Deep learning for partial volume correction in SPECT imaging

Théo Kaprélian

N° d'ordre NNT : 2025ISAL0013

THESE de DOCTORAT DE L'INSA LYON,
membre de l'Université de Lyon
Ecole Doctorale N°160
Électronique, Électrotechnique, Automatique (EEA)

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

Soutenue publiquement le 21/02/2025, par :

Théo Kaprélian

Deep Learning for Partial Volume Correction in SPECT Imaging

Devant le jury composé de :

COMTAT	Claude	Ingénieur-Chercheur CEA (HDR)	Université Paris-Saclay, Orsay,France	Rapporteur
READER	Andrew	Professor	School of Biomedical Engineering and Imaging Sciences, King's College London, UK	Rapporteur
BUVAT	Irène	Directrice de Recherche CNRS	Université Paris-Saclay, Orsay,France	Examinatrice
TRAN-GIA	Johannes	Heisenberg-Professor	University Hospital Würzburg, Germany	Examinateur
SARRUT	David	Directeur de Recherche CNRS	INSA Lyon	Directeur de thèse
ETXEBESTE	Ane	Maître de Conférence	INSA Lyon	Co-Encadrante de thèse

Département FEDORA – INSA Lyon - Ecoles Doctorales

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
ED 206 CHIMIE	**CHIMIE DE LYON** **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**
ED 341 E2M2	**ÉVOLUTION, ÉCOSYSTÈME, MICROBIOLOGIE, MODÉLISATION** **http ://e2m2.universite-lyon.fr** Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 **secretariat.e2m2@univ-lyon1.fr**	**Mme Sandrine CHARLES** Université Claude Bernard Lyon 1 UFR Biosciences Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69622 Villeurbanne CEDEX **e2m2.codir@listes.univ-lyon1.fr**
ED 205 EDISS	**INTERDISCIPLINAIRE SCIENCES-SANTÉ** **http ://ediss.universite-lyon.fr** Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 **secretariat.ediss@univ-lyon1.fr**	**Mme Sylvie RICARD-BLUM** Laboratoire 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**
ED 34 EDML	**MATÉRIAUX DE LYON** **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**
ED 160 EEA	**ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE** **https ://edeea.universite-lyon.fr** Sec. : Philomène TRECOURT 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**
ED 512 INFOMATHS	**INFORMATIQUE ET MATHÉMATIQUES** **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 **direction.infomaths@listes.univ-lyon1.fr**
ED 162 MEGA	**MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE** **http ://edmega.universite-lyon.fr** Sec. : Philomène TRECOURT Tél : 04.72.43.71.70 Bâtiment Direction INSA Lyon **mega@insa-lyon.fr**	**M. Etienne PARIZET** INSA Lyon Laboratoire LVA Bâtiment St. Exupéry 25 bis av. Jean Capelle 69621 Villeurbanne CEDEX **etienne.parizet@insa-lyon.fr**
ED 483 ScSo	**ScSo**[1] **https ://edsciencessociales.universite-lyon.fr** Sec. : Mélina FAVETON Tél : 04.78.69.77.79 **melina.faveton@univ-lyon2.fr**	**M. Bruno MILLY** (INSA : J.Y. TOUSSAINT) Univ. Lyon 2 Campus Berges du Rhône 18, quai Claude Bernard 69365 LYON CEDEX 07 Bureau BEL 319 **bruno.milly@univ-lyon2.fr**

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

Résumé

La tomographie par émission monophotonique (TEMP) est une modalité d'imagerie nucléaire notamment utilisée en oncologie pour la détection et le suivi des tumeurs. La quantification par TEMP est également devenue cruciale pour la dosimétrie dans le cadre des Radiothérapies Internes Vectorisées (RIV) au ^{177}Lu-PSMA ou ^{177}Lu-DOTATATE. Dans cette thèse, nous explorons deux sujets: 1) une méthode Quasi Monte Carlo (QMC) qui pourrait potentiellement accélérer les simulations 2) la correction des effets de volume partiel (EVPs) des images TEMP par apprentissage profond.

Les simulations Monte Carlo (MC) jouent un rôle essentiel en médecine nucléaire, notamment pour le calcul de dose, la conception de nouveaux systèmes TEMP et l'évaluation des algorithmes de reconstruction. Bien que très réalistes et précises, les simulations MC sont coûteuses en calcul et lentes à converger.

La première contribution de cette thèse est l'investigation de l'usage des séquences quasi-aléatoires pour l'échantillonnage dans les simulations MC pour la médecine nucléaire. Les séquences quasi-aléatoires sont conçues pour remplir l'espace d'échantillonnage de manière plus homogène que les générateurs de nombres pseudo-aléatoires couramment utilisés. Une suite de Sobol a été utilisée dans Geant4/Gate pour échantillonner les positions, directions et l'énergie des gamma primaires, ainsi que leurs premières interactions Compton. Bien que les premiers résultats ont montré des réductions d'erreur prometteuses allant jusqu'à 40% pour le suivi des particules gamma dans le vide, d'autres expériences, notamment dans l'eau, ont montré que lorsque le nombre de dimensions requises augmentait, le gain induit par l'échantillonneur quasi-aléatoire se réduisait à seulement quelques pourcents.

Une autre limite de la dosimétrie basée sur l'image TEMP est que leur résolution spatiale reste limitée, en raison des EVP. Les EVPs sont une sous-estimation ou surestimation de l'intensité des voxels dans les images reconstruites. Généralement, les EVPs se manifestent par le flou des bords des petites régions ou par la contribution des régions environnantes à l'activité estimée d'un objet. Leur cause principale est la réponse du collimateur-détecteur, définie par sa fonction d'étalement du point (PSF), qui peut être corrigée en la modélisant dans les algorithmes de reconstruction itératifs. Cette méthode n'étant pas suffisante, la deuxième contribution de cette thèse est PVCNet-sino, une méthode de correction de EVPs basée sur l'apprentissage profond, entraînée sur un grand jeu de données de sinogrammes synthétiques pour corriger la PSF sur les projections, avant reconstruction. Les données d'entraînement ont été générées à l'aide de scanner de patients utilisés pour créer 10 000 sources d'activité aléatoires, projetées analytiquement 1) avec modélisation de la PSF et du bruit et 2) sans PSF. Un 3d-Unet a été entraîné

pour inverser l'effet de la PSF sur ces sinogrammes. La méthode proposée ne repose pas sur une segmentation de l'image, à l'inverse de la plupart des méthodes de correction de EVPs. Après reconstruction, PVCNet-sino permet une meilleure réduction des EVPs par rapport à d'autres méthodes de correction et a été évaluée sur des données simulées, des expériences sur fantôme et des données de patients.

La correction avec PVCNet-sino a également été comparée à d'autres approches basées sur l'apprentissage profond: une dans le domaine image (PVCNet-img) et une approche opérant dans les deux domaines (PVCNet-sino-img). Les résultats indiquent que l'entraînement à la fois sur les projections et sur l'image mène à la meilleure correction, visuellement et quantitativement. Cependant, toutes les méthodes basées sur l'apprentissage profond proposées dans cette thèse ont montré des difficultés à corriger les petites lésions. A l'avenir, ce travail pourrait améliorer la quantification basée sur la TEMP dans le cadre des RIV et être une alternative aux méthodes basées sur la segmentation.

Abstract

Single Photon Emission Computed Tomography (SPECT) is a nuclear imaging modality used in oncology for tumour detection and staging. Quantification with SPECT has emerged as a key tool for dosimetry during Targeted Radionuclide Therapies (TRT) such as ^{177}Lu-PSMA or ^{177}Lu-DOTATATE. In this thesis we explore two topics: 1) a Quasi Monte Carlo (QMC) method that could potentially speed up simulations and 2) deep learning based Partial Volume Correction (PVC) for SPECT images.

Monte Carlo (MC) simulations play a critical role in Nuclear Medicine for various tasks, such as dose computation, new SPECT system design and reconstruction algorithm evaluation. While highly realistic and accurate, thanks to the large numbers of particles and interactions simulated, MC simulations are computationally expensive and slow to converge.

The first contribution of this thesis is the exploration of the potential use of Quasi Random Sequences in MC simulations for Nuclear Medicine as a Variance Reduction Technique to reduce their computation cost. Quasi-Random sequences are designed to fill the sampling space more homogeneously compared to commonly used pseudo-random number generators. This approach, denoted QMC simulations, uses a quasi-random Sobol sampler, that we implemented in the Geant4 code, used by the Gate software, to sample primary particles positions, directions, energy, as well as the first Compton interactions of photons with matter. While initial results demonstrated promising error reductions up to 40% for gamma particle tracking in a vacuum, further experiments showed that changing the material to water drastically increased the number of required dimensions and thus reduced the gain induced by the quasi-random sampler to only a few percent.

Another limitation to SPECT-based dosimetry is that the quantitative accuracy of ^{177}Lu-SPECT images remains limited, due to Partial Volume Effects (PVE) that need to be compensated. PVEs are inaccuracies in reconstructed image counts that are no longer proportional to the true activity. Typically, PVEs manifest as blurred edges on small regions (spill-out) or contribution of surrounding regions to the estimated activity in an object (spill-in). The main cause of the finite spatial resolution in SPECT, leading to PVEs, is the collimator-detector response, defined by its Point Spread Function (PSF), which can be modeled in iterative reconstruction algorithms to be corrected. However, PSF modeling during reconstruction is not enough to correct PVEs. To address this, the second contribution of this thesis is PVCNet-sino, a Deep Learning based PVC method, trained on a large synthetic sinogram dataset to correct for the depth-dependent PSF on projections, before reconstruction. The training dataset was generated using real patient CT scans used to create

10,000 random activity maps. These activity maps were analytically forward-projected 1) with PSF modeling and noise and 2) without PSF. A 3d-Unet was trained to revert the effect of the PSF on these sinograms. Importantly, our proposed method did not rely on an image segmentation mask. After reconstruction, PVCNet-sino demonstrated improved PVE reduction compared to other state-of-the-art PVC techniques. This PVC method was evaluated on simulated data, phantom experiments and real patient data. The PVCNet-sino correction was also compared to other deep learning based approaches, such as image-to-image training and correction (PVCNet-img) as well as to a dual domain approach (PVCNet-sino-img). Results of this comparison indicate that training both in the projection and the image domains leads to the best visual and quantitative improvement of SPECT images. However, all deep learning based methods proposed in this thesis showed difficulty to correct small lesions. This work could potentially improve SPECT-based quantification in TRT workflows and be an alternative to segmentation-based methods.

Remerciements

3 ans, 4 mois et 17 jours. Ça a été long et court, passionnant et fatiguant, très souvent épanouissant et parfois frustrant de ne pas pouvoir faire tout ce qu'on aurait voulu faire. Bien sûr ce travail a été le fruit de nombreux échanges avec pleins de gens à qui je voudrais dédier les quelques prochaines lignes.

Mes premiers remerciements vont tout naturellement à mes encadrants: David et Ane. Merci à tous les deux de m'avoir fait confiance, de m'avoir conseillé, guidé, relu, corrigé et encouragé. J'ai pris conscience tout au long de ma thèse de la chance que j'avais de vous avoir comme encadrants, non seulement pour votre expertise scientifique très complémentaires mais aussi pour vos qualités d'écoute et de confiance. En particulier, merci David pour ton optimisme qui m'a fait le plus grand bien lors de mes moments de doutes, pour tes idées, tes conseils et tes encouragements. Merci à Ane, car je crois avoir été un des tout premier doctorant que tu as encadré! Merci pour ton expertise de la physique, tes conseils, tes remarques très pertinentes. Malheureusement, tu as dû t'absenter lors de la fin de ma thèse, mais tu avais la meilleure des raisons, je te souhaite beaucoup de bonheur!

Je voudrais également remercier chaleureusement les membres du jury qui ont accepté d'évaluer mon travail de thèse: merci à Irène Buvat et Johannes Tran-Gia d'avoir endossé le rôle d'examinateur et merci à Andrew Reader et Claude Comtat pour votre relecture attentive de mon manuscrit. Vos travaux m'ont accompagné tout le long de ma thèse et c'est donc une grande chance pour moi de conclure cette aventure par votre participation à mon jury.

Bien sûr une thèse n'est pas qu'un travail individuel. Beaucoup de collègues se sont succédé dans notre équipe de CREATIS du Centre Léon Bérard et ont contribué à faire de ma thèse ce qu'elle est, de part nos échanges scientifiques mais aussi informels. Merci donc à Albert, Antoine, Aurélien, Axel, Bastien, Corentin, Eduardo, Hoa, Jean-Michel, Jean-Noël, Joey, Laure, Luca, Maxime, Mélanie, Moritz, Simon, Thomas et à celles et ceux que j'oublie. Mention spéciale à Thomas, pour sa patience et son aide précieuse! Merci à vous pour nos moments conviviaux, débats, parties d'échecs, BBQ, soirées...

Merci à mes amis sans qui la vie serait quand même beaucoup moins marrante. Je pense à Anne-Laure, Bédig, Fréddy, Léo GB, Léo P, Minass, Martin, Romain, Sophie, Théo, Thomas, Seb et Yann, mais j'oublie sûrement beaucoup de monde.

Merci à mes parents envers qui je suis extrêmement reconnaissant pour leur soutient et leurs encouragements de toujours, à mes grand parents et enfin à mes sœurs Eugénie et Juliette pour tout ce qu'elles m'apportent.

Enfin, comment ne pas remercier Ambre: merci pour ton soutient sans faille lors de ces longs moments à travailler et rédiger, je ne sais pas comment

j'aurais pu mener cette thèse à son terme sans toi. J'en profite pour te remercier infiniment pour tout ce que tu m'apportes au quotidien !

Merci à toutes et tous et merci à toi Lecteur, Lectrice...

Contents

List of Figures

List of Tables

Accronyms

AI	Artificial Intelligence
BRDF	Bidirectional Reflectance Distribution Functions
CDF	Cumulative Distribution Function
CNN	Convolutional Neural Network
CT	Computed Tomography
CF	Correction Factor
CZT	Cadmium Zinc Telluride
DL	Deep Learning
DEW	Dual Energy Window
FBP	Filtered BackProjection
FWHM	Full Width at Half Maximum
GAN	Generative Adversarial Network
GPU	Graphics Processing Unit
HU	Hounsfield Unit
ITS	Inverse Transform Sampling
LD	Low-Discrepancy
LEHR	Low Energy High Resolution
MC	Monte Carlo
MEGP	Medium Energy General Purpose
MIRD	Medical Internal Radiation Dose
MLEM	Maximum Likelihood Expectation Maximization
MRI	Magnetic Resonance Imaging
NEMA	National Electrical Manufacturers Association
NMAE	Normalized Mean Absolute Error
NRMSE	Normalized Root Mean Squared Error
OAR	Organ at Risk
OSEM	Ordered-Subset Expectation Maximization
PET	Positron Emission Tomography
PSF	Point Spread Function
PSMA	Prostate-Specific Member Antigen
PRNG	Pseudo-Random Number Generator

PVC	Partial Volume Correction
PVE	Partial Volume Effect
QMC	Quasi-Monte Carlo
RM	Resolution Modeling
ROI	Region of Interest
RTK	Reconstruction Toolkit
SPECT	Single Photon Emission Computed Tomography
TAC	Time-Activity Curve
TBR	Target-to-Background Ratio
TEW	Triple Energy Window
TRT	Targeted Radionuclide Therapy
VAA	Volume Activity Accuracy

Introduction

In 2022, 20 million new cancer cases were diagnosed worldwide, and 9.7 million people died from this disease [75]. Unfortunately, the number of cases is predicted to increase by more than 75% until 2050, because of population growth, aging, exposure to risk factors (tobacco, alcohol, air pollution...) but also thanks to improved diagnosis. Cancer designates diseases characterized by the uncontrolled growth of some abnormal cells which can proliferate in surrounding tissues and, in the case of *metastatic* tumours, spread to other parts of the body. Depending on the cancer type, stage and location, different cancer treatments exist and can be administered to a patient such as chemotherapy, surgery, radiation therapy or immunotherapy. Targeted Radionuclide Therapies (TRT) are now a treatment option proposed in nuclear medicine departments, in the case of some inoperable metastatic cancers.

TRT consists in injecting to the patient a radiopharmaceutical, for example ^{177}Lu-PSMA or ^{177}Lu-DOTATATE, whose aim is to bind to tumour cells and irradiate them. One advantage of ^{177}Lu-based TRT is that ^{177}Lu not only emits β^- particles, useful for therapy, but also γ-rays, which can be used for imaging with Single Photon Emission Computed Tomography (SPECT). SPECT/CT images thus allow to estimate the biodistribution of the tracer in the patient between each injection and assess the efficiency of the treatment by computing the amount of absorbed dose in each tissue. SPECT-based dosimetry is important because in the context of ^{177}Lu-labeled TRTs, Organs at Risks (OAR) may also show significant radiopharmaceutical uptake. While TRT with ^{177}Lu are mostly standardized, SPECT-based dosimetry is gaining a lot of interest for treatment personalization, in which the goal is to maximize the dose received by tumours while sparing healthy tissues.

In the context of nuclear medicine and TRT in particular, Monte Carlo (MC) simulations are a key tool for several research purposes such as for establishing dose-response relationships, evaluating new SPECT systems or image reconstruction algorithms. However, their main drawback is their computational cost, due to the high number of particles tracked, physical processes involved and complexity of the geometry. The principle of MC simulations is to randomly sample a high number of events according to their pre-defined probability distributions, from particles emission by radioactive decay to their interactions with matter. Consequently, MC simulation strongly rely on *pseudo-random* number generators (PRNG), designed to produce sequences of numbers uniformly distributed in $]0, 1[$ that mimic true randomness, *i.e.* independency between samples. Because samples are not correlated to each other, the uniformity of a given sample generated by a PRNG cannot be guaranteed. On the other hand, *quasi-random* number generators produce correlated samples such that the sampled space is known

to be sampled uniformly. Using these samplers, the integration error and computation time are expected to be reduced. Quasi Monte Carlo (QMC) simulations have found very practical applications in various domains such as in Computer Graphic Rendering. The first contribution of this thesis was to explore the interest of such quasi-random samplers in MC simulations for nuclear medicine, in particular using the Gate software, based on Geant4.

One of the major limitation of personalized dosimetry is the low quantitative accuracy of SPECT images, due to several effects that need to be corrected, including attenuation, scatter and spatial resolution. The poor spatial resolution of SPECT images is mainly due to the collimator-detector response of gamma cameras, leading to *Partial Volume Effects* (PVEs). Recent breakthroughs in Artificial Intelligence, and specially Deep Learning, have found a very wide range of applications for SPECT imaging, including image reconstruction, attenuation correction or scatter estimation. The second objective of this thesis is to propose a deep-learning based Partial Volume Effect Correction (PVC) method *PVCNet-sino*, trained in the sinogram domain to reduce the effect of the collimator response on projections, before reconstruction. Another objective of this work was to compare different training strategies, including image-domain training (*PVCNet-img*) and a dual domain approach (*PVCNet-sino-img*).

This manuscript is structured as follows. Chapter 1 is dedicated to expose key concepts and state-of-the-art about the clinical context, SPECT imaging and Monte Carlo simulations. Chapter 2 is divided into two Sections presenting two different contributions: Quasi Monte Carlo simulations for medical physics and fast Monte Carlo simulations with deep learning. In Chapter 3, we provide a review on Partial Volume Effects in SPECT, describing their main causes and consequences. Importantly, this Chapter also aims at presenting state-of-the art approaches to correct PVE in SPECT, from conventional PVC algorithms to the most recent deep learning based methods. Chapter 4 contains most of the methodological aspects of our contribution for PVC in SPECT, including our training database creation and deep learning models training. Importantly, we also describe in this Chapter the datasets we used for evaluation of our methods, including analytical test datasets, Monte Carlo simulated patient, phantom acquisition and patient data. Results of the *PVCNet-sino* model are presented and discussed. Finally, in Chapter 5, we compared the sinogram-based PVC to image-domain and dual domain training strategies.

This PhD thesis was part of the project named MoCaMed (Advanced Monte Carlo Methods for Medical Physics), funded by the French National Research Agency (ANR-20-CE45-0025). The ambition of this project was to design, develop and evaluate advanced sampling and Deep Learning methods to improve and use MC simulations in medical physics. This projects involved the collaboration of three labs including the Laboratoire de traitement de l'information médicale (LaTIM, Brest, France), the Laboratoire d'Informatique en Image et Systèmes d'information (Lyon, France) and Centre de Recherche en Acquisition et Traitement de l'Image pour la Santé (CREATIS, Lyon, France). This work was granted access to the HPC resources of IDRIS under the allocation 2019-101203 made by GENCI (Jean Zay computing center).

Chapter 1

State of the Art

Sommaire

1.1 Clinical Background

1.1.1 Targeted Radionuclide Therapy

1.1.1.1 Principle

Targeted Radionuclide Therapy (TRT) is a type of cancer treatment administered in nuclear medicine departments that has gained significant interest in recent years. This type of treatment in particularly suited for inoperable or metastatic tumours. It consists in injecting intravenously a radiopharmaceutical made of three components: a *vector* whose role is to bind to tumour cells and to be internalized in tumours, a *radioisotope*, treating the tumours by delivering radiations resulting in DNA damages, and a *chelator*, linking the vector to the radioisotope.

TRT are administered for tumours over-expressing a certain type of receptors, for which a vector exists. For example, neuro-endocrine tumours over-express somatostatin receptors that are well-targeted by the ^{177}Lu-DOTATATE radiopharmaceutical, while ^{177}Lu-PSMA (for Prostate-Specific Membrane

Antigen) treatment are an option for metastatic castration-resistant prostate cancers.

1.1.1.2 Difference with external beam radiotherapy

TRT is very different from more common external radiation radiotherapy treatments, which consist in irradiating the tumour with high energy x-ray beams from different positions. In this case, treatment radiation is delivered from a linear accelerator, which targets a localized region of the patient including the tumour, during a short period of time. Additionally, radiotherapy treatments follow a treatment plan (beam geometry, number of treatments, dose delivered per treatment) optimized to maximize the dose delivered to the tumour while minimizing the dose delivered to healthy tissues, specially organs at risks. On the other hand, TRT is a systemic treatment, delivering dose to responding regions which can be heterogeneously distributed in the whole body. Moreover, the evolution of the treatment between each injection is an important issue for TRT since $\approx$ 70% of the administered activity is evacuated 12 hours after the injection (through urine or radioactive decay) [115].

All these differences with external radiotherapy make TRT treatment plans more difficult to optimize. At present, the delivered dose is standardized to four injections of 7.4 GBq with 8 weeks between each injection for ^{177}Lu-DOTATATE, and six injections of 7.4 GBq every six weeks for ^{177}Lu-PSMA.

1.1.1.3 Benefits and main issues

Benefits in terms of progression-free and overall survival of ^{177}Lu-DOTATATE and ^{177}Lu-PSMA treatments were shown by the NETTER-1 Phase III [226] and VISION [207] clinical trials respectively. The success of these studies have raised hope for inoperable cancer treatments and lead to their authorization by public health authorities.

However, one main issue with TRT is that some healthy organs can also have physiological uptake because receptors targeted by the vector are not only present on tumour cells but also on kidneys, spleen and liver for ^{177}Lu-DOTATATE [145, 251], and also on salivary glands for ^{177}Lu-PSMA [48, 89]. Some Organs at Risks (OARs) must then be defined and spared as much as possible. The metric to evaluate and optimize is the *Absorbed Dose*, expressed in gray (Gy), measuring the amount of energy deposited by radiations per unit of mass:

$$\text{Dose}(Gy) = \frac{\text{Energy deposited}(J)}{\text{Mass}(kg)} \tag{1.1}$$

The objective of TRT is to maximize the dose absorbed by tumours while minimizing the dose received by healthy tissues. Safety guidelines usually define maximal tolerated dose thresholds based on kidneys uptake [218]. However, the evaluation of dose-response and dose-toxicity relationships is still an ongoing issue [97], specially concerning the inter- and intra-patient

response variability [241]. One major advantage of using ^{177}Lu-labeled radiopharmaceuticals is that ^{177}Lu does not only emits high energy β^- particles for therapy but also medium energy photons, allowing to estimate the distribution of radiotracer in the patient thanks to Single Photon Emission Computed Tomography (SPECT) combined with Computed Tomography (CT) acquisitions. Thus, the absorbed dose in each organ and tumour can be assessed after each injection by performing between one and three SPECT/CT acquisitions.

1.1.2 Workflow of ^{177}Lu based TRT

1.1.2.1 Patient selection for TRT

Patient selection for TRT is an essential step which consists in assessing that tumours effectively have receptors that will bind to the radiopharmaceutical that will eventually be used for therapy. For example, before prescribing a ^{177}Lu-PSMA treatment to a patient, the good response of the tumour cells is assessed by performing a ^{68}Ga-PSMA PET/CT acquisition. The ^{68}Ga-PSMA injection is not intended for therapy, just for imaging and patient selection. This pre-treatment ^{68}Ga-PSMA PET distribution was shown to be well correlated to patient responses to following TRT [245].

1.1.2.2 Treatment workflow

As mentioned earlier, TRT are standardized to four injections of 7.4 GBq every 8 weeks for ^{177}Lu-DOTATATE and six injections of 7.4 GBq every six weeks for ^{177}Lu-PSMA. Between each injection, it is important to assess the effectiveness of the treatment on tumours and its safety with respect the dose absorbed by OARs. This can be done by performing one or several SPECT/CT acquisitions between each injection, leveraging the photon-emitting property of ^{177}Lu. From SPECT images, Time Activity Curves (TAC) can be derived for each Region of Interest (ROI), delineated from the co-registered CT image. From the TAC, several dosimetry calculation softwares are commercially available to compute the absorbed dose in each ROI [156]. While not performed in every center and not mandatory in France, dosimetry between injections is an important tool to decide whether or not a patient is responding to the therapy and if following injections may be useful or not. It is also an essential mean to establish dose-response and dose-toxicity relationships that are still not fully understood. The overall treatment workflow is depicted in Figure 1.1.

1.1.3 Challenges

While ^{177}Lu-labeled TRT are standardized and have shown their efficacy, several studies have highlighted the variability of patients response in the case of ^{177}Lu-DOTATATE treatments [226, 19]. As a consequence, they concluded that, for some patients, the injected activity could be increased to maximize the dose delivered to tumours while still sparing healthy tissues, paving the way to personalized TRT [47, 73]. Treatment individualization includes variable injected activity, number of injections and time between each injection.

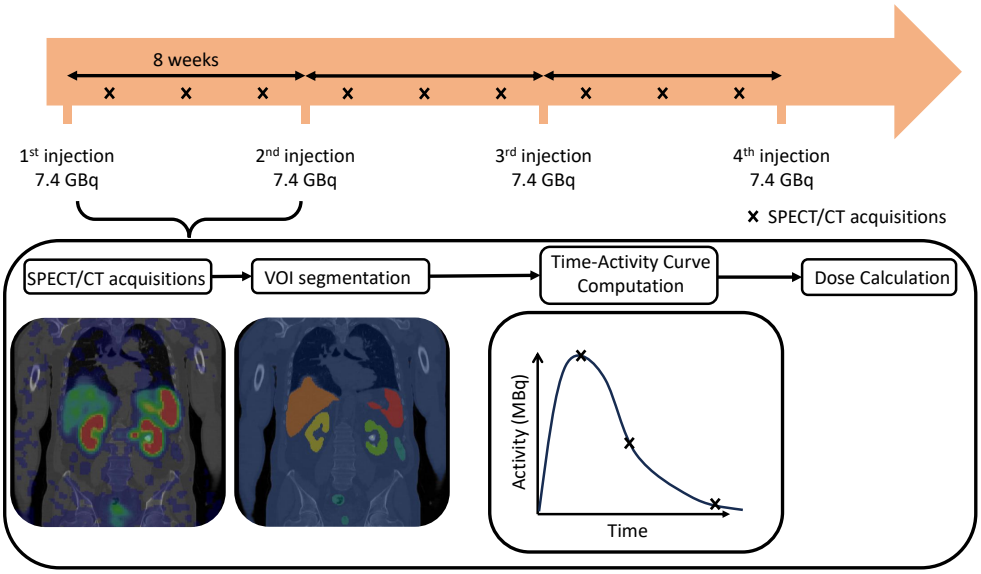


FIGURE 1.1 – Typical workflow of a Targeted-Radionuclide-Therapy with ^{177}Lu.

Whether it is for treatment response assessment or individualization, SPECT-based dosimetry is playing an increasingly central role in TRT workflows. Many different workflows were implemented. Laure Vergnaud dedicated her PhD thesis [240] to dosimetric workflows comparison and optimization and proposed an adaptive workflow with respect to the number of SPECT/CT acquisitions for ^{177}Lu-DOTATATE therapies [243, 240, 241]. In all existing approaches however, the quantitative accuracy and resolution of SPECT images is one of the bottlenecks for accurate dosimetry, specially for voxel-based dosimetry [146, 31]. Compared to CT scans, Magnetic Resonance (MR) images or even Positron Emission Tomography (PET), the resolution of SPECT images is much worse (more than 12 mm for most SPECT systems). The next section aims at describing SPECT imaging systems, from acquisition to reconstruction, and sources of inaccuracies in resulting SPECT images.

1.2 SPECT imaging

SPECT imaging is a *functional* imaging modality, in opposition to *anatomical* imaging modalities such as CT scans or MRI, whose objective is to obtain a 3D image of a previously injected radiopharmaceutical distribution in the patient. A SPECT system typically consists in a rotating gamma camera that collects several 2D projections of γ photons emitted by the radionuclide in the patient. Then, a *tomographic reconstruction* algorithm uses these projections, forming the *sinogram*, to reconstruct the 3D biodistribution in the patient. The overall emission, detection and reconstruction process in depicted in Figure 1.2.

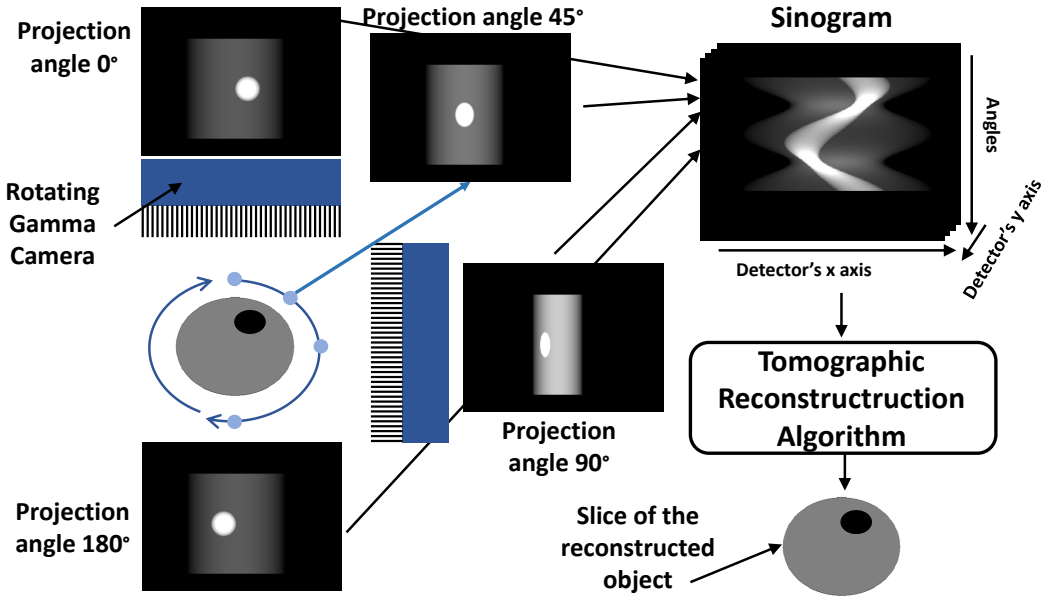


FIGURE 1.2 – General overview of the SPECT detection and reconstruction process.

1.2.1 Radioisotopes

As mentioned in Section 1.1, Lutetium-177 (^{177}Lu) is a radioistope that can be used for SPECT imaging purposes. Several other photon-emitter radionuclides can be used: Gallium-67 (^{67}Ga), Technetium-99m (^{99m}Tc), Indium-111 (^{111}In), Iode-123 (^{123}I) or Thallium-201 (^{201}Tl). In this section and in the following Chapters, we will focus on ^{177}Lu and ^{99m}Tc.

^{99m}Tc is the meta-stable daughter of the Molybdenum-99 (^{99}Mo) radionuclide and decays into ^{99}Tc. The transmutation from ^{99}Mo to ^{99m}Tc (Equation (1.2)) happens through a β^- emission process, meaning that a neutron in the nucleus is converted into a proton and an electron, the latter being the β^- particle. ^{99m}Tc then decays into a more stable ground state through an isomeric transition:

$$^{99}\text{Mo} \xrightarrow{\beta^-} {}^{99m}\text{Tc} \xrightarrow{\gamma} {}^{99}\text{Tc} \tag{1.2}$$

The isomeric transition results in an γ-ray emission in 87.87% of cases (with 140.5 keV energy in 98.6% of cases and 142.6 keV in 1.4%). The other 13.13% of disintegrations result in an internal conversion (the energy is transferred to eject an orbital electron). ^{99m}Tc has a half-life of approximately 6 hours. Since almost only low-energy photons are emitted by ^{99m}Tc, it leads to few deposited energy, which is why it is one of the main radionuclide used for imaging purposes in nuclear medicine [121], in addition to its ease to produce and affordable price.

^{177}Lu has a half life of 6.7 days and decays either to a stable ground state or an excited state of Hafnium-177 (^{177}Hf). During its decays, ^{177}Lu emits β^- particles with a continuous energy spectra with probabilities/$E_{\beta(\text{mean})}$/$E_{\beta(\text{max})}$ of: 79.3%/ 149.4 keV /498.3 keV, 9.1%/111.7 keV /385.4 keV and 11.64%/47.66

keV /177 keV. Main γ photon energies are 113 keV (6.2%) and 208 keV (11%) [44]. The main advantage of ^{177}Lu-labeled radiopharmaceuticals is that they enable both therapy and imaging with the same radionuclide, paving the way for the *theranostic* (combination of therapy and diagnostic) approach described in Section 1.1. The main drawback is their expensiveness, supply difficulties and the fact they are still emerging.

Decay characteristics of ^{99m}Tc and ^{177}Lu are summarized in Table 1.1.

Radionuclide	Half life	Photon Energy Peak(s)
^{99m}Tc	6 hours	140.5 keV (87%)
^{177}Lu	6 days	113 keV (6.2%) / 208 keV (11%)

TABLE 1.1 – Main characteristics of ^{99m}Tc and ^{177}Lu radionuclides.

1.2.2 Gamma Camera

The gamma camera, or *Anger scintillation camera*, was invented in 1957 by Hal Anger to convert incident photons into a digital signal. This section, inspired by [37], aims at providing a brief overview of main components contained in a gamma camera *i.e* a collimator, a scintillation crystal, a light guide and an array of photo-multiplier tubes (Figure 1.3).

1.2.2.1 Main components

Collimator

In order to reconstruct the emission position of photons detected by the gamma camera, their incident direction must be known. The very inefficient but practical method employed to know their incoming direction is to place an absorptive collimator on the detector front: only γ photons crossing the detector orthogonally to the detector plane are accepted to pass trough the scintillation crystal, the others being absorbed by the collimator material. A typical material used for collimation is lead, perforated by hexagonal holes with a diameters of less than 3 mm. The type of collimator presented in Figure 1.3 and considered in this thesis have parallel holes. However, several other holes geometry can be used, depending on the context: pinhole collimators for very small organs magnification, diverging collimators when the detector's size is small compared to the object to be imaged and converging holes for close enough object magnification [37].

This collimation process is highly inefficient since only a very small fraction of the emitted γ rays (which are emitted in all directions) is utilized by the detector: between approximately 5 and 20 out of 10,000 emitted γ are detected by a gamma camera. The geometry of the collimator follows a noise-resolution trade-off and must be chosen carefully depending on the context: if the collimator is very restrictive on the incoming angle selection, the uncertainty on detected counts origin will be low (high resolution) but very few counts

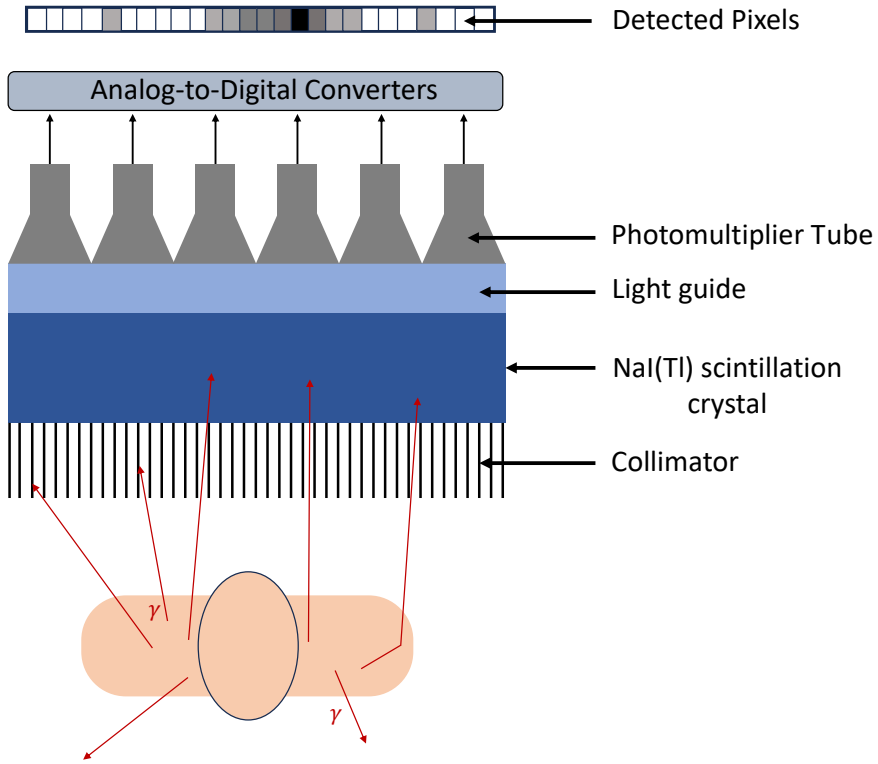


FIGURE 1.3 – Illustration of a Gamma Camera.

will contribute to the image, leading to a noisy estimate (low efficiency), and *vice versa*.

The uselessness of a large majority of photons is one of the main cause of the poor image quality in SPECT. An alternative to absorptive collimation, currently still in development, is the *Compton Camera* [61, 69], using a forced Compton scattering interaction to determine the incident photon direction, based on the knowledge of its initial and final energy.

Scintillation crystal

A scintillation crystal is usually made of NaI(Tl) and is aimed at converting incoming γ-rays into visible light photons. At the density and atomic number of the NaI(Tl) scintillators and typical γ rays energy encountered (100-250 keV), the dominant type of interaction for γ-rays is the Photo-Electric Effect: the incoming γ-ray is absorbed in the scintillator material and a *photo-electron* is emitted, electrons in the material are excited and after de-excitation, light photons are emitted. The quantity of light photons exiting the scintillator is proportional to the incoming γ photon energy. The crystal thickness also follows a noise-resolution trade-off: the thicker the crystal, the more spatial spreading can occur (affecting the *intrinsic* spatial resolution), but also the

more likely it is for the photon to be detected. Typical crystal thickness values are comprised between 6 mm and 15 mm.

Photomultiplier tubes

The low intensity light emitted by the scintillator is then converted into electric signal by a *Photomultiplier (PM) tube*: the visible light interacts with a photocathode which releases electrons, guided and amplified in the tube. Typically, the number of electrons from the photocathode is multiplied by 10^6. Once the resulting electron beam reaches the anode, the electric impulse is converted into a numeric signal by an analog-to-digital converter.

CZT-based detectors

Cadmium Zinc Telluride (CZT) detectors [239] are a new type of detector that has gained increasing interest for SPECT imaging. The scintillator and photomultiplier tubes are replaced by a CZT semi-conductor, directly converting γ photons into electric signal, leading to a better energy resolution, intrinsic spatial resolution, detection efficiency and counting rate performances than gamma cameras [95, 244].

However, despite the growing interest they are generating, CZT detectors will not be considered in this thesis, since most existing SPECT systems are still equipped with a NaI(Tl) scintillator associated with PM tubes.

1.2.3 Tomographic Image Reconstruction

Once several (typically 60 or 120) 2D projections are acquired, the problem of tomographic image reconstruction is to reconstruct a 3D image of the object. It is an inverse problem for which the formulation, modeling and solution must consider the physics of the detection process. This section is inspired by the following works: [63, 28, 37, 237, 187].

In the following, we will consider a 2D-formulation of the problem for simplicity: an object $f(x)$ for $x \in \mathbb{R}^2$ is imaged by a rotating 1D detector. Measured 1D projections, forming the sinogram, are denoted $p(r, \theta)$, where θ is the detector's direction and r the pixel position in the polar coordinate system $(\vec{u}_\theta, \vec{v}_\theta)$, as illustrated in Figure 1.4.

1.2.3.1 Analytical reconstruction

In 1917, Johann Radon proved that an object can be perfectly reconstructed if an infinite number of projections are acquired. With the apparition of CT imaging in the 70's, his work was applied in the first tomographic reconstruction algorithms. The measurements $p(r, \theta)$ can be modeled by the *Radon transform* of the object f, defined for $r \in \mathbb{R}$ and $\theta \in [0, 2\pi[$ as:

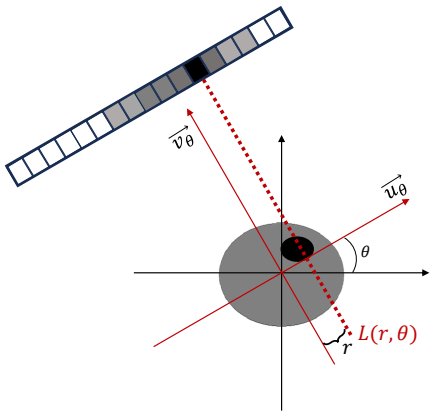


FIGURE 1.4 – Coordinate system of the Radon Transform.

$$\begin{aligned} p(r,\theta) = \mathcal{R}[f](r,\theta) &= \int_{x \in L(r,\theta)} f(x)dx \\ &= \int_{\mathbb{R}} f(r\vec{u}_\theta + t\vec{v}_\theta)dt \end{aligned} \tag{1.3}$$

where $L(r,\theta)$ is the line integral from the (r,θ) pixel locations, defined as:

$$L(r,\theta) = \{x \in \mathcal{R}^2 |\ x \cdot \vec{u}_\theta = r\} \tag{1.4}$$

The Radon transform $\mathcal{R}[f]$ can be seen as a *forward* projection of the object f, *i.e.* as the operator that maps the image domain (the object to be measured) to the projection domain (measures obtained by the detector).

