI : [10.1118/1.595958](https://doi.org/10.1118/1.595958).
- Kang, E., W. Chang, J. Yoo et J. Ye (juin 2018). « Deep Convolutional Framelet Denoising for Low-Dose CT via Wavelet Residual Network ». In : *IEEE Transactions on Medical Imaging* 37.6, p. 1358-1369. ISSN : 1558-254X. DOI : [10.1109/TMI.2018.2823756](https://doi.org/10.1109/TMI.2018.2823756).

- Kang, E., J. Min et J. Ye (2017). « A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction ». In : *Med. Phys.* 44.10, e360-e375.
- Kao, P.-Y., S. Shailja, J. Jiang, A. Zhang, A. Khan, J. Chen et B. Manjunath (2020). « Improving Patch-Based Convolutional Neural Networks for MRI Brain Tumor Segmentation by Leveraging Location Information ». English. In : *Frontiers in Neuroscience* 13. ISSN : 1662-453X. DOI : [10.3389/fnins.2019.01449](https://doi.org/10.3389/fnins.2019.01449).
- Karras, T., T. Aila, S. Laine et J. Lehtinen (fév. 2018). « Progressive Growing of GANs for Improved Quality, Stability, and Variation ». en. In : *arXiv :1710.10196 [cs, stat]*.
- Kawata, S. et O. Nalcioglu (juin 1985). « Constrained Iterative Reconstruction by the Conjugate Gradient Method ». In : *IEEE Transactions on Medical Imaging* 4.2, p. 65-71. ISSN : 1558-254X. DOI : [10.1109/TMI.1985.4307698](https://doi.org/10.1109/TMI.1985.4307698).
- Kellgren, J. et J. Lawrence (déc. 1957). « Radiological assessment of osteo-arthritis ». eng. In : *Annals of the Rheumatic Diseases* 16.4, p. 494-502. ISSN : 0003-4967.
- Kingma, D. et J. Ba (2015). « Adam : A Method for Stochastic Optimization ». In : ICLR, San Diego.
- Klein, O. et Y. Nishina (nov. 1929). « Über die Streuung von Strahlung durch freie Elektronen nach der neuen relativistischen Quantendynamik von Dirac ». de. In : *Zeitschrift für Physik* 52.11-12, p. 853-868. ISSN : 0044-3328. DOI : [10.1007/BF01366453](https://doi.org/10.1007/BF01366453).
- Konno, T. et M. Iwazume (juil. 2018). « Icing on the Cake : An Easy and Quick Post-Learnig Method You Can Try After Deep Learning ». In : *arXiv :1807.06540 [cs, stat]*.