The adjoint operator of the Radon transform is, for $x \in \mathbb{R}^2$:

$$\mathcal{R}^\star[p](x) = \int_0^{2\pi} p(x \cdot \vec{u}_\theta, \theta)d\theta \tag{1.5}$$

$\mathcal{R}^\star$ can be seen as a *backprojection* of the measured data p, mapping detected counts to the image domain. However, the backprojection is not the inverse of the Radon transform, since measured counts in p are uniformly backprojected on their line integral, while originating from different positions in this line. Moreover, since a projection is the sum of object values along the line integral $L(r,\theta)$, the contribution of samples is proportional to $1/r$ where r is their distance to the center. A simple backprojection thus leads to a blurred version of the true object f, and is thus not a suitable approach for image reconstruction.

This blurring can be removed in the frequency domain with the filtered-backprojection reconstruction algorithm. As a reminder, the n-dimensional Fourier Transform of function h is, for all $n \in \mathbb{N}^\star$ and $k \in \mathbb{R}^n$:

$$\mathcal{F}_n[h](k) = \int_{\mathbb{R}^n} h(k)e^{-2i\pi k \cdot x}dx \tag{1.6}$$

The Central slice theorem, or Fourier slice theorem, states that the 1D Fourier transform of a projection of object f obtained at angle θ is equal to a slice of the 2D Fourier transform of the object f [37, 187, 28], that is:

$$\mathcal{F}_1\left[p(\cdot,\theta)\right](k) = \mathcal{F}_2\left[f\right](k\vec{u}_\theta) \tag{1.7}$$

The object to reconstruct f can thus be written for all $x \in \mathbb{R}^2$:

$$\begin{array}{rcl} f(x) & = & \mathcal{F}_2^{-1}\left[\mathcal{F}_2\left[f\right]\right](x) \\ & = & \int_{\mathbb{R}^2} \mathcal{F}_2\left[f\right](\nu) e^{2i\pi x\cdot\nu} d\nu \\ & \overset{\nu=k\vec{u}_\theta}{=} & \int_0^{2\pi}\int_{\mathbb{R}} \mathcal{F}_2\left[f\right](k\vec{u}_\theta) e^{2i\pi kx\cdot\vec{u}_\theta}|k| dk d\theta \\ & \overset{(1.7)}{=} & \int_0^{2\pi}\int_{\mathbb{R}} \mathcal{F}_1\left[p(\cdot,\theta)\right](k) e^{2i\pi kx\cdot\vec{u}_\theta}|k| dk d\theta \\ & = & \mathcal{R}^\star\left[\hat{p}\right](x) \end{array} \tag{1.8}$$

where:

$$\begin{array}{rcl} \hat{p}(\theta,r) & = & \mathcal{F}_1^{-1}\left[|\cdot| \times \mathcal{F}_1\left[p(\cdot,\theta)\right](\cdot)\right](r) \\ & = & (h \otimes p(\cdot,\theta))(r) \end{array} \tag{1.9}$$

with h such that:

$$\mathcal{F}_1\left[h\right](k) = |k| \tag{1.10}$$

$\hat{p}$ is the projection *filtered* with the Ramp filter h. This means that the *exact* object f can be recovered from the measured projections by taking the backprojection of the filtered projections. This is the filtered-backprojection (FBP) algorithm.

FBP is easy to implement and fast to execute. It is thus a very popular reconstruction algorithm, specially for CT reconstruction. However, the exact object recovery provided by this mathematical background relies on several hypothesis including the completeness of the acquired data and absence of noise. The tomographic reconstruction inverse problem is ill-posed because of the non-uniqueness of the solution and sensitivity to noise. This is particularly an issue for SPECT image reconstruction because of the low number of acquired projections, high noise, poor spatial resolution and presence of several degrading physical effects (attenuation, scatter, collimator response, motion...). FBP thus leads to noisy reconstructed images with artefacts. Several other filters than the Ramp filter have been studied and used [143] to improve the noise-resolution trade-off of the FBP algorithm.

Iterative reconstruction algorithms, presented in the next section, relying on a statistical and detection model closer to reality are more commonly used for SPECT reconstruction.

1.2.3.2 Iterative reconstruction

For iterative reconstruction methods, the tomographic reconstruction problem is first discretized the following way, with p the detector measurements and f the discretized true activity distribution:

$$\forall j \leq M\ ,\ p_j = \sum_{i \leq N} a_{ij} f_i \tag{1.11}$$

where M is the number of detected pixels (in all positions) and N is the number of voxels in the image to reconstruct. The matrix $A = (a_{ij})_{i,j} \in \mathbb{R}^{N \times M}$ is the *system matrix*. This means that the coefficient a_{ij} is the probability that a photon emitted in voxel i will be detected in pixel j. The matrix A being very large and difficult to compute, its direct inversion to obtain f is not a conceivable solution.

Maximum Likelihood Expectation Maximization (MLEM)

The solution provided by the Maximum Likelihood Expectation Maximization (MLEM) algorithm is to solve Equation (1.11) with a statistical model of the detection process.

The objective of the MLEM method is to find f maximizing $P(f|p)$, *i.e.* such that the probability that the true activity is f, knowing the measured projections p, is maximal. In other words, the objective is to find the most likely solution considering the measured data with their statistical properties. Bayes' theorem allows to invert this likelihood:

$$P(f|p) = \frac{P(p|f)P(f)}{P(p)} \tag{1.12}$$

To maximize $P(f|p)$ with respect to f, the denominator $P(p)$ will not have any effect. $P(f)$ is the *a priori* knowledge on the solution distribution. In the first formulation of MLEM, no *a priori* on the solution is considered. The optimization problem thus only consists is maximizing $P(p|f)$, which, considering the independence of measurements, can be written:

$$P(p|f) = \prod_{j=1}^{M} P(p_j|f) \tag{1.13}$$

The counting statistic of nuclear counting systems follows a Poisson distribution. Each $P(p_j|f)$ thus follows a Poisson distribution with an expectation λ_j, where $\lambda_j = \sum_{i \leq N} a_{ij} f_i$:

$$P(p_j|f) = \frac{\lambda_j^{p_j} e^{-\lambda_j}}{p_j!} \tag{1.14}$$

Taking the logarithm of the likelihood, we obtain:

$$\log(P(p|f)) = \sum_{j=1}^{M} p_j \log(\lambda_j) - \lambda_j - \log(p_j!) \tag{1.15}$$

Once again, since $\log(p_j!)$ is constant with respect to f, maximizing $P(p_j|f)$ is equivalent to find the maximum of $\mathcal{L}$:

$$\begin{aligned}\mathcal{L}(f) &= \sum_{j=1}^{M} \left(p_j \log(\lambda_j) - \lambda_j\right) \\ &= \sum_{j=1}^{M} \left(p_j \log(\sum_{i \leq N} a_{ij} f_i) - \sum_{i \leq N} a_{ij} f_i\right)\end{aligned} \tag{1.16}$$

The optimization problem can be formalized in the following way:

$$f^\star = \arg\max_{f, f \geq 0} \mathcal{L}(f) \tag{1.17}$$

A necessary condition on $f^\star$ to be a solution of Equation (1.17) is that all partial derivatives of $\mathcal{L}$ vanish in $f^\star$, *i.e.* for all $i \leq N$:

$$\begin{aligned}&\frac{\partial \mathcal{L}}{\partial f_i}(f^\star) = 0 \\ \Leftrightarrow &\sum_{j=1}^{M} \left(\frac{p_j a_{ij}}{\sum_{i'=1}^{N} a_{i'j} f^\star_{i'}} - a_{ij}\right) = 0\end{aligned} \tag{1.18}$$

It can be shown that $\mathcal{L}$ is concave, which implies that the solution of Equation (1.18) is the global maxima of $\mathcal{L}$. However, there is no closed-form solution for this set of non-linear equations, so the solution must be estimated iteratively.

An estimation of the solution of Equation (1.17) is provided by the Maximum Likelihood Expectation-Maximization (MLEM) iterative algorithm [213, 238, 118]. The iterative solution can be derived using concave surrogate functions [119]. If we can find a function $Q \in \mathcal{C}^1$ that satisfies:

$$\begin{cases} Q(f^{(n)}, f^{(n)}) = \mathcal{L}(f^n) \\ \forall f, Q(f, f^n) \leq \mathcal{L}(f) \end{cases} \tag{1.19}$$

and if we choose f^{n+1} such that:

$$f^{n+1} = \arg\max_{f, f \geq 0} Q(f, f^n) \tag{1.20}$$

then, the $(f^n)_n$ sequence will converge toward the solution of Equation (1.17) [45, 187, 63]. The idea behind this iterative derivation is that the surrogates Q should be chosen such that solving Equation (1.20) becomes a straightforward task.

A surrogate function Q that satisfies [1] Equation (1.19) is:

$$Q(f, f^n) = \sum_{i=1}^{N} Q_i(f_i, f^n) \tag{1.21}$$

1. This can be shown by noticing that h_j is concave

with

$$Q_i(f_i, f^{(n)}) = f_i^{(n)} \sum_{j=1}^{M} \frac{a_{ij}}{\sum_{i'=1}^{N} a_{i'j} f_{i'}^{(n)}} h_j \left(\frac{f_i}{f_i^{(n)}} \sum_{i'=1}^{N} a_{i'j} f_{i'}^{(n)} \right) \tag{1.22}$$

and $h_j : t \mapsto p_j \log t - t$. Solving $\frac{dQ_i(f_i, f^{(n)})}{df_i} = 0$ finally provides the iterative scheme of the MLEM algorithm:

$$f_i^{(n+1)} = \frac{f_i^{(n)}}{\sum_{j=1}^{M} a_{ij}} \sum_{j=1}^{M} \frac{a_{ij} p_j}{\sum_{i'=1}^{N} a_{i'j} f_{i'}^{(n)}} \tag{1.23}$$

Convergence of the MLEM algorithm is guaranteed [45] by the fact that the likelihood is monotonically increasing with the number of iterations. Additionally, since MLEM updates (Equation (1.23)) are purely multiplicative from an iteration to the next with a non-negative factor, the non-negativity of the solution is ensured. The algorithm is usually initialized with $f^{(0)}$ being a constant image with a value of 1 in each voxel.

A more intuitive interpretation of the MLEM algorithm can be formulated in terms of forward-projections ($\mathfrak{F}$) and back-projections ($\mathfrak{B}$) as follows:

$$f^{(n+1)} = \frac{f^{(n)}}{\mathfrak{B}(\mathbb{1}_M)} \mathfrak{B} \left(\frac{p}{\mathfrak{F}(f^{(n)})} \right) \tag{1.24}$$

where $\mathbb{1}_M$ is the vector of projection bins containing ones in all pixels. Equation (1.24) highlights that at each iteration, the current estimate $f^{(n)}$ is multiplied by a corrective factor which is the normalized ($\mathfrak{B}(\mathbb{1}_M)$) backprojection of the ratio between the measures (p) and the forward projected current estimate $f^{(n)}$.

As we will see in the next sections, this iterative model-based algorithm is particularly well-suited for SPECT reconstruction as it allows to take into account several physical effects that need to be corrected and can thus be modeled in forward and back-projection operators $\mathfrak{F}$ and $\mathfrak{B}$.

Ordered-Subset MLEM (OSEM)

The main drawback of the MLEM algorithm is that it converges very slowly. The Ordered-Subset MLEM (OSEM) algorithm [90] allows to accelerate its convergence. The idea is the same in MLEM, except that projections are divided into several subsets, each of which is subsequently used to update the image. Suppose projections are divided into N_s subsets $S_1, ...S_{N_s}$, the intermediate update $f_i^{(n,m+1)}$ is:

$$f_i^{(n,m+1)} = \frac{f_i^{(n,m)}}{\sum_{j \in S_m}^{M} a_{ij}} \sum_{j \in S_m} \frac{a_{ij} p_j}{\sum_{i'=1}^{N} a_{i'j} f_{i'}^{(n,m)}} \tag{1.25}$$

and $f^{(n+1,1)} = f^{(n,N_s+1)}$. If $N_s = 1$, OSEM is equivalent to MLEM. Convergence of OSEM is much faster by a factor close to N_s because the image is updated more often. However, unlike MLEM, the convergence of the OSEM algorithm with more than one subset is no longer guaranteed. In practice, images reconstructed with OSEM tend to be more noisy than MLEM images when the number of iterations increases. For this reason, the OSEM algorithm is often stopped way before convergence. This reconstruction algorithm has become the reference reconstruction algorithm for SPECT and is implemented in many commercially available SPECT system's software.

1.2.4 Physical Effects Corrections

In this Section, the main physical effects that degrade the image quality will be detailed, as well as their corresponding correction methods.

1.2.4.1 Attenuation Correction

Photons emitted in the patient may interact with matter and transfer their energy either partially or totally. At the γ-ray energies and patient tissue densities, the main interaction types for photons are the Photo-electric Effect and Compton Scattering. Compton scattering occurs when a photon collides with an atom with enough energy to eject an orbiting electron of the atom. Compton scattering has three effects: 1) an electron is ejected from the atom 2) the incident photon is deviated and 3) it loose part of its energy. It is the dominant possible interaction type at the emission energy considered (100–200 keV). Lower energy photons, deviated by one or more Compton scattering interaction(s), may undergo a Photo-electric Effect: the photon transfers all its energy to an atom to eject an electron. The photon is thus absorbed. A photon may undergo multiple Compton scattering interactions, loosing each time part of its energy, until it is eventually absorbed.

Depending on the density of the material crossed and photons energy, absorption may be more or less important. Photons absorption must be corrected because less photons will be reconstructed in high attenuation areas, not because the activity is lower in the region but because emitted photons will be more likely absorbed by the material. The *attenuation* property of a material is quantified by its *linear attenuation coefficient* μ expressed in cm^{-1} defined such that if a mono-energetic beam with I_0 intensity passes through a x cm thick material, the resulting beam intensity will be:

$$I(x) = I_0 e^{-\mu x} \tag{1.26}$$

Linear attenuation coefficient μ depend on the energy of the beam, density and atomic number Z of the material crossed. A CT scan aims at reconstructing the distribution of attenuation coefficients in the patient with X-ray transmission measurements. CT images are usually expressed in Hounsfield Units (HU), defined as:

$$\mathrm{HU}(x) = 1000 \times \frac{\mu(x, E_{\mathrm{CT}}) - \mu_{\mathrm{water}}(E_{\mathrm{CT}})}{\mu_{\mathrm{water}}(E_{\mathrm{CT}})} \tag{1.27}$$

where $\mu(x, E_{\mathrm{CT}})$ is the reconstructed attenuation of material at position x and $\mu_{\mathrm{water}}(E_{\mathrm{CT}})$ the attenuation coefficient of water both at the energy of the X-ray tube E_{CT}. This normalisation results in HU values of -1000 and 0 for air and water respectively. To be used for attenuation correction of SPECT images, these HU values must be converted to the energy corresponding to the energy of the radionuclide used, E_{SPECT}. This can be done with a bi-linear transformation [8, 171]:

$$\mu(x, E_{\mathrm{SPECT}}) = \begin{cases} \mu_{\mathrm{water}}(E_{\mathrm{SPECT}}) + \frac{\mathrm{HU}(x)}{1000}\left(\mu_{\mathrm{water}}(E_{\mathrm{SPECT}}) - \mu_{\mathrm{air}}(E_{\mathrm{SPECT}})\right) & \mathrm{HU} < 0 \\ \mu_{\mathrm{water}}(E_{\mathrm{SPECT}}) + \frac{\mathrm{HU}(x)\mu_{\mathrm{water}}(E_{\mathrm{CT}})}{1000}\left(\frac{\mu_{\mathrm{bone}}(E_{\mathrm{SPECT}}) - \mu_{\mathrm{water}}(E_{\mathrm{SPECT}})}{\mu_{\mathrm{bone}}(E_{\mathrm{CT}}) - \mu_{\mathrm{water}}(E_{\mathrm{CT}})}\right) & \mathrm{HU} \geq 0 \end{cases} \tag{1.28}$$

Chang's Attenuation Correction

Initially proposed by Chang [33] for an uniform attenuation and extended with non-uniform CT-derived attenuation [252], Chang's correction applies a voxel-wise post-reconstruction multiplicative factor to reconstructed SPECT images. This Attenuation Correction Factor (ACF) can be derived with the following equation:

$$\mathrm{ACF}(x) = \left(\frac{1}{M}\sum_{j=1}^{M} \exp\left(-\int_{L_j(x)} \mu(l, E_{\mathrm{SPECT}})dl\right)\right)^{-1} \tag{1.29}$$

where M is the number of projection bins and $L_j(x)$ the line integral between the position x and the detector bin $j \leq M$.

The corrected image $f^{\mathrm{AC}}(x)$ is then obtained by multiplying the uncorrected image $f(x)$ with $\mathrm{ACF}(x)$.

Reconstruction based Attenuation Correction

A more realistic way to correct the attenuation is to model it in the system matrix used in the MLEM/OSEM algorithm, as proposed by [78, 235]. Instead of being discretized versions of the Radon (Equation (1.3)) and adjoint Radon transforms (Equation (1.5)), the forward projection operator is derived from the *attenuated* Radon transform. Keeping the same notations as in Section 1.2.3.1, the attenuated projection is calculated as follows:

$$p(r, \theta) = \mathcal{R}_\mu[f](r, \theta) = \int_{\mathbb{R}} f(r\vec{u}_\theta + t\vec{v}_\theta) e^{-\int_t^{+\infty} \mu(r\vec{u}_\theta + t'\vec{v}_\theta)dt'} dt \tag{1.30}$$

This reconstruction-based attenuation correction method is known to result in better quantitative accuracy compared to post-reconstruction methods [235]. It is available in most SPECT/CT reconstruction softwares. However, some limitations remain. First, the SPECT-CT images co-registration can lead to misalignment errors. Moreover, the CT scan acquired to perform attenuation correction is often a low-dose CT scan, which allows to limit the dose due to X-ray radiations, but tends to introduce some artifacts in the μ-map. In the following, attenuation correction (AC) will refer to this reconstruction-based attenuation correction.

1.2.4.2 Scatter Correction

Scattered photons, resulting from Compton scattering, may exit the patient and reach the detector. If the energy lost by such a photon is high, it may fall outside the photopeak energy window and thus be excluded from detection. However, photons with low energy loss can still be detected and contribute to the photopeak energy window, due to the finite energy resolution of the gamma camera. These detected scattered photons must be removed from projections since their initial direction is generally not orthogonal to the detector plane and thus carry incorrect information that degrades the reconstructed image quality. The most widespread method to correct for photon scattering is to not only acquire photons around the emission peak but also additional energy windows, referred to as *scatter* windows.

Dual-Energy Window (DEW)

In the Dual-Energy Window (DEW) method [100], a lower-energy *scatter* window is acquired, in addition to the *photopeak* energy window. This is illustrated in Figure 1.5 with a ^{99m}Tc source. The proportion of scattered photons detected in the photopeak is estimated thanks to the scatter window measurement in each pixel. Estimated scattered events can then be removed by subtracting the scattered window, pixel-by-pixel. If $P_{\text{photopeak}}$ is the vector of projection bins acquired in the photopeak window, and P_{scatter} the corresponding vector for the scatter window, the scatter-corrected sinogram $P_{\text{SC-DEW}}$ is obtained with:

$$P_{\text{SC-DEW}} = P_{\text{photopeak}} - kP_{\text{scatter}} \tag{1.31}$$

The factor k needs to be estimated for each radionuclide, SPECT system and acquisition parameters, including photopeak and scatter window widths.

Triple Energy Window (TEW)

A more accurate method for removing detected scattered events proposed by [162], is to acquire three energy windows: the photopeak window ($P_{\text{photopeak}}$) , as well as a lower (P_{lower}) and an upper adjacent energy window (P_{upper}). The corrected sinogram $P_{\text{SC-TEW}}$ is then:

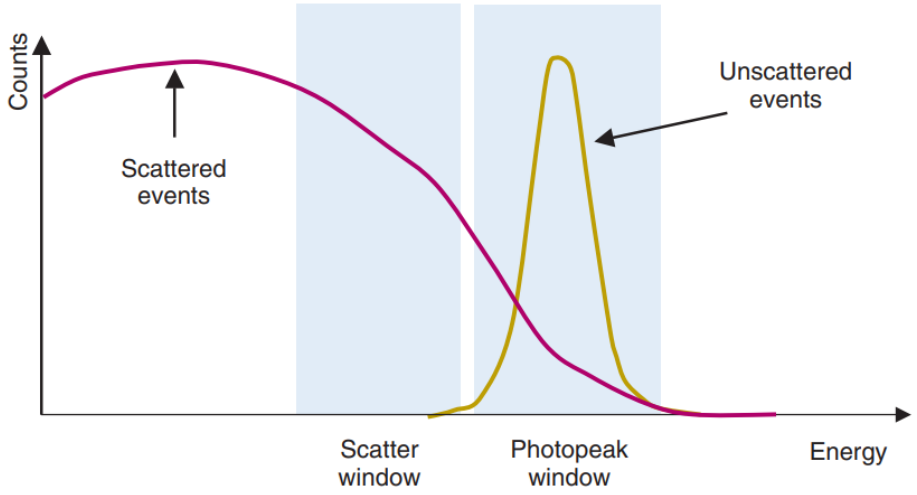


FIGURE 1.5 – Illustration of the Dual-Energy Window (DEW) method for scatter correction. The scatter window is used to estimate the amount of scatted events in the photopeak window. These scattered events are removed following Equation (1.31). This illustration was extracted from [37].

$$P_{\text{SC-TEW}} = P_{\text{photopeak}} - \frac{w_{\text{photopeak}}}{2}\left(\frac{P_{\text{lower}}}{w_{\text{lower}}} + \frac{P_{\text{upper}}}{w_{\text{upper}}}\right) \tag{1.32}$$

where w_{lower}, $w_{\text{photopeak}}$, w_{upper} are the widths (in keV) of each window. Unlike for the DEW method, there is no need to tune any parameter k for the TEW method. However, lower and upper energy window widths must be chosen considering, once again, a noise-resolution trade-off: narrow upper and lower windows provide better estimate of the scattered events detected in the photopeak window but lead to noisier projections, specially since the correction is done pixel-by-pixel and since the number of counts in the adjacent window is generally low.

Review of different scatter correction methods, for example based on Monte Carlo simulated scatter distribution, analytically derived or modeled in model-based iterative reconstruction methods, can be found in [137, 93].

1.2.4.3 Resolution Recovery

The typical spatial resolution of a gamma camera is comprised between 10 and 20 mm, mainly because of the collimator : due to the noise-resolution trade-of, collimator holes accept a relatively large range of solid angles, as illustrated in Figure 1.6. The collimator-detector response thus leads to a spatial blurring of detected counts in projections, as well as in reconstructed images, commonly referred to as Partial Volume Effects (PVEs). Causes of PVEs and a review of existing correction methods will be thoroughly detailed in Chapter 3.

For now, we can remark in Figure 1.6 that the more an object is far from the detector, the more blurry it will appear in projections. This implies that

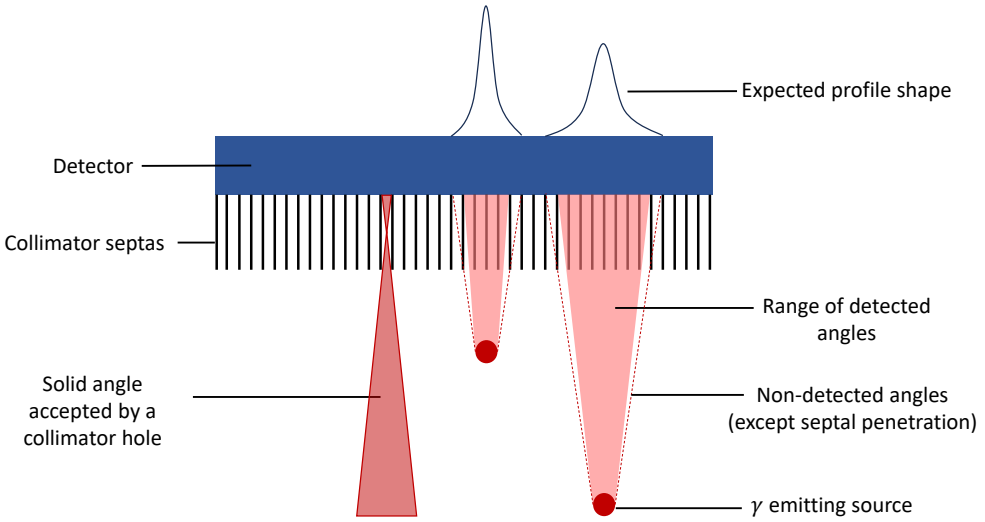


FIGURE 1.6 – Illustration of collimator-induced blurring.

the spatial resolution of a gamma camera is non-stationary. Spatial resolution can be quantified by computing the Full-Width at Half Maximum (FWHM) of a the projection obtained with a point source. The correspond projection profile obtained at a certain source-depth is the point-spread function (PSF). The PSF of a SPECT detector is often modeled by a Gaussian, whose FWHM linearly increases with the source-to-detector distance d:

$$\text{FWHM}(d) = ad + b \tag{1.33}$$

If FWHM parameters a and b are known, the collimator response can be partially corrected by incorporating the PSF blurring in the system matrix used in the MLEM/OSEM algorithm. This approach was developed by [235, 263, 261, 262] and has become a standard method to improve the image resolution [177, 180, 85]. This method is known as Resolution Modeling (RM) or Resolution Recovery.

In [261, 262], a numerical implementation of RM in forward and backprojectors is detailed as a slice-by-slice blurring and attenuation algorithm. To compute a forward projection of current image estimate f, it is first rotated by the angle of the desired projection angle and resampled with the same grid as the initial volume. Then, slices of f parallel to the detector $(f_k)_k$ are iteratively summed with the previous iteration, attenuated and blurred, starting from the farthest slice ($k = 0$):

$$\begin{cases} S_0 = (f_0 \times A_0) \otimes h_0 \\ A_k = e^{-\mu_k l_k} \\ S_{k+1} = \Big((S_k + f_{k+1}) \times A_{k+1} \Big) \otimes h_{k+1} \end{cases} \tag{1.34}$$

where μ_k is the slice k of the attenuation map, l_k is the slice thickness, h_k is the blurring kernel of slice-k and $\otimes$ denotes a 2d-convolution. A derivation of

the kernel weights of h_k with respect to slice k-to-detector distance, a, b (from Equation (1.33)) and slice thickness is provided in [261]. In the following, all forward-projections with RM will refer to this implementation, named *Zeng* in the open source reconstruction software RTK [185] and detailed in [187].

The main drawbacks of RM, including its insufficiency to correct PVE as well as resulting noise and ring artefacts amplification, will be illustrated and discussed in Section 3.2.1.

1.2.4.4 Motion Correction

One hypothesis for SPECT images reconstruction is the stationarity of the radionuclide distribution in the patient during the acquisition. If 60 projection angles are acquired during 20 seconds each, the total acquisition time is around 20 minutes. Thus, the stationarity hypothesis does not hold for short half life radionuclides but the decay can be corrected easily [259]. A more challenging issue concerning the stationarity hypothesis arises due to unavoidable patient motion resulting in spatially blurred images. The magnitude and direction of patient motion can be detected either from an additional motion detection captor [55, 151], or from projections directly [196, 187].

1.2.4.5 Dead-Time Correction

Another issue is the *dead time* of detectors [37], *i.e.* the minimal time between two events allowing them to be both detected. If the dead time or the activity is too high, some events will be lost or distorted. This is particularly an issue for post-TRT injection acquisitions, since the injected activity is typically of several GBq. Dead-Time correction factors can be determined to limit this issue [32, 72].

1.2.5 Quantification for dosimetry

1.2.5.1 Calibration Factor

After reconstruction with FBP or MLEM, the unit of reconstructed images is in number of counts. To convert reconstructed image counts into activity concentration (kBq/mL), a calibration factor (CF) must be determined from a calibration phantom measurement with known activity with the following Equation:

$$\mathrm{CF}\left(\frac{\mathrm{cps}}{\mathrm{kBq/mL}}\right) = \frac{\text{mean number of counts in FOV / acquisition time}}{\text{injected activity (kBq/mL)}} \tag{1.35}$$

Usually, the calibration phantom is chosen to be a large water cylinder filled with uniform activity, as recommended by [140]. Since the calibration phantom is large and uniform, the resulting calibration image should not be affected by Partial Volume Effects. The resulting CF is specific to the SPECT system, radionuclide, acquisition parameters and reconstruction algorithm.

1.2.5.2 SPECT-based dosimetry

As mentioned earlier in Section 1.1, the ultimate goal of post-TRT injection SPECT imaging is to estimate the dose absorbed by each ROI trough dosimetry. Once several SPECT images have been acquired, reconstructed and converted into activity concentrations, ROIs have to be delineated. The Time-Activity Curve (TAC) of each region can then be computed (see Figure 1.1) and fitted with a mono/bi/tri-exponential, depending on the number of SPECT acquisition performed [91]. The fitted TAC is then integrated to compute the total activity in a volume. Several dosimetry methods exist to compute the absorbed dose in each region from the cumulative activity.

References [68, 91, 155] provide good reviews of existing dosimetry methods, which is out of the scope of this section. We briefly mention here two methods: Monte Carlo Simulations and the MIRD methodology.

The reference and most accurate dosimetric method is Monte Carlo (MC) simulations, which essentially consist in tracking a large number of particles (photons, electrons ...) and their stochastic interactions with matter. The deposited dose of each interaction is stored in each voxel to provide the final dose map. Because of the patient-specificity of MC simulations, the high number of tracked particles and simulated interactions, the physical relevance of MC simulation is obtained at the cost of long computation time which makes them not easily usable in clinical routines. MC simulations are more extensively used for research or as reference methods to be compared to. Available MC softwares for dosimetry include Gate [99, 206], GAMOS [6] or TOPAS [174]. For region-based dose estimation, acceptable computation times can be reached (less than 10 min), as used in [242, 203]. Section 1.3 will provide an overview of MC simulations in a more general context.

In the Medical Internal Radiation Dose (MIRD) framework [140, 50] for organ-based S-value dosimetry, the absorbed dose in a target region is the sum of all neighbor regions activities multiplied by pre-determined "S-values" (in mGy/MBq/s), representing the mean absorbed dose by the target region due to radiations from source regions. All regions are alternatively considered as source and target. S-values depend on the radionuclide, patient characteristics, tissue density and source-to-target distance. The same formalism can be applied on a voxel-basis in the voxel-based S-value method.

A review and discussion of dosimetric workflows implemented in clinical routines is provided in [241].

1.3 Monte Carlo Simulations

1.3.1 History and principle

The invention of the Monte Carlo simulation method is attributed to Stanislaw Ulam and John von Neumann in 1946 while working on the well-known Manhattan Project in Los Alamos [153]. The objective of the method is to estimate an unknown quantity by sampling random realizations of possible outcomes. The key property of a MC simulation is that, if the samples are

truly random, the quantity estimated with the sample will converge toward the unknown quantity, as the number of samples increases. Monte Carlo simulations are a very powerful tool that has found applications in many fields. Typically, when an equation cannot be solved directly with an analytical form, the solution can be approximated with a statistical inference.

The Monte Carlo method can be illustrated as a method to compute the following d-dimensional integral, with $d \in \mathbb{N}^\star$:

$$I(f) = \int_{[0,1]^d} f(u)du \tag{1.36}$$

The Monte Carlo estimate of $I(f)$ is :

$$Q_n(f) = \frac{1}{n}\sum_{i=1}^{n} f(u_i) \tag{1.37}$$

where $(u_i)_{i\leq n}$ are random variables, independently sampled from the uniform distribution in $[0,1]^d$. The two key properties of this MC estimator are:

1. It is an unbiased estimator of $I(f)$, *i.e.* $\forall n, \mathbb{E}(Q_n(f)) = I(f)$.
2. The order of magnitude of integration error is $O(\frac{1}{\sqrt{n}})$. More precisely, it can be shown [29] with the central limit theorem that:

$$\frac{Q_n(f) - I(f)}{\sigma/\sqrt{n}} \xrightarrow[n\to+\infty]{} \mathcal{N}(0,1) \tag{1.38}$$

where $\sigma = \left(\int_{[0,1]^d}(f(u) - I(f))^2 du\right)^{1/2}$, the limit is in the sens of distributions, and $\mathcal{N}(0,1)$ is the standard normal distribution.

Because of this $O(\frac{1}{\sqrt{n}})$ error, the convergence of the MC method is slow. However, its main advantage is that this convergence rate does not depend on the problem's dimensionality, d. Therefore, MC integration is rather employed in complex settings.

In the context of medical physics, MC simulations can be used for several purposes, including: dose computation in the context of external beam radiotherapy [208] or for TRT optimization (see Section 1.2.5), design of new medical imaging systems [7], to test novel image correction techniques or more generally for experiment *vs* simulation agreement validation.

In this context, a MC simulation essentially consists in:

— Particle transportation through matter, whether it is photons, electrons, positrons, protons..., by tracking their position, direction, velocity, energy etc...

— Simulation of their interactions with matter: Compton scattering, photoelectric effect, pair production and Raleigh scattering mainly. Each interaction may generate secondary particles, which have also to be tracked. The realization of an interaction is randomly sampled according to several parameters (type and density of the material crossed, energy

of the particle...) thanks to physical models and pre-computed cross-sections.

— Recording of the quantity of interest while particles interact. Typically, it can be the dose in each voxel or detected counts in a gamma camera.

Several Monte Carlo codes are available for this purpose. Two types of codes can be distinguished:

— *General-purpose* codes such as Geant4 [1], MCNP [24] or PENELOPE [10] which provide a general framework for particle tracking, physics models and geometry definitions. These codes are used for a very wide range of applications, from High Energy Physics experiments to space applications.

— *Specific* Monte Carlo codes, based on a general-purpose code, which provide a more used-friendly approach for a limited number of MC simulation types. Codes designed for Emission tomography and radiation therapy include Gate [99, 206], GAMOS [6] or TOPAS [174]. SIMIND [139] is an autonomous MC code specifically designed for SPECT imaging applications.

In this thesis, all Monte Carlo simulations will be performed with Gate, which is based on Geant4. The next section provides a brief overview of some key principles behind Gate simulations.

1.3.2 Geant4 and Gate

Geant4 [1] is a general-purpose Monte Carlo simulation code developed at CERN, first released in 1998 and coded in C++. In addition to providing particles definitions, geometries and models for many physical processes, the architecture of a Geant4 simulation is divided into several classes. First, a *run* is an experiment for which the geometry and the set of possible physical processes (the *physics list*) are fixed. A run is a collection of *events*, each one starting with a primary particle creation. An event is the collection of the primary particle and its secondaries. Each particle evolves into a *track*, from its creation to its death. A track ends when the particle is absorbed or exits the simulation "world". The basic unit of a track is the *step*, which is the transport of the particle between two points. A new step occurs either because the particles exits a volume to enter another one, or because a physical interaction occurred.

Built on top of Geant4, initially for emission tomography simulations, Gate is an open-source code, developed by the OpenGATE collaboration which released its first version in 2004 [99]. Since then, several versions of Gate have been released [204, 206] to progressively include several imaging modalities other than PET and SPECT, including X-ray imaging and Ion imaging. Dosimetry modules also enable external beam radiotherapy and radionuclide therapy simulations.

In this thesis, MC simulations were performed with the new Gate 10 (beta) version [2], in which the user interface is fully python-oriented.

2. https://github.com/OpenGATE/opengate

1.3.3 Variance Reduction Techniques

As mentioned several times earlier, Monte Carlo Simulations are known for their slow and time consuming convergence. Given Equation (1.37), the uncertainty associated to a simulation can be defined by an estimator of the variance:

$$s(n) = \sqrt{\frac{Q_n(f^2) - Q_n(f)^2}{n-1}} \tag{1.39}$$

An obvious solution to decrease the variance would be to increase the number of samples n. However, doing so would also increase the computational cost of the simulation. Thus, a more appropriate quantity to compare different simulations is the *efficiency* $\epsilon(n)$ defined by also taking into account the simulation time $T(n)$:

$$\epsilon(n) = \frac{1}{s(n)^2 T(n)} \tag{1.40}$$

Any attempt to reduce the computational cost of a MC simulation should thus increase the efficiency of simulations, *i.e.* reduce the variance or the simulation time, or both [208]. Such methods are denoted Variance Reduction Techniques (VRT).

We describe here two of the most well known VRT for particle transport.

1.3.3.1 Particle Splitting

Particle Splitting is a VRT in which a particle, known to contribute to the final output, is replaced by N_{split} independent particles. For example in photon-radiotherapy simulations, when an electron produces a bremsstrahlung photon in the LINAC head, this photon can be split into $N_{\text{split}} = 10$ photons with different directions but the same energy as the original one. To keep the simulation unbiased, a weight of $w = \frac{1}{N_{\text{split}}}$ is assigned to each of the N_{split} photons. This technique reduces the simulation time because the tracking of only one electron results in 10 photons, thus sparing some electron simulation time.

1.3.3.2 Russian Roulette

In the Russian Roulette (RR) technique, particles that are known to scarcely contribute to the output can be killed with a predefined probability $p_{\text{RR}} \in]0, 1[$. Concretely, for each of these types of particles, a random number u is uniformly sampled between 0 and 1 and if $u > p_{\text{RR}}$, the particle is killed and its tracking time is spared. Surviving particles' weight are multiplied by a factor $w = \frac{1}{p_{\text{RR}}}$.

1.3.3.3 Production Cuts

A Gate simulation requires the user to define *production cuts*, *i.e.* threshold distances, bellow which secondary particles will not be created. For example, if production cuts are set to 1 cm, any secondary particle without enough energy to travel 1 cm will not be simulated. This is not a Variance Reduction Technique *per se*, but allows the user to reduce the simulation complexity depending on the context.

Many other Variance Reduction Techniques exist, such as forced interaction [208] (forcing interactions in a pre-defined volume of interest) or track length estimators [249] (in which photon energy deposition is scored in each voxel crossed between two interactions). VRT allow to save resources and computation time but should be used with care because precise knowledge of the system and physics involved is required, especially to preserve the unbiasedness of the simulation.

1.4 Conclusion

In this Chapter, we reviewed the main notions about the clinical context of this thesis, that is ^{177}Lu-based Targeted Radionuclide Therapies and the increasing importance of quantitative accuracy of SPECT image in the treatment workflow. We also reviewed the SPECT image detection and reconstruction process. These reminders were useful for Chapter 3 which is a literature review on Partial Volume Effects (PVE) and Partial Volume Correction (PVC) as well as for contributions of Chapters 4 and 5, in which Deep Learning based PVC methods will be proposed. We also set the basis of Monte Carlo simulations and Variance Reduction Techniques in the context of particles transport because the next Chapter will be dedicated to the exploration of a VRT based on Quasi-Random number sampling in this context.

Chapter 2

Advanced Monte Carlo simulations in Nuclear Medicine

Sommaire

2.1 Quasi Random Sampling in Geant4/Gate

Note: This work was conducted in collaboration with Bastien Doignies, a PhD student from the LIRIS team (UMR CNRS 5205, Lyon) and his supervisors Victor Ostromoukhov, David Coeurjolly, Julie Digne and Nicolas Bonneel as part of the Mocamed ANR project. The majority of the work presented here is a joint effort and has been previously presented at the IV Geant4 International User Conference at the physics-medicine-biology frontier [105].

2.1.1 Introduction

As explained in Section 1.3, the main limitation of Monte Carlo simulations in medical physics is their computation cost, increasing with their complexity which depends on the voxel size in the patient, the list of the considered physical interactions, the user-defined *cuts* (see Section 1.3.3), the number

of primary particles to track, the complexity of the detector... As an example, a full simulation of a SPECT acquisition for a patient to which 100 MBq of ^{99m}Tc were injected would involve to simulate $100 \times 10^6(\text{Bq}) \times 15(\text{s}) \times 120(\text{number of projection angles}) = 18 \times 10^{10}$ primary photons. These simulations can last several days (when not parallelized) and consume more than 10 kWh [1]. Hence, there is a strong need to reduce the computational cost of Monte Carlo simulations. Some established Variance Reduction Techniques (VRT) designed for this purpose were described in Section 1.3.3. However, even using these VRT, Monte Carlo simulations remain time consuming and a gain of even a few percents of error reduction can have significant impact in practice.

The computational cost of MC simulations is due to the wide range of physical events, materials and geometries taken into account during the simulation and the number of particles tracked. Considering the stochastic nature of physical events like disintegration, diffusion or interaction in particle transport (Compton and Rayleigh scattering, Photoelectric Effect, pair production ...), the strength of Monte Carlo simulations is to simulate a large number of random events leading to a globally relevant statistical behavior. Increasing the number of simulated events leads to a more accurate simulation output, due to a reduction of statistical uncertainty. For this reason, sampling of pseudo-random numbers is somehow the fuel of Monte Carlo simulations.

The exploratory work described in this Section aims at reducing the statistical uncertainty of a MC simulation, or equivalently, their computation time, by optimizing the way random numbers are sampled. While MC codes for medical physics such as Geant4/Gate unanimously use *pseudo-random* number generators, we propose to investigate whether or not it may be more efficient to use *quasi random* sequences. Unlike pseudo-random number generators, quasi random sequences introduce some correlation between samples so that the sampled space is filled more homogeneously and rapidly than with independent samples. Quasi randomness has found very efficient practical applications in various domains, such as in Computer Graphic Rendering [217, 172, 173] and finance simulations [123]. To our knowledge, Quasi Random Sampling (QRS) has not been studied before for Monte Carlo particle tracking.

Changing the random number sampler of Gate and Geant4 is not straightforward, neither from a theoretical point of view, nor from a practical one. The dependency between samples, needed to guarantee sampling homogeneity, requires to define *dimensions*. However, in this context, there is no trivial definition of what a dimensions refers to. Intuitively, the three random numbers required to sample primary particles positions should correspond to different dimensions than those used to sample the energy of secondary photons created after a Compton interaction. As we will see in Section 2.1.3, QMC samplers tend to behave badly in high-dimensional contexts (number of dimensions ≥ 50), which implies that the *dimension* definition can have a critical impact on the results. Moreover, a large framework like Geant4 (1 million lines of code, developed and maintained at CERN since more than 20 years and used by thousands of people) is a complex source code. Even

1. http://www.idris.fr/jean-zay/calcul-empreinte-carbone.html

though changing the pseudo random number generator is easy for uncorrelated sequence generators, the Geant4 structure is not designed for correlated sequences.

Section 2.1.3 is an introduction to quasi random samplers. Section 2.1.4 describes our investigations for quasi-random sampling in the context of SPECT simulation and dose calculation with Gate [206]. Results are presented in Section 2.1.5, followed by some discussion elements in Section 2.1.6.

2.1.2 Random sampling

2.1.2.1 Inverse Transform Sampling (ITS)

In order to sample points according to any (known) probability distribution, the method that is generally employed is the Inverse Transform Sampling (ITS) [122]. If we want to sample from the random variable $\mathbf{X} : \Omega \to \mathbb{R}$, whose probability distribution is p, then its Cumulative Distribution Function (CDF) takes value in [0,1] and is defined by :

$$\mathrm{CDF}(x) = \mathbb{P}(\mathbf{X} \leq x) = \int_{-\infty}^{x} p(t)dt \tag{2.1}$$

Then, to sample an outcome of the random variable $\mathbf{X}$, we need to sample from a uniformly distributed random variable $\mathbf{U} \sim \mathcal{U}([0,1])$. If we denote u an outcome of $\mathbf{U}$, then x, the outcome of $\mathbf{X}$, is defined as the inverse of the cumulative taken in u :

$$x = \mathrm{CDF}^{-1}(u) \tag{2.2}$$

In practice, the discretization of the CDF and its inverse allows to easily compute x. By this method, the sampling of any distribution can be reduced to the sampling of a uniform random variable in [0,1]. The method can be generalized for multivariate variables: for $\mathbf{X}$ taking values in $\mathbb{R}^d$, we need to sample $\mathbf{U} \sim \mathcal{U}([0,1]^d)$.

2.1.2.2 Pseudo-random number generation

The classical way of sampling uniformly distributed random numbers is to use a *Pseudo Random Number Generator* (PRNG). These samplers are based on some deterministic algorithm, often defined by a recurrent relation and thus determined by their first term, the *seed*. The produced sequence may not only produce numbers that "look" random but must also have some mathematical guaranties and pass numerical tests. One of the most used PRNG is the *Mersenne Twister* [150] with a period of $M = 2^{19937} - 1$. Despite their deterministic nature, pseudo-random numbers tend to mimic "true" randomness, that is, each value of the sequence can be considered as an independent realization from the others and the whole sequence approaches a *white noise* process.

PRNG are very popular in Monte Carlo simulations because of their simplicity to implement, to use and to interpret but also for their robustness and reproducibility. For example, Geant4 proposes a large choice of random

engines, including the Mersenne Twister [150], RANSHI [79], RANMAR [148] and RANCU [120]. See [98] for a review of several *pseudo-random generators* and their use in Monte Carlo simulations.

However, the main drawback of pseudo-random generators comes from their whitenoise aspect: as each sample is independent from the others, each subspace always has a non-negligible probability to be over-sampled (or under-sampled). Clearly, this probability diminishes as the number of samples increases but there is no means to guarantee any homogeneity of a given sampled pointset. Thus, the convergence rate of a Monte Carlo simulation with white noise sampling is always $O(\frac{1}{\sqrt{N}})$, as shown in Section 1.3. The advantage is that this convergence rate holds regardless of the number of dimensions or smoothness of the function to integrate. However, one aspect that limits the convergence to $\frac{1}{\sqrt{N}}$ is the independency of samples. As demonstrated for other high-dimensional integration problems [51], this convergence rate can be reduced by using quasi-random (also called low-discrepancy) sequences.

2.1.3 Quasi-Random sampling

2.1.3.1 Discrepancy of a Point Set

The goal of Quasi Monte Carlo is to increase this convergence rate by introducing dependency between samples, forcing the sequence to sample the space as homogeneously as possible. In order to quantify the *homogeneity* of a point set $P_n = \{x_1, ..., x_n\}$, the *discrepancy* measure was first defined in [160] as:

$$\mathcal{D}(P_n) = \sup_{B \subset [0,1]^d} \left| \frac{\#\{x_i \in B, i \leq n\}}{n} - |B| \right| \tag{2.3}$$

where B is any rectangular solid in $[0,1]^d$ with its sides parallel to the axes, *i.e.* B is of the form $\prod_{i=1}^{d} [a_i, b_i[$ with $\forall i \leq d, 0 \leq a_i, b_i \leq 1$. The *discrepancy* is a measure of distance between the distribution of a given point set and the uniform distribution. An illustration of the Discrepancy is shown in Figure 2.1. A "good" pointset in terms of its discrepancy is a pointset whose points are homogeneously distributed on the sampling space.

Importantly, the Koksma-Hlawka inequality is given in Equation (2.4). It underlines why discrepancy is a good measure of the quality of the point set used for Monte Carlo estimation:

$$\left| \frac{1}{n} \sum_{i=1}^{n} f(x_i) - \int_{I^d} f(u) du \right| \leq V(f) \mathcal{D}(P_n) \tag{2.4}$$

where $V(f)$, the total variation of f, does not depend on the point set used for MC estimation. Equation (2.4) indicates that reducing the discrepancy of the pointset P_n used to evaluate the integral of f, can reduce the integration error.

2.1.3.2 Low Discrepancy Sequences

Low discrepancy (LD), or Quasi-Random sequences, are defined as sequences converging with a $O(\frac{(\ln n)^d}{n})$ rate, which is better than the whitenoise sequence convergence rate of $O(\frac{1}{\sqrt{n}})$, at least for large enough values of n.

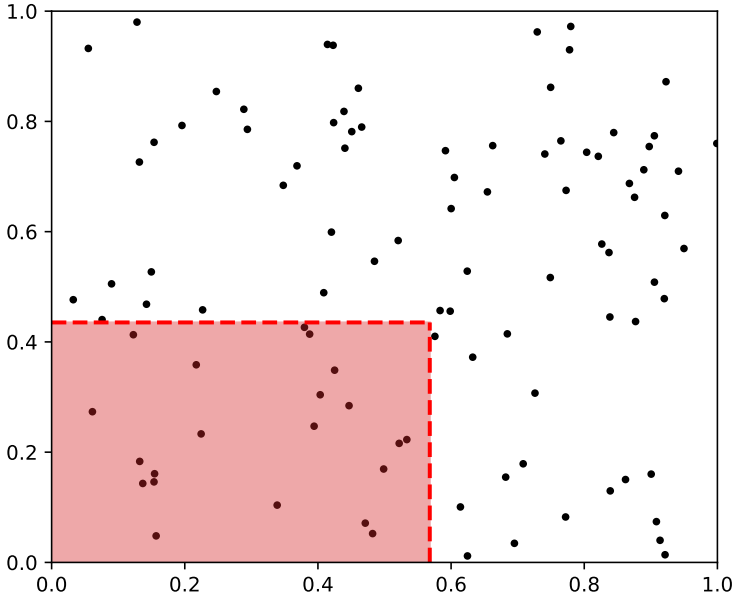


FIGURE 2.1 – 2d illustration of the discrepancy $\mathcal{D}(P)$ of the point set: it is the maximum, over each varying red rectangles, of the difference between the proportion of point inside the box and its area. Homogeneously distributed point sets result in low-discrepancy.

Several LD sequences have been proposed in the past: Halton sequence [80], Hammersley set [81] or Sobol sequence [222]. Recent works [172, 173] designed advanced samplers for Computer Graphic applications. In Figure 2.2, four sampling strategies can be visually compared in two dimensions. Whitenoise sampling is performed with a Mersenne twister. Low Discrepancy Blue Noise (LDBN), Sobol and Halton sequences are correlated sequences. It is clearly visible in Figure 2.2 that quasi-random sequences have strong homogeneous properties compared to whitenoise, allowing to better cover the sampling space, but also that their very regular patterns can be source of artifacts.

2.1.3.3 Sobol' Sequence

The Sobol' Sequence [222, 51] is a very popular low discrepancy sequence introduced in 1967. It is an algebraic sampler, meaning that each point is produced by a matrix multiplication with its index. To compute the ith sample

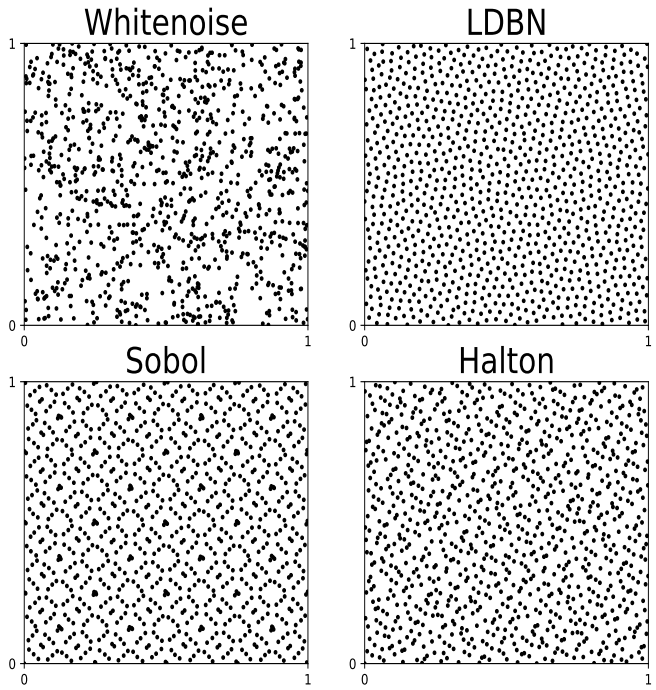


FIGURE 2.2 – 2D comparison of four sampling strategies for N=1024 points: Whitenoise, Low Discrepancy Blue Noise (LDBN), Sobol and Halton sequences.

of the Sobol sequence in dimension d, i is written in binary form as follows:

$$i = \sum_{k=0}^{m-1} a_k(i) 2^k \tag{2.5}$$

This decomposition of the index is used to compute, for each dimension $j \leq d$, the following vector by multiplication with a pre-determined matrix M_j, referred to as a *generating matrix*:

$$\begin{pmatrix} b_{j,0}(i) \\ b_{j,1}(i) \\ \vdots \\ b_{j,m-1}(i) \end{pmatrix} = M_j \cdot \begin{pmatrix} a_0(i) \\ a_1(i) \\ \vdots \\ a_{m-1}(i) \end{pmatrix} \tag{2.6}$$

The j-th dimension of the i-th point of the Sobol sequence is then given by:

$$x_{i,j} = \sum_{k=0}^{m-1} b_{j,k}(i) 2^{-k-1} \tag{2.7}$$

Generating matrices $(M_j)_{j \in \mathbb{N}}$ are well chosen binary upper triangular matrices. The classical Sobol' sequence starts with matrices M_0 being the identity matrix and M_1 the binary Pascal's triangle. While other matrices can be explicitly constructed, the process is not straightforward. Readers might refer to [102, 122, 173] for more information. This construction ensures that the points are distributed uniformly within the unit domain.

This sequence has the *low discrepancy* property, *i.e.* converges with a $O(\frac{(\ln n)^d}{n})$ rate, for any number of points and dimension. In practice, for computational efficiency, the Sobol sequence is not directly computed with Equation (2.7) but recursively in the Grey Code algorithm, detailed in [122]. The generative matrices construction relies on primitive polynomials and initialization numbers that have demonstrated good sequence properties [102]. The advantage of the Sobol' sequence is that it can be used in high dimensions, since only new primitive polynomials of increasing degree are needed.

In one dimension, the first numbers of the Sobol sequence are: 0, $\frac{1}{2}$, $\frac{3}{4}$, $\frac{1}{4}$, $\frac{3}{8}$, $\frac{7}{8}$, $\frac{5}{8}$, $\frac{1}{8}$...

As explained before and shown on Figure 2.2, Sobol' sequence is deterministic and seems to output a very regular pattern. For this reason, a last step is generally applied: the *scrambling*. It is a random transformation applied to a deterministic LD sequence which preserves the LD nature while changing samples locations. The most popular scrambling technique is Owen' scrambling [166].

2.1.3.4 The Curse of dimensionality

While being low discrepancy in any dimension, some projections of the generated point set on certain pairs of dimensions show very poor uniformity, especially for high dimensions, as illustrated in Figure 2.3. Although

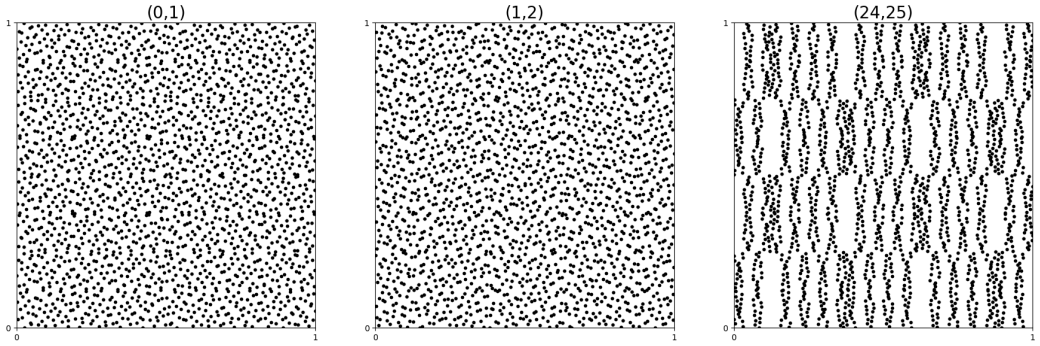


FIGURE 2.3 – 2d projections of the Sobol' sequence with N=2048 points. While dimensions (0,1) and (1,2) show good uniformity, dimensions (24,25) exhibit very weak homogeneity.

the benefits of using LD sequences in low dimensional simulations are no longer in question, their application to high dimensional problems is still not straightforward. In [172], a modified version of the Sobol sequence was proposed in order to mitigate this high dimensional issue.

2.1.3.5 Usage in path tracing for Computer Graphics and differences with MC for medical physics

Quasi Monte Carlo have found successful application in a large number of domains. One notable example is rendering : a path tracing technique that aims to produce images that are indistinguishable from a photograph of the same scene. The outline of the method reads as follows. It starts with the sampling of a primary ray, in the image plane, that will be tracked around a scene we wish to render. Then, an interaction location is computed. At the interaction location, depending on the material of the hit object, the ray can be either absorbed or reflected. Reflected rays are randomly sampled from the Bidirectional Reflectance Distribution Functions (BRDF) which can be highly diffuse for matte materials, glossy or unidirectional for mirrors. The scattered ray will continue its path until it is absorbed or escapes the scene. Along the whole trajectory, the color is accumulated and reported to the pixel the ray originated from.

The outline is very similar to Monte Carlo simulations in medical physics. However some key differences must be highlighted. The goal of the two methods are indeed different : in nuclear medicine the aim is to produce physically accurate results while path tracing prioritizes pleasing images. As such, the way the particles (or rays) are tracked in standard methods is different. Medical imaging simulations start from a source to reach detectors to simulate exact physical processes. In path tracing, the rays are tracked from the detectors (camera) to the source (lights) because less time will be spent tracking rays that won't reach the camera and the noise due to random

selection will be reduced [2]. Interactions are also computed very differently. In path tracing, simple models are used and most of their outcome has close form inverse [18], thus making the sampling from them very easy via ITS. In contrast, in MC for Nuclear Medicine, the number considered particle types is higher (photons, electrons, positrons ...), interactions models can be complex and produce secondary particles. Moreover, several properties of each particles are dynamically scored (energy, velocity, polarization...). Depending on the context, some samplers may be more adapted than others. For example in computer graphic rendering, as the regularity of LD sequences can be source of aliasing to which the human eye is particularly sensitive, the visual quality of the result is sometimes as important as quantitative metrics [2]. On the other side, for samplers used in financial or medical applications, the main issue is to increase the accuracy of the simulation in very high dimensions [122].

2.1.4 Contribution: investigation of QMC usage in Geant4/Gate

2.1.4.1 Challenges

Theoretical properties of quasi random pointsets offer great asymptotic improvements over traditional Monte Carlo methods. However, when the dimension grows, the bound becomes more and more pessimistic. Owen [165] showed that better integration error are to be expected only when $n > d^d$, which for $d = 15$ is not in the reach of nowadays computers. This effect is highly mitigated by the concept of effective dimension [219, 51, 165]. To illustrate the idea, consider the function $(x, y) \mapsto x + 10^{-15}y, (x, y) \in [0, 1]^2$. While the function is 2-dimensional, one can argue that y has a very small impact on the function, and thus the integration error. Empirically, in a dose simulation with 10^7 simulated photons of 2 MeV, the random number generator is called from more than 5000 locations. In this context, the effective dimension determination can be out of the reach of QMC methods.

Another challenge of using quasi-random sequences in Gate simulations comes from the fact that the sampling of secondary particles for most type of interaction is done through the *rejection method* instead of ITS (Section 2.1.2.1). This method is used when the CDF of the distribution to sample is difficult to compute and inverse. This sampling method reads as follows:

1. Sample x uniformly in Ω and u in $[0, 1]$.
2. If $u < p(x)$, accept x as a sample; otherwise, go back to 1.

Rejection sampling is more to difficult to handle with quasi-random sequences because it uses an unknown number of iterates, meaning that the number of dimension can not be known in advance. Moreover, the number of needed sample can differ from particle to particle and can grow arbitrarily, worsening the dimensionality problem. QMC simulations with rejection sampling have been studied by [264] who showed that the distribution of samples generated

2. This "reverse" simulation mode is also possible for SPECT MC simulations, referred to as Forced Detection [30] but it is much less frequently employed than in Computer Graphic rendering.

with Rejection-QMC methods converges still faster to the desired distribution than purely random samples.

The last issue is rather technical than theoretical. Since LD samplers introduce correlations between samples depending on the index and dimension of the needed random number, a quasi random engine must be able to index samples and distinguish how each of their dimensions is used. Despite being very modular, Geant4 does not provide any interface to implement such methods nor query the required information.

In the following, we propose an implementation that never modifies Geant4 source code. This is important because Geant4 is a well tested toolkit and altering its source code is error-prone.

2.1.4.2 Dimension and index definitions

We considered the use of the Sobol' sequence, which, in dimension d, outputs the sequence $(x_i^1, x_i^2, ...x_i^d)_{i\in\mathbb{N}}$. In our implementation, a *dimension* is uniquely defined by:

— The location of the code where the random number generator is called.

— The interaction number of the current *track*.

— The rejection loop iteration, if the call originates from a rejection loop.

With this definition, since it is not possible to know in advance the maximal number of interaction per track and rejection loop needed, the number of needed dimensions is unknown *a priori*. This is a problem since the dimension d of the Sobol' sequence needs to be defined before the simulation starts. Hence, we need to define, for each type of interaction (Compton, Rayleigh, Photo-Electric effect...) the maximal number of interactions a particle can undergo and for each part of the code where a rejection loop is used, the maximal number of rejection iterations. Beyond these thresholds, all the interactions/loop iterations will be sampled with a PRNG.

To illustrate our dimension definition, we consider the case where only Compton interactions are to be sampled with a Quasi-Random sequence. We define d_{primary} as the number of calls to the random number generator for primary particles generation. For each photon emitted isotropically from a given source distribution:

— 5 random numbers are needed to determine its initial position and direction ($d_{\text{primary}} = 5$).

— For each process (*i.e.* interaction type) defined in the physics list, the length to the next interaction is sampled according to cross-sections, with 1 random number. The process with the shortest length is selected.

— If the selected process is a Compton interaction:

  — One random number is used to select the atom with which the interaction occurs.

  — One random number is needed to compute the new direction of the scattered photon.

— 3 random numbers are needed in a rejection loop for the scattered photon energy.
— 5 random numbers, also in a rejection loop, for the ejected electron information choice.

If we restrict the number of Compton interactions sampled with QMC to n_{Compton} and the maximal number of rejection loops for the photon and the electron to n_γ and n_{e^-}, the maximal number of QMC dimensions needed will be $d_{\text{primary}} + n_{\text{Compton}}(3 + 3n_\gamma + 5n_{e^-})$. For every new photon, a new index of the sequence is used. In case all the dimensions are not used, *i.e.* less than n_{Compton} occur or the two rejection loops end before n_γ and n_{e^-} iterations are reached, the corresponding dimensions are left unused. The definition of dimensions is illustrated in Figure 2.4.

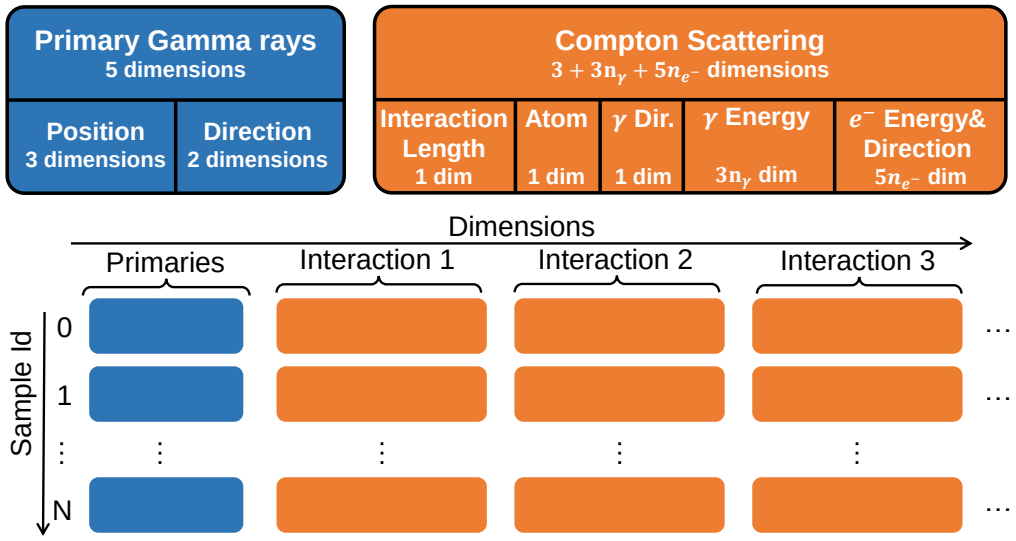


FIGURE 2.4 – Illustration of the dimension definition and usage when Compton Scattering is the only considered process for quasi-random sampling. This figure was adapted from the PhD thesis of Bastien Doignies [52].

2.1.4.3 Implementation

PRNG in Geant4 [1] are implemented via few functions and maccros wrapping with the CLHEP library [144]. Ultimately, every pseudorandom number is generated through an object deriving from the `CLHEP::HepRandomEngine` abstract class. The class has two functions of interest for sampling: `flat` and `flatArray`. Both methods generate pseudorandom numbers in the unit interval, the difference being how many points are generated by the call. To introduce QMC in Geant4, we slightly modify this interface to introduce a new parameter from the C++20 <`source_location`> library.

The <`source_location`> library provides access to certain information about the source code. Available information include file name, function signature, line and column. Through function signature, it is possible to recover the name, the class and the namespace when applicable. Previously,

those information were only available with maccros expansions of `__LINE__`, `__FILE__ and __PRETTY_FUNCTION__`. The library offers a better alternative as those information are now accessible by adding a default parameter to functions. Interestingly, adding a default parameter does not require modifying calls to the function, hence this simple addition does not require modifying Geant4 source code in itself.

The exact calling context information acquired with the aforementioned method is not enough to deduce the exact dimension. One particle may undergo multiple times the same interaction. In that case, the pseudorandom number generator is asked from the same location in the source code, but the implied dimension should be different. To accommodate this situation, it is sufficient to track the number of time an interaction occurs. This information is not available directly in Geant4 but can be easily computed by setting up callback on every steps.

Still, the information obtained mostly depends on the calling location of the pseudorandom number generator. In some cases, the call is done in utility function and it becomes harder to track the context. One example is simply sampling from alternative - eg normal - distribution. The code is abstracted by the CLHEP library in dedicated methods and the call location is therefore not informative. Two main solution exist to solve this issue. The first one involves getting the call stack. C++23 introduces the `<stacktrace>` library for this purpose. However the standard is not officially released and compilers do not have full support for this module. Moreover, the stacktrace might not be accurate enough, especially since Geant4 is viewed as library from Gate perspective and the necessary information about calls are not available if compiled in release mode. The second solution involves wrapping Geant4 physical processes. By exposing manually the process and function name when possible, the information loss is mitigated. This is the chosen solution for this project.

The described processes have a substantial cost. Sobol' points are extremely fast to compute and can be even faster than some common PRNG [150]. In itself, the overhead of the `<source_location>` is negligible. However, parsing the information and maintaining the current state is not. Moreover, a callback must be installed on every particle step during a simulation to detect particle interactions. This is why the proposal here is a feasibility experimentation and do not correspond to a realistic use case.

2.1.4.4 Experiments

We designed two toy experiments performed in different settings to show the interest of Quasi Random sequences compared to PRNG in Monte Carlo simulations for nuclear medicine.

SPECT simulation

The first experiment was the simplified simulation of the emission part of a SPECT imaging acquisition. We considered 6 spherical gamma sources of

diameter 10, 13, 17, 22, 28 and 37 mm, isotropically emitting 140 keV gammas (similar to ^{99m}Tc) with a sphere-to-background ratio of 20:1 (see Figure 2.6). Production cuts were set to 1mm.

A first simulation was performed in vacuum, in order to consider a minimal number of random numbers needed for the simulation. Indeed, as no interaction can occur, the number of random dimensions needed is limited to 5 dimensions used for sampling primary particles, *i.e.* their initial position and direction. In the QMC simulation of this experiment all 5 dimensions were sampled with the Sobol sampler. This simulation is performed to highlight the effect and interest of QMC in this context compared to analog MC, for different number of primary particles.

The second simulation was similar to the previous one but in a phantom filled with water instead of vacuum. In this situation, photons should therefore interact with matter through Compton scattering mainly but also Photoelectric effect. Through these interactions, photons deposit energy to matter and create secondary electrons, thus increasing the number of random dimensions. As explained above, using a Quasi Random Sequence in this context raises several difficulties considering the number of possible interactions, number of steps and usage of rejection loops. For this reason, we first restrict the *physics list* so that only Compton scattering can occur, which is the main possible interaction for gammas at this energy. Only the random numbers needed for primary photons emission (position and direction) and for the first Compton interaction ($n_{\text{Compton}} = 1$) were sampled with the Sobol sampler and $n_{\gamma} = n_{e^-} = 2$, reaching 24 QMC dimensions. All other random numbers were sampled with conventional white noise in the QMC simulation.

The output of each SPECT simulation is a 3D image reconstructed as follows: each particle exiting the phantom is saved in a spherical *phase space* [205] with all its information (particle type, position $\overrightarrow{p_{\text{out}}}$, direction $\overrightarrow{d_{\text{out}}}$, energy E_{out} and time Δt between emission and arrival in the phase space). Once the simulation is finished, the 3D image is reconstructed by back-projecting in the volume each photon of the phase space in straight line : the position $\overrightarrow{p}$ of the voxel in which the particle will be added is determined according to the following formula :

$$\overrightarrow{p} = \overrightarrow{p_{\text{out}}} - c\Delta t\overrightarrow{d_{\text{out}}} \tag{2.8}$$

where c is the speed of light. This straightforward reconstruction allows to compare images similar to SPECT images as the output of the simulation, without having to simulate all the detection process of a real Gamma Camera.

Both experiments were performed with an increasing number of primaries, from 10^6 to 2×10^7. The reference MC simulation was obtained with 10^8 primaries and was used to compute Normalized Root Mean Squared Error (NRMSE) with output images.

Dose simulation

Finally, a dose computation experiment was performed in a waterbox containing a 1 cm thick lung slice and a 3 cm thick bone slice. A 2 MeV photon cone beam was placed on one side of the box in order to irradiate the slices. An illustration of the simulation geometry is shown in Figure 2.5. The output of the simulation was a voxelized dose map computed by the Dose Actor of Gate. The reference MC simulation was the output dose map obtained with 2×10^8 incident photons. We used the *G4EmStandardPhysics option4* physics list for all simulations. Since the source is a mono-energetic focused beam, sampled from a disc, only two random numbers were needed for primaries generation (only 2d position, thus $d_{\text{primary}} = 2$). We compared the convergence speed of the simulation in three different setups: an analog MC simulation, a QMC simulation with $n_{\text{Compton}} = n_{\gamma} = n_{e^-} = 0$ (only d_{primary} QMC dimensions) and another QMC simulation with $n_{\text{Compton}} = n_{\gamma} = n_{e^-} = 3$ (total QMC dimensions = 83). The simulation with $N = 2 \times 10^6$ was performed 6 times for both MC and QMC (83 dimensions) settings to compute uncertainties and efficiencies.

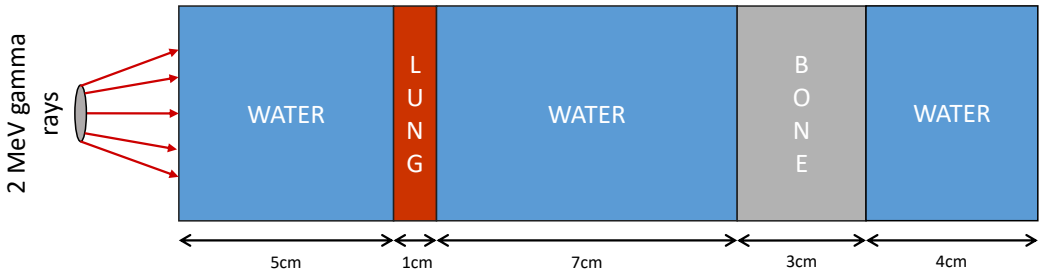


FIGURE 2.5 – Dimensions of the box in which the dosimetry simulations were performed. Gamma rays were emitted from a 2d disk with a determined beam direction. A voxelized dose map was scored in the volume made of 5 layers with different materials: water (5 cm), lung (1 cm), water (7 cm), bone (3 cm) and water (4 cm).

2.1.5 Results

SPECT simulation

Source and output images for the SPECT experiment in vacuum with $N = 2 \times 10^6$ primaries are shown in Figure 2.6. Clearly, the QMC output shows much more homogeneous spheres, showing a faster convergence to the simulation reference, which is the input source here. This faster convergence is also highlighted by Figure 2.7 showing between 15% and 40% error reduction. This toy experiment shows that in this simple case, choosing a LD sampler instead of a conventional PRNG can drastically reduce the computation time needed for a specific error tolerance.

Unfortunately, changing the simulation medium from vacuum to water annihilated almost all benefits of QMC. The MC and QMC outputs shown

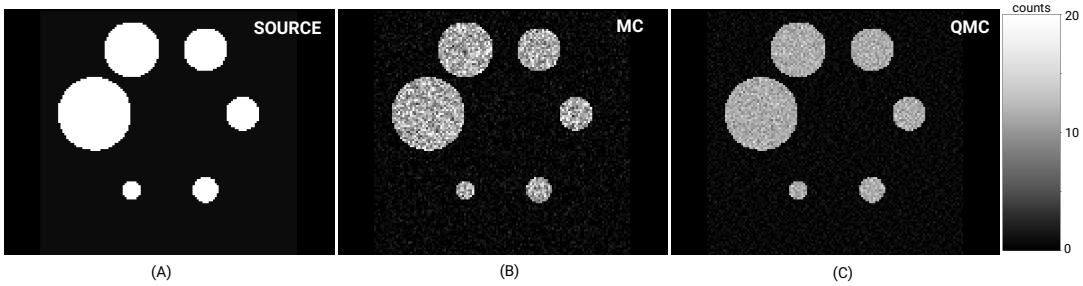


FIGURE 2.6 – (A) Six (voxelized) spherical sources with increasing radius. (B) and (C): Reconstructed images after MC and QMC simulations simulations with $N = 2 \times 10^6$ primaries in vacuum (experiment 1).

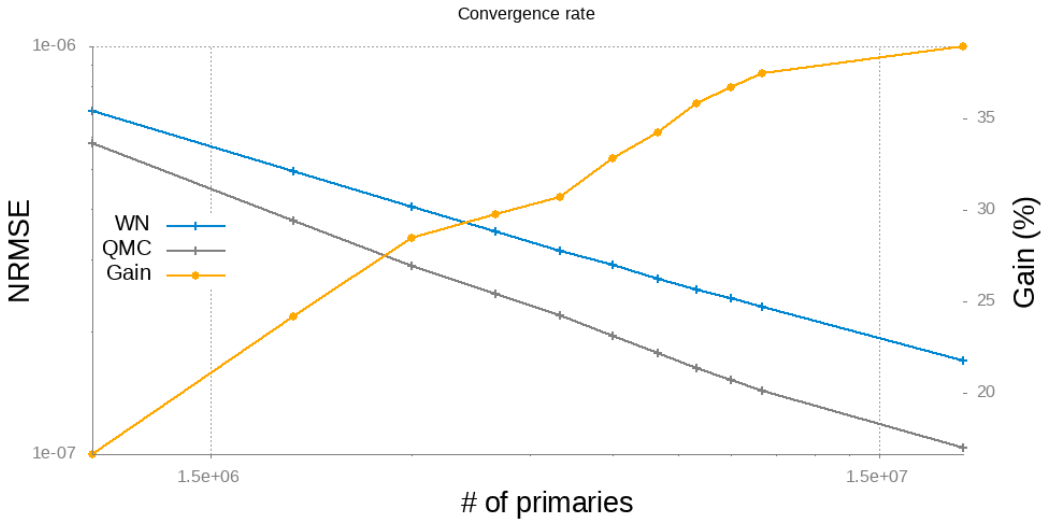


FIGURE 2.7 – Convergence plot of NRMSE of MC compared to QMC simulations in vacuum.

in Figure 2.8 are no longer distinguishable and visually seem to have comparable noise level. Convergence plot in Figure 2.9 shows between 2% and 5% improvement. However, this error reduction seems to linearly increase with the number of primaries and thus could be much more significant if the error tolerance was lower.

This difference with the simulation in vacuum can be explained by several factors. First, as mentioned, since the convergence speed of LD sequences is $O(\frac{(\ln n)^d}{n})$, increasing the number of QMC dimensions d slows down the convergence speed. Secondly, since all photons do not undergo a Compton scattering, some dimensions are left unused at the end of the simulation, while LD sequences are meant to be used entirely, otherwise the "unbiaisedness" is not guaranteed any more. Finally, since photons can undergo more than one Compton scattering, for which random numbers were sampled with white noise, the effect of the LD sequence on the first scatter may be insufficient.

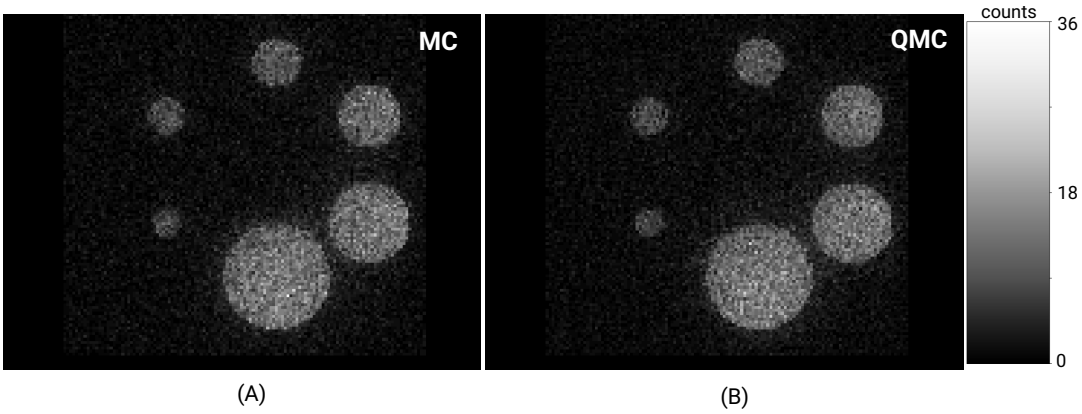


FIGURE 2.8 – Simulations results for the SPECT experiment results with $N = 4 \times 10^6$ primaries. *Left.* The simulation using White Noise samples. *Right.* The simulation using Sobol' samples.

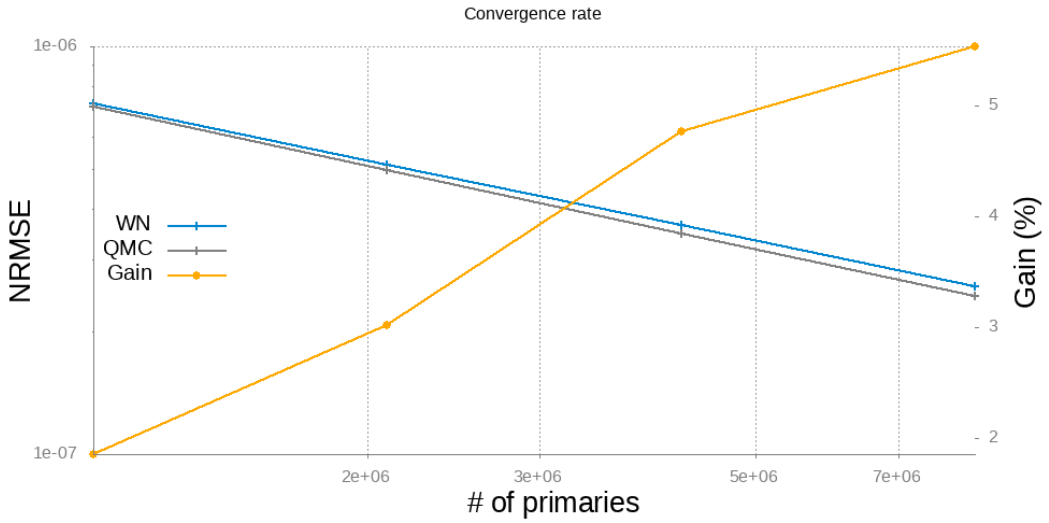


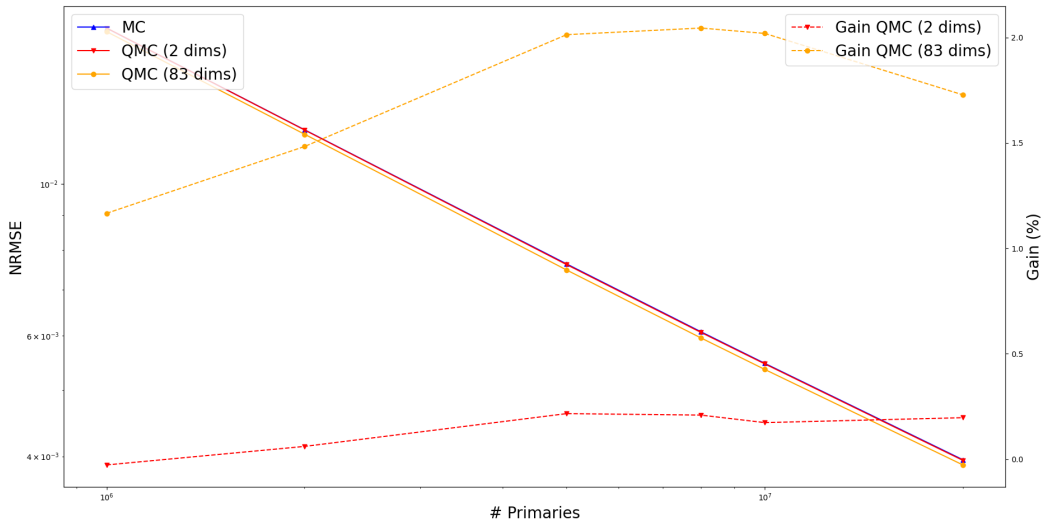
FIGURE 2.9 – Convergence plot of NRMSE of MC compared to QMC simulations in water.

Dose simulation

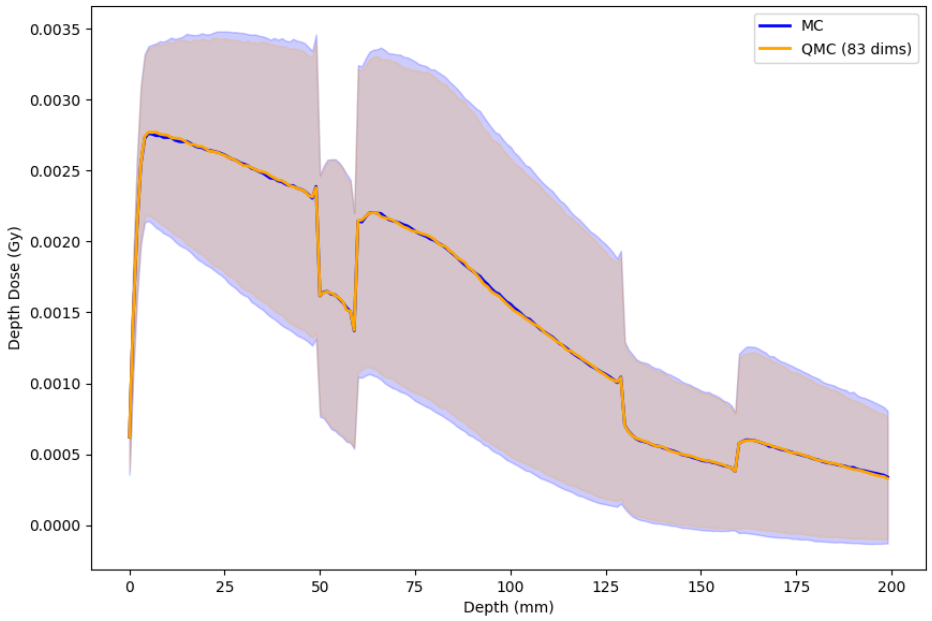
For the dose computation simulations, convergence rates can be compared in Figure 2.10a. The integrated dose profile and uncertainty for $N = 2 \times 10^6$ is shown in Figure 2.10b. These results show that both QMC methods allow to slightly reduce the NRMSE by a few percents. The uncertainty on the depth dose curves in Figure 2.10b, computed over 6 realizations of the $N = 2 \times 10^6$ simulation, is also reduced in the QMC simulation, specially in locations where the dose is the highest.

However, the magnitude of the NRMSE reduction always remains lower than 2%, which is not satisfying, specially considering the increased computation cost, due to dimensions parsing, associated to QMC simulations. For the simulation with $N = 2 \times 10^6$ primaries, the efficiency of the MC simulation

was 20.98 while being close to 20.83 for the QMC simulation (with 83 dimensions). This means that the uncertainty reduction introduced by the usage of LD sequences did not sufficiently compensate the increased computation time in this case.



(A) Convergence plot of NRMSE of the dose simulation, for MC compared to two QMC strategies



(B) Corresponding dose profiles for $N = 2 \times 10^6$ integrated along x and y dimensions and the associated uncertainty computed over 6 independent realizations of each simulation.

FIGURE 2.10 – Comparison of MC and QMC convergence and depth dose curves in the Dose computation experiment.

Even though this error reduction is not satisfying for a VRT, this result

shows that the way random dimensions are defined can influence the convergence speed of the simulation and still improve the incertitude by a few percents.

2.1.6 Discussion, perspectives and conclusion

This exploratory work focused on using LD sequences to sample random numbers used during MC simulation in medical physics. These low-discrepancy sequences are meant to introduce correlation between samples to force homogeneity in the sampling space and thus reduce the integration error. We show that is is possible to implement a QMC framework into Geant4 and Gate software, without requiring source code modification of the toolkit, that lead to encouraging results for very simple test cases. Once implemented, this unbiaised VRT is easy to use and can lead to 40% error reduction for a SPECT simulation in vacuum. However, in higher dimensions, the benefit of the method is minor and highly depends on the dimensions chosen to be sampled with QMC, which seems to be simulation-dependent.

This also shows that future works on QMC for medical physics should focus on ways to determine which dimensions influence the most the output of a specific simulation in order to sample those dimensions with LD sequences.

This work has been discontinued to prioritize other research tasks within the scope of this PhD project. Although initial findings were promising, subsequent tests did not show sufficient improvement. Given the time required to continue this line of investigation, it was deemed outside the initial scope of the PhD. However, we hope that the initial findings could be revisited in the future if time and resources permit.