- Krizhevsky, A., I. Sutskever et G. Hinton (mai 2017). « ImageNet classification with deep convolutional neural networks ». In : *Communications of the ACM* 60.6, p. 84-90. ISSN : 0001-0782. DOI : [10.1145/3065386](https://doi.org/10.1145/3065386).
- Le Cun, Y. et F. Fogelman-Soulié (1987). « Modèles connexionnistes de l'apprentissage ». fre. In : *Intellectica* 2.1, p. 114-143. DOI : [10.3406/intel.1987.1804](https://doi.org/10.3406/intel.1987.1804).
- Lee, H., J. Lee, H. Kim, B. Cho et S. Cho (mar. 2019). « Deep-Neural-Network-Based Sinogram Synthesis for Sparse-View CT Image Reconstruction ». en. In : *IEEE Transactions on Radiation and Plasma Medical Sciences* 3.2, p. 109-119. ISSN : 2469-7311, 2469-7303. DOI : [10.1109/TRPMS.2018.2867611](https://doi.org/10.1109/TRPMS.2018.2867611).
- Lefkimmiatis, S. (juin 2018). « Universal Denoising Networks : A Novel CNN Architecture for Image Denoising ». en. In : *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT : IEEE, p. 3204-3213. ISBN : 978-1-5386-6420-9. DOI : [10.1109/CVPR.2018.00338](https://doi.org/10.1109/CVPR.2018.00338).
- Li, Z., I. Chun et Y. Long (avr. 2020). « Image-Domain Material Decomposition Using an Iterative Neural Network for Dual-Energy CT ». In : *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. ISSN : 1945-8452, p. 651-655. DOI : [10.1109/ISBI45749.2020.9098590](https://doi.org/10.1109/ISBI45749.2020.9098590).
- Liu, D., B. Wen, X. Liu, Z. Wang et T. Huang (avr. 2018). « When Image Denoising Meets High-Level Vision Tasks : A Deep Learning Approach ». In : *arXiv :1706.04284 [cs]*. arXiv : 1706.04284.
- Lloyd, S. (1982). « Least Squares Quantization in PCM ». en. In : p. 129-137.
- Long, Y. et J. Fessler (août 2014). « Multi-Material Decomposition Using Statistical Image Reconstruction for Spectral CT ». In : *IEEE Transactions on Medical Imaging* 33.8, p. 1614-1626. ISSN : 1558-254X. DOI : [10.1109/TMI.2014.2320284](https://doi.org/10.1109/TMI.2014.2320284).
- Louppe, G. et P. Geurts (2012). « Ensembles on Random Patches ». en. In : *Machine Learning and Knowledge Discovery in Databases*. Lecture Notes in Computer Science.