2.2 Fast MC simulations with Deep Learning

As mentioned earlier, a Monte Carlo Simulation of a SPECT acquisition can be computationally very expensive. Such simulations can be decomposed into two distinct steps. The first step is the tracking of particles into the patient, from their emission to their exit, including interaction with matter. The second step consists in tracking exiting photons during their interaction with the gamma camera (collimation, scintillation, photo-multiplication and digital conversion). We present here two deep-learning based methods, one for each step, that can be used to drastically reduce the MC computation time.

Those methods were designed and developed in our team during the past years and, while I was not the leader of this project, I actively participated to this development. In particular, I performed various evaluations and re-implemented the "standalone" version (independent of Gate, see Section 2.2.5) of the method, with full pytorch continuity.

2.2.1 Generative Adversarial Networks (GANs) for particle generation in SPECT

Particle tracking inside the patient is very inefficient because very few of the emitted photons will effectively exit the patient (due to absorption) and few (0.01%) exiting gammas will be detected by the gamma camera (due to collimation). When the patient and source are fixed, the output of the tracking inside the patient may theoretically be stored in a large file of several GB, containing all photons exiting the patient and reaching a sphere. For each of these photons, their position, direction, energy and the time elapsed since their emission are stored in this file, usually referred to as a *phase space file*. Any other useful information about exiting particles can be stored. Then, if the acquisition system changes, the same *phase space file* can be used as the input source, thus saving the patient tracking time.

This time-saving trick still has some limitations. First, the size of these *phase space files* can rapidly grow if a good statistically relevance is needed, making them difficult to store, process and use. Moreover, if, for some reason, the source radionuclide distribution inside the patient has to change, a completely new *phase space file* has to be generated.

These limitations can be leveraged by training a Generative Adversarial Network (GAN). The idea is to train a neural-network G, the *generator*, to output particles with the same distribution as the ones contained in a previously generated *phase space*. More precisely, the GAN is made *conditional* (condGAN) with respect to the primary particles properties (position, direction, energy), which allows to use it to generate any source distribution. We consider here only the case where the source is ^{99m}Tc (considered mono-energetic, 140.5 keV). We present here the usage of condGANs for SPECT simulation acceleration, published in [202, 197], but this idea was first proposed for photon beam modeling [200] and was also extended afterwards to PET imaging in [201].

2.2.2 GAN training

Given a patient CT image, the training *phase space* was generated from a uniform ^{99m}Tc source distribution inside its body contour, with $\simeq 10^7$ isotropically generated gammas. This uniform source is used only for training, in order to simulate every possible source location. Each particle in the phase space was represented by a 12-dimensional vector $(\boldsymbol{p_0}, E_0, \boldsymbol{p}_{\text{exit}}, \boldsymbol{d}_{\text{exit}}, E_{\text{exit}}, \Delta t)$, where $\boldsymbol{p_0}$ and E_0 are the emission position and energy, $\boldsymbol{p}_{\text{exit}}, \boldsymbol{d}_{\text{exit}}$ and E_{exit} are the exit position, direction and energy and Δt is the time elapsed between the particle emission and arrival in the phase space sphere. Since the considered radionuclide is ^{99m}Tc, all E_0 are equal to 140.5 keV but a different energy spectrum could be considered for a different source. Absorbed particles were artificially added to the *phase space* with a zero energy and the position obtained after projection in the phase space sphere.

This training dataset is used to train a generator G to output a realistic exiting photons distribution, given a noise vector z of dimension d, an input emission position and an initial energy. The GAN training framework also

involves the training of a Discriminator D, trained to distinguish between real and fake samples. D takes as input the full vector of emission and exit information of a particle and outputs a real number that should be high if the sample really comes from the training dataset (*real* sample) and low if the sample was generated by G (*fake* sample).

$$G : \begin{cases} \mathbb{R}^{d+4} & \rightarrow & \mathbb{R}^8 \\ (z, \boldsymbol{p_0}, E_0) & \mapsto & x_{\text{fake}} = (\boldsymbol{p}_{\text{exit}}, \boldsymbol{d}_{\text{exit}}, E_{\text{exit}}, \Delta t) \end{cases} \tag{2.9}$$

$$D : \begin{cases} \mathbb{R}^{12} & \rightarrow & \mathbb{R} \\ (\boldsymbol{p_0}, E_0, \boldsymbol{p}_{\text{exit}}, \boldsymbol{d}_{\text{exit}}, E_{\text{exit}}, \Delta t) & \mapsto & y \end{cases} \tag{2.10}$$

For each condition vector $c = (\boldsymbol{p_0}, E_0)$ from the training dataset and corresponding to $x_{\text{real}} = (\boldsymbol{p}_{\text{exit}}, \boldsymbol{d}_{\text{exit}}, E_{\text{exit}}, \Delta t)$ contained in the training dataset, D is trained such that $D(c, x_{\text{real}})$ is a high real value and $D(c, x_{\text{fake}})$ is a low real value, where $x_{\text{fake}} = G(z, c)$ and z is sampled from a d-dimensional standard normal distribution. As its name might suggest, the discriminator D is trained to distinguish between *real* and *fake* samples. On the other hand, the generator G, is trained to fool the discriminator, that is, G is trained such that for $x_{\text{fake}} = G(z, c)$, $D(c, x_{\text{fake}})$ is high. These requirements lead to the following training loss definitions:

$$\begin{array}{rcl} \mathcal{L}_D & = & \mathbb{E}_{z \sim \mathcal{N}(0,1)^d}[D(c, G(z, c))] - \mathbb{E}_{x_{\text{real}} \sim \mathbb{P}_{\text{real}}}[D(c, x_{\text{real}})] \\ \mathcal{L}_G & = & -\mathbb{E}_{z \sim \mathcal{N}(0,1)^d}[D(c, G(z, c))] \end{array} \tag{2.11}$$

where $\mathcal{L}_G$ and $\mathcal{L}_D$ are G and D training losses and $\mathbb{P}_{\text{real}}$ is the true exiting particles distribution from the training dataset.

The conditional GAN training approach is illustrated in Figure 2.11.

Both G and D were Multi-Layer Perceptrons (MLP) containing respectively L_G and L_D hidden layers with 500 neurons, followed by a LeakyReLu activation function. In the following, we used $L_G = 3$ and $L_D = 2$. We chose $d = 15$ for the noise vector z dimension. G and D parameters were optimized simultaneously with two Adam optimizers, with 4 D updates and 1 G update at every iteration. The training batch size was set to 2×10^4, the learning rate to 4×10^{-4} and the number of epochs to 200.

GAN training can sometimes be quite unstable and difficult to control. At the end of a successful training process, the Discriminator D should not be able to differentiate between *real* and *fake* samples, meaning that L_D is close to zero and that the Generator G is well trained to output realistic samples.

2.2.3 Results

Once G is trained for a specific CT image, it can be used to generate as many photons as desired by sampling a random position $\boldsymbol{p_0}$ from the desired source distribution and a random vector $z \hookrightarrow \mathcal{N}(0,1)^d$ used to compute $x_{\text{fake}} = (\boldsymbol{p}_{\text{exit}}, \boldsymbol{d}_{\text{exit}}, E_{\text{exit}}, \Delta t) = G(z, \boldsymbol{p_0}, E_0 = 140.5\text{keV})$. During inference, the source distribution can be any distribution.

As an example, we trained a condGAN on the NEMA IEC phantom and compared an analog MC simulation with a condGAN simulation. As described earlier,

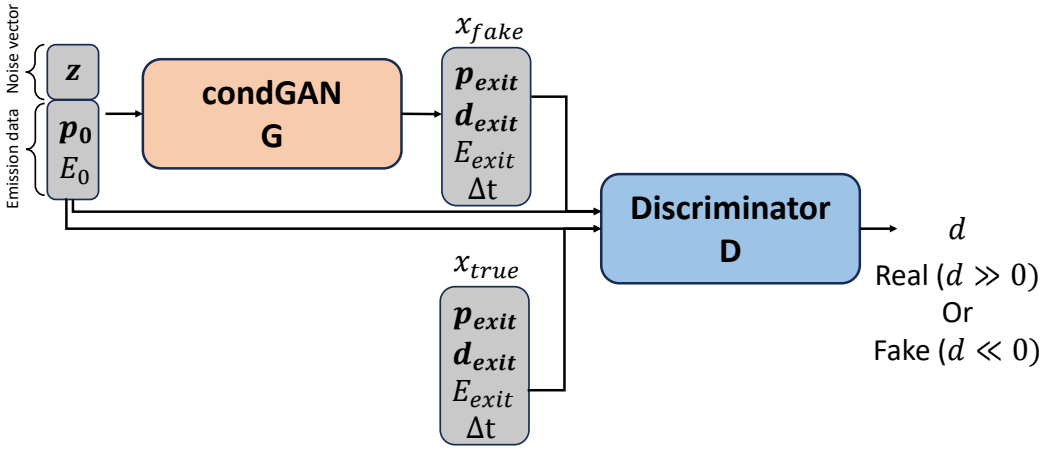


FIGURE 2.11 – Illustration of the conditional GAN (condGAN) training scheme. The condGAN G takes as input a noise vector z and the primary γ emission data information ($\boldsymbol{p_0}$, E_0) and outputs a realistic vector of the corresponding exiting γ information ($\boldsymbol{p}_{\text{exit}}$, $\boldsymbol{d}_{\text{exit}}$, E_{exit}, Δt) The Discriminator D is trained to distinguish between *true* particles (from the training phase space) and *fake* samples (from G).

the training dataset was obtained with a uniform source inside the patient. For the evaluation, the ^{99m}Tc source was composed of six spheres and a cylinder placed in a background with a Target to Background Ratio (TBR) of 10. The source image shown on the top left of Figure 2.12. The number of generated primaries was 1.8×10^7 for both analog MC and condGAN simulation. In the MC simulation, photons reaching the *phase space* were stored. Resulting *phase spaces* were used to reconstruct two 3D-images by "backprojecting" each sample in an *ideal* way, as described in Equation (2.12), where c is the speed of light. In Figure 2.12, images obtained this way for the MC and the condGAN simulations can be compared, as well as two profiles. Both images are very similar, as well as the profiles, showing that the condGAN was effectively able to learn the target photon distribution. Moreover, the ratios between the number of counts in each sphere/cylinder in the GAN simulation and the MC simulation was always comprised between 0.95 and 1.02, showing good agreement between both *phase spaces*. The only visual limitation is that the edges of the spheres obtained with the GAN seem less sharp than in the MC simulation, due to a slight error in number of absorbed/scattered photons estimation.

$$p_{\text{backproj}} = p_{\text{exit}} - c\,\Delta t\, d_{\text{exit}} \tag{2.12}$$

Another condGAN was trained with a *phase space* generated from a patient's CT image. Once again, the training dataset was obtained with a uniform source inside the patient. For evaluation, three ^{99m}Tc source spheres of 52, 36 and 30 mm diameter were placed in a background, with TBRs of respectively 50, 20 and 40.

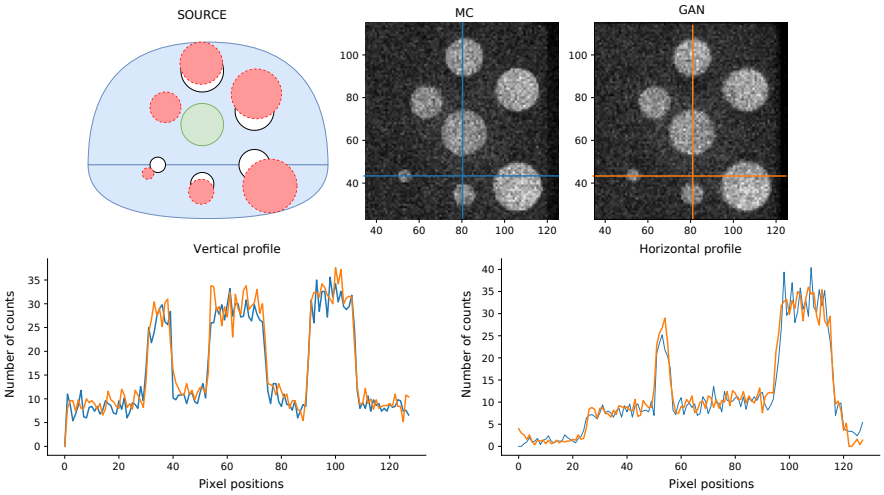


FIGURE 2.12 – Top: slice of the source image (left) and of ideally reconstructed images from MC (middle) and GAN (right) simulations. Bottom: vertical (left) and horizontal (right) profiles drawn from the corresponding MC (blue) and the GAN-generated (orange) images. This figure was extracted from [197].

Here, the complete SPECT simulation was performed with 120 projections, 10^9 gammas per angle and an isocenter-to-detector distance of 40 cm. The modeled SPECT system was the GE Discovery 670 with NaI(Tl) crystal and a Low Energy High Resolution (LEHR) collimator, with four heads. Two SPECT acquisition were performed one from a full analog MC simulation of the SPECT acquisition and one with the condGAN replacing the tracking in the patient only. Both sinograms were then used to reconstruct a 3D-image with the RTK software [185], and the OSEM algorithm with 10 iterations and 15 subsets, attenuation correction, double energy window scatter correction and Resolution Modeling [262]. Reconstructed images are shown in Figure 2.13 and show once again very good agreement.

2.2.4 Discussion

The previous sections showed that the condGAN-generated phase spaces were very similar to the ones generated by an analog MC simulation. However, the main advantage of using a GAN is their computational speedup. The number of Particles Per Second (PPS) that can be generated by a condGAN is about 8×10^5, while being near 7×10^3 for the MC case, leading to a $\times 114$ computational speedup. This speedup can even be higher if the complexity of the analog MC simulation increases, for example if the source or CT voxel sizes were smaller, while the speed of a condGAN simulation is constant.

One limitation is that training a GAN can be difficult, unstable or even fail. Moreover, even though the GAN-generated distribution does not perfectly match the target MC distribution, we argue that using such GANs for SPECT simulations is generally not limiting the accuracy of the simulation, considering the poor spatial resolution of SPECT. One last limitation is that a condGAN is trained for a specific CT,

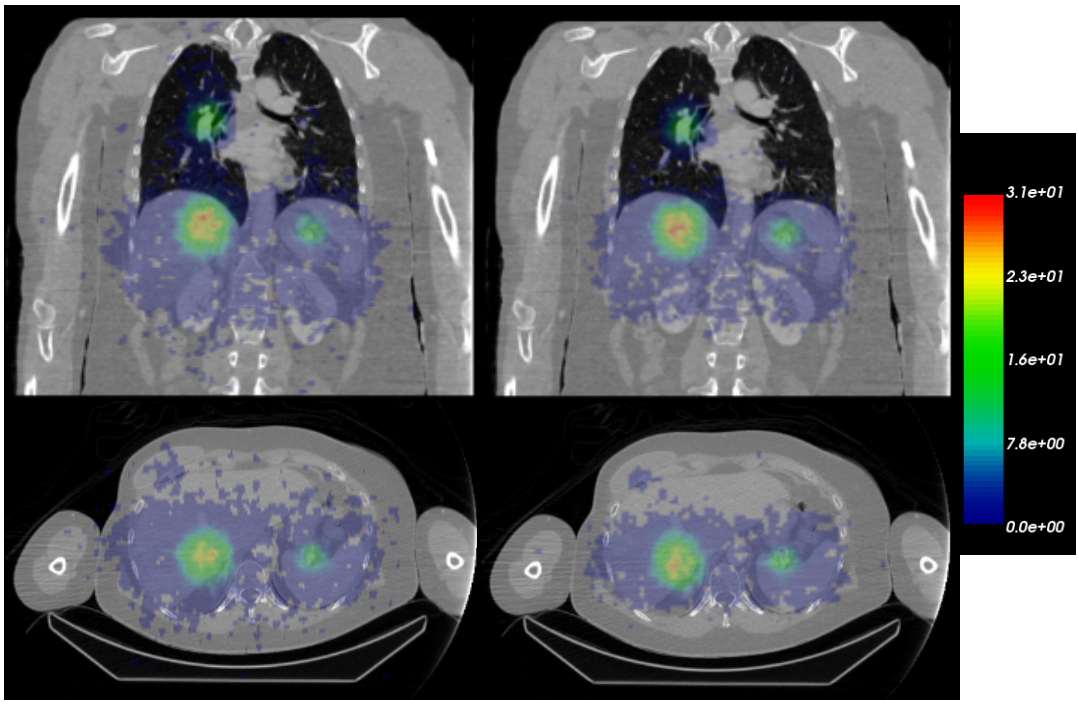


FIGURE 2.13 – Coronal (top) and transverse slices (bottom) of images reconstructed with projections obtained from the reference Monte Carlo Simulation (left) and with projections obtained from the photons generated by the conGAN (right). This figure was extracted from [197].

meaning that a new training *phase space* generation and training would be required if the phantom or patient was changed.

2.2.5 Contribution: GAN sources and nn-ARF combined

GANs can also be combined with an Angular Response Function (ARF) modeling a given detector response to replace and accelerate the second step of the MC simulation, *i.e.* tracking inside the detector. An ARF can either be a tabulated histogram, as in [223], fitted on a large dataset to compute detection probabilities of incoming photons based on their solid angle and energy, or it can be modeled by a fully connected neural network (nn-ARF), as proposed in [199]. This former approach drastically accelerates tracking inside the detector. An nn-ARF has to be trained once for each detector, collimator and desired energy window ranges, but can then be used regardless of the patient and source.

Long and computationally expensive Monte Carlo simulations with particle tracking in the patient and detector can then be replaced by the application of two successive neural networks, acting as a standalone forward model. Figure 2.14 illustrates this method. One question that arose is what to do between both neural networks? Three options are possible:

Option 1 Once a batch is generated by the GAN, a GATE event is created for each of the exiting particles. The classical MC engine is used to track each particle until they eventually reach a detector plane. Then, position, direction and energy of

intersecting particles are used as input to apply the nn-ARF and generate the projection.

Option 2 In a majority of cases, tracking between the output of the GAN and the nn-ARF detector is just a straight line ray tracing in air. For this reason, another option is to not use the GATE engine to create and track each exiting particles, but projecting them on each detector plane using the *Numpy* library.

Option 3 Finally, we proposed an implementation of Option 2. where the *Numpy* ray tracing is replaced by *PyTorch* operations. The advantage is that CPU to GPU memory transfers are no longer required between emission and detection, taking full advantage of GPU acceleration.

Pseudo-code for Options 2 and 3 is detailed in Algorithm 1. The code is available in the opensource collaboration Gate [206] [3]

To illustrate the potential of the full *Pytorch* implementation, we performed SPECT simulations of a patient with its CT image and a voxelized ^{99m}Tc source made of three spheres (with 28mm, 52mm and 76mm diameters). Three SPECT Intevo Gamma camera heads were modeled at 0°, 100°and 230°angle positions. The reference (ANALOG) MC simulation was compared to its version with a nn-ARF replacing the tracking inside detectors, to the GAN+nn-ARF within the Gate framework (Option 1) and to standalone GAN+nn-ARF implementations with *Numpy* (Option 2) and *PyTorch* (Option 3). In each simulation, 1 MBq of activity during 30 seconds were simulated, with one thread.

Results in terms of number of Particles per Seconds (PPS) tracked are detailed in Table 2.1. Using only the nn-ARF increases the number PPS tracked by 20%, but, as discussed in [199], the gain in terms of efficiency *i.e.* considering the noise reduction is much higher (between 10 and 3000 in [199]). This is because, unlike in Analog MC Simulations, each particle reaching the detector plane contributes to the projection, even slightly. Replacing the tracking in the voxelized patient by a GAN resulted in number of PPS multiplied by a factor 11.2 for the Option 1, 27 for Option 2 and 96 for Option 3. This shows the great interests of GANs, in particular when full usage of GPU acceleration and tensor operations is leveraged.

3. See https://github.com/OpenGATE/opengate and https://github.com/OpenGATE/gaga-phsp

Algorithm 1 Pseudo-code of the standalone implementation of SPECT simulation with GAN and nn-ARF

Inputs: Voxelized source distribution f, Number of particles to create N, number of projections N_{p}, batch size B, conditional GAN G, nn-ARF d
Output: Projections P_1, $P_2 \dots P_{N_{\text{p}}}$
$N_{\text{current}} \leftarrow 0$
while $N_{\text{current}} < N$ **do**
    $\text{batch}_{\text{init}} \leftarrow$ sample from f
    $\text{batch}_{\text{exit}} \leftarrow G(\text{batch}_{\text{init}})$
    $(b_1, b_2, \dots, b_{N_{\text{p}}}) \leftarrow$ Ray-Tracing and intersection with detector planes of $\text{batch}_{\text{exit}}$
    for i=1 **to** N_{p} **do**
        $P_i \leftarrow P_i + d(b_i)$
    end for
end while

Method	Particles per Seconds (PPS)
ANALOG	10k
ANALOG+nn-ARF	12k
GAN+nn-ARF (Option 1)	112k
GAN+nn-ARF (Option 2)	273k
GAN+nn-ARF (Option 3)	960k

TABLE 2.1 – Comparison of number particles per seconds (PPS) tracked in SPECT simulation between a reference MC simulation (ANALOG), tracking in detectors replaced by nn-ARFs (ANALOG+nn-ARF), both tracking steps replaced by a GAN and a nn-ARF (GAN+nn-ARF) within the Gate framework (Option 1), with *Numpy* ray tracing between both networks (Option 2) and with *PyTorch* ray tracing.

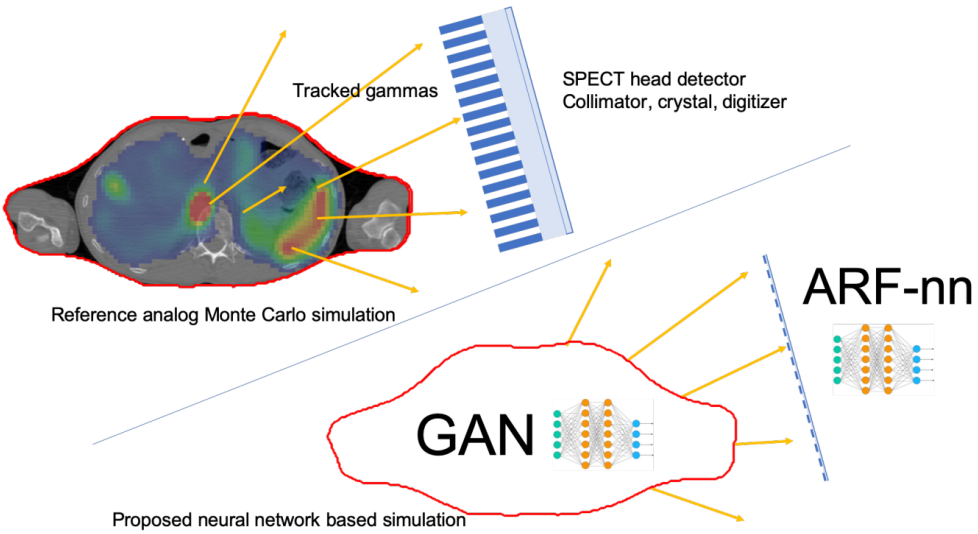


FIGURE 2.14 – Illustration of the GARF method, replacing analog Monte Carlo simulations (top) by a GAN that generates exiting gammas and a nn-ARF taking them as inputs to generate the projection (bottom). This figure was extracted from [202].

Chapter 3

PVEs and state of the art of PVC in SPECT

Sommaire

This Chapter aims at defining what are PVEs in SPECT, their causes and consequences, as well as to review state-of-the-art approaches for Partial Volume Correction (PVC).

3.1 Partial Volume Effects

3.1.1 Introduction

Definition

As described in Section 1.2.4, several physical effects need to be corrected either before, during or after reconstruction, in order to obtain a reconstructed image with accurate activity estimation. Attenuation can be corrected quite accurately by using CT-derived linear attenuation coefficients in the system matrix used for reconstruction (see Section 1.2.4.1). Scatter can be estimated and corrected with DEW [100] or TEW [162] methods before reconstruction if the appropriate energy windows are acquired (see Section 1.2.4.2). However, even tough Resolution Modeling (RM), presented in Section 1.2.4.3, is useful to compensate the effect of collimator blurring and for resolution recovery, it is still insufficient. Indeed, even after RM, the estimated activity in the 3D image is still not proportional to the true activity in the object, specially for objects of small size.

Visually and quantitatively, PVEs are usually characterized by the fact that a relatively high amount of counts are reconstructed outside the boundaries of the object they should be in, leading to underestimation of the activity in this object and overestimation in the neighbor object. In the first definition of Partial Volume Effects in Emission Tomography, Hoffman, Huang and Phelps [87] describe PVEs as the "depression of apparent isotope concentration in an image for objects equal in size to the instrument resolution (FWHM)". In fact, this underestimation of activity concentration occurs for all objects with size lower than $2\times$ FWHM [186, 37, 146]. Since this resolution can be different in different axes, the *resolution volume* is defined as a volume with size $(2 \times \text{FWHM}_x \text{ mm}) \times (2 \times \text{FWHM}_y \text{ mm}) \times (2 \times \text{FWHM}_z \text{mm})$. For objects with bigger size, their estimated activity will be proportional to their true activity.

Consequences

PVEs have a significant effect in various tasks of clinical settings in oncology. First, PVEs can be a confounding factor, since very different uptake and size can lead to similar SUV values in the reconstructed images [146], thus altering the staging and diagnosis based on the SPECT image. PVEs not only alter visual analysis of SPECT images but also their quantitative accuracy. Several studies have shown that uncorrected images tend to underestimate SUV values due to Partial Volume Effects [20, 212, 85, 65, 146, 74, 77, 194]. The quantitative inaccuracies of SPECT images may be one of the reasons why personalized ^{177}Lu-based Targeted-Radionuclide Therapies (TRT) is still not fully effective in clinical routine. As described earlier in Section 1.2.5, individual treatment response assessment and treatment planning make use of several ^{177}Lu-SPECT/CT images between injections to compute Time-Activity Curves (TAC) in Organs at Risk (OAR) and in tumour tissues. Then, one of the several dosimetric methods is employed to compute the absorbed dose in each ROI and draw personalized dose-response/dose-toxicity relationships. As a consequence, the limited spatial resolution of SPECT images has a major impact on dosimetric estimates [31, 147, 181, 146, 130] leading to between 63% to 99% underestimation of absorbed dose in the example of [146].

Illustration

PVEs are illustrated in a toy experiment in Figure 3.1 where 6 spherical sources with 10/13/17/22/28/37 mm diameters and with the same activity concentration were projected with the RTK software [185] with PSF modeling ($\text{FWHM}(d) = 2.75 + 0.076d$ where d is the source-to-detector distance in mm), without attenuation, scatter nor noise. The image was reconstructed with the OSEM algorithm (10 iterations, 8 subsets), with and without RM during reconstruction.

Clearly, without RM, the smaller the sphere, the more the maximum intensity was underestimated and the sphere was blurred. The spreading of estimated activity outside the boundaries of the object, characteristic of PVEs, is called *spill-out* effect and leads to underestimation of activity inside the object and over-estimation of activity in the background. In presence of background activity, some counts belonging to

the background can be reconstructed inside an object of interest, leading to over-estimation in this object. These effects are often designed as *spill-out* effects. The term *spill-over* designs both *spill-in* and *spill-out* effects.

One key metric to quantitatively estimate PVEs is the Recovery Coefficient (RC), first defined in [87] by Equation (3.1). RC can be computed for objects of interests as well as for background estimation. A RC close to 1 means almost no PVEs. RCs for the six different spheres were computed and shown in Figure 3.2.

$$\mathrm{RC} = \frac{\text{Mean Estimated Activity}}{\text{True Activity}} \tag{3.1}$$

Figures 3.1 and 3.2 also show that using RM during reconstruction partially succeeds to deblur the image and to mitigate *spill-out* effects both visually and quantitatively. However, the three largest spheres of Figure 3.1 exhibit an heterogeneous pattern with over-estimation of activity at the edges of the spheres and under-estimation in its center. This well documented effect is known as *ringing artifacts* or *Gibbs artifacts* and tends to be amplified when using RM with a high number of iterations. These Gibbs artifacts are due to PSF modeling (RM) during forward and backprojections of the MLEM algorithm and to limited sampling in the image space, both acting as low-pass filters truncating the high frequencies needed to reconstruct sharp objects like the spheres boundaries in Figure 3.1. Without the inclusion of the high frequency signal, these oscillations appear as the number of iteration increases, so that the mean value in the object is consistent with the observed data [221, 260, 170].

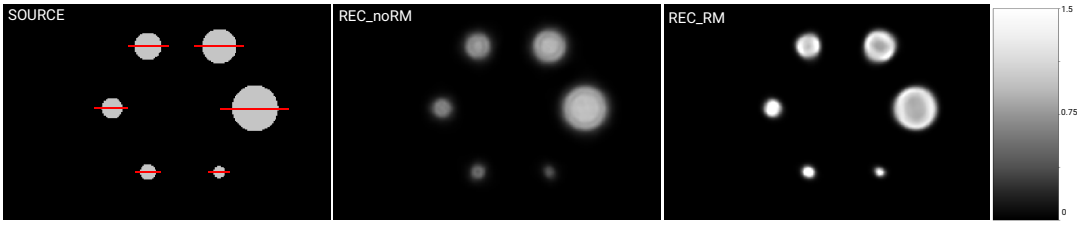
One simple strategy to reduce Gibbs artifacts is to blur the image (during or after reconstruction) to remove these oscillations [220]. Many other regularization techniques were also proposed such as Total Variation [192, 169], quadratic [45] or Bayesian [76]. Another strategy is to use a narrower PSF for RM than the "true" or estimated one. All these strategies can successfully reduce or even remove Gibbs artifacts but at the cost of poorer spatial resolution.

3.1.2 Causes of Partial Volume Effects

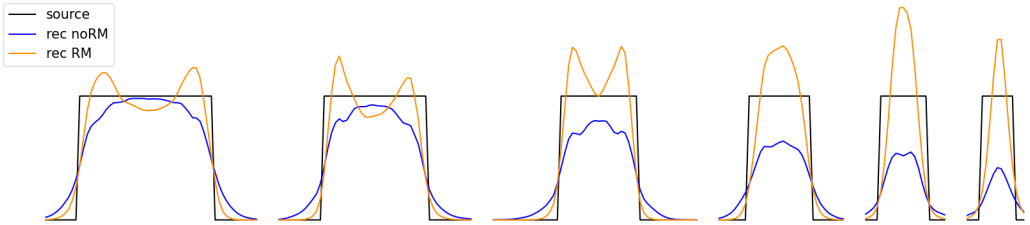
Partial Volume Effects are due to the limited spatial resolution of SPECT detectors. Gamma camera's response can be divided into two main effects: the intrinsic resolution and the collimator response. In both cases, the resolution of a detector is defined as the full width at half maximum (FWHM) of a profile obtained by imaging a point source. The image of the point source is designed as the Point Spread Function (PSF).

Intrinsic Spatial Resolution

The intrinsic spatial resolution refers to the spatial blurring due to the detector components (NaI crystal, light guide and PM tube) without considering the collimator response. It is mainly due to two effects: misplacement of γ detection due to scattering in the crystal and statistical fluctuations due to the number of photons reaching the PM tubes. The latter effect is often considered as negligible, since low or medium energy photons emitted by a ^{99m}Tc or ^{177}Lu source have a low probability



(A)



(B)

FIGURE 3.1 – Simulation of six spheres of ^{177}Lu, projected with PSF and reconstructed with and without Resolution Modeling (RM). No attenuation, scatter nor noise was considered in this simulation. Source and reconstructed images (A) and horizontal profiles of each image (B). The red lines in the source image in (A) correspond to the profiles drawn in (B).

to undergo a Compton interaction in NaI crystals. The former effect is the main cause of intrinsic resolution. It depends on incoming γ-ray energy E_0 since the number of photons reaching the PM tubes is directly related to the energy deposited in the crystal scintillator. Thus, the intrinsic resolution $\text{FWHM}_{\text{intrinsic}}$ is proportional to $\frac{1}{\sqrt{E_0}}$. The intrinsic resolution linearly improves with a decrease in crystal thickness due to reduced spatial spreading of scintillation light, although this comes at the cost of reduced efficiency. This trade-off between resolution and efficiency leads to crystal thicknesses chosen between 6 mm and 25 mm, leading to an intrinsic resolution comprised between 3 mm and 5 mm and a detection efficiency between 70% and 99% (at 140keV).

Collimator Resolution

The main causes of PVEs and poor spatial resolution in SPECT originate from the limitations due to the collimator placed in front of the detector. The role of the collimator is to select only the γ-rays coming with an angle orthogonal to the detector plane. All the other incoming γ-rays should be absorbed by the collimator, usually made of a lead block full of longitudinal holes, as it can be seen in Figure 3.3. The collimator is necessary to be able to map detected photons to a line integral in the patient and then reconstruct the 3D emission image from sinogram, but it is also very inefficient since a large majority of γ-rays reaching it are absorbed (usually more than 99%). Generally, collimator holes have an hexagonal shape but can also be rounded. Depending on the context (*i.e.* requirements in terms of field of view, resolution , efficiency), different types of collimator can be used: parallel, pinhole,

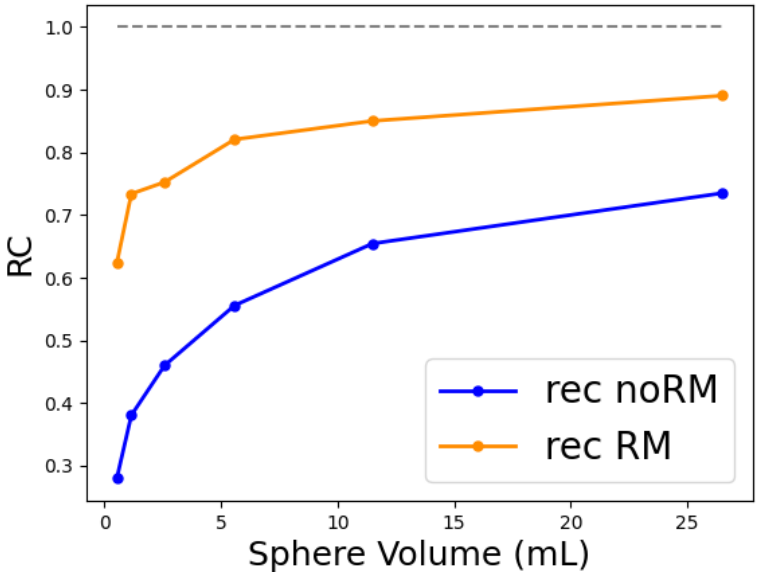


FIGURE 3.2 – Recovery Coefficients (RC) of the images of Figure 3.1a for all sphere volumes.

convergent and divergent. In the following, we will only consider hexagonal parallel holes collimators.

Ideally, the projected image of a γ-emitting point source should also be a point but in reality, it has a gaussian shape, as illustrated in one dimension in Figure 3.4. Two main factors cause this blurring.

The first one is potential septal penetration, *i.e.* a γ-ray passing through a septa of the collimator while it should be absorbed in it. This probability depends on collimator material, the energy of the interacting γ-rays and from the septal thickness. Lead is usually the preferred material since it has good γ-rays absorption properties (high Z and density) and is cheap. Depending on the type of radionuclide used (low, medium or high energy), the septal thickness t is chosen such that septal penetration probability is lower than 5%. *Low-energy* collimators can thus have a small septal thickness (near 0.16 mm) and be used with source under 150keV, while *medium-energy* ones (until 400keV) have a septal thickness close to 1.14 mm.

The other reason why a point source results in a blurred sphere is because of the collimator geometry. Indeed, since the length of the collimator holes is limited, the collimator only accepts a certain range of incoming solid angles. The narrower this range, the higher the resolution, resulting in a lower FWHM. It can be observed from Figure 3.4 that the FWHM would decrease with an increased hole length l, a decreased whole diameter h or a decreased source-to-detector distance d. The FWHM of the collimator can be expressed as a function of these parameters with Equation (3.2):

$$\mathrm{FWHM}_{\mathrm{colli}}(d) = h + \frac{h}{l_{\mathrm{eff}}} d \tag{3.2}$$

where $l_{\mathrm{eff}} = l - \frac{2}{\mu}$ is the *effective length* crossed by a γ-ray and μ is the linear

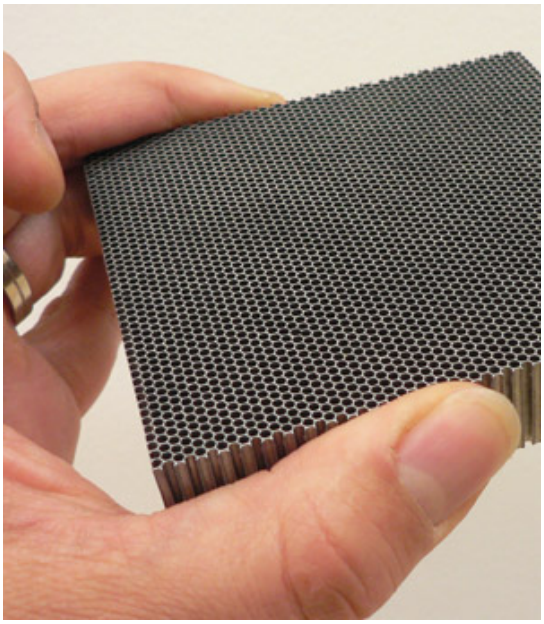


FIGURE 3.3 – Example of a parallel-hole lead collimator with hexagonal holes. This image was obtained from [64].

attenuation coefficient (in mm^{-1}) of the collimator material. Equation 3.2 might suggest a way to improve the detector's resolution: using very thin and long holes. However, using such a restrictive collimator would significantly decrease its efficiency ϵ_{colli}, *i.e.* the percentage of γ-ray emitted that pass through the collimator, thus leading to a noisier image. The efficiency can be expressed as:

$$\epsilon_{\text{colli}} = \left(K \frac{h^2}{l_{\text{eff}}(h+t)} \right)^2 \tag{3.3}$$

where K is a constant depending on the hole shapes ($K = 0.26$ for hexagonal holes) and t is the septal thickness. From Equations (3.2) and (3.3), it is clear that $\epsilon_{\text{colli}} \propto \text{FWHM}^2_{\text{colli}}$ meaning that improving the collimator's resolution can only be achieved at the cost of reduced efficiency, leading to a noisier image.

This noise-resolution trade-off is the reason why several types of collimators exist: *high-resolution*, *high-sensitivity*, and *general-purpose* collimators. Typically, *low-energy high resolution* (LEHR) collimators are used when the radionuclide is ^{99m}Tc, while *medium-energy general purpose* (MEGP) collimators are well suited for ^{177}Lu sources (see their photon emission peaks in Table 1.1). The characteristics of these two types of collimator are compared in Figure 3.5 and in Table 3.1.

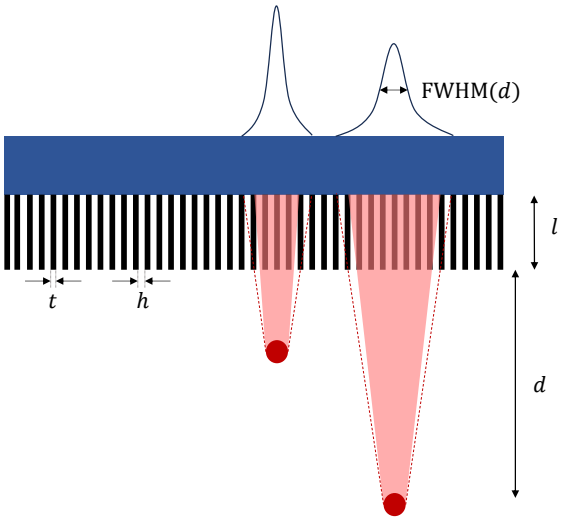
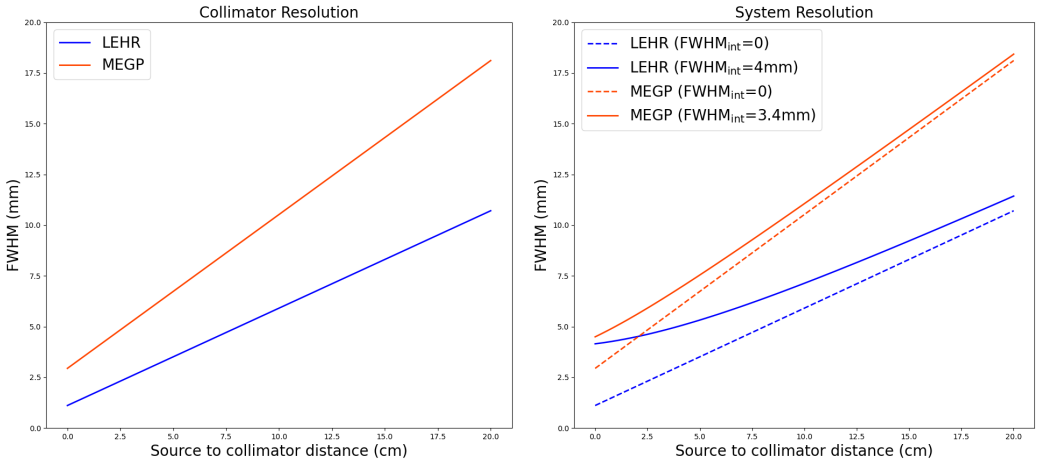


FIGURE 3.4 – Profiles and full width at half maximum (FWHM) of two point source placed at different distance from a parallel-hole collimator. d is the source-to-collimator distance, l the length of the holes, t the septal thickness and h the diameter of the holes.



(A) (B)

FIGURE 3.5 – Collimator (A) and System (B) Full width at Half maximum (FWHM) as a functions of source to collimator distance for low-energy high-resolution (LEHR) and medium-energy general-purpose (MEGP) collimators. Parameters of each collimator are listed in Table 3.1.

	LEHR	**MEGP**
Radionuclide	^{99m}Tc	^{177}Lu
Hole shape	Hexagonal	Hexagonal
Hole length	24.05 mm	40.64 mm
Septal Thickness	0.16 mm	1.14 mm
Hole Diameters across the flats	1.11 mm	2.94 mm
Collimator Resolution (FWHM) at 10 cm	5.9 mm	10.5 mm
Efficiency	1.18×10^{-4}	2.02×10^{-4}

TABLE 3.1 – Geometric specification of the Low Energy High Resolution (LEHR) and Medium Energy General Purpose (MEGP) collimators available with the Siemens Intevo Bold SPECT/CT system and their corresponding collimator resolution and efficiency derived from Equations (3.2) and (3.3)

As shown in one dimension in Figure 3.4, the collimator-induced PSF can be modeled by a Gaussian function. When the collimator and detector are planar, the 2D PSF can be expressed as a function of the distance d:

$$\mathrm{PSF}_{\mathrm{colli}}(x, y, d) = \frac{1}{2\pi\sigma^2_{\mathrm{colli}}(d)} e^{-\frac{x^2+y^2}{2\sigma^2_{\mathrm{colli}}(d)}} \tag{3.4}$$

where $\sigma_{\mathrm{colli}}(d)$ is related to the FWHM with the relation:

$$\sigma_{\mathrm{colli}}(d)) = \frac{\mathrm{FWHM}_{\mathrm{colli}}(d)}{2\sqrt{2\ln 2}} \tag{3.5}$$

System resolution

The combination of intrinsic resolution and collimator response defines the system resolution $\mathrm{FWHM}_{\mathrm{sys}}$. Since the convolution of two Gaussian functions with variance σ_1^2 and σ_2^2 is also a Gaussian function with variance $\sigma_1^2 + \sigma_2^2$, the system resolution at distance b from the detector is given by the Equation (3.6):

$$\mathrm{FWHM}_{\mathrm{sys}}(d) = \sqrt{\mathrm{FWHM}^2_{\mathrm{int}} + \mathrm{FWHM}_{\mathrm{colli}}(d)^2} \tag{3.6}$$

It can be seen from Figure 3.5b that, at common source-to-collimator distances (between 5cm and 30cm), the system resolution $\mathrm{FWHM}_{\mathrm{sys}}(b)$ can be considered a linear function, since the intrinsic resolution is typically much lower than the collimator resolution.

With a one-dimensional detector and a two dimensional object f, the full detection model is the attenuated and blurred Radon transform expressed in [11] and [179]. With the same notations as in Section 1.2.3, the attenuated and blurred Radon transform in SPECT reads as follows, $\forall r \in \mathbb{R}, \theta \in [0, 2\pi]$:

$$\begin{gathered} p_{\mu,h}(r,\theta) = \int_{-\pi/2}^{\pi/2} w(\delta)\, p_\mu(R\sin\delta + r\cos\delta, \theta+\delta)\, d\delta \\ \text{where } w(\delta) = h(R\tan\delta, R) \end{gathered} \tag{3.7}$$

where, p_μ is the attenuated Radon transform, given in Equation (1.30), R is the detector-to-isocenter distance and $r \mapsto h(r, d)$ is the system's PSF at distance d. The idea behind this formula is that the $L(r, \theta)$ (see Equation (1.4)) is not the only line integral contributing to the pixel $p_{\mu,h}(r, \theta)$, but instead, all line integrals $L(R\sin\delta + r\cos\delta, \theta + \delta)$ do (at least for small values of the angle δ), weighted by $w(\delta)$, depending on the PSF. A complete proof is given in [179].

3.1.3 Experimental PSF determination

In this section we describe some experiments and simulations we performed in order to determine intrinsic and collimator resolutions for a LEHR collimator with ^{99m}Tc and for a MEGP collimator with ^{177}Lu source. In the following, we only

considered the simple Gaussian and symmetric PSF model from [37] described by Equations (3.2) and (3.4).

LEHR collimator with ^{99m}Tc source

SPECT measurements with the Symbia Intevo Bold SPECT/CT scanner were performed with two detector heads with NaI crystal, a LEHR collimator, a 256×256 detection matrix with 2.3976×2.3976 mm^2 size and a photopeak energy window centered around 140.5 keV (+/- 15%). A ^{99m}Tc point-like source of 1.63 MBq was placed at the center of the imaging bed. The acquisition was performed in air. Several planar acquisitions were conducted with varying source-to-detector distances (5, 10, 15, 25, 30 cm). For each 2D projection, horizontal and vertical FWHM were calculated by fitting the summed profiles in each direction with a Gaussian function. The resulting standard deviation of the fitted Gaussian was used to compute the FWHM with Equation (3.5), as shown in Figure 3.7. Values of the FWHM at different distances of the point source were used to fit a linear function with the least square method, leading to the following result :

$$\mathrm{FWHM}_{\mathrm{experimental}}(d) = 0.043d + 3.527 \tag{3.8}$$

Monte Carlo (MC) simulations of the same experimental settings, source and SPECT system were also performed with the GATE Software [206]. Details of the Symbia Intevo Bold SPECT/CT scanner model validation can be found in [26]. Similar FWHM measurements lead to the following fitted linear function :

$$\mathrm{FWHM}_{\mathrm{GATE}}(d) = 0.041d + 2.993 \tag{3.9}$$

Using the so-called *Zeng* forward-projector implemented in the RTK Software [185], we also performed analytical simulations of a voxelized point source in air, placed at different distances from the detector's front. This forward-projection operator was applied with depth-dependent PSF modeling, whose parameters were set to the ones derived from our measurements (Equation (3.8)).

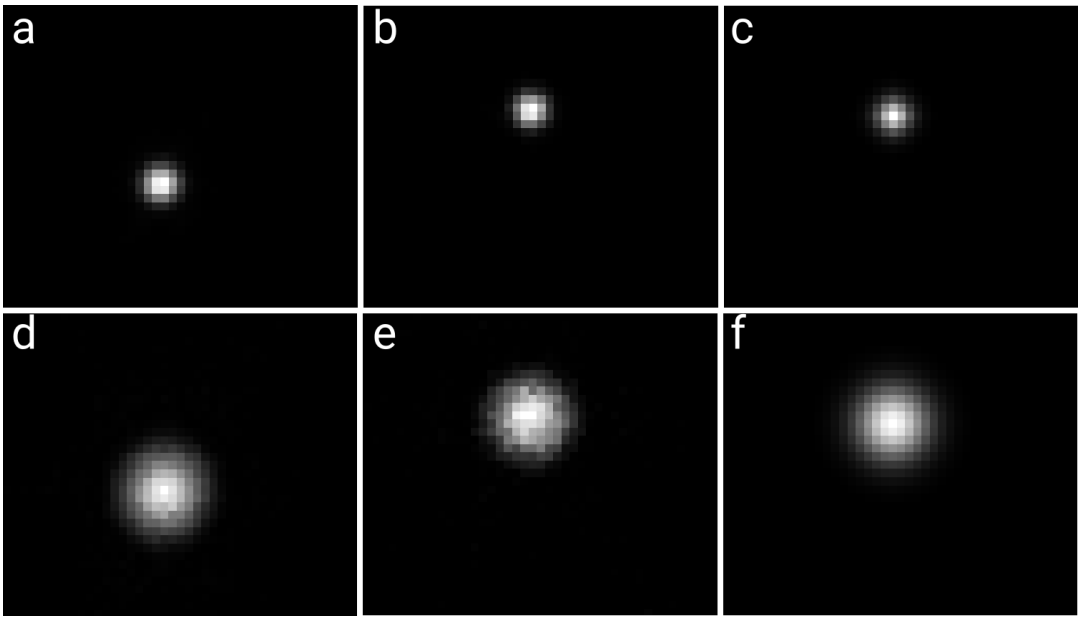
Finally, since the geometrical dimensions of the LEHR collimator are known (see Table 3.1), the analytical geometric response was derived from Equation (3.2). Moreover, the specifications from the manufacturer included the intrinsic resolution of the detector for a ^{99m}Tc source ($\mathrm{FWHM}_{\mathrm{intrinsic}} = 3.9$mm), leading to the following equation for $\mathrm{FWHM}_{\mathrm{Analytical}}$:

$$\mathrm{FWHM}_{\mathrm{Analytical}}(d) = \sqrt{3.9^2 + (0.048d + 1.11)^2} \tag{3.10}$$

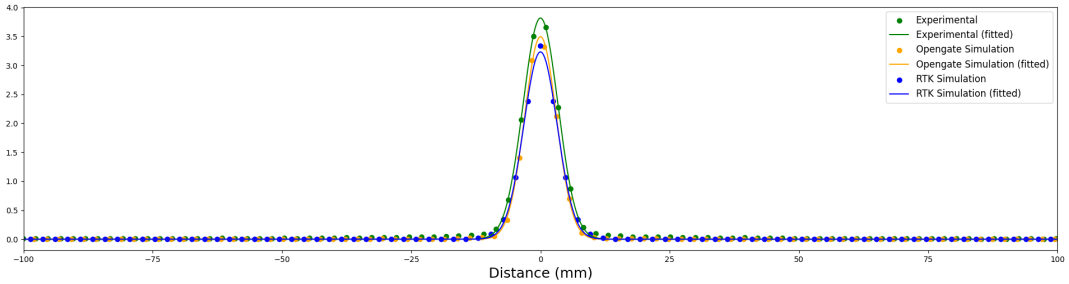
Projections obtained at 10 and 30 cm are shown in Figure 3.6a and the corresponding horizontal profiles in Figure 3.6b and 3.6c. Projections obtained at the same distance look very similar, showing good agreement between simulations (Monte Carlo and analytical) and experimental point source measurements. Moreover, all curves fit very well with the assumed Gaussian shape. A small mismatch can however be observed between the tails of the profile obtained experimentally and the other curves (including its gaussian fit). A more complex model taking more

specifically into account collimator scatter and septal penetration (like in [40]) would be required to model the tails of the PSF.

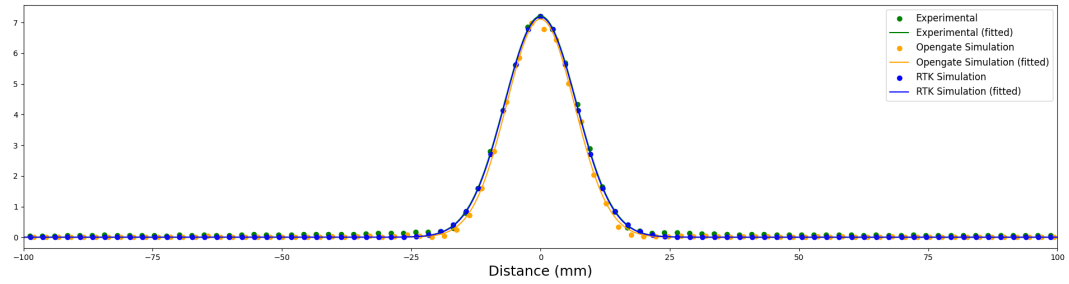
FWHM values as functions of the source-to-detector distance can be compared in Figure 3.7. All curve show good agreement with experimentally derived FWHM values, specially regarding their slope. A small systematic mismatch can be observed: both Monte Carlo and analytical FWHMs were 1.5mm lower than the measured resolution.



(A) Projections obtained with a point source placed at 10cm/30cm from the detector with experimental measures (a)/(d), a Monte Carlo simulation (b)/(e) and analytical forward-projection with RTK (c)/(f). All projections were scaled to have the same maximal value for visual comparison.



(B) Horizontal profiles of projections (a), (b) and (c) obtained at a 10cm distance and the corresponding fitted Gaussian functions. All profiles were centered to allow visual comparison.



(C) Horizontal profiles of projections (d), (e) and (f) obtained at a 30cm distance and the corresponding fitted Gaussian functions. All profiles were centered to allow visual comparison.

FIGURE 3.6 – Comparison of experimental, Monte Carlo simulations and analytically forward-projected ^{99m}Tc point source with LEHR collimator at distance 10cm and 30cm and corresponding profiles and fitted Gaussian functions.

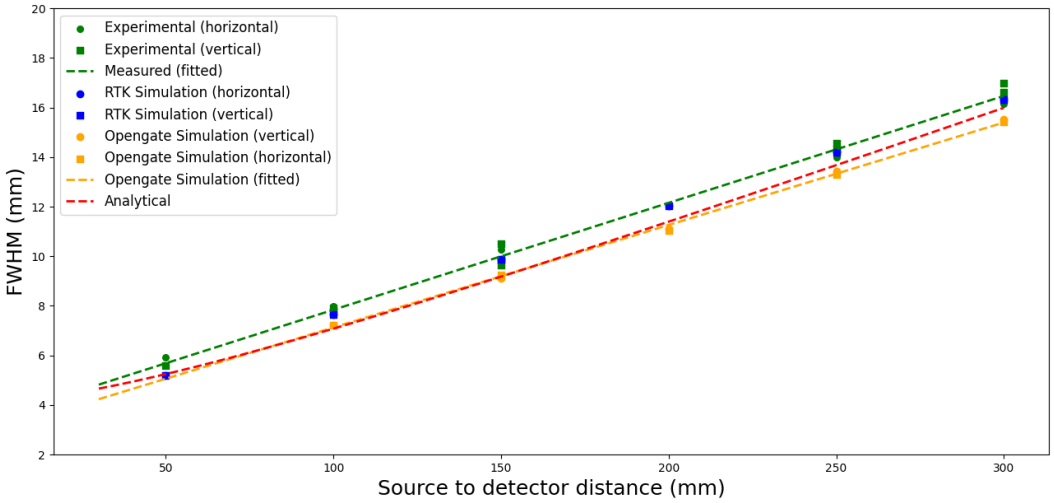


FIGURE 3.7 – Full width at half maximum (FWHM) versus source-to-detector distance obtained from experimental measures, Monte Carlo simulations and analytically forward-projected ^{99m}Tc point source with LEHR collimator.

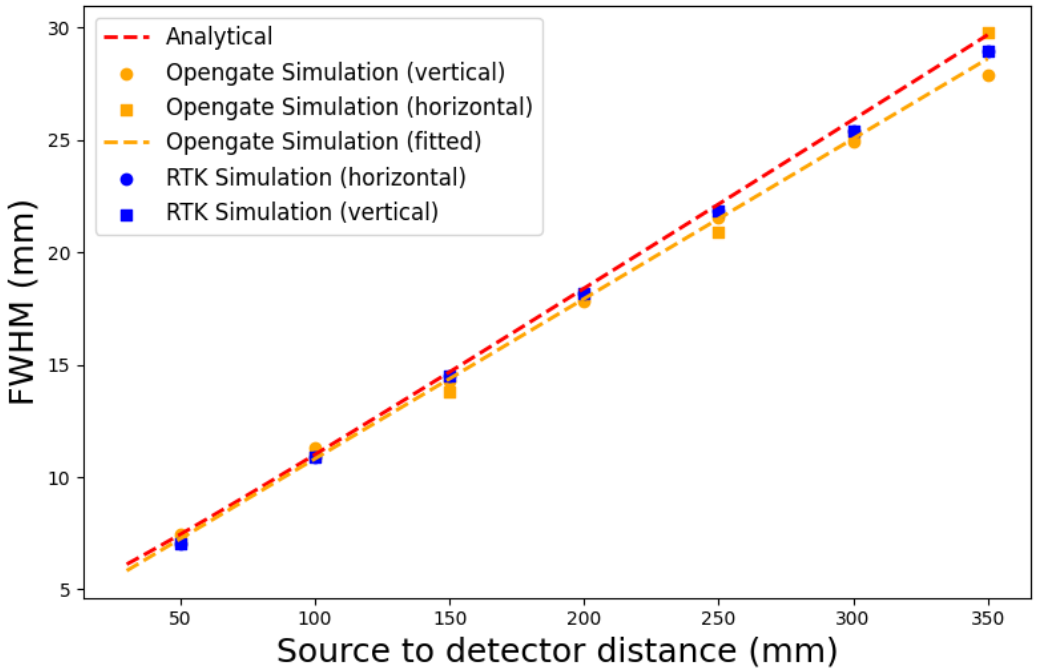


FIGURE 3.8 – Full width at half maximum (FWHM) versus source-to-detector distance obtained from Monte Carlo simulations and from analytically forward-projected ^{177}Lu point source with MEGP collimator. The analytical derivation of the FWHM from Equation (3.12).

MEGP collimator with ^{177}Lu source

The exact same GATE Monte Carlo simulation was performed for the MEGP collimator (see Table 3.1) and a point ^{177}Lu source. With these simulated data, the following coefficients were found for PSF modeling:

$$\mathrm{FWHM}_{\mathrm{GATE}}(d) = 0.071d + 3.704 \tag{3.11}$$

The parameters from Equation (3.11) were used to analytically forward-project a point source with RTK.

Finally, the intrinsic resolution for a 208keV source was derived from the intrinsic resolution given by the manufacturer for a 140keV source, since we know that it is proportional to $\frac{1}{\sqrt{E}}$. This lead to $\mathrm{FWHM}_{\mathrm{intrinsic}} = 3.2$ mm. Combined with analytical derivation from geometrical properties of the MEGP collimator (see Equation (3.2) and Table 3.1), we found the following analytical expression for the system resolution :

$$\mathrm{FWHM}_{\mathrm{Analytical}}(d) = \sqrt{3.2^2 + (0.076d + 2.94)^2} \tag{3.12}$$

Results for the measured FWHM in both cases can be found in Figure 3.8. Unfortunately, experimental data with ^{177}Lu point source were not available with ^{177}Lu source and MEGP collimator.

3.1.4 Conclusion

Partial Volume Effects are likely the most significant factor affecting SPECT image quantification, and their underlying causes are well understood. Here we reviewed the established equations allowing the characterization of the final FWHM. We computed and evaluated the parameters of the Gaussian models for the SPECT system at our institution, a Siemens Symbia Intevo Bold, for two radionuclides, using experimental data, Monte Carlo simulations (GATE), and analytical simulations (RTK). While the GATE model is still being refined in the team, it has been deemed sufficient for the purposes of this thesis. In the following, we are reviewing methods to mitigate PVEs.

3.2 State-of-the-art of PVC

A large number of Partial Volume Correction (PVC) techniques were proposed since the 70's, with the ultimate goal of making SPECT a truly quantitative imaging modality. Several reviews and comparisons of existing PVC methods were proposed [92, 230, 231, 86, 186, 56, 180, 228, 253, 3, 194]. The most complete and cited review on the subject was published by Erlandsson *et al* [56] in 2012. In this section, we propose to summarize the main PVC methods proposed in the literature.

In this Section, the problem will be formulated as follows. The measured sinogram p is used to reconstruct the 3D SPECT image y, whose relation to the true object activity x is described by:

$$y = \mathcal{N}(h \otimes x) \tag{3.13}$$

where h is the 3D-PSF of the system, $\otimes$ denotes the 3D convolution operator and $\mathcal{N}$ describes the noise process. Usually, the noise distribution is considered either Gaussian or Poisson. The PVC task can be assimilated to a deconvolution task, where the ultimate goal is to recover the true activity x from the measured one y. Usually, the PSF h is assumed to be known. Alternatively, the PVC task can be performed by taking the measured sinogram p as input and trying to estimate x during the reconstruction.

The difficulty to solve this inverse problem comes from its ill-posedness [28], meaning that it doesn't have a unique solution and/or the solution is not stable with respect to the measurement y. To solve this issue, all PVC methods have to include either a better model description or additional information.

3.2.1 Deconvolution methods

The first category of PVC includes all methods that just rely on the PSF knowledge to deconvolve the reconstructed image affected by PVEs. This deconvolution can be performed during or after tomographic reconstruction.

3.2.1.1 Resolution Modeling (RM)

Resolution Modeling (RM), which consists in including the depth-dependent PSF into the system matrix used for forward and back-projections in the MLEM

algorithm [235, 67, 263, 261, 94, 177], was already described in Section 1.2.4. RM was proven to be useful in improving the quantitative accuracy of SPECT images [177, 180, 85]. However, as mentioned earlier and illustrated in Figure 3.1, the main drawbacks of RM are its insufficiency to reduce PVEs and the noise amplification and Gibbs artifacts it gives rise to [56, 180, 161], especially as the number of iteration increases. For these reasons, it is generally not even considered a PVC method because of its insufficiency [56] to reduce PVEs.

One way to improve the method is to increase the realism of the PSF used during reconstruction. Indeed, the PSF model and resolution definition in Equations (3.2) and (3.4) rely on four hypothesis: 1) the PSF is assumed to be Gaussian 2) with a FWHM linearly increasing with source-to-detector distance 3) the resolution is the same in vertical and horizontal axis and 4) the detector response only depends on geometrical and physical properties of the detector. More realistic PSF models that do not rely on these hypotheses were used by [234, 12, 71, 70, 13], as well as object dependant optimized PSFs [21]. Moreover, with high energy isotopes such as ^{131}I, septal penetration and collimator scatter are no longer negligible and can lead to "star-shaped" tails in the detector response, which may need to be modeled in the PSF [40].

Another strategy is to use a sub-optimal PSF during RM or post-reconstruction smoothing [161] to reduce Gibbs artifacts at the cost of poorer resolution. Finally, during [104] or post-reconstruction [260] regularization can also help to reduce ringing artifacts.

3.2.1.2 Restoration or Inverse Filtering

From Equation (3.13), a very tempting operation is to neglect the noise process and apply a Fourier transform ($\mathcal{F}$) to both sides of the equation, and estimating x by applying the inverse Fourier transform ($\mathcal{F}^{-1}$), which leads to the following Equation :

$$\hat{x} = \mathcal{F}^{-1}\left(\frac{\mathcal{F}(y)}{\mathcal{F}(h)}\right) \tag{3.14}$$

However, this method to estimate x is never used in practice because $\mathcal{F}(h)$ rapidly vanishes in the frequency domain, meaning that even a very small amount of noise in the measured image y would be highly amplified in the estimated $\hat{x}$. The idea of restoration filtering is to apply a low-pass filter before the inverse Fourier transform is applied in Equation (3.14). The Wiener filter deconvolution follows this idea:

$$\hat{x} = \mathcal{F}^{-1}\left(\frac{|\mathcal{F}(h)|^2}{|\mathcal{F}(h)|^2 + \frac{1}{\mathrm{SNR}}}\frac{\mathcal{F}(y)}{\mathcal{F}(h)}\right) \tag{3.15}$$

where SNR is the Signal-to-Noise ratio, defined as the ratio between the object power spectrum and the noise power spectrum. The Metz filter was also proposed to restore signal [111]. Both method showed good results to improve image quality [111, 112, 154].

However, as pointed out in [236] and [143], the choice of the appropriate filter, as well as the associated parameters such as cut-off frequencies, have to be made

considering the type of task, noise level and object that is imaged, which is a time-consuming optimization task. This is because the optimal restoration filtering is always a trade-off between noise and resolution. In that regard, [114] concluded that even if restoration filtering is faster, RM leads to a better noise/resolution trade-off. Finally, exact knowledge of neither the PSF h (which is usually simplified by a spatially invariant Gaussian) nor the SNR (estimated from *a priori* knowledge) is achievable in practice.

3.2.1.3 Iterative Deconvolution methods

Well known iterative deconvolution methods, such as the Richardson-Lucy (RL) [184, 142] or the reblurred Van Cittert (RVC) [41, 16] algorithms were adapted and used for PET and SPECT image resolution enhancement. As an example, the RL algorithm, assuming a Poisson noise model, can be described as follows [16, 230, 228]:

$$\hat{x}_{k+1} = \hat{x}_k \left(h \otimes \frac{y}{h \otimes \hat{x}_k} \right) \tag{3.16}$$

where $\hat{x}_k$ is the corrected image after k iterations, initialized with $\hat{x}_0 = y$. Like RM, both RL and RVC suffer from noise amplification while the number of iteration increases. Thus, a regularization is needed, for example with Total Variation (TV) [230], again at the cost of poorer resolution.

3.2.2 Anatomical Methods

While PET and SPECT images only provide functional and metabolic information, their quantitative accuracy can be enhanced by incorporating anatomical information obtained from a co-registered CT or MRI. Since these anatomical images usually have a much better spatial resolution, ROIs can be segmented directly from them. Assuming the activity to be homogeneous in each anatomical region, regional masks can then be used to correct PVEs. In this section, all methods make use of segmented regional masks $(R_i)_{i=1\dots M}$, where M is the number of ROIs.

3.2.2.1 Resolution Recovery (RR) or Recovery Coefficient (RC) methods

The very first PVC method was proposed in 1979 by [87] and consists in using precomputed RC (see Equation (3.1)) for different object shapes and sizes. Then, $\hat{x}_{R_i}$, the corrected the mean activity in region R_i, is estimated by dividing the mean activity in region R_i by the corresponding RC_i:

$$\hat{x}_{R_i} = \frac{\text{Mean measured activity in region } R_i}{\mathrm{RC}_i} \tag{3.17}$$

Even tough this correction method is quite straightforward, the methodology to compute RC values is not. When the PSF h is known, RC_i can be computed analytically, considering that y_{R_i}, the mean measured activity in region R_i, can be written:

$$\begin{aligned} y_{R_i} = \frac{1}{\int_{R_i} dt} \int_{R_i} y(t)\,dt &= \frac{1}{\int_{R_i} dt} \int_{R_i} \int_{R_i} x(t)h(t,t')\,dt'dt \\ &\approx \frac{\hat{x}_{R_i}}{\int_{R_i} dt} \int_{R_i} \int_{R_i} h(t,t')\,dt'dt \end{aligned} \tag{3.18}$$

Then, $\mathrm{RC}_i := \frac{1}{\int_{R_i} dt} \int_{R_i} \int_{R_i} h(t,t')\,dt'dt$.

However, in practice, RC are most often determined experimentally with sphere phantoms [250, 50]. The RC curve is fitted as a function of object size or volume and used to correct PVEs in any other shaped object. However, the amount of PVEs depends on plenty of other parameters and so do RCs. As shown by [259] on simulated spheres, RC values depend on various acquisition and reconstruction parameters, such as TBR, sphere size and position, reconstruction voxel size and number of OSEM updates. In this study, higher RC values were obtained for bigger sphere diameters, smaller reconstruction voxel size, higher TBR, higher number of updates and higher off-center positions (specially for low number of updates).

Thus, in order to increase the realism of experimentally derived RC, an increasing number of studies estimate PVEs on organ-shaped phantoms [233, 65, 77, 194]. All these studies show that RCs obtained on spherical shapes cannot be applied to correct more complex organ-like shapes, such as kidneys.

As suggested by [4] and [158], and applied by [77] and [194], RC parametrization by the Surface Area-to-Volume ratio of the object to quantify results in a more robust RC-based correction when applied to objects with various shapes.

One last uncertainty related to RC-based correction is that there is no consensus on the method to employ to delineate the ROI used both to compute RC_i and the region R_i in Equation (3.17). For RC calculation on sphere inserts, some use CT-based contours of the inserts, some other reduce these contours by a small margin while some method apply a thresholding method on the SPECT image directly. In [244], authors showed that the thresholded-SPECT method was leading to significantly higher RC values than CT-based. ROI delineation is still an open question that is particularly an issue for tumour delineation since small lesions are often not visible on CT images. Lesion delineation is one of the weaknesses of all the anatomy-based methods presented in this section because PVEs affect both object volumes and image intensities, while ROI delineation is performed before PVC.

RC-based PVC is currently the most used PVC method in clinical routine and the most recommended one [50, 138]. However, it should be used with caution for objects with non-spherical shapes and complex shapes and RCs should be recomputed whenever an acquisition or reconstruction parameter is changed.

3.2.2.2 Post-reconstruction methods

Geometric Transfer Matrix (GTM)

In the RC-method, only spill-out is corrected in each region, without considering any spill-in coming from background or from an adjacent hot region. In the Geometric Transfer Matrix (GTM) method, all spill-over possibilities are considered in the correction. The method was proposed by [84, 191]. In this method, contribution

of other regions in region R_i are taken into account by adapting Equation (3.18) as follows:

$$\begin{array}{rcl} y_{R_i} & = & \sum_{j=1}^{M} \hat{x}_{R_j} \frac{1}{\int_{R_i} dt} \int_{R_i} \int_{R_j} h(t,t')\,dt\,dt' \\ \Longleftrightarrow \quad \mathbf{Y} & = & A\hat{\mathbf{X}} \\ \Longleftrightarrow \quad \hat{\mathbf{X}} & = & A^{-1}\mathbf{Y} \end{array} \tag{3.19}$$

where M is the number of segmented ROIs in the object, $\mathbf{X}$ is the M-dimensional vector of estimated true mean activity in region R_i defined by $\hat{\mathbf{X}}_i = \hat{x}_{R_i} \approx \frac{1}{\int_{R_i} dt} \int_{R_i} x(t)\,dt$. A is a $M \times M$ matrix and $\mathbf{Y}$ is a M-dimensional vector defined, for all $i, j \leq M$, by:

$$\begin{cases} A_{i,j} = \frac{1}{\int_{R_i} dt} \int_{R_i} \int_{R_j} h(t,t')\,dt\,dt' \\ \mathbf{Y}_i = y_{R_i} = \frac{1}{\int_{R_i} dt} \int_{R_i} y(t)\,dt \end{cases} \tag{3.20}$$

Coefficient $A_{i,j}$ of matrix A can be understood as the spill-over contribution of region R_j to region R_i, *i.e.* as the matrix of cross-talk factors [56, 194]. $\mathbf{Y}$ is the vector of uncorrected mean values in each region. As mentioned in [117], the matrix inversion in Equation (3.19) can be performed with Singular Value Decomposition of matrix A.

Labbé

The Labbé method [116, 117] is very similar to the GTM method, except that coefficients $A_{i,j}$ and $\mathbf{Y}_i$ are not computed over the exact regions R_i and R_j like in Equation (3.20) but on regions convolved with the PSF h:

$$\begin{cases} A_{i,j} = \frac{1}{\int_{R_i} dt} \int \left[\int_{R_i} h(t,t')\,dt' \right] \left[\int_{R_j} h(t,t'')\,dt'' \right] dt \\ \mathbf{Y}_i = y_{R_i} = \frac{1}{\int_{R_i} dt} \int \left[\int_{R_i} h(t,t')\,dt' \right] dt \end{cases} \tag{3.21}$$

As demonstrated by [194], this slight modification of the GTM method leads to a more robust correction.

However, the main drawback of both GTM and Labbé methods is that they are region-based method, assuming uniform activity in each region and do not output a voxelized corrected image but a piece-wise constant image, which is not very likely to correspond to the true activity distribution in the patient.

Multi-target Correction (MTC)

The Multi-target Correction (MTC) method [60] is an extension of the Müller-Gärtner method [159]. It is a voxel-based correction technique correcting the voxels of each region sequentially and independently. To be applied, the MTC requires a first estimation of mean corrected region values, stored in the vector $\mathbf{c} = (c_1, \ldots c_M)$.

Voxels in region R_i are corrected by considering other regions R_j as already corrected and constant (with value c_j). For each voxel t of the image:

$$\hat{x}(t) = \sum_{j=1}^{M} \mathbb{1}_{R_j}(t) \frac{y(t) - \sum_{i \neq j} c_i \int_{R_i} h(t,t')\,dt'}{\int_{R_j} h(t,t')\,dt'} \tag{3.22}$$

$\mathbb{1}_R$ is the indicator function of region R, *i.e.*:

$$\mathbb{1}_R(t) = \begin{cases} 1 & t \in R \\ 0 & t \notin R \end{cases} \tag{3.23}$$

The vector $(c_1, \ldots c_M)$ can be computed by taking the mean regional values in the uncorrected image y, but a pre-calculation using GTM or Labbé methods is often preferred.

A more recent version of this method was proposed by [57] and described in [198], without any need of a prior estimation of mean region values $\mathbf{c}$, but can be applied only to a single target VOI. This method is designed as the Single Target Correction (STC) method.

Region-Based Voxel-wise correction (RBV)

The Region-Based Voxel-wise (RBV) correction was proposed by [229] as an extention of the Yang [257] method. Once again, a first estimation of mean corrected region values (the vector $\mathbf{c} = (c_1, \ldots c_M)$) is required. These regional values are used to create a piece-wise constant image, with value c_i in region R_i: $s(t) = \sum_{i=1}^{M} c_i \mathbb{1}_{R_i}(t)$. The corrected image is then given by:

$$\hat{x}(t) = y(t) \frac{s(t)}{(s \otimes h)(t)} \tag{3.24}$$

Here again, the initial vector $\mathbf{c}$ is usually obtained from GTM or Labbé methods.

Iterative Yang (iYang)

The iterative Yang (iYang) algorithm is essentially an iterative version of the RBV method, where for all $k \geq 0$:

$$\begin{cases} s_k(t) = \displaystyle\sum_{i=1}^{M} \mathbb{1}_{R_i}(t) \frac{\displaystyle\int_{R_i} x_k(t')\,dt'}{\displaystyle\int_{R_i} dt'} \\ x_{k+1}(t) = y(t) \dfrac{s_k(t)}{(s_k \otimes h)(t)} \end{cases} \tag{3.25}$$

with $\hat{x}_0$ usually initialized with $\hat{x}_0(t) = y(t)$. At each iteration k, s_k is a piece-wise constant image where the value in each region R_i is the mean of the current estimate x_k computed over R_i. Usually, the algorithm converges after less than 10 iterations [228].

Discussion

GTM, Labbé, MTC, RBV and iYang are the most popular PVC methods both in PET and SPECT and were often compared to each other [229, 56, 15, 228, 194]. All these methods are implemented in the "PETPVC" open-source toolbox [228], in addition to RL and RVC deconvolution methods. Importantly, as described and implemented in [228], several PVC methods can be combined with each other by first choosing a region-based method (GTM or Labbé), then applying a voxel-based anatomical method (MTC, RVB or iYang) followed by a deconvolution method (RL or RVC) with the appropriate PSF. The MTC method can also be applied iteratively, as proposed by [134].

The main drawback of these methods is that they rely on a ROI segmentation. Besides the laboriousness of the segmentation task, this step represents a source of errors in the PVC. In case region masks are obtained from the CT, the correction is prone to registration mismatches [194]. Moreover, a new segmentation needs to be performed for each SPECT acquisition since the volume of ROIs can evolve during treatment [108]. In case a region mask is obtained from the uncorrected SPECT image, for lesions for example, there is no reference for the delineation method (see Section 3.2.2.1). Moreover, its contours cannot be delineated precisely, otherwise no PVC would be needed.

The assumption of uniformity within a ROI is also questionable. To address this issue, [56] proposed to replace the piece-wise constant image s_k of the iYang algorithm (see Equation (3.26)) by a fitted hyper-plan in each region, for each iteration. Thinner segmentations within an organ can also be needed when the uptake is expected to be different in different parts of the ROI, for example with renal cortex and medullary compartments [233].

Very often, the PSF used in post-reconstruction methods is assumed to be stationary for simplicity. This hypothesis is relatively true in PET but not in SPECT where the resolution depends on the object activity, attenuation and position [133, 259, 74, 127]. In every presented PVC method, a depth-dependent PSF h can be modeled. Instead of modeling the resolution loss process by a 3D convolution, a more realistic strategy, developed by [43], is to forward-project and reconstruct the ROI-based templates. This way, the non-linear characteristics of the tomographic reconstruction algorithm are taken into account in the correction.

3.2.2.3 Reconstruction-based methods

Projection-based

A correction method similar to iYang was also proposed to be applied iteratively in the projection domain either with FBP [58] or with OSEM [59]. For the OSEM algorithm, the projection-based correction reads as follows:

$$\begin{cases} s_k(t) = \sum_{i=1}^{M} \mathbb{1}_{R_i}(t) \dfrac{\int_{R_i} x_k(t')\,dt'}{\int_{R_i} dt'} \\ c_k = \dfrac{\mathfrak{F}(s_k)}{\mathfrak{F}_{\mathrm{RM}}(s_k)} \\ x_{k+1} = x_k \dfrac{1}{\mathfrak{B}(\mathbf{1})} \mathfrak{B}\left(\dfrac{c_k p}{\mathfrak{F}(x_k)}\right) \end{cases} \tag{3.26}$$

where p is the set of measured projections, $\mathfrak{F}$ and $\mathfrak{F}_{\mathrm{RM}}$ respectively denote forward-projection without and with Resolution Modeling (RM), while $\mathfrak{B}$ represents the back-projection operator without RM.

Perturbation-based

PVC methods presented in Section 3.2.2.2 generally assume a position-invariant PSF, while this assumption is generally not true in SPECT. Ideally, a PSF specific to the imaged object and reconstruction settings should be used in these PVC methods. The *perturbation-based* method was proposed by [225] to achieve this by adding a noise-free projection of a point source to the acquired projection data. The PSF at the location of the point source is then obtained by taking the difference between the image reconstructed from the perturbed sinogram and the one reconstructed from the acquired sinogram. This method was applied in [216, 20, 178] for more realistic template generation. This method can also be used to model the PSF in the GTM method [54] (p-GTM method) or in the STC method [74].

Regularization-based

As mentioned in Section 1.2.3.2, a regularization can be incorporated in iterative reconstruction algorithms such as MELM. When this regularization makes use of a high resolution anatomical image modality (CT or MRI) to constrain the emission image to be homogeneous inside anatomical regions, this regularization can be considered as a PVC technique. Many studies have explored the integration of anatomical information for regularization in PET and SPECT [62, 22, 39, 46, 14]. For example, the Bowsher prior [22] encourages similarity between neighbouring voxels if their CT intensities are close. The advantage of this method is that it does not need the anatomical image (MR or CT) to be segmented into VOIs, but the output image tends to over-smoothed.

In SPECT, an edge-preserving strategy was proposed in [49] by including a regularization term of the form:

$$\begin{array}{rcl} \mathcal{R}(x) & = & \beta \sum_i \sum_{j \in \mathcal{N}_i} w_{i,j} (x_i - x_j)^2 \\ w_{i,j} & = & 1 \text{ if } |l_i - l_j| \leq 0.1 \text{ otherwise } 0 \end{array} \tag{3.27}$$

where β is the strength of the regularization, $\mathcal{N}_i$ is the set of neighbour voxels of voxel i, $w_{i,j}$ is the weighting parameter, the image $(l_i)_i$ is the label image obtained by down-sampling the ROI-segmented CT image to the SPECT resolution. Compared to Bowsher's prior, this method requires the CT to be previously segmented into VOIs. With this approach, the estimated activity distribution in each VOI is smoothed while edges are preserved.

In [147], authors exploit the better resolution of a previously acquired ^{68}Ga-PET image to enhance the resolution of ^{177}Lu-SPECT images.

However, the main issue with regularization methods using an external modality (whether it is PET, MR or CT) to enhance SPECT images is that they assume some kind of similarity between images that inherently carry different information about the patient. For example, some lesions may be invisible on CT images while having

significant uptake in the SPECT image. The assumed uniformity inside a CT-derived VOI might also erroneously smooth the activity distribution inside this region.

3.2.3 Deep Learning methods

Artificial Intelligence, and Deep Learning in particular, has emerged during the last decade as a new way to solve very different tasks such as speech-to-text transcription, image inpainting, drug discovery, chess playing or face recognition. The main reason of the big interest given to deep learning algorithms comes from their complete new paradigm: while a classical algorithm is generally based on the knowledge of the task to solve and on pre-defined rules, AI algorithms are data-driven, *i.e.* are previously trained on a large dataset before being applied to new unseen data. If we take the example of chess playing, a classical algorithm like *minimax* calculates the best move over all possibilities until a certain depth. A data-driven approach would use knowledge of many other chess games to determine the one leading to the best winning probabilities.

More formally, designing a deep learning model trained in a supervised setting requires:

— To have a training database $\mathcal{D}_{\text{training}} = \{(x_i, y_i), i \leq N\}$ containing N pairs of inputs ($x_i \in \mathcal{X}$) and targets ($y_i \in \mathcal{Y}$). In the context of face detection in natural images, the training dataset should contain various images with faces and the location of each face in each image.

— A neural network architecture h_θ which is nothing more than a function $h_\theta : \mathcal{X} \mapsto \mathcal{Y}$ depending on its parameters θ sometimes called *weights*. The "deep" qualifier of deep learning methods, refers to the large number of parameters contained in θ, ranging from a few thousands to several billions.

— A loss function $\mathcal{L}$ defining a distance between elements of $\mathcal{Y}$.

Once these elements are defined, the training process consists in finding the optimal set of parameters $\widehat{\theta}$ such that $\mathcal{L}(y, h_{\widehat{\theta}}(x))$ is minimal for all $(x, y) \in \mathcal{D}_{\text{training}}$. Usually, this optimization problem is solved by initializing $\widehat{\theta}$ with random weights and by applying the following steps for all pairs $(x, y) \in \mathcal{D}_{\text{training}}$:

1. compute the loss $L := \mathcal{L}(y, h_{\widehat{\theta}}(x))$.
2. compute $\nabla_{\widehat{\theta}} L$, the gradient of the loss function with respect to the weights $\widehat{\theta}$ thanks to the chain rule (this is the *backpropagation*).
3. update $\widehat{\theta}$ with a gradient descend algorithm.

One of the most widely used algorithm to optimize and update the weights θ, is the Adam optimizer [113], which is a stochastic gradient descent algorithm with gradient momentum and adaptive learning rate (depending on the gradient magnitude). In practice, updating θ only once per data of the training dataset is never sufficient, this is why it is performed several times. One *epoch* corresponds to all training data having been through the network/loss/backpropagation. During each epoch, one *iteration* corresponds to one θ update. A θ update can be performed once for a *batch* of samples (x, y), the gradient loss being average over the batch. Weights of the network are then updated N/B times per epoch, B being the *batch size*.

The strength of the gradient descent update at every iteration is controlled by an hyper-parameter called the *learning rate*, generally smaller than 10^{-2} and decreased as the number of epochs grows.

During training, the ability of the network to generalize on data on which it was not directly optimized is tracked on a *validation dataset*. The resulting error on this validation dataset and its evolution while the number of epoch increases indicates if training converges or if the model is over-fitting (*i.e.* the features and noise of the training dataset were learned so closely that the network is not able to generalize to unseen data).

The challenge of training a deep learning model often arises from the high number of hyper-parameters to tune (model architecture, learning rate, batch size, normalization steps, training loss...). These hyper-parameters are typically chosen to minimize the validation error. However, even though the chosen network has never directly "learned" from this validation dataset, the hyper-parameters were chosen to optimize the performances of the network on this particular dataset. Therefore, it is highly recommended to use a third dataset for final evaluation: the *test dataset*.

A large choice of architectures is possible for h_θ, from Multi-Layer Perceptrons (MLP) proposed in 1958 [190] to more recent convolutional neural networks (CNN), including U-Nets [189], Res-Nets [83] or Transformers [163]. Figure 3.9 illustrates several popular deep learning architectures. The U-net architecture consists of several *encoding* layers made of successive convolutional layers in combination with pooling layers, non-linear activation function and normalization layers, which reduce the spatial information and increase the number of features. Similarly, encoding layers are followed by *decoding* layers, progressively retrieving the initial spatial dimensions while the number of features reduces. The U-net architecture is also characterized by the inclusion of *skip-connections* between encoding and decoding layers, allowing the network to propagate high resolution information though the entire network. U-nets are particularly suited for medical image analysis because of their ability to capture relevant features.

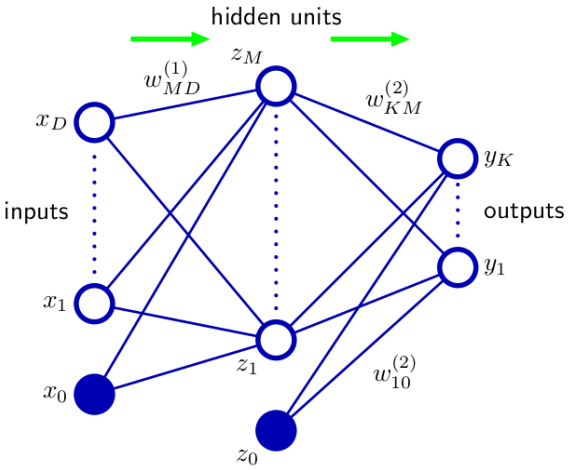
The choice of the appropriate loss function is also critical, since it is the metric that is optimized by the gradient descent algorithm. It has to be chosen carefully to match the task to be resolved. In medical imaging, popular loss functions include Mean Squared Error (pixel-wise L_2 norm between output and ground truth), Mean Absolute Error (pixel-wise L_1 norm), structural similarity (SSIM loss), perceptual loss (L_2 norm between output features of a pre-trained network such as VGG [103]) or task-specific loss (such as Dice for image segmentation). However, structural losses such as SSIM and perceptual loss are more suited for natural images than medical images. The L_2 norm has the advantage of strongly penalizing outliers but tends to over-smooth the output. The L_1 loss function promotes sparsity in the output, which enables the network to recover sharp edges and enhance image resolution.