- Berlin, Heidelberg : Springer, p. 346-361. ISBN : 978-3-642-33460-3. DOI : [10.1007/978-3-642-33460-3_28](https://doi.org/10.1007/978-3-642-33460-3_28).
- McCann, M., K. Jin et M. Unser (2017). « Convolutional Neural Networks for Inverse Problems in Imaging : A Review ». In : *IEEE Signal Processing Magazine* 34.6, p. 85-95.
- McCulloch, W. et W. Pitts (déc. 1943). « A logical calculus of the ideas immanent in nervous activity ». en. In : *The bulletin of mathematical biophysics* 5.4, p. 115-133. ISSN : 1522-9602. DOI : [10.1007/BF02478259](https://doi.org/10.1007/BF02478259).
- Mory, C., B. Sixou, S. Si-Mohamed, L. Boussel et S. Rit (nov. 2018). « Comparison of five one-step reconstruction algorithms for spectral CT ». eng. In : *Physics in Medicine and Biology* 63.23, p. 235001. ISSN : 1361-6560. DOI : [10.1088/1361-6560/aaeaf2](https://doi.org/10.1088/1361-6560/aaeaf2).
- Mozumder, M., A. Hauptmann, I. Nissilä, S. Arridge et T. Tarvainen (avr. 2021). « A model-based iterative learning approach for diffuse optical tomography ». In : *arXiv :2104.09579 [physics]*.
- Murphy, K. (2022). *Probabilistic Machine Learning : An introduction*. MIT Press.
- Oei, E. H., J. Tiel, W. Robinson et G. Gold (2014). « Quantitative Radiologic Imaging Techniques for Articular Cartilage Composition : Toward Early Diagnosis and Development of Disease-Modifying Therapeutics for Osteoarthritis ». en. In : *Arthritis Care & Research* 66.8, p. 1129-1141. ISSN : 2151-4658.
- Ozguner, O., A. Dhanantwari, S. Halliburton, G. Wen, S. Utrup et D. Jordan (jan. 2018). « Objective image characterization of a spectral CT scanner with dual-layer detector ». en. In : *Physics in Medicine & Biology* 63.2, p. 025027. ISSN : 0031-9155. DOI : [10.1088/1361-6560/aa9e1b](https://doi.org/10.1088/1361-6560/aa9e1b).
- Paszke, A., S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga et A. Lerer (2017). « Automatic Differentiation in PyTorch ». In : *NIPS Autodiff Workshop*.
- Penrose, R. (juil. 1955). « A generalized inverse for matrices ». en. In : *Mathematical Proceedings of the Cambridge Philosophical Society* 51.3, p. 406-413. ISSN : 1469-8064, 0305-0041. DOI : [10.1017/S0305004100030401](https://doi.org/10.1017/S0305004100030401).
- Persson, M., B. Huber, S. Karlsson, X. Liu, H. Chen, C. Xu, M. Yveborg, H. Bornefalk et M. Danielsson (oct. 2014). « Energy-resolved CT imaging with a photon-counting silicon-strip detector ». en. In : *Physics in Medicine and Biology* 59.22, p. 6709-6727. ISSN : 0031-9155. DOI : [10.1088/0022-3727/59/22/6709](https://doi.org/10.1088/0022-3727/59/22/6709).
- Peyrin, F. et K. Engelke (2012). « CT Imaging : Basics and New Trends ». In : *Handbook of Particle Detection and Imaging*. Springer Berlin Heidelberg, p. 883-915.
- Poludniowski, G., G. Landry, F. DeBlois, P. Evans et F. Verhaegen (sept. 2009). « SpekCalc : a program to calculate photon spectra from tungsten anode x-ray tubes ». en. In : *Physics in Medicine and Biology* 54.19, N433-N438. ISSN : 0031-9155. DOI : [10.1088/0031-9155/54/19/N01](https://doi.org/10.1088/0031-9155/54/19/N01).
- Pourmorteza, A., R. Symons, V. Sandfort, M. Mallek, M. Fuld, G. Henderson, E. Jones, A. Malayeri, L. Folio et D. Bluemke (avr. 2016). « Abdominal Imaging with Contrast-enhanced Photon-counting CT : First Human Experience ». eng. In : *Radiology* 279.1, p. 239-245. ISSN : 1527-1315. DOI : [10.1148/radiol.2016152601](https://doi.org/10.1148/radiol.2016152601).
- Purves, D., G. Augustine, D. Fitzpatrick, W. Hall, A.-S. LaMantia, J. McNamara et S. Williams (2005). *Neurosciences*.
- Radon, J. (déc. 1986). « On the determination of functions from their integral values along certain manifolds ». In : *IEEE Transactions on Medical Imaging* 5.4, p. 170-176. ISSN : 1558-254X. DOI : [10.1109/TMI.1986.4307775](https://doi.org/10.1109/TMI.1986.4307775).
- Rajendran, K., C. Löbker, B. Schon, C. Bateman, R. Aamir, N. J. A. de Ruiter, A. I. Chernoglazov, M. Ramyar, G. Hooper, A. Butler, T. Woodfield et N. Anderson