3.2.3.1 Deep Learning in Nuclear Medicine

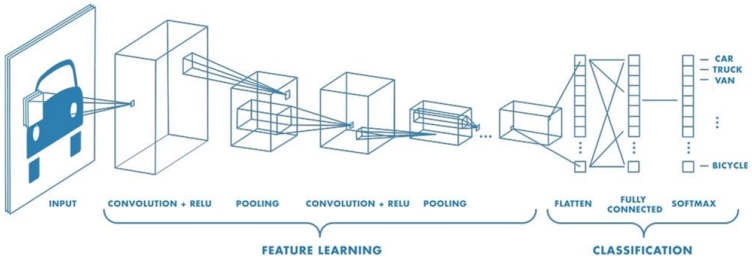
With the success of Deep Learning in many fields, medical imaging has naturally become a major application domain of such methods. Main applications of AI for medical imaging are image segmentation [96] and automatic disease (or lesion)

detection, whether it is on CT, MRI, PET or SPECT images. Many recent reviews have enumerated and discussed deep learning applications for PET and SPECT imaging [164, 246, 5, 135, 210, 167, 9, 101]. Main applications of deep learning for SPECT include:

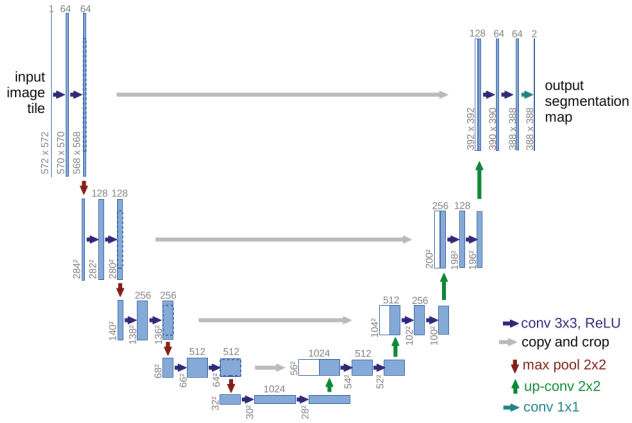
— image reconstruction [246, 38, 209, 211, 129, 132]
— denoising [183, 136, 168]
— scatter correction [254, 110]
— attenuation correction [214]
— intermediate view generation [193, 125]
— super-resolution [36]
— partial volume correction [255, 124]
— simulation [199, 202, 197]
— dosimetry [130]



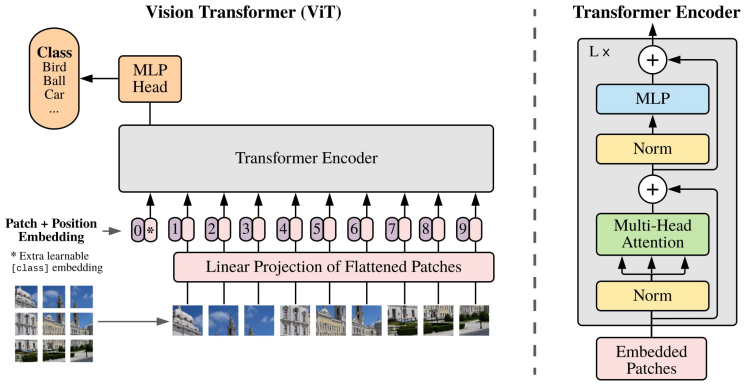
(A) Illustration of a Multi Layer Perceptron (MLP) made of Fully Connected Layers. Extracted from [17]



(B) Illustration of a Convolutional Neural Network (CNN) combined with a MLP for image classification. Extracted from [**<empty citation>**]



(C) Illustration of a 2D U-Net with four encoding and decoding blocks. Extracted from the original paper [189].



(D) Illustration of a Vision Transformer for image classification. Extracted from [53].

FIGURE 3.9 – Illustrations of four popular architectures: (A) Multilayer Perceptron, (B) Convolutional Neural Network, (C) U-Net (D) Vision Transformer.

3.2.3.2 Deep Learning for PVC in PET/SPECT

Very few works specifically focus of PVC for SPECT but several approaches that try to enhance SPECT or PET images resolution can be considered as PVC methods. In this section, we enumerate some works and approaches proposed in the literature to enhance SPECT or PET image resolution or to mitigate PVEs with deep learning.

Learned Direct reconstruction
The first type of approach is referred to as *learned direct reconstruction*. In [38], [209] and [211], authors trained a convolutional neural network to map raw SPECT sinograms to target activity distributions. Since the training dataset was obtained from simulations and that target activity distributions were the true activity maps thus unaffected by PVEs, this method can be considered as both as a reconstruction and PVC method. However, all of these methods were trained on 2D images.

Learned Regularizer
An other approach, investigated by [256, 129, 132] is to apply a trained regularizer to reduce RM-induced artefacts. For example in [132], authors opted for an unrolled-OSEM approach by training a neural network to denoise the reconstruction update at each iteration (using also the segmented CT information), whose output is used for regularization. This work showed good noise and error reduction property but depended on some hyper-parameters and on the quality of the CT-based segmentation used.

Super-Resolution
Super-resolution methods consist in inferring the over-sampled sinogram for the input raw sinogram in order to recover smaller object structures and more accurate activity. Deep learning based super-resolution was explored by [88, 66] to enhance image quality of PET images by converting low resolution sinograms (obtained with large crystal) into high resolution sinograms (with smaller pixel detector size). In [88], for better generalization properties, authors used a transfer learning approach: their "deep residual sinogram super-resolution network" was first trained on a large analytically simulated data, then on a smaller Monte Carlo simulated database, and finally on few pre-clinical data with ground truth obtained by down-sampling the acquired sinograms. The super-resolution framework was also explored by [36] for SPECT sinograms interpolation from 128×64 to 128×128 pixels.

Blind Deconvolution
Blind deconvolution methods were also explored, specially in PET. These methods try to revert the effect of the PSF, without *a priori* knowledge on the PSF parameters. This approach was followed by [128] in the case of intra-operative TOF-PET imaging. A first training stages consisted in training two networks: one for object-specific PSF estimation and the other one for TOF backprojection image correction. At inference time, a second training stage was performed only with a consistency loss ensuring that the estimated PSF convolved with the output of the image correction-network is close to the input image. Their method showed good ability to recover small spheres

but was only evaluated on simulated data. In [195], authors trained a deep learning model both for super-resolution and blind deconvolution for PSMA-PET.

Post-reconstruction correction

The most natural and popular approach remains post-reconstruction methods. For example [42] trained a 3D-CNN to reduce RM-induced artefacts and enhance resolution after reconstruction. In [232], authors used a 3D-CNN specifically trained to enhance tumour quantification in PET. Cropped PET images around tumours are used as inputs of the network. Results show good generalization properties to various PSF but were only evaluated on simulated data. In ^{177}Lu-SPECT, [126, 124] trained a 3D U-net to correct images reconstructed with RM. These methods have in common that even tough they used real patients anatomy (MRI or CT images) in their training dataset, their ground truth activity maps were artificially generated with random TBRs, lesion shapes and locations. These activity maps were given as input of an analytical simulation software that imitated the image (and PVEs) formation process. The main benefit from these deep learning methods trained on synthetic data is that superior performance compared to other state-of-the art methods can be achieved, without the need of a segmentation mask at inference time.

On the other hand, [149] for brain PET and [255] for cardiac SPECT trained their post-reconstruction PVC networks using real patient data. In these cases, the only available ground truth for training and evaluation was to apply another state-of-the-art PVC —RBV for [149], iYang for [255]. The main advantage of these approaches is that *domain shift* (*i.e.* differences in training and evaluation dataset distributions) is minimized. Moreover, no segmentation was needed during inference, which is a significant advantage compared to anatomy based methods used as references. However, real patient training datasets are often limited in size, which increases the risk of over-fitting during training. Additionally, the quality of the ground truth is questionable due to limitations discussed in Section 3.2.2.2.

In [27, 23], authors discuss the best practices and pitfalls to avoid while developing a deep learning algorithm for Nuclear Medicine. Key points include:

— Generalizability: Domain shift, biased training datasets and over-fitting must be avoided in order to ensure that the trained model can be applied to unseen, new data. This is a challenge both for models trained on large synthetic datasets and for those trained on limited real-world datasets.

— Reproducibility and transparency: Every step from data collection to training details should be explicitly described, including data normalization, model architecture and learning rate.

— Explainability/interpretability: deep learning models are often viewed as "black boxes". However, in the medical context, it is essential to have at least a basic understanding of why the model produces certain outputs. Interpretability includes identifying which parts of the input and of the model have the greatest impact on accuracy and robustness.

3.3 Conclusion

To conclude, PVEs are still an unresolved issue which is important to address, specifically in the context of radionuclide therapies in which the importance of quantitative accuracy of the image is crucial.

Several PVCs have been proposed during the last decades, but no consensus has been reached on the standard correction. While deconvolution-based methods are often insufficient or lead to artifact and noise amplification, anatomy-based methods rely on a region delineation task, impracticable in clinical routines. Current investigations on the use of AI to reduce PVEs are still in progress through very different approaches. In the next Chapter, we will detail and evaluate our proposed deep-learning-based PVC method, which operates in the projection-domain, before reconstruction.

Chapter 4

Contribution: PVC on SPECT sinograms with supervised Deep Learning, trained with synthetic data

Sommaire

As discussed in Chapter 3, many Partial Volume Correction (PVC) methods have been proposed, but none have reached enough consensus to be widely applied in clinical routines. While deconvolution based methods (Section 3.2.1) are often insufficient, enhance noise and artifacts and rely on several hyper-parameters to be tuned, anatomy-based methods (Section 3.2.2) rely on a tedious or unavailable

(tumours) segmentation step and an often simplified image degradation model. On the other hand, data-driven methods such as Deep Learning models for PVC in Emission Tomography (Section 3.2.3), especially for PET, are becoming increasingly popular for their ability to learn to correct the PSF. However, from the various reviews of the literature on the subject [164, 246, 5, 135, 210, 167, 9, 101], it is clear that PVC for SPECT is still an under-investigated application of deep learning methods. Additionally, all the proposed methods are post-reconstruction correction methods [255, 124] while the PVEs formation process essentially appears because of collimator-detector response (Section 3.1.2).

The objective of this Chapter is to describe, evaluate and discuss our proposed Deep Learning-based PVC, named *PVCNet-sino*. The originality of the method is that it operates in the sinogram domain before reconstruction. To counteract the lack of ground truth data to train our deep learning method, a large analytically generated dataset was used for training. The data generation process is described and discussed in Section 4.2. More details on the training framework are outlined in Section 4.3. To ensure our method generalizes well on unseen, more realistic and complex data, PVCNet-sino is evaluated on a test dataset, a phantom acquisition, Monte Carlo Simulations and real patient data. Evaluation dataset descriptions are provided in Section 4.4 followed by corresponding results and discussion in Section 4.6 and Section 4.7. Finally, ongoing and future directions for PVCNet improvement and validation are discussed in Section 4.8. Part of this work was presented at the *17th International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine* [107].

4.1 Motivations

Partial Volume Effects (PVEs) in SPECT are due to object and depth-dependent collimator-detector response, causing blurring in the sinograms that need to be compensated before, during, or after reconstruction. Based on the identified causes of PVEs and the advantages and drawbacks of various PVC methods, the requirements of any new PVC method should be the following:

1. No segmentation required at inference time: As discussed in Section 3.2.2, PVC methods relying on a region segmentation mask are difficult to apply in practice because of the laboriousness (and uncertainty) of the segmentation task and the risk of SPECT/CT registration errors. Even though AI-based automatic segmentation —like the software TotalSegmentator [248]— is becoming more and more efficient at delineating organs on CT images, lesion segmentation on SPECT images remains a much more difficult and unresolved task. Ideally, the segmentation task should arise *after* image correction, not before.
2. The quantitative accuracy both on organs and lesions should be enhanced, without introducing excessive noise or artifacts.
3. The method should take into account the non-linear, object- and position-dependent nature of PVEs and generalize well to unseen new and complex data.

Taking into account these requirements, our proposed method consists of generating a large simulated dataset for training that contains realistic inputs and

high-quality ground truths. The main benefit of using a simulated dataset for training resides in the ability to create a much larger dataset than what is commonly achievable with real-world data, both in terms of dataset size and bias and quality of ground truths. However, when using a simulated dataset for training, it is crucial to ensure the realism and unbiasedness of the dataset. Otherwise, the model may learn patterns that do not generalize well to real-world data, leading to reduced accuracy and reliability.

The other particularity of our proposed method is that we propose to correct for the collimator-detector PSF in projections directly, before reconstruction. This choice is motivated by the fact that, as described in Section 3.1.2, PVEs are intrinsically related to the collimator response, operating primarily in the projection domain. Correcting PVEs on sinograms directly allows to conserve the maximal amount of information, while post-reconstruction PVC methods use already truncated data. Moreover, it allows to be (mostly) independent of the chosen reconstruction algorithm, provided it does not include additional PVC method. Our proposed sinogram-based correction will be compared with an image-domain training strategy in the next Chapter.

The overall dataset generation training and inference frameworks that we proposed are described in Figure 4.1. Patient CT were used to create random activity sources, projected into different sinograms. There sinograms were then used in the supervised training stage to train the PVCNet-sino deep learning model. At test and inference time, PVCNet-sino was applied to new raw projections, acquired with the same SPECT/CT system as the one modeled in the training dataset, but with different patient CTs and activity distributions.

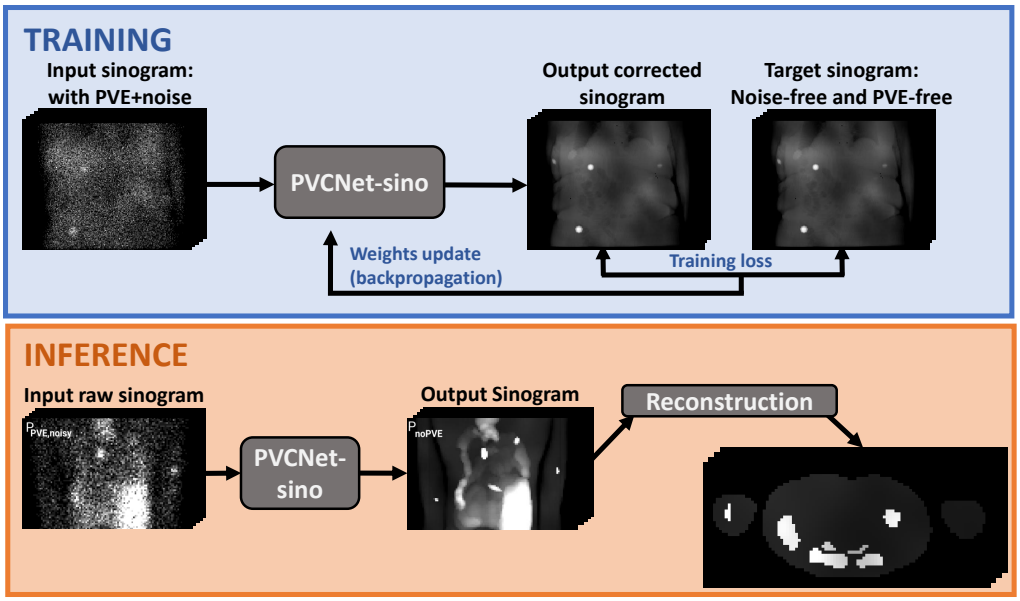


FIGURE 4.1 – Pipeline of the proposed method. During training 3D input noisy sinograms were used as inputs of PVCNet-sino. Weights of the network were updated by computing the training loss between output and target partial volume-free sinograms.

4.2 Training Dataset

4.2.1 Patient and activity distributions

Low dose CT scans from 19 patients and one phantom were obtained from clinical routine at the Leon Bérard cancer center (Lyon, France). CT scans were acquired with a voltage comprised between 110kVp and 140kVp and a X-ray tube current between 15mAs and 160mAs, depending on the scanned region. Half of the patient's scans were centered around their chest, while the other half was centered around their pelvis. CT images contained $512 \times 512 \times N_z$ voxels of size $0.9765\text{mm} \times 0.9765\text{mm} \times v_z\text{mm}$, where N_z varied between 193 and 401 and v_z between 1mm and 2mm.

Each of these CT images was segmented —using the automatic segmentation software TotalSegmentator [248]— into the following regions (if present in the CT): body contour, liver, kidneys, spleen, gallbladder, stomach, pancreas, small bowel, colon, duodenum and urinary bladder. The choice to keep these regions was made according to the main ROIs usually considered in ^{177}Lu-PSMA therapy in which physiological uptake is expected.

Then, 3D random activity maps were created with the following pipeline:

1. For each new activity map to create, a random CT scan was chosen.
2. Activity in the whole body was set to 1.
3. Each of the previously enumerated organs was selected with a probability of 0.5. If selected, a random TBR for this organ was randomly and uniformly chosen in the ranges reported in Table 4.1 derived from the literature [48, 251].
4. A random number of lesions is chosen between 0 and 10. Each of these lesions was an ellipsoid placed at a random position chosen inside the body and with a random radius in each direction comprised between 4 mm and 32 mm, to which a random rotation was applied. TBR assigned to each lesion randomly varied between 8 and 120.
5. Each of the previously generated region mask (organs or lesions) was multiplied by a different gradient function F given by Equation (4.1) to mimic heterogeneity.
6. The resulting random 3D activity map was scaled such that the total activity was comprised between 200 MBq and 1 GBq and the acquisition time was 15 seconds per projection.

The heterogeneity within each region (step 5) is obtained by sampling $u_{i,j,k}$ from a standard normal distribution and $\phi_{i,j,k}$ uniformly between $-\pi$ and π and computing $F(x,y,z)$ given by:

$$F(x,y,z) = 1 + \sum_{\substack{i,j,k=-M \\ (i,j,k)\neq(0,0,0)}}^{M} \frac{u_{i,j,k}}{i^2+j^2+k^2} \cos\left(2\pi(ix+jy+kz)/p + \phi_{i,j,k}\right) \tag{4.1}$$

where M is the maximal frequency and p the period of the series. M was chosen equal to 5 in the background (body) and 8 in organs and lesions. The period p was equal to 500 mm to cover the entire patient in each dimension. This heterogeneity

ROI	min TBR	max TBR
body	1	1
kidneys	20	60
spleen	20	60
liver	6	20
urinary bladder	6	20
gallbladder	2	8
stomach	2	8
pancreas	2	8
small bowel	2	8
colon	2	8
duodenum	2	8
lesions	8	120

TABLE 4.1 – Range of Target to Background Ratios (TBR) for several regions of interest used in the generated training database.

is illustrated in Figure 4.2. Negative voxels after multiplication by F were assigned the value 0 and each region was normalized such that their mean value was still the TBR chosen in step 3.

14 CTs (including the phantom) were used to generate 10,000 activity distributions for the training dataset, 2 patient CTs for the validation dataset (500 activity distributions) and 4 other patient CTs for the test dataset (1,000 activity distributions). Resulting activity maps — denoted I_{src} in the following— were at the same resolution (voxel size) as their corresponding CT image in order to avoid committing an inverse crime [157, 215]. Four examples of CT images superimposed with random synthetic activity distributions are shown in Figure 4.3.

4.2.2 Sinograms generation

Emission and detection were analytically simulated by forward-projection operators implemented in the reconstruction software RTK [185]. More specifically, we used the "Zeng" ray-tracing method which consists in rotating and resampling the volume before summing to obtain each projection [262, 187]. This projector was chosen for its ability to model attenuation and depth-dependent PSF blurring, as described in [261] and Equation (1.34).

CT images were converted into attenuation maps —denoted I_{att}— by converting HU values to linear attenuation coefficients for photons of 208keV, considering only the highest photon emission peak of ^{177}Lu (see Equation (1.28)).

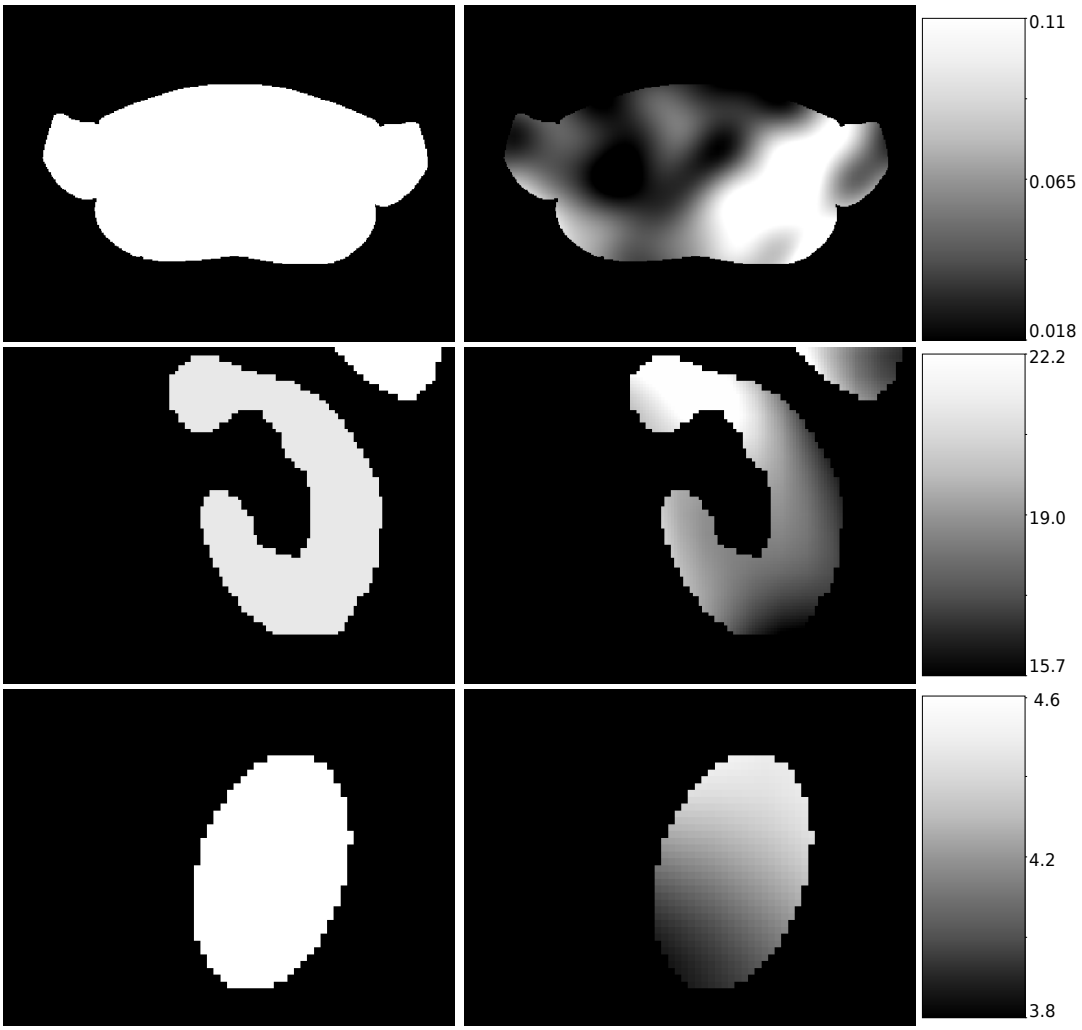


FIGURE 4.2 – Illustration of the heterogeneity on background (top), a kidney (middle) and on a synthetic lesion (bottom). Units are arbitrary at this stage.

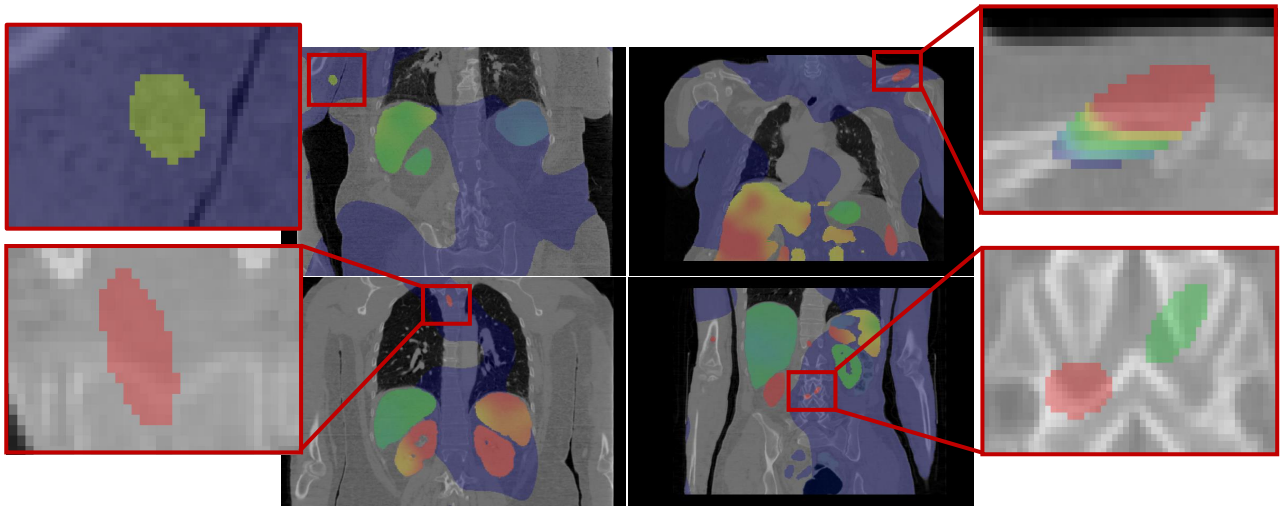


FIGURE 4.3 – Examples of initial real-life CT images with random synthetic activity distributions.

For each source/attenuation pair, several sinograms were obtained by forward-projection, all of them obtained with a circular orbit with a radius of 280 mm, 120 projection angles, 128×128 pixels of size 4.7952 mm $\times$ 4.7952 mm:

— The target P_{noPVE} sinogram was obtained by forward-projecting the activity distribution, without any PSF modeling during the forward-projection:

$$P_{\text{noPVE}} = A\, I_{\text{src}} \tag{4.2}$$

where A is the system matrix modeling the detector orbit and attenuation in the patient.

— PVEs were modeled in the P_{PVE} sinogram, obtained by forward-projecting the source with depth-dependent PSF modeling as described in [261]. We used the PSF parameters of the Siemens Intevo Bold SPECT/CT derived by MC simulations in Section 3.1.3 (Equation 3.11) for the MEGP collimator with ^{177}Lu source (see Table 3.1 for collimator specifications).

$$P_{\text{PVE}} = A_{\text{PSF}}\, I_{\text{src}} \tag{4.3}$$

where A_{PSF} is the system matrix modeling the detector orbit, attenuation in the patient and convolution by the depth-dependent PSF.

— In order to simulate the detection process, Poisson noise was applied to P_{PVE} to get the realistic noisy sinogram $P_{\text{PVE,noisy}}$:

$$P_{\text{PVE,noisy}} \sim \text{Poisson}\left(P_{\text{PVE}}\right) \tag{4.4}$$

— We used $P_{\text{PVE,noisy}}$ to reconstruct (again with RTK) a first estimation of the image, I_{rec10}, with OSEM and RM (10 iteration, 8 subsets, 4.7952 mm isotropic voxel size). Finally, this volume was forward-projected (without PSF modeling) to obtain $P_{\text{rec10,fp}}$, which was used as an additional input. The idea behind the use of $P_{\text{rec10,fp}}$ is that, as argued by [82], pre-processing on SPECT projections for denoising can lead to information loss, specially in low-count settings and for small or cold sources. The noisy and blurry raw sinogram $P_{\text{PVE,noisy}}$ may not be sufficient to recover object shapes and activity, as shown later in the ablation study. On the other hand, $P_{\text{rec10,fp}}$ is a denoised version of the raw sinogram and contains 3D tomographic information as well as PSF knowledge from RM during reconstruction.

— The attenuation map was also forward-projected to get P_{attmap} (without PSF modeling in the projection).

In total, for each of the 10,000 samples of the training dataset, 5 corresponding sinograms were stored: P_{noPVE}, P_{PVE}, $P_{\text{PVE,noisy}}$, $P_{\text{rec10,fp}}$ and P_{attmap}. As a reminder, 3 images were also stored: I_{src}, I_{att} and I_{rec10}. Sinograms of a training data example are shown in Figure 4.4.

The total disk size of the 10,000 samples of the database reached 250 GB.

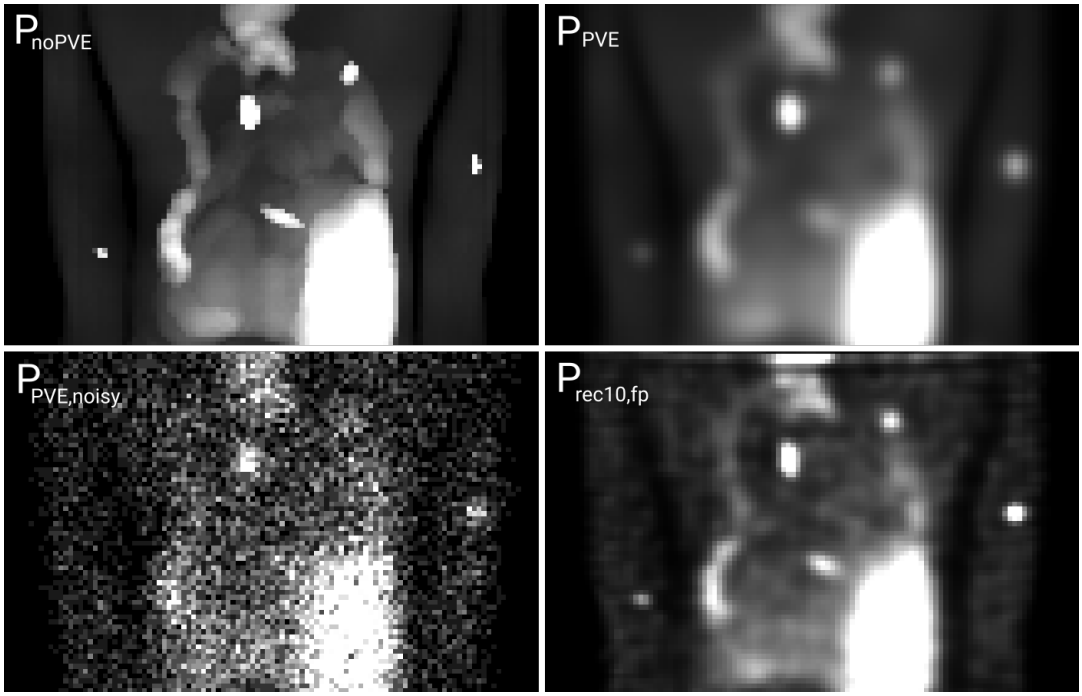


FIGURE 4.4 – Projections generated from the CT and synthetic activity: 1) P_{noPVE} without Partial Volume Effect (ground truth), 2) P_{PVE} with PVEs 3) $P_{\text{PVE,noisy}}$ with PVEs and noise, the input data, 4) $P_{\text{rec10,fp}}$ the forward projection of the reconstructed image.

4.3 Training

4.3.1 Optimization framework

The proposed neural network architecture uses sinograms as input to correct them from Partial Volume Effects.

First, all sinograms were padded circularly in the projection number dimension, expanding from (120,128,128) to (128,128,128) pixels. This was done by replicating the last four and first four projections at the beginning and end of the sinogram. The rationale behind this padding is that although the first and last projection numbers are spatially distant in the sinogram, they carry close information from the object being imaged.

Since the noise-free sinogram was available for each sample, Poisson noise was applied to P_{PVE} with Equation (4.4) at each iteration independently during training, reducing the risk of over-fitting.

All input sinograms P were normalized by their maximal value (Equation 4.5) and output sinograms $\widehat{P}$ were scaled so that the total number of counts in each projection i was the same as in the input raw sinogram (Equation 4.6).

$$P^{\text{normalized}} = \frac{P}{\max(P)} \tag{4.5}$$

$$\begin{aligned}
&\forall \text{ projections } 1 \leq i \leq 128, \\
&1 \leq j,k \leq 128, \\
&\widehat{P}^{\text{de-normalized}}[i,j,k] = \widehat{P}[i,j,k] \frac{\sum_{j',k'} P_{\text{PVE,noisy}}[i,j',k']}{\sum_{j',k'} \widehat{P}[i,j',k']}
\end{aligned} \tag{4.6}$$

We developed and trained two neural networks $h_{\text{denoising}}$ and $h_{\text{deblurring}}$, both having a 3-channels input of size (3,128,128,128) and a one-channel sinogram output. $h_{\text{denoising}}$ and $h_{\text{deblurring}}$ were applied successively, as described in Equation (4.7). Since the output sinogram is not expected to be very far from the input $P_{\text{rec10,fp}}$ sinogram, residual connections were added to learn only the difference with this sinogram. A ReLu activation function was finally applied to force positivity of the corrected sinogram.

The training loss $\mathcal{L}_{\text{sino}}$ is defined in Equation (4.8) such that $h_{\text{denoising}}$ was trained with P_{PVE} as target and $h_{\text{deblurring}}$ with P_{noPVE} as target, both with a L_1 loss function. We denote $\widehat{P_{\text{PVE}}}$ (resp. $\widehat{P_{\text{noPVE}}}$) the output of $h_{\text{denoising}}$ (resp. $h_{\text{deblurring}}$). The architectures of both networks will be described in the next section.

$$\begin{aligned}
\widehat{P_{\text{PVE}}} &= \text{ReLu}\left(h_{\text{denoising}}(P^{\text{normalized}}_{\text{PVE,noisy}}, P^{\text{normalized}}_{\text{attmap}}, P^{\text{normalized}}_{\text{rec10,fp}}) + P^{\text{normalized}}_{\text{rec10,fp}}\right) \\
\widehat{P_{\text{noPVE}}} &= \text{ReLu}\left(h_{\text{deblurring}}(\widehat{P_{\text{PVE}}}, P^{\text{normalized}}_{\text{attmap}}, P^{\text{normalized}}_{\text{rec10,fp}}) + P^{\text{normalized}}_{\text{rec10,fp}}\right)
\end{aligned} \tag{4.7}$$

$$\mathcal{L}_{\text{sino}} = ||P_{\text{PVE}} - \widehat{P_{\text{PVE}}}^{\text{denormalized}}||_1 + ||P_{\text{noPVE}} - \widehat{P_{\text{noPVE}}}^{\text{denormalized}}||_1 \tag{4.8}$$

Figure 4.5 illustrates the training framework. The complete application of normalization followed by $h_{\text{denoising}}$, $h_{\text{deblurring}}$ and de-normalization is denoted $h_{\text{PVCNet-sino}}$ for simplicity.

4.3.2 Networks

The architecture for both $h_{\text{denoising}}$ and $h_{\text{deblurring}}$ was chosen to be the 3D U-net architecture. This type of networks has been very popular during the last decade [189], specially in medical imaging. It is appreciated for its performances, simplicity to use and implement and interpretability. A U-net consists in two distinct parts: an *encoder* and a *decoder*. The *encoder* part consists in several convolutional blocks, each one halving spatial size of the input, for example with Max Pooling operations, and doubling the number of feature channels. The goal of the encoder is to transform relevant spatial information into abstract feature information. Once the lowest spatial size is reached (the *bottleneck*), the *decoder*, which is also made of convolutions, up-samples spatially the output of each previous block and halves the number of feature channels to retrieve the initial spatial size and the desired number of output channels. The particularity of the U-net is that it also contains "skip connections" which consists

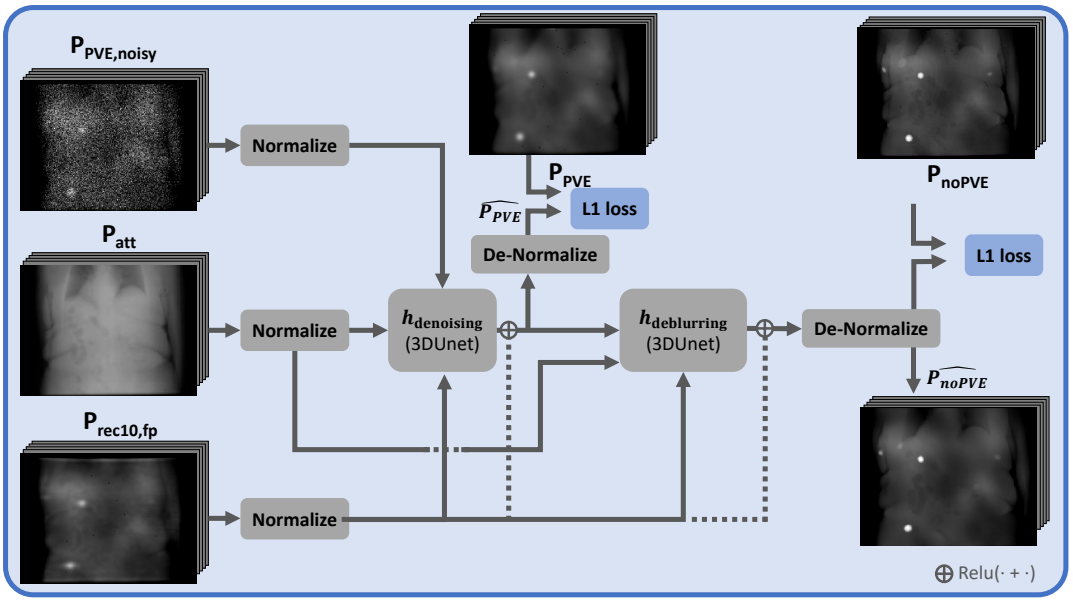


FIGURE 4.5 – Illustration of the $h_{\text{PVCNet-sino}}$ pipeline. $P_{\text{PVE,noisy}}$, P_{attmap} and $P_{\text{rec10,fp}}$ are concatenated in the channel dimension and given as inputs to $h_{\text{denoising}}$ after normalization. The output is concatenated with P_{attmap} and $P_{\text{rec10,fp}}$ and given as input to $h_{\text{deblurring}}$. Residual connections are drawn in dashed lines.

in copying the output of each encoder block to the corresponding decoder block. The goal of skip connections is to conserve relevant spatial information that may be lost by successive pooling operations. An illustration of the U-net architecture we used is shown in Figure 4.6. N_{ch} refers to the number of channels after the first convolution. All convolutions were done with 3x3x3 kernels. We used *dropout* [224] during training to limit over-fitting.

U-net was not the only architecture we considered. In the ablation study presented later (Sections 4.4.1.1 and 4.6.1.1), we replaced U-net architectures by CNNs with 8 layers. The CNN architecture is illustrated in Figure 4.7.

4.3.3 Training details

Training was performed using the PyTorch library, in its 2.2.0 version. 200 epochs were performed on 4 GPUs thanks to HPC resources of GENCI–IDRIS ("Jean Zay" CNRS cluster). We used the Adam optimizer [141], an initial learning rate of 5×10^{-4} halved every 50 epochs. Due to input size, we were only able to consider a batch size of 1. We used the *Automatic Mixed Precision* (AMP) mode for *forward* calls and losses evaluations, but not for backpropagation. The AMP allows to automatically select half (float16) or single precision (float32) for simple operations, such as convolutions, and thus reduces the computation time.

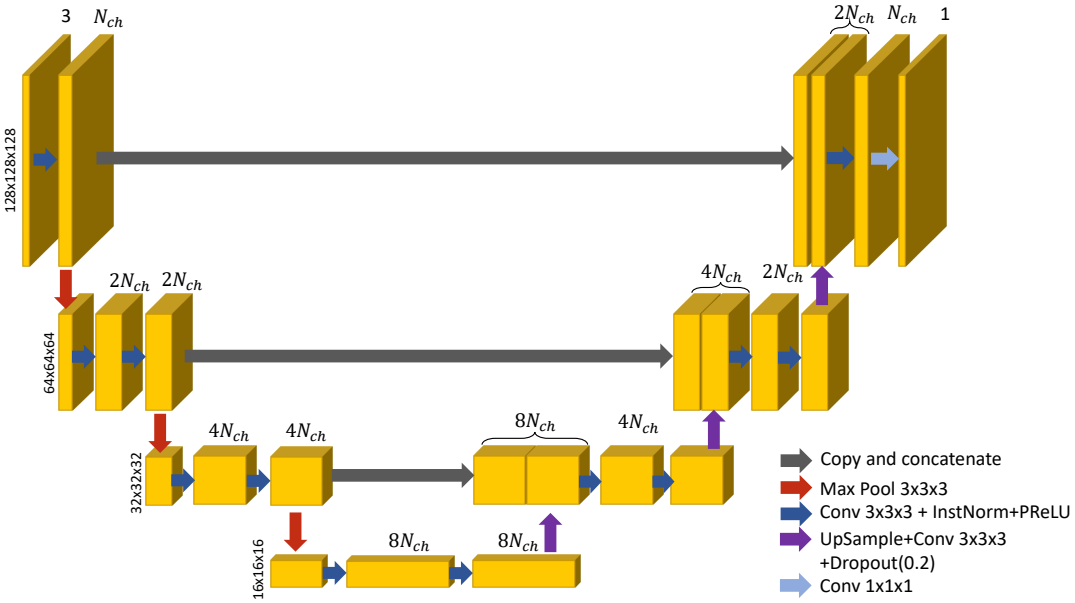


FIGURE 4.6 – Illustration of the 3D U-net we used in PVCNet-sino. The number of channels in the input is increased to N_{ch} after the first layer. It contains three encoding blocks between the input layer and the bottleneck, each doubling the number of channels and halving the number of pixels in spatial dimensions. Each part of the decoder halves the number of channels and double the spatial size of images. Skip connections are added. This architecture can be adapted to have more than three blocks in each encoder/decoder part.

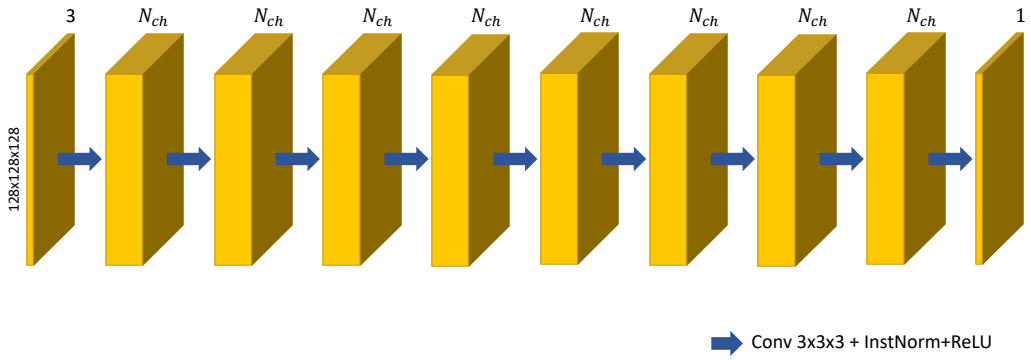


FIGURE 4.7 – Illustration of a 3D CNN used in the ablation study. The number of channels is increased to N_{ch} after the first layer. It contains 9 blocks made by successive 3d convolutions with 3x3x3 kernels, InstNorm and ReLu activation function.