- (mai 2016). « Quantitative imaging of excised osteoarthritic cartilage using spectral CT ». In : *European Radiology* 27. DOI : [10.1007/s00330-016-4374-7](https://doi.org/10.1007/s00330-016-4374-7).
- Ravishankar, S., J. Ye et J. Fessler (jan. 2020). « Image Reconstruction : From Sparsity to Data-adaptive Methods and Machine Learning ». en. In : *Proceedings of the IEEE* 108.1, p. 86-109. ISSN : 0018-9219, 1558-2256. DOI : [10.1109/JPROC.2019.2936204](https://doi.org/10.1109/JPROC.2019.2936204).
- Rit, S., M. Oliva, S. Brousmiche, R. Labarbe, D. Sarrut et G. Sharp (mar. 2014). « The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK) ». en. In : *Journal of Physics : Conference Series* 489, p. 012079. ISSN : 1742-6596. DOI : [10.1088/1742-6596/489/1/012079](https://doi.org/10.1088/1742-6596/489/1/012079).
- Roessl, E. et R. Proksa (juil. 2007). « K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors ». en. In : *Physics in Medicine and Biology* 52.15, p. 4679-4696. ISSN : 0031-9155. DOI : [10.1088/0031-9155/52/15/020](https://doi.org/10.1088/0031-9155/52/15/020).
- Ronneberger, O., P. Fischer et T. Brox (2015). « U-Net : Convolutional Networks for Biomedical Image Segmentation ». In : Springer, C., p. 234-241.
- Rosenblatt, F. (1958). « The perceptron : A probabilistic model for information storage and organization in the brain ». In : *Psychological Review* 65.6, p. 386-408. ISSN : 1939-1471(Electronic),0033-295X(Print). DOI : [10.1037/h0042519](https://doi.org/10.1037/h0042519).
- Rowlands, J. A. (nov. 2002). « The physics of computed radiography ». en. In : *Physics in Medicine and Biology* 47.23, R123-R166. ISSN : 0031-9155. DOI : [10.1088/0031-9155/47/23/201](https://doi.org/10.1088/0031-9155/47/23/201).
- Roy, K., D. Banik, D. Bhattacharjee et M. Nasipuri (jan. 2019). « Patch-based system for Classification of Breast Histology images using deep learning ». en. In : *Computerized Medical Imaging and Graphics* 71, p. 90-103. ISSN : 0895-6111. DOI : [10.1016/j.compmedimag.2018.11.003](https://doi.org/10.1016/j.compmedimag.2018.11.003).
- Rudin, L., S. Osher et E. Fatemi (nov. 1992). « Nonlinear total variation based noise removal algorithms ». en. In : *Physica D : Nonlinear Phenomena* 60.1, p. 259-268. ISSN : 0167-2789. DOI : [10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F).
- Rumelhart, D., G. Hinton et R. Williams (jan. 1986). « Learning internal representations by error propagation ». In : *Parallel distributed processing : explorations in the microstructure of cognition, vol. 1 : foundations*. Cambridge, MA, USA : MIT Press, p. 318-362. ISBN : 978-0-262-68053-0.
- Russakovsky, O., J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. Berg et L. Fei-Fei (2015). « ImageNet Large Scale Visual Recognition Challenge ». In : *International Journal of Computer Vision (IJCV)* 115.3, p. 211-252. DOI : [10.1007/s11263-015-0816-y](https://doi.org/10.1007/s11263-015-0816-y).
- Salomé, M., F. Peyrin, P. Cloetens, C. Odet, A. M. Laval-Jeantet, J. Baruchel et P. Spanne (1999). « A synchrotron radiation microtomography system for the analysis of trabecular bone samples ». en. In : *Medical Physics* 26.10, p. 2194-2204. ISSN : 2473-4209. DOI : [10.1118/1.598736](https://doi.org/10.1118/1.598736).
- Schawinski, K., C. Zhang, H. Zhang, L. Fowler et G. Santhanam (jan. 2017). « Generative Adversarial Networks recover features in astrophysical images of galaxies beyond the deconvolution limit ». In : *Monthly Notices of the Royal Astronomical Society : Letters*, slx008. ISSN : 1745-3925, 1745-3933. DOI : [10.1093/mnrasl/slx008](https://doi.org/10.1093/mnrasl/slx008).
- Schlomka, J., E. Roessl, R. Dorscheid, S. Dill, G. Martens, T. Istel, C. Bäumer, C. Herrmann, R. Steadman, G. Zeitler, A. Livne et R. Proksa (août 2008). « Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography ». eng. In : *Physics in Medicine and Biology* 53.15, p. 4031-4047.
- Shan, H., A. Padole, F. Homayounieh, U. Kruger, R. Khera, C. Nitiwarangkul, M. Kalra et G. Wang (juin 2019). « Competitive performance of a modularized deep

- neural network compared to commercial algorithms for low-dose CT image reconstruction ». en. In : *Nature Machine Intelligence* 1.6. ISSN : 2522-5839. DOI : [10.1038/s42256-019-0057-9](https://doi.org/10.1038/s42256-019-0057-9).
- Shorten, C. et T. Khoshgoftaar (juil. 2019). « A survey on Image Data Augmentation for Deep Learning ». en. In : *Journal of Big Data* 6.1, p. 60. ISSN : 2196-1115. DOI : [10.1186/s40537-019-0197-0](https://doi.org/10.1186/s40537-019-0197-0).
- Si-Mohamed, S., D. Bar-Ness, M. Sigovan, D. Cormode, P. Coulon, E. Coche, A. Vlasenbroek, G. Normand, L. Boussel et P. Douek (nov. 2017). « Review of an initial experience with an experimental spectral photon-counting computed tomography system ». In : *Nuclear Instruments and Methods in Physics Research Section A : Accelerators, Spectrometers, Detectors and Associated Equipment*. Imaging 2016 873, p. 27-35. ISSN : 0168-9002. DOI : [10.1016/j.nima.2017.04.014](https://doi.org/10.1016/j.nima.2017.04.014).
- Simonyan, K. et A. Zisserman (avr. 2015). « Very Deep Convolutional Networks for Large-Scale Image Recognition ». In : *arXiv :1409.1556 [cs]*.
- Szegedy, C., W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke et A. Rabinovich (sept. 2014). « Going Deeper with Convolutions ». en. In : *arXiv :1409.4842 [cs]*.
- Taguchi, K., E. Frey, X. Wang, J. Iwanczyk et W. Barber (2010). « An analytical model of the effects of pulse pileup on the energy spectrum recorded by energy resolved photon counting x-ray detectors ». en. In : *Medical Physics* 37.8, p. 3957-3969. ISSN : 2473-4209. DOI : <https://doi.org/10.1118/1.3429056>.
- Taguchi, K. et J. Iwanczyk (oct. 2013). « Vision 20/20 : Single photon counting x-ray detectors in medical imaging ». eng. In : *Medical Physics* 40.10, p. 100901. ISSN : 2473-4209.
- Tikhonov, A. (1943). « On the stability of inverse problems ». In : *Dokl. Akad. Nauk SSSR* 39, p. 195-198.
- Van Rossum, G. et F. L. Drake (2009). *Python 3 Reference Manual*. Scotts Valley, CA : CreateSpace. ISBN : 1441412697.
- VGG16 - *Convolutional Network for Classification and Detection* (nov. 2018). en-US.
- Wang, G., J. Ye et B. De Man (déc. 2020). « Deep learning for tomographic image reconstruction ». en. In : *Nature Machine Intelligence* 2.12, p. 737-748. ISSN : 2522-5839. DOI : [10.1038/s42256-020-00273-z](https://doi.org/10.1038/s42256-020-00273-z).
- Wang, Z., A. Bovik, H. Sheikh et E. Simoncelli (avr. 2004). « Image quality assessment : from error visibility to structural similarity ». In : *IEEE Transactions on Image Processing* 13.4, p. 600-612. ISSN : 1941-0042. DOI : [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861).
- Webb, A. et G. Kagadis (2003). « Introduction to Biomedical Imaging ». en. In : *Medical Physics* 30.8, p. 2267-2267. ISSN : 2473-4209. DOI : <https://doi.org/10.1118/1.1589017>.
- WHO | *Chronic rheumatic conditions*. Publisher : World Health Organization.
- Wisnibaugh, E., R. Paden, A. Silva et M. Humphreys (juin 2014). « Dual-energy vs conventional computed tomography in determining stone composition ». eng. In : *Urology* 83.6, p. 1243-1247. ISSN : 1527-9995. DOI : [10.1016/j.urology.2013.12.023](https://doi.org/10.1016/j.urology.2013.12.023).
- Wolterink, J., T. Leiner, M. Viergever et I. Isgum (déc. 2017). « Generative Adversarial Networks for Noise Reduction in Low-Dose CT ». eng. In : *IEEE transactions on medical imaging* 36.12, p. 2536-2545. ISSN : 1558-254X. DOI : [10.1109/TMI.2017.2708987](https://doi.org/10.1109/TMI.2017.2708987).