4.4 Evaluation Datasets

4.4.1 Analytical Test dataset

As mentioned in Section 4.2, a test dataset containing 1,000 samples was generated the same way as the training dataset but with 4 other patient CTs not used during training.

4.4.1.1 Ablation Study

This dataset was used to perform an ablation study to assess the usefulness of several parts of the $h_{\text{PVCNet-sino}}$ pipeline. A single U-Net approach, denoted $h_{\text{denoising+deblurring}}$ was directly trained such that $\widehat{P_{\text{noPVE}}} = h_{\text{denoising+deblurring}}(P^{\text{normalized}}_{\text{PVE,noisy}}, P^{\text{normalized}}_{\text{attmap}}, P^{\text{normalized}}_{\text{rec10,fp}})$ and $\mathcal{L}_{\text{sino}} = ||P_{\text{noPVE}} - \widehat{P_{\text{noPVE}}}^{\text{denormalized}}||_1$. In this setting, we trained $h_{\text{denoising+deblurring}}$ with 3, 4 and 5 blocks U-Net architectures.

In the two networks approache described in Equation (4.7) and Figure 4.5, both $h_{\text{denoising}}$ and $h_{\text{deblurring}}$ were also trained with 3, 4 and 5 blocks in the 3D-U-Nets architectures (Figure 4.6). We also trained $h_{\text{PVCNet-sino}}$ by replacing 3D U-Net models by 3D CNNs (Figure 4.7).

Finally, the PVCNet-sino framework was adapted without the $P_{\text{rec10,fp}}$ input and without the $||P_{\text{PVE}} - \widehat{P_{\text{PVE}}}||_1$ term in the loss.

All these training options and their corresponding references are summarized in Table 4.2. Once trained, each method was evaluated on sinograms of the analytical test dataset.

Method	Nb of network(s)	Architecture	Nb of blocks	First Layer Channels (N_{ch})	with $P_{\text{rec10,fp}}$	with denoising loss	Total Nb of params
1UNET-recfp-3B	1	3D U-Net	3	8	Yes	N/A	392k
1UNET-recfp-4B	1	3D U-Net	4	8	Yes	N/A	1.6M
1UNET-recfp-5B	1	3D U-Net	5	8	Yes	N/A	6.4M
2CNN-recfp-PVEloss-32ch	2	3D CNN	8	32	Yes	Yes	339k
2CNN-recfp-PVEloss-64ch	2	3D CNN	8	64	Yes	Yes	1.3M
2UNET-recfp-PVEloss-3B	2	3D U-Net	3	8	Yes	Yes	783k
2UNET-recfp-PVEloss-4B	2	3D U-Net	4	8	Yes	Yes	3.2M
2UNET-recfp-PVEloss-5B	2	3D U-Net	5	8	Yes	Yes	12.9M
2UNET-PVEloss-3B	2	3D U-Net	3	8	No	Yes	784k
2UNET-recfp-3B	2	3D U-Net	3	8	Yes	No	783k

TABLE 4.2 – All methods trained in the ablation study.

4.4.1.2 Post-reconstruction evaluation

This test dataset was also used to evaluate the performances of our method on reconstructed images. While both the error monitoring during training and ablation study were performed in the sinogram domain, we also used the test dataset to assess the ability of our proposed method to correct PVC on reconstructed images. To achieve this, the PVCNet-sino framework was first applied to all input sinograms. Corrected sinograms were then used to reconstruct volumes with OSEM algorithm (10 iterations, 8 subsets) and a $(4.7952)^3$ mm^3 voxel size, resulting in 1,000 $I_{\text{PVCNet-sino}}$ images. Since I_{rec10} (OSEM with RM from raw data) was already reconstructed during test dataset generation, it was used to compare both RM and PVCNet-sino images to ground truth activity distributions I_{src}.

4.4.2 Calibration Phantom

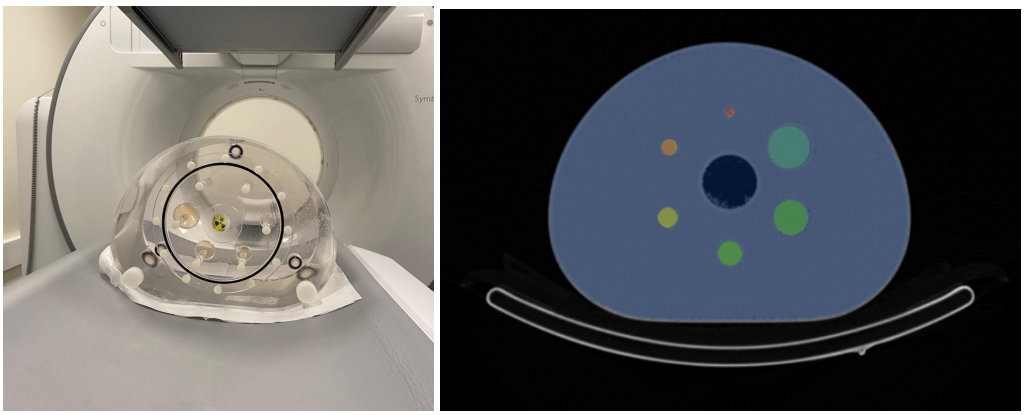
A first acquisition of a large uniform water cylinder filled with a 200 kBq/mL uniform activity was performed for calibration. The Calibration Factor (CF) was used in following experiments to convert reconstructed counts into activity concentration in kBq/mL, as recommended by [140]. The CF was computed according to Equation (1.35).

The FOV was manually delineated from a subsequent CT scan with the 3DSlicer Software [109].

4.4.3 NEMA IEC Phantom

To evaluate our proposed method on real data, we performed an acquisitions of a NEMA IEC phantom, containing 6 spheres (with 10, 13, 17, 22, 28, 37mm diameters) filled with a 1110 kBq/mL ^{177}Lu solution surrounded by a background activity of 83 kBq/mL.

Both phantoms acquisitions (Calibration and NEMA IEC phantoms) were performed with the Siemens Symbia Intevo Bold SPECT/CT system with dual head and Medium Energy (ME) collimators. 120 projections were acquired during 15 seconds each, with a 20% primary window centered around 208 keV and two surrounding 10% scatter windows. A non-circular body-contour orbiting trajectory was covered by detectors. Each projection contained 128×128 pixels with a $4.7952\text{mm} \times 4.7952\text{mm}$ size. Scatter was corrected before reconstruction and PVC, with the triple energy window method [162]. Images were reconstructed with the RTK software, using the OSEM algorithm with 8 subsets and 20 iterations, CT-based attenuation correction, no regularization, without any post-reconstruction smoothing. Spheres and background were manually delineated from the CT image with 3DSlicer, as illustrated in Figure 4.8.



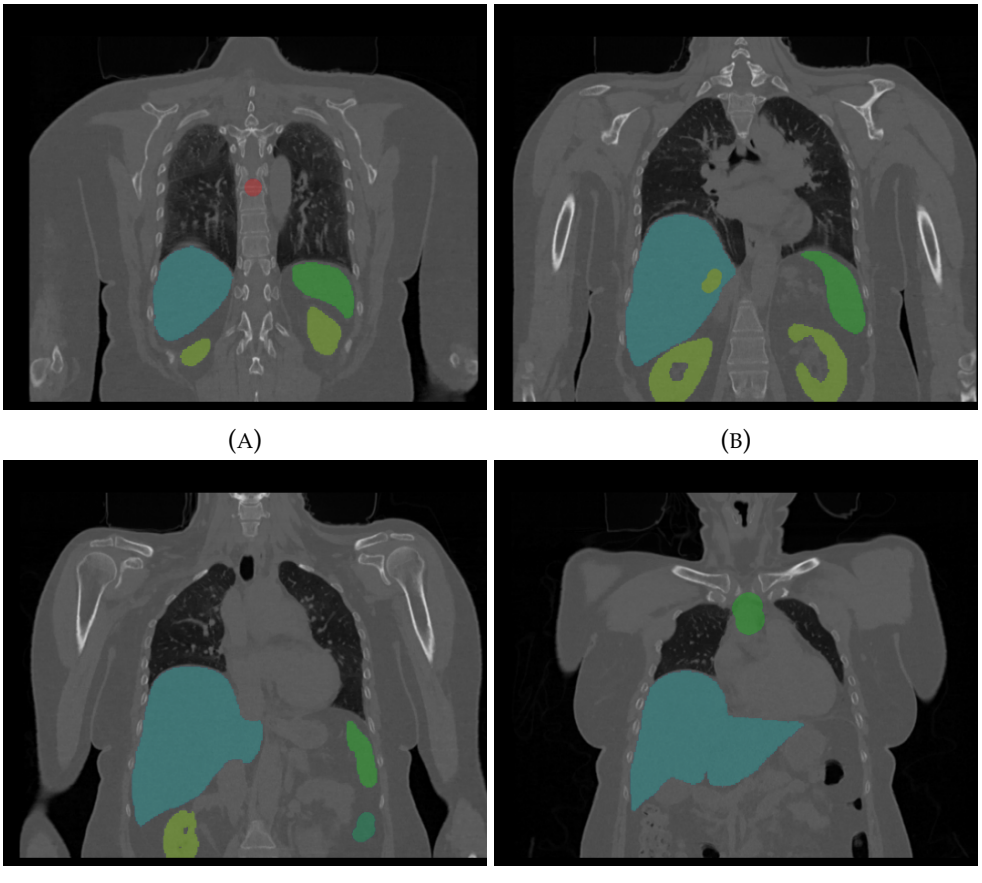
(A) (B)

FIGURE 4.8 – (A) Picture of the NEMA IEC Phantom before acquisition and (B) CT images of the phantom overlapped with spheres and background delineations.

4.4.4 Monte Carlo Simulations

A patient's CT was used to perform a MC simulation of a realistic SPECT acquisition with Gate [206]. The voxelized source image was defined by using an automatic segmentation of the CT in addition to four manually added lesions. TBR for all regions were chosen to be similar to ^{177}Lu-DOTATATE uptake distributions (see Table 4.4). Patient CT and source distribution are shown on four coronal slices in Figure 4.9. Once again, 120 projections were acquired during 15 seconds each, with a 20% primary window centered around 208 keV and two surrounding 10% scatter windows. In this case, a circular orbiting trajectory was used, with 280 mm radius. Same number of pixels, pixel size, scatter correction and reconstruction algorithm as described in Section 4.4.3 were used.

We used the Geant4 physics list "standard electromagnetic option 3" with production cuts set to 1 mm in the patient. A neural-network Angular Response Function (nn-ARF, see Section 2.2.5), trained with a model of the Siemens Symbia Intevo Bold with MEGP collimator, was used to model the detector [199]. This Variance Reduction Technique enabled us to reduce the simulation time by simulating only 300 MBq of total injected activity of ^{177}Lu, achieving approximately the same noise level as a 3 GBq simulation.



(A) (B)

(C) (D)

FIGURE 4.9 – Coronal slices of the patient CT and activity distribution used for MC simulation of the SPECT acquisition with Gate

4.4.5 Real Patient data

Finally, PVC methods were compared on a real patient treated with ^{177}Lu-DOTATATE and imaged with a SPECT/CT acquisition, performed 24 hours after a 7.4 GBq injection at the Centre Henry Becquerel of Rouen. This case was shown for visual assessment only, as no ground truth was available to quantitatively evaluate our proposed method. Organs were automatically segmented by TotalSegmentator. In order to apply iterative Young method, the five lesions contours were obtained with a region growing algorithm with a threshold value of $A_{\text{threshold}} = \frac{1}{2}(A_{\text{max,lesion}} + A_{\text{mean,background}})$ as proposed by [175].

4.4.6 Compared PVC methods

For each evaluation datasets, three PVC algorithms were compared. First, I_{RM} was reconstructed with 20 iterations of OSEM with RM, by using the same depth-dependant PSF as the one used for the synthetic dataset generation (Equation 3.11). This method is considered as the current state-of-the-art method and is used most of the time [56]. Second, the iterative Yang (iY) algorithm (see Equation 3.26) was applied to I_{RM} during 10 iterations, to obtain I_{iY}. iY was applied using the PETPVC Software [228]. The 3D-PSF used to apply iY was determined with the IEC phantom results, by fitting the three parameters of a 3D gaussian kernel k such that $||I_{\text{RM}} - k \otimes I_{\text{source}}||^2$ was minimal. We found FWHMs of 12.2 mm, 10.2 mm, 12.5 mm in the x, y and z directions respectively (where y was the rotation axis and $x - z$ the transverse plane). Third, $I_{\text{PVCNet-sino}}$ was reconstructed with the sinogram corrected by $h_{\text{PVCNet-sino}}$, 10 iterations of OSEM with 8 subsets, without RM nor regularization. All images were reconstructed with a $(4.7952\text{ mm})^3$ voxel size.

4.5 Metrics

PVC methods were compared to each other using pixel-wise metrics such as Normalized Root Mean Squared Error (NRMSE) and Normalized Mean Absolute Error (NMAE) defined by:

$$\text{NRMSE}(I^{\text{ref}}, I) = \frac{1}{\sum_i |I_i^{\text{ref}}|}\sqrt{\sum_i (I_i^{\text{ref}} - I_i)^2} \tag{4.9}$$

$$\text{NMAE}(I^{\text{ref}}, I) = \frac{1}{\sum_i |I_i^{\text{ref}}|}\sum_i |I_i^{\text{ref}} - I_i| \tag{4.10}$$

Region based metrics, such as Contrast-to-Noise Ratio (CNR), Recovery Coefficient (RC) and Contrast Recovery Coefficient (CRC) were also considered and defined, for region R, by:

$$\text{CNR}_R(I) = \frac{|\overline{I}_R - \overline{I}_{\text{background}}|}{\text{std}(I_{\text{background}})} \tag{4.11}$$

$$\text{RC}_R(I^{\text{ref}}, I) = \frac{\overline{I}_R}{\overline{I^{\text{ref}}}_R} \tag{4.12}$$

$$\mathrm{CRC}_R(I^{\mathrm{ref}}, I) = \frac{\overline{I}_R / \overline{I}_{\mathrm{background}} - 1}{\overline{I^{\mathrm{ref}}}_R / \overline{I^{\mathrm{ref}}}_{\mathrm{background}} - 1} \tag{4.13}$$

where $\bar{I}_R$ is the mean of image I in region R and $\mathrm{std}(I_{\mathrm{background}})$ is the standard deviation of image I in the background.

We also compared volume activity accuracy (VAA) values, defined by [124] as the percentage of voxels in the volume being at lower than 5% relative difference with the reference image, *i.e.*:

$$\mathrm{VAA}(I^{\mathrm{ref}}, I) = \frac{1}{\#\mathrm{voxels}} \#\left\{ i \text{ such that } \frac{|I_i^{\mathrm{ref}} - I_i|}{|I_i^{\mathrm{ref}}|} \leq 0.05 \right\} \times 100\% \tag{4.14}$$

Note: we do not use Structural Similarity Index (SSIM) [247] or Feature-based similarity index (FSIM) that are often used in publications because those metrics were designed for natural images, not medical images and we think they are not adapted to quantitative imaging [258].

4.6 Results

4.6.1 Analytical Test Dataset

4.6.1.1 Ablation Study

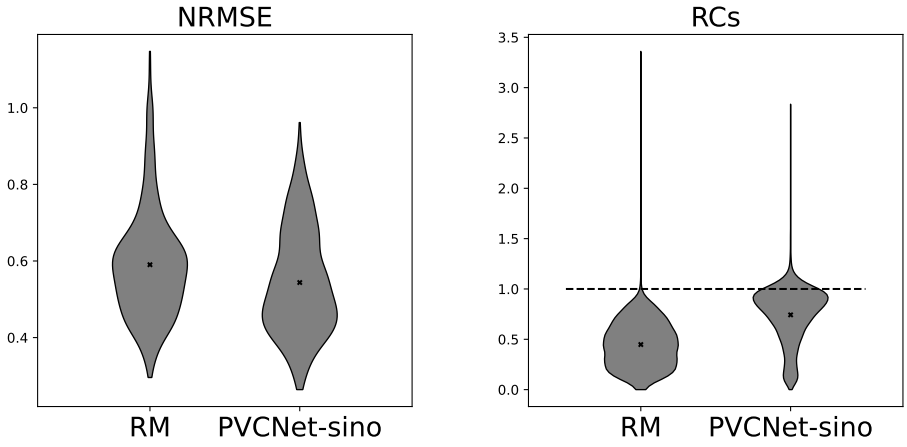
Results of the ablation study are displayed in Table 4.3. The method that obtains the best error on the test dataset is "2UNET-recfp-PVEloss-3B", *i.e.* with two U-Nets both containing 3 encoder/decoder blocks, usage of the $P_{\mathrm{rec10,fp}}$ as input and of separated denoising and deblurring losses.

Method	NMAE
1UNET-recfp-3B	$5.13 \times 10^{-2} \pm 1.57 \times 10^{-2}$
1UNET-recfp-4B	$5.56 \times 10^{-2} \pm 1.97 \times 10^{-2}$
1UNET-recfp-5B	$5.32 \times 10^{-2} \pm 1.42 \times 10^{-2}$
2CNN-recfp-PVEloss-32ch	$6.31 \times 10^{-2} \pm 1.22 \times 10^{-2}$
2CNN-recfp-PVEloss-64ch	$5.12 \times 10^{-2} \pm 1.27 \times 10^{-2}$
2UNET-recfp-PVEloss-3B	$\mathbf{4.91 \times 10^{-2} \pm 1.41 \times 10^{-2}}$
2UNET-recfp-PVEloss-4B	$5.36 \times 10^{-2} \pm 1.52 \times 10^{-2}$
2UNET-recfp-PVEloss-5B	$5.71 \times 10^{-2} \pm 2.32 \times 10^{-2}$
2UNET-PVEloss-3B	$5.29 \times 10^{-2} \pm 1.57 \times 10^{-2}$
2UNET-recfp-3B	$4.93 \times 10^{-2} \pm 1.39 \times 10^{-2}$

TABLE 4.3 – Results of each method on the 1,000 samples test dataset in terms of NMAE $\pm$ standard deviation. Characteristics of each method can be found in Table 4.2. Best method is highlighted in bold.

4.6.1.2 Post-reconstruction evaluation

After reconstruction, mean and distributions of NRMSE and RCs on the analytical test dataset are shown as violinplots in Figure 4.10. Figure 4.10a shows that the mean quadratic error obtained after reconstruction with sinograms corrected with PVCNet-sino is 8% lower than with RM. The distribution of Recovery Coefficients on lesions shown in Figure 4.10b is also very encouraging since the mean RC obtained with PVCNet-sino was 0.74 while only 0.45 for RM. Additionally, the RC distribution of PVCNet-sino shows a peak around 0.92.



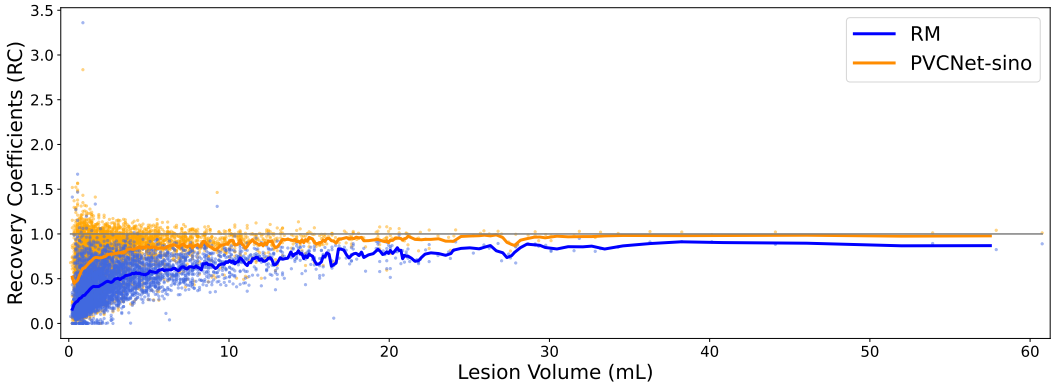
(A) (B)

FIGURE 4.10 – Violinplots of (A) NRMSE (the lower, the better) and (B) RCs (the closer to one, the better) for images reconstructed with RM and with sinograms corrected by PVCNet-sino on the analytical test dataset. RCs were computed on lesions only. The cross on each plot represents the mean value.

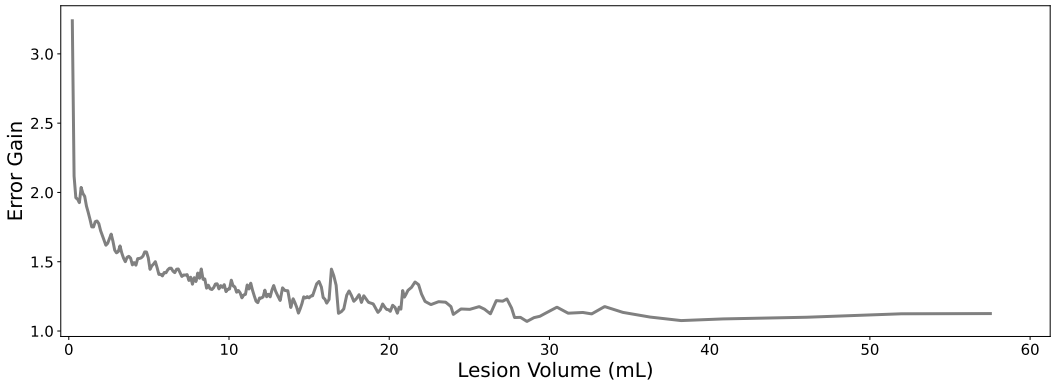
Figure 4.11a shows the distribution of RC values as a function of the lesion volume. Each point corresponds to a different synthetic lesion. As expected, both RM and PVCNet-sino performances increase with the lesion volume. Solid lines represent the moving average of the absolute error to 1 of each method ($1 - |\text{RC} - 1|$). This illustrates the overall capacity of our method to increase the SPECT image accuracy on lesions, compared to RM. A good property of PVCNet-sino on this test dataset is that the gain regarding this absolute error is superior for small structures, as illustrated by Figure 4.11b. This behavior is desirable since it is on small lesions that PVEs have the highest impact. However, it is also on these smallest lesions that the activity tended to be over-estimated by our proposed method.

4.6.2 Calibration Phantom

Reconstruction of the calibration phantom is displayed in Figure 4.12. The resulting Calibration Factor (CF) was 1.98×10^{-4} cps/kBq/mL. This factor will be used in all the following to convert reconstructed counts per seconds (cps) into activity concentrations (kBq/mL).



(A)



(B)

FIGURE 4.11 – (A) RCs on lesions as a function of their volume on the analytical test dataset for images reconstructed with RM and with sinograms corrected with PVCNet-sino. Each point represents a different lesion. Solid lines represent the moving average of each method. The ratio of these moving averages is displayed in (B) to illustrate the gain of using PVCNet-sino over RM.

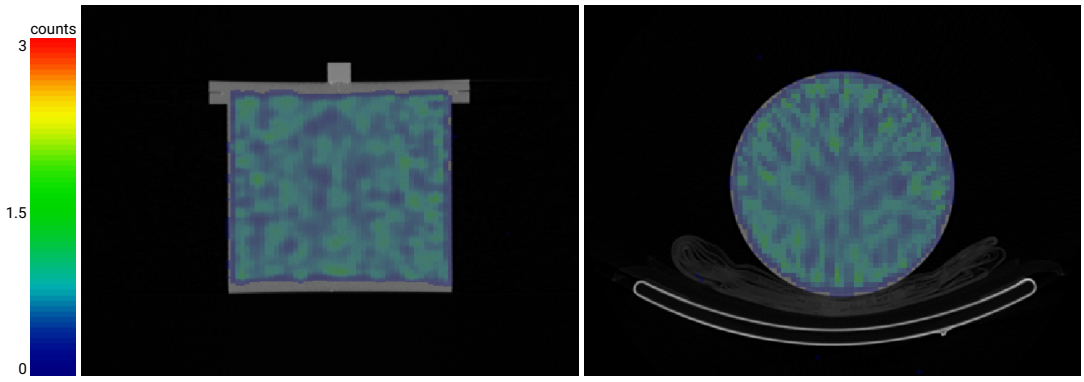


FIGURE 4.12 – Result of the calibration phantom, used to convert reconstructed counts into activity concentration with Equation (1.35)

4.6.3 NEMA IEC Phantom Acquisition

For the NEMA IEC Phantom experiment, input and corrected sinograms are shown in Figure 4.13 for visual comparison. Reconstructed images and RCs are shown in Figure 4.14. As expected, the magnitude of PVEs strongly depends on object size, specially for I_{RM}, which resulted in RC of 0.23/0.24/0.35/0.52/0.66/0.77 for the 10/13/17/22/28/37 mm diameter spheres respectively. Both iY and PVCNet-sino led to better visual and quantitative results than RM, which outputs a blurry and noisy image. I_{iY} demonstrated high image quality and activity estimation, except for the smallest sphere that was highly overestimated, with RC of 1.51/0.69/0.79/1.01/1.07/1.11. PVCNet-sino was also able to reduce PVEs in most spheres, both visually and quantitatively, with RC of 0.20/0.56/0.73/0.90/1.04/1.08. However, our proposed PVCNet-sino method was not able to do better than RM to recover the 10mm sphere, as it is still hardly distinguishable from background. Moreover, reconstruction artefacts between spheres are visible on PVCNet-sino and the spherical shape of the 22 and 28mm spheres is not completely conserved. Overall, results on the NEMA IEC Phantom show that PVCNet-sino is able to decrease PVEs both visually and quantitatively, reaching performances close to iY, without the need of any segmentation nor object-specific PSF.

In results presented in Figure 4.14, iY benefited from perfect spheres delineation and from a post-reconstruction 3D-PSF specifically determined for this object, unlike RM and PVCNet-sino. These two advantages are not achievable in practice, since tumour contours are difficult to delineate on CT scans and a pre-determined object-independent PSF is usually used to apply the iY method. To illustrate the effect of ROI delineation on the PVC, we applied the iterative Yang method to I_{RM} using different masks. First, we used the perfect spheres delineation, manually segmented on the CT image, as previously (iY image). A second mask was obtained by thresholding the reconstructed SPECT image I_{RM}, using a threshold value of $A_{\text{threshold}} = \frac{1}{2}(A_{\text{max,sphere}} + A_{\text{mean,background}})$ for each sphere, resulting in the $I_{\text{iY-threshold}}$ corrected image. Finally, a third mask was created by fitting the previous regions to spheres ($I_{\text{iY-threshold-spheres}}$ image).

Results of this comparison can be found in Figure 4.15. It is clear that using the threshold-based mask resulted in degraded image quality: segmented shapes are not spherical and are always smaller than the true spheres. This resulted in underestimated activity, with RC of 0.07/0.47/0.56/0.68/0.84/0.89. The thresholding+sphere fitting approach resulted in better activity quantification for the two biggest spheres but failed for the smallest ones (RC of 0.26/0.27/0.40/0.70/0.97/0.96). This can be explained because the mask used before fitting spheres contained quite distant voxels, resulting in too large spheres. These results show the dependence of the iY method on the mask used to delineate regions. Simply thresholding the SPECT image resulted in reduced PVEs compared to RM but significantly underestimated activity compared to iY with perfect segmentation. Introducing some *a priori* knowledge about the spherical shape of ROIs improved results only for the largest ones. On the other hand, PVCNet-sino does not need any segmentation nor *a priori* knowledge on the shapes of ROIs to correct. In the following, iY will refer to the iterative Yang method used with the perfect segmentation mask, when available, but this limitation has to be kept in mind. This limitation also applies for all region-based methods presented in Section 3.2.2.

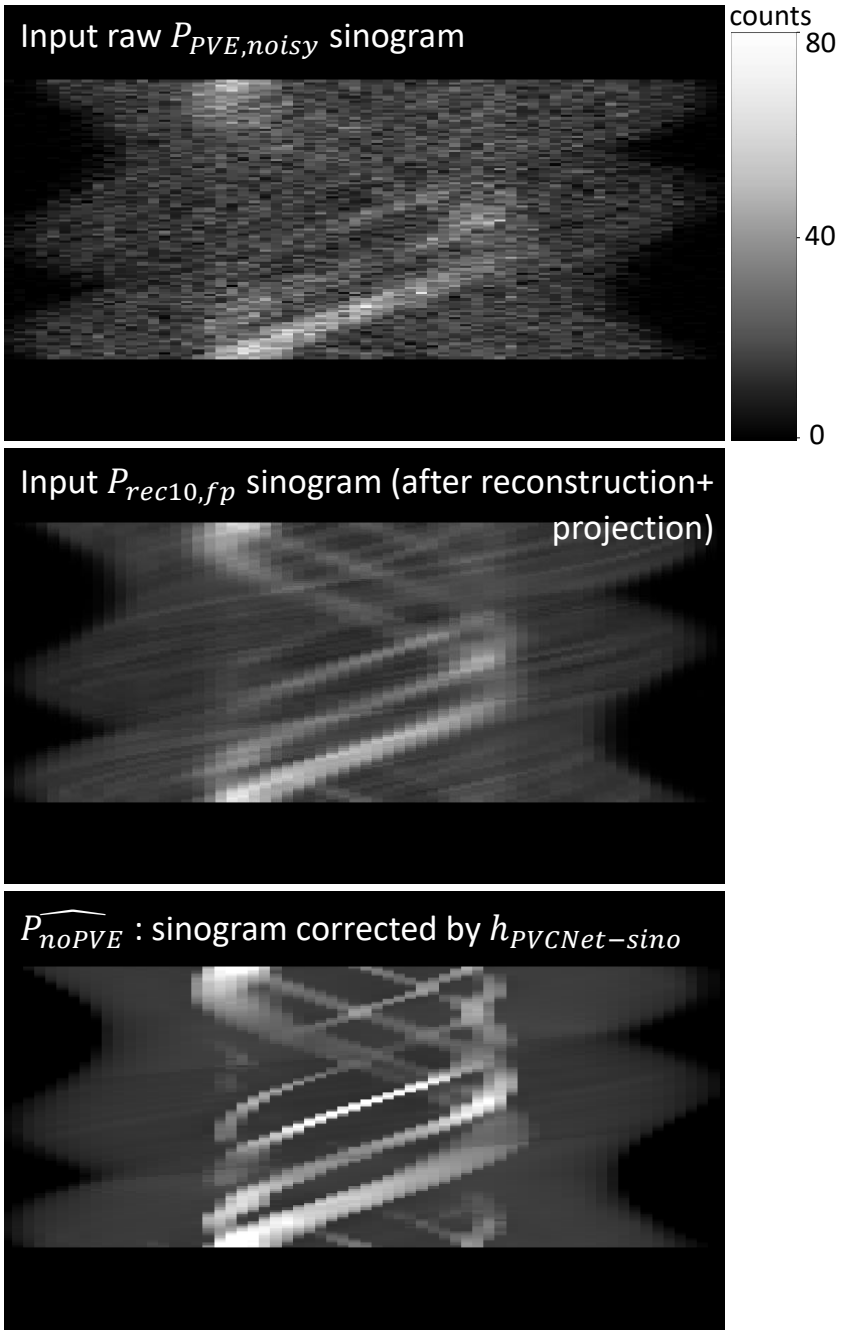


FIGURE 4.13 – Visual comparison of the input sinogram $P_{\text{PVE,noisy}}$ (after scatter correction with the Triple Energy Window method) with the sinogram obtained after 10 iterations of OSEM (with RM) and forward-projection, $P_{\text{rec10,fp}}$, used as input of $h_{\text{PVCNet-sino}}$ and the output $\widehat{P_{\text{noPVE}}}$ sinogram, corrected by $h_{\text{PVCNet-sino}}$.

4.6.4 Monte Carlo Simulations

Visual comparison of a transverse and a sagittal slice of reconstructed SPECT images of the Patient's MC simulation are shown in Figure 4.16 and Figure 4.17. Region-based RC and global error metrics are presented in Table 4.4 and Table 4.5

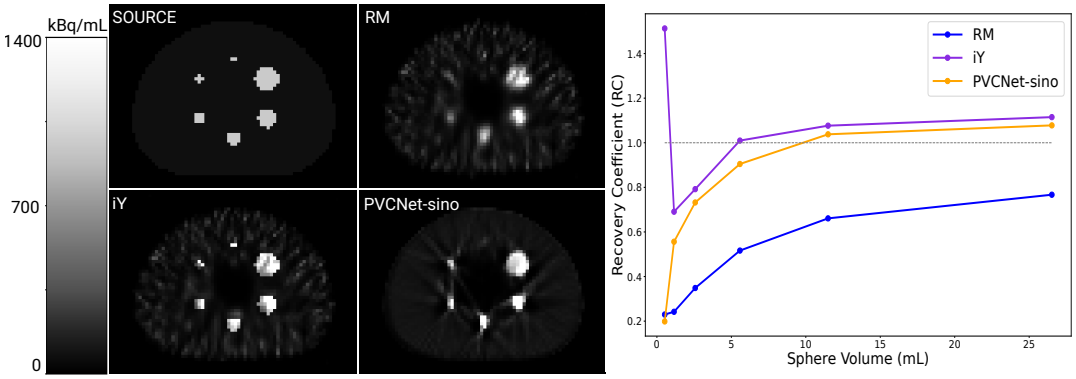


FIGURE 4.14 – Visual comparison of three PVC methods on the NEMA IEC Phantom acquired with the Siemens Intevo SPECT/CT system. Left panel shows slices of the manually segmented source and of images reconstructed with Resolution Modeling (RM), iterative Yang method (iY) and with our proposed deep-learning based method (PVCNet-sino). Right panel show recovery coefficients (RC) for each sphere and reconstruction method.

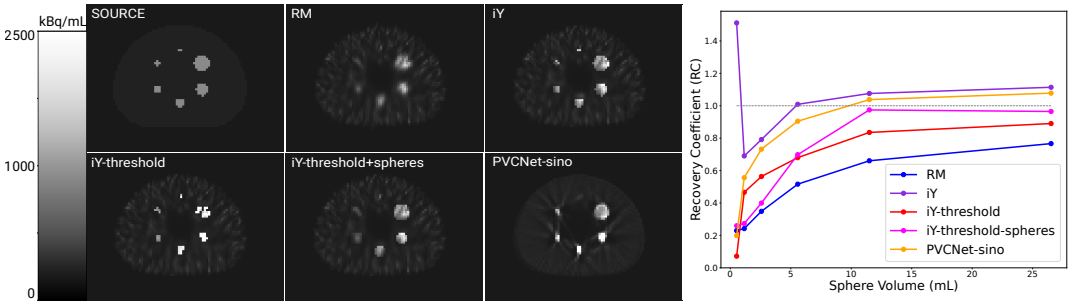


FIGURE 4.15 – Results of the NEMA IEC phantom reconstruction using different segmentation masks to apply the iterative Yang (iY) method: with the ideal sphere delineation (iY image), with the thresholded I_{RM} mask (iY-threshold) and thresolded+spheres fitting (iY-threshold-spheres). The target source activity image, I_{RM} and $I_{\text{PVCNet-sino}}$ are also shown for comparison. The right panels shows Recovery Coefficients (RC) of all methods.

respectively. Visually, both methods allowed to reduce PVEs on kidneys and liver compared to RM. iY recovered almost perfectly object shapes (which was given as input of the method) and activity, specially in lesions (with RC comprised between 0.98 and 1.15, systematically better than PVCNet-sino). PVCNet-sino resulted in RC closer to 1 for all regions except lesions, and significantly better NRMSE, NMAE and VAA than both RM and iY.

The left panel of Figure 4.18 shows the NRMSE convergence of each correction method as a function of the number of OSEM update. The PVCNet-sino iterations start at 10 iterations to account for the fact that the model uses $P_{\text{rec10,fp}}$ obtained after 10 iterations of OSEM and forward-projection. Using the sinogram corrected by *PVCNet-sino* instead of the raw sinogram in the reconstruction algorithm allowed to reduce by 28% the error after 3 additional updates. Both RM and PVCNet-sino slightly diverge after a few updates. This is due to the use of OSEM algorithm instead

	body	gallbladder	kidney left	kidney right	liver	spleen	lesion 1	lesion 2	lesion 3	lesion 4
RM	1.13 ± 1.68	1.08 ± 0.56	0.84 ± 0.51	0.86 ± 0.62	**0.98** ± 0.27	0.88 ± 0.34	0.59 ± 0.35	0.78 ± 0.13	0.56 ± 0.19	0.80 ± 0.37
iY	0.90 ± 0.60	0.82 ± 0.22	1.07 ± 0.53	1.06 ± 0.67	1.06 ± 0.20	1.06 ± 0.26	**0.97** ± 0.43	**1.06** ± 0.09	**1.15** ± 0.21	1.10 ± 0.28
PVCNet-sino	**1.05** ± 1.45	**1.10** ± 0.67	**0.94** ± 0.36	**0.95** ± 0.33	**1.02** ± 0.22	**0.96** ± 0.30	0.58 ± 0.25	0.66 ± 0.08	0.80 ± 0.24	**0.90** ± 0.19

TABLE 4.4 – Recovery Coefficients (RC) with ± standard deviation for the three compared methods in 10 different structures. True Target-to-background ratios (TBR) are displayed in parenthesis for each region. Best RC values are highlighted in bold.

	NRMSE	**NMAE**	**VAA**
RM	1.26	0.42	7.55%
iY	1.09	0.31	8.55%
PVCNet-sino	**0.90**	**0.26**	**21.24%**

TABLE 4.5 – NRMSE, NMAE and VAA for the three compared PVC. Best values are highlighted in bold.

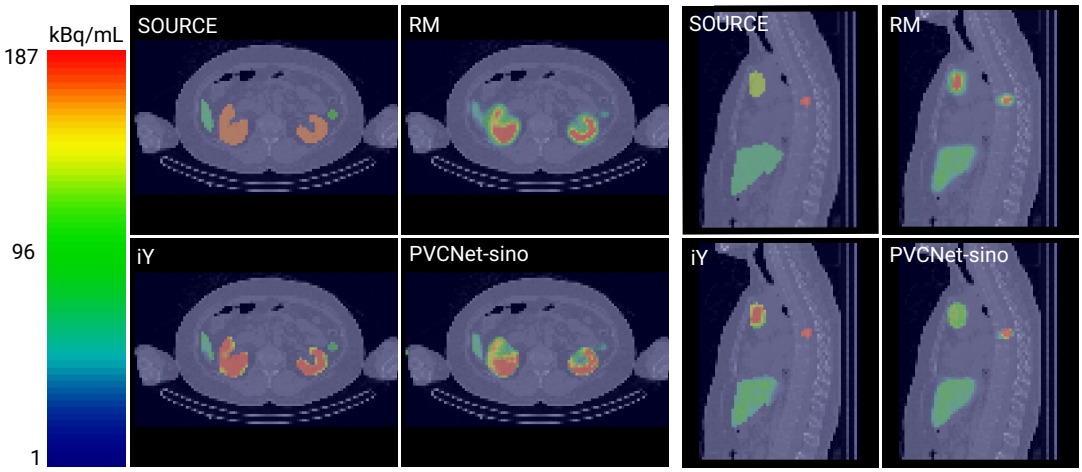


FIGURE 4.16 – Visual comparison of the true activity distribution and results of the three PVC methods on a patient MC simulation. Images reconstructed with Resolution Modeling (RM), iterative Yang method (iY) and with our proposed deep-learning based PVC method (PVCNet-sino) are shown. Axial slices are shown on the left panel and sagittal slices on the right.

of MLEM and the fact that no regularization was used in both cases. The NRMSE obtained after having applied the iY PVC method to the image reconstructed with 20 iterations (160 updates) of RM is also shown, reaching less than 14% error reduction.

In the right panel of Figure 4.18, the CNR is plotted as a function of the mean CRC_{err}. CRC_{err} is defined as $CRC_{err} = 1 - |CRC - 1|$ to account to the fact that CRC values should be close to 1. The mean value is taken over all ROIs (organs and lesions). This graph illustrates the accuracy *vs* noise trade-off to assess that the

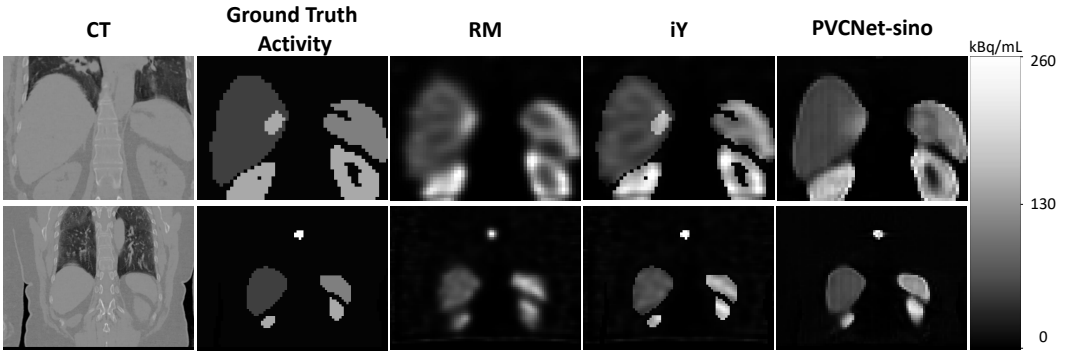


FIGURE 4.17 – Visual comparison of the true activity distribution and results of the three PVC methods on a patient MC simulation. Images reconstructed with Resolution Modeling (RM), iterative Yang method (iY) and with our proposed deep-learning based PVC method (PVCNet-sino) are shown..

better quantification properties of PVCNet-sino compared to RM, shown in Table 4.4 and Table 4.5, is not reached at the cost of poorer noise properties. Compared to RM, PVCNet-sino allowed to increase both the quantitative accuracy and the Contrast-to-Noise ratio (CNR), which is very encouraging. Final mean CRC_{err} with PVCNet-sino is even slightly better than with the iterative Yang method, with 0.82 and 0.81 respectively. However, iY reaches a three times higher CNR than PVCNet-sino, meaning more homogeneous background correction and higher uptakes in organs.

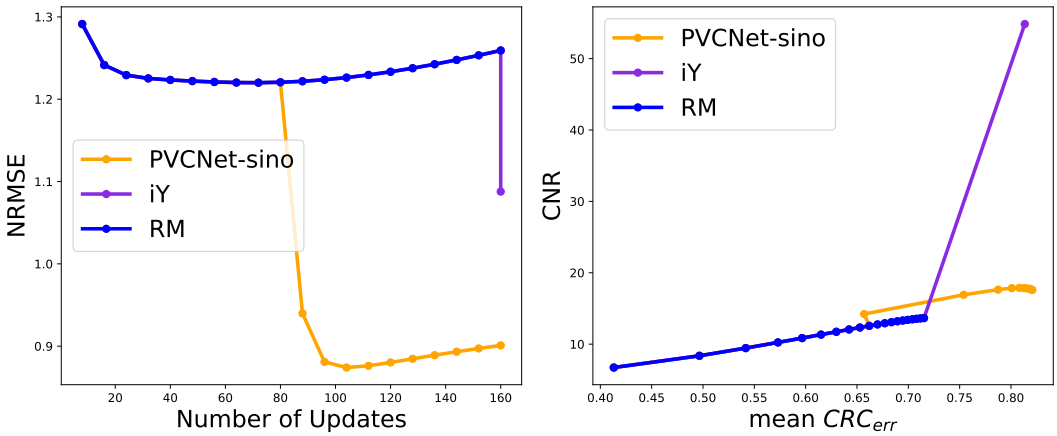


FIGURE 4.18 – Convergence of the NRMSE with respect to number of OSEM updates for RM, PVCNet-sino and iY (left) and evolution of CNR *vs.* mean CRC_{err} trade-off (right).