- Wu, D., K. Kim, G. El Fakhri et Q. Li (déc. 2017). « Iterative Low-Dose CT Reconstruction With Priors Trained by Artificial Neural Network ». eng. In : *IEEE transactions on medical imaging* 36.12, p. 2479-2486. ISSN : 1558-254X. DOI : [10.1109/TMI.2017.2753138](https://doi.org/10.1109/TMI.2017.2753138).
- Würfl, T., F. Ghesu, V. Christlein et A. Maier (2016). « Deep Learning Computed Tomography ». en. In : *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016*. Sous la dir. de S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal et W. Wells. Lecture Notes in Computer Science. Cham : Springer International Publishing, p. 432-440. ISBN : 978-3-319-46726-9. DOI : [10.1007/978-3-319-46726-9_50](https://doi.org/10.1007/978-3-319-46726-9_50).
- Xu, C., M. Persson, H. Chen, S. Karlsson, M. Danielsson, C. Svensson et H. Bornefalk (fév. 2013). « Evaluation of a Second-Generation Ultra-Fast Energy-Resolved ASIC for Photon-Counting Spectral CT ». In : *IEEE Transactions on Nuclear Science* 60, p. 437-445. DOI : [10.1109/TNS.2012.2228276](https://doi.org/10.1109/TNS.2012.2228276).
- Yang, X., M. Kahnt, D. Brückner, A. Schropp, Y. Fam, J. Becher, J.-D. Grunwaldt, T. Sheppard et C. Schroer (mar. 2020). « Tomographic reconstruction with a generative adversarial network ». en. In : *Journal of Synchrotron Radiation* 27.2, p. 486-493. ISSN : 1600-5775. DOI : [10.1107/S1600577520000831](https://doi.org/10.1107/S1600577520000831).
- You, J., G. Zeng et Z. Liang (oct. 2005). « FBP Algorithms for Attenuated Fan-Beam Projections ». eng. In : *Inverse Problems* 21.5, p. 1801. ISSN : 0266-5611. DOI : [10.1088/0266-5611/21/5/C01](https://doi.org/10.1088/0266-5611/21/5/C01).
- Zeng, G. (2010). « Fan-Beam Image Reconstruction ». en. In : *Medical Image Reconstruction : A Conceptual Tutorial*. Sous la dir. de G. L. Zeng. Berlin, Heidelberg : Springer, p. 49-66. ISBN : 978-3-642-05368-9. DOI : [10.1007/978-3-642-05368-9_3](https://doi.org/10.1007/978-3-642-05368-9_3).
- Zhang, K., W. Zuo, Y. Chen, D. Meng et L. Zhang (juil. 2017). « Beyond a Gaussian Denoiser : Residual Learning of Deep CNN for Image Denoising ». en. In : *IEEE Transactions on Image Processing* 26.7, p. 3142-3155. ISSN : 1057-7149, 1941-0042. DOI : [10.1109/TIP.2017.2662206](https://doi.org/10.1109/TIP.2017.2662206).
- Zhang, W., H. Zhang, L. Wang, X. Wang, X. Hu, A. Cai, L. Li, T. Niu et B. Yan (2019). « Image domain dual material decomposition for dual-energy CT using butterfly network ». en. In : *Medical Physics* 46.5, p. 2037-2051. ISSN : 2473-4209. DOI : <https://doi.org/10.1002/mp.13489>.
- Zheng, X., S. Ravishankar, Y. Long et J. Fessler (juin 2018). « PWLS-ULTRA : An Efficient Clustering and Learning-Based Approach for Low-Dose 3D CT Image Reconstruction ». eng. In : *IEEE transactions on medical imaging* 37.6, p. 1498-1510. ISSN : 1558-254X. DOI : [10.1109/TMI.2018.2832007](https://doi.org/10.1109/TMI.2018.2832007).
- Zhou, Y., R. Chellappa, A. Vaid et B. Jenkins (juil. 1988). « Image restoration using a neural network ». In : *IEEE Transactions on Acoustics, Speech, and Signal Processing* 36.7, p. 1141-1151. ISSN : 0096-3518. DOI : [10.1109/29.1641](https://doi.org/10.1109/29.1641).
- Zhu, B., J. Liu, S. Cauley, B. Rosen et M. Rosen (mar. 2018). « Image reconstruction by domain-transform manifold learning ». eng. In : *Nature* 555.7697, p. 487-492. ISSN : 1476-4687. DOI : [10.1038/nature25988](https://doi.org/10.1038/nature25988).
- Zimmerman, K., G. Sharma, A. Parchur, A. Joshi et T. Schmidt (fév. 2020). « Experimental investigation of neural network estimator and transfer learning techniques for K-edge spectral CT imaging ». eng. In : *Medical Physics* 47.2, p. 541-551. ISSN : 2473-4209. DOI : [10.1002/mp.13946](https://doi.org/10.1002/mp.13946).
- Çiçek, A. Abdulkadir, S. Lienkamp, T. Brox et O. Ronneberger (2016). « 3D U-Net : Learning Dense Volumetric Segmentation from Sparse Annotation ». en. In : *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*. Lecture

Notes in Computer Science. Cham : Springer International Publishing, p. 424-432. ISBN : 978-3-319-46723-8. DOI : [10.1007/978-3-319-46723-8_49](https://doi.org/10.1007/978-3-319-46723-8_49).



FOLIO ADMINISTRATIF

THESE DE L'UNIVERSITE DE LYON OPEREE AU SEIN DE L'INSA LYON

NOM : BUSSOD
(avec précision du nom de jeune fille, le cas échéant)

DATE de SOUTENANCE : 8/12/2021

Prénoms : Suzanne

TITRE : Développement d'algorithmes d'apprentissage profond pour la tomographie spectrale :
Application à l'arthrose du genou

NATURE : Doctorat

Numéro d'ordre : 2021LYSEI092

Ecole doctorale : Electronique, Electrotechnique, Automatique

Spécialité : Traitement du Signal et de l'Image

RESUME :

La tomographie spectrale est une nouvelle modalité d'imagerie à rayons X qui permet d'acquérir des données avec une dimension énergétique. Cela est possible grâce aux détecteurs à comptage de photons qui classent les photons en fonction de leur énergie. En exploitant cette dimension énergétique, nous pouvons estimer les matériaux constituant l'objet ou reconstruire des images mono-énergétiques.

La décomposition en matériaux est un problème inverse non linéaire et mal posé. De nombreuses méthodes de décomposition de matériaux ont été développées. Certaines sont basées sur l'inversion d'un modèle physique. Toutefois, les algorithmes d'apprentissage profond peuvent également être utilisés pour résoudre des problèmes inverses. Des travaux récents en tomographie ont montré que ceux-ci peuvent améliorer la qualité des reconstructions et sont plus rapides.

Le but de la thèse est de développer des algorithmes d'apprentissage profond pour la décomposition de matériaux et la reconstruction d'images mono-énergétiques en tomographie spectrale, et de les évaluer par rapport à des méthodes basées sur des techniques d'optimisation. Nous nous sommes particulièrement focalisés sur l'application de la tomographie spectrale à l'arthrose du genou sans produit de contraste.

Des genoux humains excisés ont été scannés à l'aide d'un prototype de scanner spectral mais aussi par tomographie synchrotron mono-énergétique, afin d'avoir une image de référence. Des fantômes de genoux ont été générés à partir de ces volumes synchrotron. Puis, les projections spectrales sont simulées en utilisant des modèles de scanners spectraux. Nous avons ensuite développé un algorithme d'apprentissage profond basé sur U-net afin d'effectuer la décomposition de matériaux. Les résultats sont comparés à un algorithme de Gauss Newton régularisé. Nous reconstruisons également les images mono-énergétiques. Les méthodes sont comparées en calculant l'erreur quadratique moyenne et l'indice de similarité. Finalement, les méthodes basées sur le modèle et d'apprentissage sont appliquées sur des données réelles issues du prototype clinique de scanner spectral.

MOTS-CLÉS : Apprentissage profond, Tomographie spectrale, Décomposition de matériaux, Images mono-énergétiques, Arthrose du genou

Laboratoire (s) de recherche : CREATIS, CNRS UMR 5220 – INSERM U1294 – Université Lyon 1 – INSA Lyon – Université Jean Monnet Saint-Etienne.

Directeur de thèse: Françoise Peyrin

Président de jury : Philippe Douek

Composition du jury :

Rodet, Thomas	Professeur des Universités, ENS Cachan	Rapporteur
Visvikis, Dimitris	Directeur de recherche, INSERM, Brest	Rapporteur
Douek, Philippe	Professeur des universités, Université Lyon 1	
	Praticien hospitalier, HCL	Examineur
Chappard, Christine	Chargée de recherche, INSERM, Paris	Examinatrice
Boursier, Yannick	Maître de conférence, Université Aix/Marseille	Examineur
Peyrin, Françoise	Directrice de recherche, INSERM, Lyon	Directrice de thèse
Ducros, Nicolas	Maître de conférence, INSA, Lyon	Co-encadrant
Abascal, Juan	Ingénieur R&D, Carestream dental, Paris	Invité