4.6.5 Patient data

Figure 4.19 shows a coronal and a transverse slice of each reconstructed images of the patient treated with ^{177}Lu-DOTATATE. Once again, since the segmentation was available, iY was the most able to recover objects contours, even tough the image

appears quite noisy. PVCNet-sino resulted in an image with better kidneys and liver contour definition, as visible in the first line as well as a visual enhancement of tumours close to iY. However, artifacts at the edges of regions and similarity with the blurry RM image (specially in the second line) still remain. Ground truth activity distribution was not available in this case for quantitative comparison.

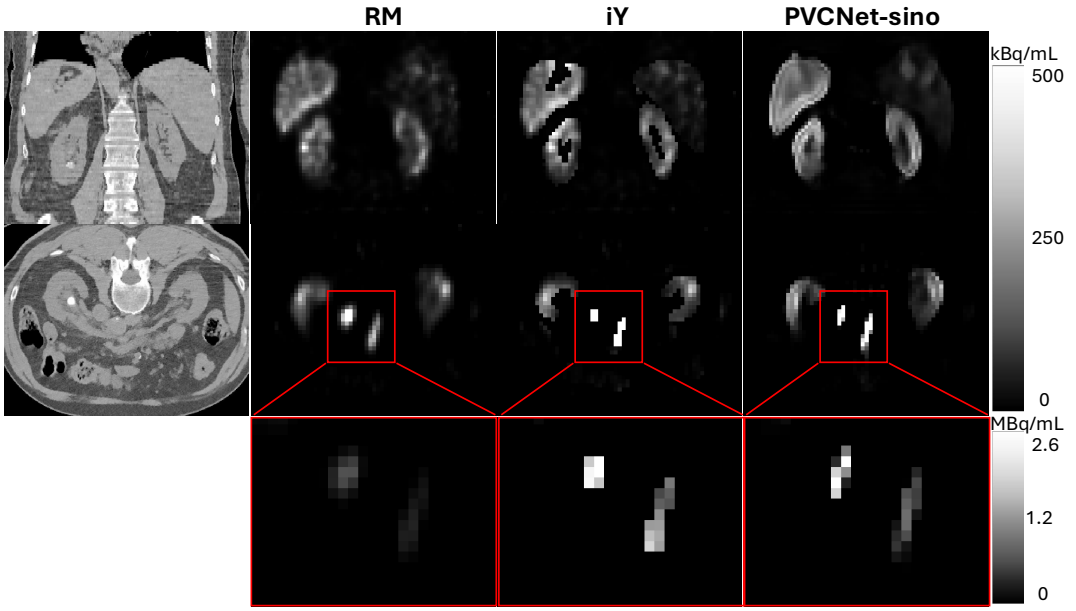


FIGURE 4.19 – Visual comparison of reconstructed images of a real patient data, treated with ^{177}Lu-DOTATATE. From left to right columns: attenuation map, image corrected with Resolution Modeling (RM), image corrected with the iterative Yang (iY) method, image corrected with PVCNet-sino. First two lines show a coronal and a tranverse slice while the third line is a zoom on two lesions visible in the second line.

4.7 Discussion

Collimator-induced Partial Volume Effects are one of the main factors limiting SPECT-based quantification. Existing approaches to reduce PVEs, such as RM and iY, are either insufficient or require the a priori knowledge of regions contour. In this chapter, we proposed an alternative deep learning approach to correct PVEs in the sinogram domain. We developed a method to generate a large training database of sinograms generated by analytical simulations made of 10,000 data. The $h_{\text{PVCNet-sino}}$ correction pipeline (see Figure 4.5) contained two 3D U-Nets trained using those synthetic sinograms to reduce noise and correct PVEs respectively. We showed the method was able to generalize to unseen synthetic data, to MC simulations, phantom experiments and real patient data.

The choice to train a deep-learning model to correct PVEs on SPECT sinograms was never explored to our knowledge, while post-reconstruction image domain approaches are more popular [124, 255]. The lack of sinogram-based PVC method

may be surprising since PVEs are primarily due to the collimator's geometric response function, an effect clearly highlighted in the projections. Moreover, unlike post-reconstruction methods, sinogram-based PVC can be applied independently of the tomographic reconstruction algorithm used and from its parameters. Another advantage of sinogram-based PVC is that the tomographic reconstruction is slightly faster since no more resolution modeling has to be applied during reconstruction. However, challenges of sinogram-based learning include the fact that training a *convolutional* neural network on a sinogram may seem surprising and counter-intuitive: while convolution kernel parameters are optimized to exploit local information, an object appears at several very "distant" locations in a sinogram. Moreover, applying a correction on the sinogram as a pre-processing step before reconstruction has to be performed carefully since the data consistency of the measure has to be preserved.

We showed in the ablation study in Section 4.6.1.1 that sinogram-domain learning was possible provided sufficient information was given as input to $h_{\text{PVCNet-sino}}$. In particular, using the sinogram obtained after reconstruction with RM and forward-projection ($P_{\text{rec10,fp}}$) helped to reduce the error by more than 7% on a test dataset. The interest of using $P_{\text{rec10,fp}}$ as input can be explained by the fact when trying to solve an inverse problem with deep learning, using an approximation of the inverse as a first estimate is beneficial, as argued by [164]. Indeed, taking back notations of Equations 4.2 and 4.3, the forward model of the inverse problem we tried to solve in the projection domain was to estimate the sinogram unaffected by collimator/detector response from the blurred and noisy sinogram. This inverse problem can be formulated as $P_{\text{PVE,noisy}} = A_{\text{PSF}} \tilde{A}^{-1} P_{\text{noPVE}}$, where the $\tilde{A}^{-1}$ matrix denotes an estimation of the inverse of A, without PSF modeling and A_{PSF} the forward projection with PSF modeling. Thus, an estimation of P_{noPVE} can be written $P_{\text{noPVE}} = A\,\tilde{A}_{\text{PSF}}^{-1} P_{\text{PVE,noisy}} \approx A\,I_{\text{rec10}} = P_{\text{rec10,fp}}$, since I_{rec10} is reconstructed using the PSF in its system matrix. It can be seen from Figure 4.4 and Figure 4.13 that $P_{\text{rec10,fp}}$ is an already denoised and corrected version of the raw $P_{\text{PVE,noisy}}$ sinogram. This is also the reason why residual units were added after $h_{\text{denoising}}$ and $h_{\text{deblurring}}$ (see Figure 4.5) since their output is expected to be close from $P_{\text{rec10,fp}}$.

Preservation of the consistency on the number of counts was imposed by the normalization (Equations 4.5 and 4.6). Usage of the denoising loss in Equation (4.7), motivated by the usage of as much knowledge as possible during training and task decomposition, helped to improve the performances of the method, but very slightly– less than -0.5% error reduction in the ablation study. Additionally, since conventional data augmentation techniques such as translations and rotations cannot be simply applied on sinograms, we proposed to apply Poisson noise realization at each training iteration.

We have shown that *PVCNet-sino* was able to reduce PVEs both visually and quantitatively, specially for the biggest objects: on spheres with 22, 28 and 37 diameters on the IEC acquisition (RCs between 0.90 and 1.08) and on organs of the patient simulation (RCs between 0.94 and 1.10). However, despite very encouraging results on the test dataset on small lesions (Figure 4.11), the ability of the *PVCNet-sino* method to correct lesions in more realistic context is still limited (RC between 0.58 and 0.90 for the patient's simulation), specially compared to the iterative Yang method (RCs comprises between 0.98 and 1.15 on lesions of the patient's simulation). This is particularly problematic since it is on small lesions that PVEs have highest

impact.

The main difficulty of training a deep learning model to correct PVEs is the lack of real ground truth data. This is an issue both for training and evaluation. This limitation was circumvented by the use of an analytically generated dataset to train our $h_{\text{PVCNet–sino}}$ pipeline and with phantom experiment and MC simulation for the evaluation. However, the ability of our proposed method to correct PVEs in a clinical context is still to be demonstrated.

4.8 Perspectives

While the sinogram based *PVCNet-sino* approach presented in this Chapter was shown to provide good results compared to RM and iY methods, other deep-learning methods could be investigated. For example, as mentioned in the discussion, image-based training may seem more natural and straightforward to train and apply. Training a PVC model to correct only on a specific patient could provide better results and be an alternative to the search of generalizability. Also, the usefulness of using $P_{\text{rec10,fp}}$ as input to reduce PVEs on sinograms suggests that combining image and sinogram domain approaches may be the path to follow. Finally, unsupervised methods, leveraging the lack of ground truth data, may be a solution to deep-learning based PVC in SPECT.

Some of these perspectives will be investigated in Chapter 5.

Chapter 5

Contribution: alternative PVC methods

Sommaire

This Chapter investigates deep-learning PVC methods that are alternatives to the *PVCNet-sino* method proposed in Chapter 4. Those contributions are proof-of-concepts showing potential advantages of different approaches for PVC within the same deep learning framework, using the same datasets than in the previous chapter. This work was presented at the *2024 IEEE Nuclear Science Symposium, Medical Imaging Conference, and Room-Temperature Semiconductor Detectors Symposium* [106].

5.1 Image Domain Training

Post-reconstruction image domain training is certainly the most natural way to enhance image resolution with deep learning. In this section, a unique convolutional neural network $h_{\text{PVCNet-img}}$ is trained in a supervised setting to estimate the true activity, corrected from PVE, from an uncorrected image.

All necessary inputs and targets were already available from the dataset generated in Section 4.2:

— I_{src} is the ground truth activity image to recover. The only difference with the I_{src} from Section 4.2 is that this source image was re-sampled from the CT resolution (with ≈ 1 mm^3 voxel size) to the SPECT voxel size ($4.7952 \times 4.7952 \times 4.7952$ mm^3). This re-sampling was necessary to standardize all image sizes and so that inputs/targets have the same dimension.

— I_{rec10} the image reconstructed with 10 iterations of OSEM (8 subsets, Resolution Modeling during reconstruction and $4.7952 \times 4.7952 \times 4.7952$ mm^3 voxel size) will directly serve as input in this section.

— The attenuation map I_{att} was used as an additional input of $h_{\text{PVCNet-img}}$. To match the target activity map I_{src} dimensions, it was also re-sampled from the CT voxel grid to the SPECT image grid.

An example of $I_{\text{src}}/I_{\text{rec10}}/I_{\text{att}}$ from the training dataset is shown in Figure 5.1.

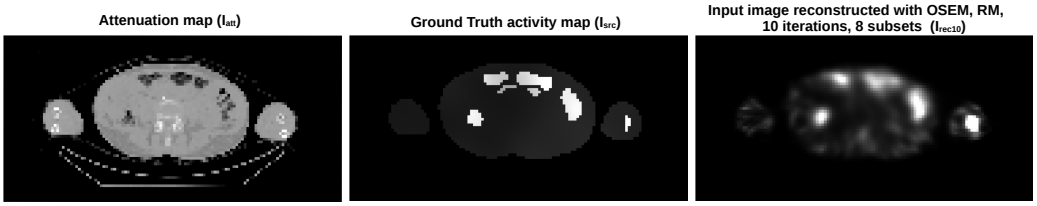


FIGURE 5.1 – Example of I_{att}, I_{src} and I_{rec10} images, from the training dataset.

Data augmentation was performed during training by randomly rotating each $I_{\text{att}}/I_{\text{src}}/I_{\text{rec10}}$ triplets at each epoch, as well as switching their axes.

All input/output images were padded to have $128 \times 128 \times 128$ voxels. Similarly as in Section 4.3.1, images were normalized such that the input values were in the $[0, 1]$ interval (see Equation (5.1)) and the output total activity was conserved (Equation (5.3)). The image-to-image neural network $h_{\text{PVCNet-img}}$ was applied with a residual unit (Equation (5.2)) for the same reasons as for $h_{\text{PVCNet-sino}}$, *i.e.* because the output is expected to be close from I_{rec10} and to ensure positivity of the corrected image.

$$I^{\text{normalized}} = \frac{I}{\max(P)} \tag{5.1}$$

$$\widehat{I}_{\text{PVCNet-img}} = \text{ReLu}\left(h_{\text{PVCNet-img}}(I_{\text{att}}^{\text{normalized}}, I_{\text{rec10}}^{\text{normalized}}) + I_{\text{rec10}}^{\text{normalized}}\right) \tag{5.2}$$

$$\widehat{I}_{\text{PVCNet–img}}^{\text{de-normalized}} = \frac{\widehat{I}_{\text{PVCNet–img}}}{\sum \widehat{I}_{\text{PVCNet–img}}} \sum I_{\text{rec10}} \tag{5.3}$$

$$\mathcal{L}_{\text{img}} = ||I_{\text{src}} - \widehat{I}_{\text{PVCNet–img}}||_1 \tag{5.4}$$

The architecture of the $h_{\text{PVCNet-img}}$ convolutional neural network was a 3d U-net (see Figure 4.6). After hyper-parameter tuning, N_{ch} was also chosen to be equal to 8, the batch size was set to 8, and the learning rate was 1×10^{-4} halved every 50 epochs.

5.2 Dual Domain Training

In the sinogram-to-sinogram correction proposed in Chapter 4, the ablation study in Section 4.6.1.1 showed that using the sinogram obtained after reconstruction with RM and forward projection helped the model to reduce PVEs. This finding suggests that mixing reconstruction domain and sinogram domain learning could enhance the results.

In [131], authors compared deep learning methods in the projection domain, image domain and in both domains for sparse-view CT image reconstruction and concluded that combining both domains leads to the best NRMSE/SSIM results. In the context of self-supervised PET image reconstruction, one of the results of [182] is also that dual domain learning leads to better generalization properties. For low-dose and few-angle cardiac SPECT reconstruction, this dual domain approach was also iteratively used in [35, 34].

In this Section, we follows this idea to combine sinogram domain learning (Chapter 4) and image domain learning (Section 5.1). However, including 3D-tomographic reconstruction during training is not possible with the RTK software and, if it was, is expected to be computationally expensive, due to the necessity to compute gradients across OSEM iterations. For this reason, the pre-trained network $h_{\text{PVCNet-sino}}$ (trained in Chapter 4) is first applied to the input sinograms, the same way as in Chapter 4. Then, the output corrected sinogram $\widehat{P_{\text{noPVE}}}$ was used to reconstruct the $I_{\text{PVCNet-sino}}$ image. This *PVCNet-sino* application and reconstruction step were done only once for each sample of the training/validation/test datasets. An image-to-image model was trained similarly as in Section 5.1 but with three input images: I_{att}, I_{rec10} and the additional $I_{\text{PVCNet-sino}}$ image. This framework is denoted PVCNet-sino-img and is described in the following equations:

$$\widehat{I}_{\text{PVCNet-sino-img}} = \text{ReLu}\big(h_{\text{PVCNet-sino-img}}\big(I_{\text{att}}^{\text{normalized}}, I_{\text{rec10}}^{\text{normalized}}, I_{\text{PVCNet-sino}}^{\text{normalized}}\big) + I_{\text{PVCNet-sino}}^{\text{normalized}}\big) \tag{5.5}$$

$$\mathcal{L}_{\text{sino-img}} = ||I_{\text{src}} - \widehat{I}_{\text{PVCNet-sino-img}}||_1 \tag{5.6}$$

Note that in this Section, weights of the sinogram-to-sinogram network $h_{\text{PVCNet-sino}}$ are fixed and only the parameters of $h_{\text{PVCNet-sino-img}}$ are trained.

5.3 Training details

For comparison, the architecture of both $h_{\text{PVCNet-img}}$ and $h_{\text{PVCNet-sino-img}}$ were the 3d U-net depicted in Figure 4.6) with $N_{\text{ch}} = 8$. Both were trained with the same hyper-parameters during 100 epochs, with a batch size of 8, a learning rate of 1×10^{-4} and the Adam optimizer. Data augmentation was performed during training by applying random rotations (between -5° and 5° in each direction).

	Inputs	**Target**	**Architecture**	**Number of parameters**
PVCNet-sino	$(P_{\text{PVE,noisy}}, P_{\text{attmap}}, P_{\text{rec10,fp}})$	P_{noPVE}	2 U-nets (3 blocks, $N_{\text{ch}} = 8$)	784k
PVCNet-img	$(I_{\text{att}}, I_{\text{rec10}})$	I_{src}	1 U-net (3 blocks, $N_{\text{ch}} = 8$)	400k
PVCNet-sino-img	$(I_{\text{att}}, I_{\text{rec10}}, I_{\text{PVCNet-sino}})$	I_{src}	1 U-net (3 blocks, $N_{\text{ch}} = 8$)	400k

TABLE 5.1 – Summary of main training details for the three deep learning models.

5.4 Evaluation datasets

In this section, we compare the performance of PVCNet-sino with PVCNet-img and PVCNet-sino-img, all trained with the same dataset but respectively only in the sinogram domain, only in the image domain and on both. PVC methods were first compared on the analytical test dataset containing 1,000 samples (see Section 4.4.1). This dataset was used to compute Normalized Root Mean Squared Errors (NRMSE, Equation (4.9)) and Recovery Coefficients (RC, Equation (3.1)) on lesions. In addition to these metrics that give information on the local and global quantitative accuracy of uncorrected/corrected images, lesion detectability may also be important. For this reason, lesions on each reconstructed images were segmented using the same threshold as in Section 4.4.5 and Section 4.6.3, *i.e.*: $A_{\text{threshold}} = \frac{1}{2}(A_{\text{max,sphere}} + A_{\text{mean,background}})$. Once a mask of each lesion was segmented on each reconstructed image, the Dice coefficient, defined for two regions A and B in Equation (5.7), was used to quantify similarity of the segmented mask with true lesion contours. Values closer to 1 indicate better segmentation.

$$\text{Dice(A,B)} = 2\frac{|A \cap B|}{|A| + |B|} \tag{5.7}$$

The Monte-Carlo simulation of the patient was also used for evaluation. Quantitative accuracy was evaluated with RC (see Equation (4.12)) and NMAE (Equation (4.10)) in each ROI, including the four manually added lesions. In addition to NRMSE, NMAE and VAA calculation for global error metric comparison, background noise was evaluated in each reconstructed image by computing the standard deviation of the estimated activity over 10 spheres of 4.5 cm diameter, all located in the background. We are aware that this threshold-based segmentation is probably not optimal in clinical situations, but it provides here a common reference for the three compared methods.

5.5 Results

NRMSE and RCs (only on lesions) can be compared in Figure 5.2. The violin plot shows their distribution over the 1,000 samples. Both image domain training approaches lead to better mean squared error than RM and PVCNet-sino with means of 0.407 for PVCNet-sino-img, 0.428 for PVCNet-img, 0.544 for PVCNet-sino and 0.590 for RM. However, in terms of RC on synthetic lesions, PVCNet-sino leads to values closer to 1 than other methods with a mean RC of 0.732 while the mean value for PVCNet-sino-img was 0.628 and only 0.480 for PVCNet-img and 0.432 for RM. The potential gain in RC is not obvious here: PVCNet-sino shows less variability around 1 but with some over-estimations compared to PVCNet-img which is only marginally better than RM. PVCNet-sino-img seems to be a good compromise between both.

Additionally, Dice coefficients are displayed in Figure 5.2c. Segmentations on images corrected with PVCNet-sino-img were the most accurate with a mean Dice of 0.606, while only 0.554 with PVCNet-img, 0.527 with RM and 0.478.

For results on the MC simulation, four coronal slices of reconstructed images of the patient's activity distribution are presented in Figure 5.3 and Recovery Coefficients for ROIs in Table 5.2. Visually, PVCNet-sino-img shows better inter-region correction than PVCNet-img, which is visible in the first two rows of Figure 5.3 between the spleen and the left kidney or between the liver and the right kidney. This is also highlighted on the profile shown in Figure 5.4b. Importantly, when only PVCNet-img and PVCNet-sino-img are compared quantitatively with RC values, the dual domain approach allows a better quantification for 8 over 10 ROIs.

Nevertheless, deep-learning based correction methods show difficulty to correct some small lesions. For example, the lesion located in the liver (Lesion 2, visible in the second row of Figure 5.3) is discarded by all PVCNet methods which output RC even lower than the image reconstructed with RM. It is also the case on Lesion 1, except for PVCNet-sino-img correction. On lesions 3 and 4, all deep-learning based methods allow to significantly increase RCs compared to RM. Lesion 4, visible on top of the 4^{th} row in Figure 5.3, is better corrected by PVCNet-img in terms of RC but the profile in Figure 5.4a shows that its size is under-estimated while the activity is over-estimated. This effect is not visible when only RC are compared. For this reason, local NMAE are displayed in Table 5.4, showing disparity between methods. On Lesion 4, PVCNet-sino and PVCNet-sino-img resulted in the lowest NMAE and PVCNet-img is the highest while its RC was the closest to 1.

In terms of global metrics such as NRMSE, NMAE and VAA, detailed in Table 5.3, PVCNet-sino-img resulted in the best image quality. Both image domain training approaches allowed to significantly decrease the background noise level compared to RM, iY and PVCNet-sino.

5.6 Discussion

In this Chapter, we compared the previously described PVCNet-sino method, trained in the sinogram domain, with other deep learning approaches trained in the image domain. Two models were trained: PVCNet-img, only with images corrected with RM as input, and PVCNet-sino-img, trained with an additional channel being the image reconstructed with projections corrected by the PVCNet-sino model. The overall objective of this Chapter was to evaluate which type of input allowed to better correct PVEs, reduce noise while being generalizable.

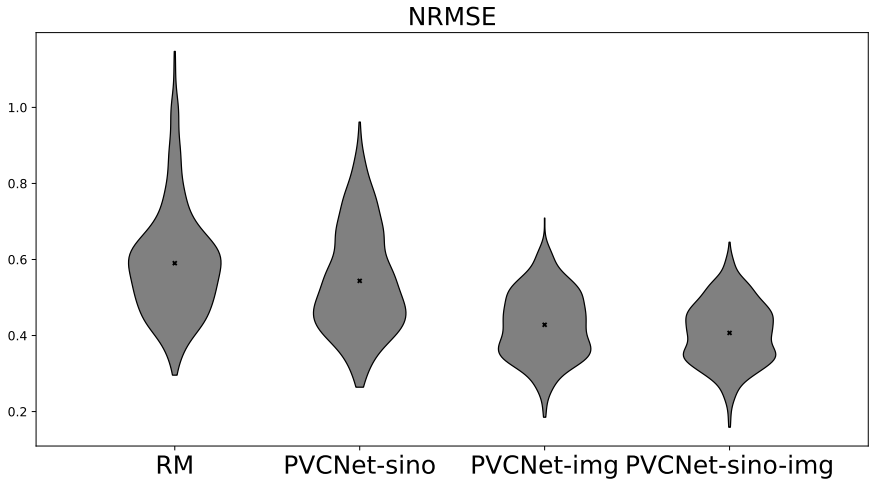
On both analytical and MC simulated datasets, better global errors (NRMSE, NMAE, VAA) and noise reduction were achieved by image domain training methods (PVCNet-img and PVCNet-sino-img) compared to the sinogram-only training (PVCNet-sino) as shown in Figure 5.2a and Table 5.3. This behavior can be explained by the fact that in image-to-image training, optimization and evaluation are performed on the same domain, while an additional reconstruction step was required after applying the PVCNet-sino model. Moreover both PVCNet-img and PVCNet-sino-img were optimized with a L1 loss, *i.e.* NMAE, also related to VAA, which can explain their relative better performances concerning these metrics.

However, both models that included sinogram information, PVCNet-sino and PVCNet-sino-img, allowed better lesion quantification on the test dataset with respectively 37.8 % and 27.1 % higher mean RC than PVCNet-img. On the MC simulation however, results were not so clear: while PVCNet-sino-img showed

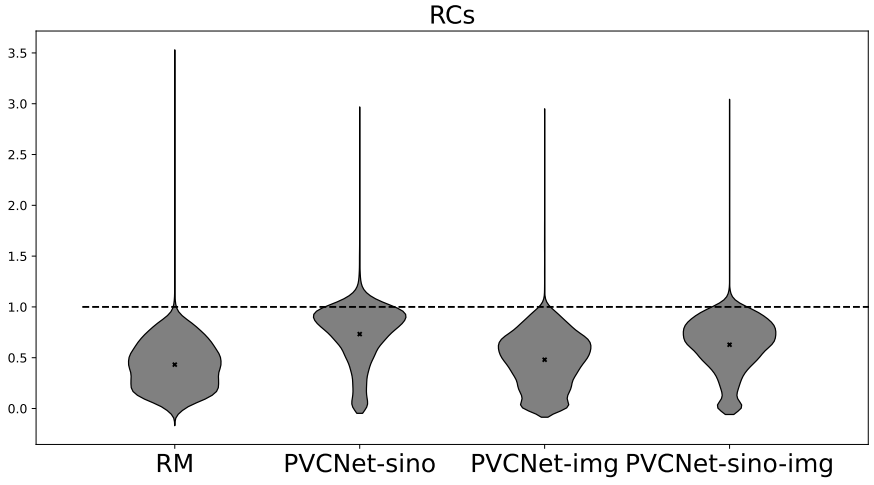
visually better corrected images and profiles as well as RCs closer to one on most ROIs, NMAE on largest volumes such as kidneys, liver and spleen were lower with PVCNet-img. This contradiction can be explained by the over-smoothed texture of PVCNet-img, as visible in Figure 5.3 and 5.4b leading to low NMAE in these regions but over-estimated background activity, which is not the desired behavior of a PVC method.

The Dice score comparison in Figure 5.2c is of importance since PVE affect the activity distribution of tumours and thus the ability to segment them on SPECT images. This segmentation is necessary for all personalized dosimetric methods, as reminded in Section 1.2.5.2. Assuming that the segmentation mask of lesions is available before any PVC is applied, such as in the iY method, is thus unrealistic. As a reminder, none of the proposed model require a segmentation mask at inference time. In this context, we showed that the deep-learning approach leading to best segmentation outcome was PVCNet-sino-img and that it could thus be used to reduce PVE as a pre-segmentation step.

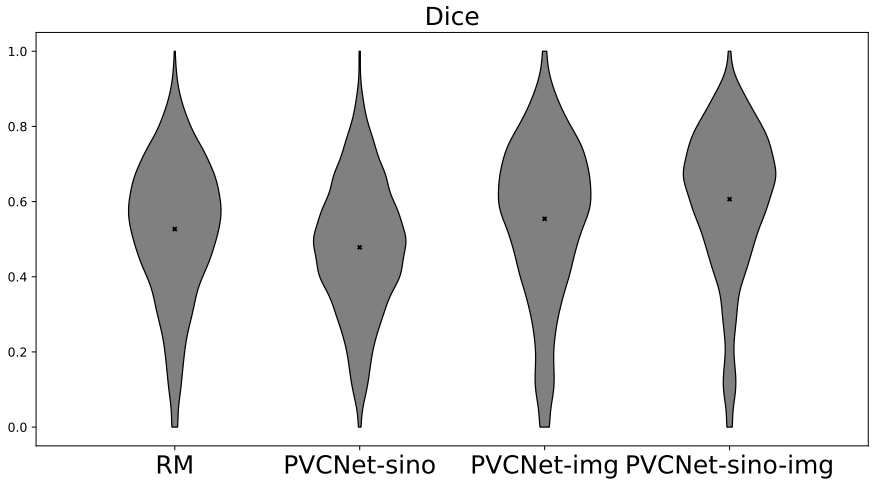
Overall, considering all the studied metrics and datasets, we argue that the most suitable alternative is the dual domain training, benefiting both from raw sinogram domain information — useful for distance-dependent PSF correction, object detection and quantification— and from more direct image-to-image correction — leading to better noise properties and global error reduction. However, as discussed, each method have their weaknesses. For example, lesion correction with PVCNet-sino-img is not satisfying since some lesions still may be discarded by the model. For this reason, the proposed deep-learning based PVC methods are still not ready to be used in clinical settings with a high enough level of confidence. Still, we believe some improvements of the dual domain approach can be imagined. We detail these perspectives in the next Section.



(A)



(B)



(C)

FIGURE 5.2 – Violin plots comparison of RM, PVCNet-sino, PVCNet-img and PVCNet-sino-img on the 1,000 sample test dataset in terms of (A) Normalized Root Mean Squared Error (NRMSE), (B) Recovery Coefficients (RC) on lesions and (C) Dice coefficient after lesions segmentation with the threshold-based method.

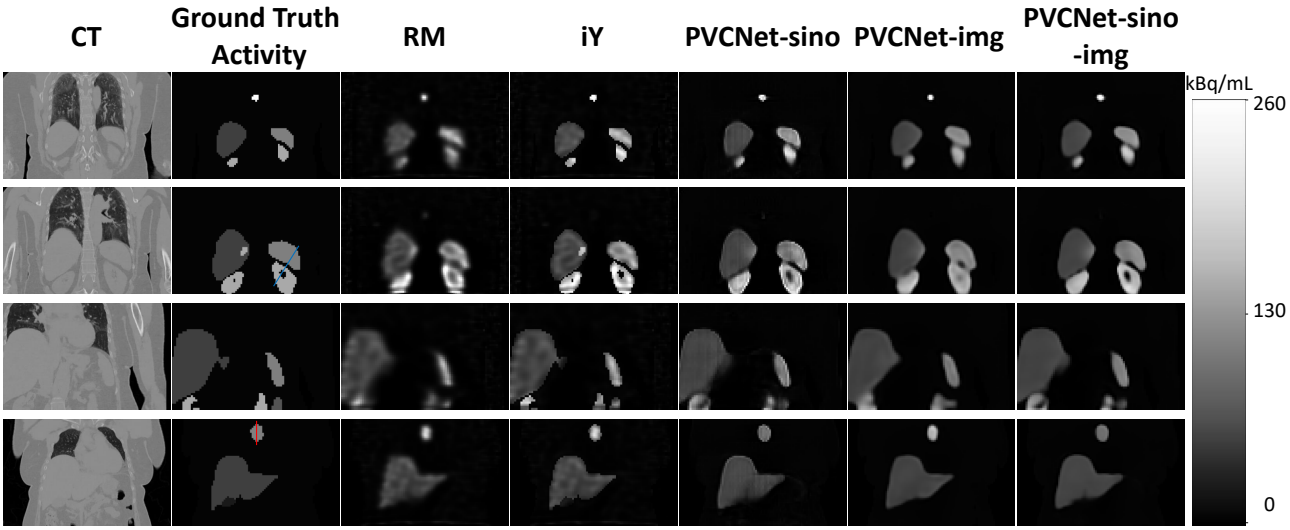


FIGURE 5.3 – Visual comparison of the true activity distribution and results of the five PVC methods on a patient MC simulation. Images reconstructed with Resolution Modeling (RM), iterative Yang method (iY) and with our proposed deep-learning based PVC methods PVCNet-sino, PVCNet-img and PVCNet-sino-img are shown.

	body	gallbladder	kidney left	kidney right	liver	spleen	lesion 1	lesion 2	lesion 3	lesion 4
RM	1.13 ± 1.68	1.08 ± 0.56	0.84 ± 0.51	0.86 ± 0.62	**0.98** ± 0.27	0.88 ± 0.34	0.59 ± 0.35	0.78 ± 0.13	0.56 ± 0.19	0.80 ± 0.37
iY	0.90 ± 0.60	0.82 ± 0.22	1.07 ± 0.53	1.06 ± 0.67	1.06 ± 0.20	1.06 ± 0.26	**0.97** ± 0.43	**1.06** ± 0.09	**1.15** ± 0.21	1.10 ± 0.28
PVCNet-sino	1.05 ± 1.45	1.10 ± 0.67	0.94 ± 0.36	0.95 ± 0.33	1.02 ± 0.22	0.96 ± 0.30	0.58 ± 0.25	0.66 ± 0.08	0.80 ± 0.24	0.90 ± 0.19
PVCNet-img	1.08 ± 1.91	1.23 ± 0.92	0.87 ± 0.16	0.89 ± 0.16	1.03 ± 0.18	0.91 ± 0.21	0.54 ± 0.34	0.57 ± 0.05	0.75 ± 0.36	**0.97** ± 0.44
PVCNet-sino-img	**1.02** ± 1.59	**0.96** ± 0.79	**0.96** ± 0.30	**0.95** ± 0.29	1.04 ± 0.18	**0.97** ± 0.23	0.63 ± 0.27	0.61 ± 0.05	0.84 ± 0.22	0.87 ± 0.13

TABLE 5.2 – Recovery Coefficients (RC) with ± standard deviation in 10 different structures for the five compared methods: Resolution Modeling (RM), iterative Yang method (iY) and our proposed deep-learning based PVC methods PVCNet-sino, PVCNet-img and PVCNet-sino-img. Best RC values are highlighted in bold.

5.7 Perspectives

Even though visual and quantitative improvements were obtained in terms of PVE reduction by the proposed PVCNet-sino-img method, some concern remain about its generalizability to real world data, its data consistency preservation properties and ability to detect and correct small structures. To address these issues,

	NRMSE	NMAE	VAA	Background std
RM	1.26	0.42	7.55%	1.86
iY	1.09	0.31	8.55%	1.85
PVCNet-sino	0.90	0.26	21.24%	1.11
PVCNet-img	0.88	0.22	26.30 %	**0.37**
PVCNet-sino-img	**0.85**	**0.21**	**28.33 %**	0.39

TABLE 5.3 – NRMSE, NMAE and VAA for the five compared PVC. Best values are highlighted in bold.

	body	gallbladder	kidney left	kidney right	liver	spleen	lesion 1	lesion 2	lesion 3	lesion 4
RM	0.68 ± 1.54	0.48 ± 0.31	0.40 ± 0.35	0.48 ± 0.41	0.21 ± 0.17	0.30 ± 0.19	0.48 ± 0.24	0.22 ± 0.13	0.44 ± 0.19	0.37 ± 0.20
iY	0.44 ± 0.41	**0.22** ± 0.17	0.36 ± 0.39	0.46 ± 0.49	0.16 ± 0.13	0.22 ± 0.15	**0.37** ± 0.23	**0.09** ± 0.06	0.23 ± 0.12	0.23 ± 0.19
PVCNet-sino	0.40 ± 1.4	0.58 ± 0.34	0.28 ± 0.24	0.25 ± 0.22	0.15 ± 0.16	0.22 ± 0.21	0.43 ± 0.23	0.34 ± 0.08	0.23 ± 0.21	0.16 ± 0.14
PVCNet-img	0.37 ± 1.87	0.88 ± 0.37	**0.15** ± 0.15	**0.14** ± 0.13	**0.11** ± 0.14	**0.14** ± 0.18	0.47 ± 0.33	0.43 ± 0.05	0.36 ± 0.25	0.38 ± 0.23
PVCNet-sino-img	**0.32** ± 1.55	0.70 ± 0.37	0.22 ± 0.22	0.20 ± 0.22	0.12 ± 0.14	**0.14** ± 0.18	**0.37** ± 0.26	0.39 ± 0.05	**0.21** ± 0.18	**0.14** ± 0.13

TABLE 5.4 – Normalized Mean Absolute Errors (NMAE) with ± standard deviation in 10 different structures for the five compared methods: Resolution Modeling (RM), iterative Yang method (iY) and our proposed deep-learning based PVC methods PVCNet-sino, PVCNet-img and PVCNet-sino-img. Best NMAE values are highlighted in bold.

we mention in this Section some possible improvements of the PVCNet-sino-img method.

5.7.1 Improvements of PVCNet-sino-img

Dual Domain optimization during training

In the dual domain approach proposed in Section 5.2, sinogram and image domain training were performed as two distinct steps: once $h_{\text{PVCNet-sino}}$ was trained to correct synthetic sinograms, an additional network $h_{\text{PVCNet-sino-img}}$ was trained using both the image reconstructed with PSF modeling (RM) and the image reconstructed with projections corrected with $h_{\text{PVCNet-sino}}$. A logical next step is to train these two neural networks jointly, *i.e* to perform the following operations at each training iteration:

1. Apply the $h_{\text{PVCNet-sino}}$ network to compute the corrected sinogram $\widehat{P_{\text{noPVE}}}$ according to Equation (4.7).
2. Compute the sinogram-domain loss $\mathcal{L}_{\text{sino}}$ (Equation (4.8)).

3. Reconstruct the image $I_{\text{PVCNet-sino}}$ with the estimation of the corrected sinogram $\widehat{P_{\text{noPVE}}}$.
4. Apply the $h_{\text{PVCNet-sino-img}}$ network to compute the estimation of the corrected image $\widehat{I}_{\text{PVCNet-sino-img}}$ according Equation (5.5) (using the image $I_{\text{PVCNet-sino}}$ reconstructed at the previous step as input).
5. Compute the image-domain loss $\mathcal{L}_{\text{sino-img}}$ (Equation (5.6)).
6. Update the weights of $h_{\text{PVCNet-sino}}$ and $h_{\text{PVCNet-sino-img}}$ to minimize the double domain loss: $\mathcal{L} = \lambda_{\text{sino}}\mathcal{L}_{\text{sino}} + \lambda_{\text{sino-img}}\mathcal{L}_{\text{sino-img}}$.

The main difficulty of this double domain training pipeline is that a reconstruction step (step 3) is embedded in the optimization process, requiring it to be differentiable with PyTorch. This reconstruction could either be performed with a simple filtered backprojection or with OSEM iterations (at the cost of increased memory requirement). The differentiability is not a theoretical issue but a numerical implementation one, since forward and backprojectors implemented in most reconstruction softwares (RTK [185], PyTomography [176], STIR [227], CASToR [152]) are not differentiable. To our knowledge, only the MIRTorch [129] library provides PyTorch-differentiable SPECT projectors, with PSF modeling.

Data consistency fine-tuning

One concern that remains with many deep learning approaches in medical imaging is the eventuality of neural network hallucinations. There is always a risk that a trained model, when applied for inference on a new dataset interprets noise as signal and enhances structures that are not present in the measured sinogram. Even though we did not observe such behavior neither in the test dataset nor in simulated or experimental data, the risk remains. We put emphasis in the normalization/-denormalization process, both in the sinogram domain and the image domain to conserve number of counts between the input and the estimated output (see Equations (4.6) and (5.3)) but this forced total activity conservation may not be sufficient to conserve consistency with measured data. Inspired by self-supervised methods proposed for PET image reconstruction by [182], we could imagine to fine-tune the pre-trained PVCNet-sino-img model such that the output of the network is consistent with a given measured data. For a given measured sinogram p, the fine-tuning procedure would read as follows for iterations $0 \leq k \leq N$:

1. Apply the $h_{\text{PVCNet-sino-img}}$ model to the usual inputs:

$$\widehat{I}^{(k)}_{\text{PVCNet-sino-img}} = \text{ReLu}\left(h_{\text{PVCNet-sino-img}}\left(I^{\text{normalized}}_{\text{att}}, I^{\text{normalized}}_{\text{rec10}}, I^{\text{normalized}}_{\text{PVCNet-sino}}\right) + I^{\text{normalized}}_{\text{PVCNet-sino}}\right) \tag{5.8}$$

2. Forward-project the current image estimation with CT-derived attenuation and PSF modeling:

$$\widehat{p}^{(k)} = \mathfrak{F}_{\text{RM}}\left(\widehat{I}^{(k)}_{\text{PVCNet-sino-img}}\right) \tag{5.9}$$

3. Compute a Poisson Loss (negative log likelihood loss, see Equation (1.16)) between projected and measured projections:

$$\mathcal{L} = \widehat{p}^{(k)} - p \log\left(\widehat{p}^{(k)}\right) \tag{5.10}$$

4. Updates weights of $h_{\text{PVCNet-sino-img}}$ to minimize $\mathcal{L}$.

We implemented this fine-tuning algorithm but results will not be exposed here since we did not achieve satisfying results.

Single-Patient training

The proposed deep learning based PVC method aimed at reaching generalizability, both in terms of possible patient anatomy distribution and activity distribution. However, the same data generation and training pipelines could be imagined but for a single patient:

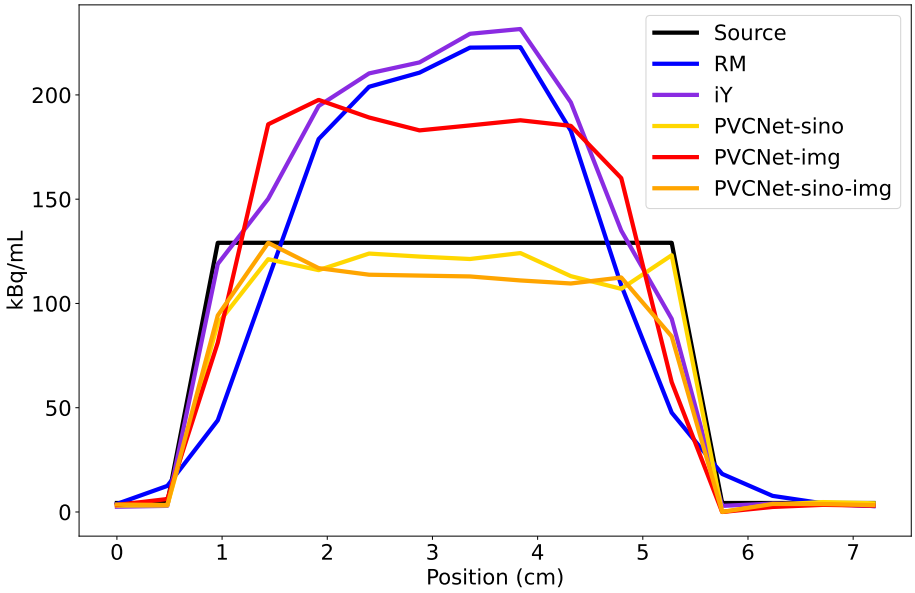
1. The SPECT/CT data of a patient is acquired with the Siemens-intevo MEGP system.
2. The patient anatomy is segmented using the CT image. Only organs where physiological fixation is expected need to be segmented (for example those in Table 4.1).
3. A simulated training dataset is generated specifically for this patient's anatomy, but with variable activity distribution, similarly as in Section 4.2.
4. One of the PVCNet model (PVCNet-sino, PVCNet-img or PVCNet-sino-img) can be trained or fine-tuned on this specific dataset.
5. Once trained this model can be applied to corrected the measured patient's data.

The expected gain of this training strategy is the specificity of the network with respect to the patient's anatomy, which have great influence on PVEs. In this setting generalizability to other patients is no longer seeked but an increased accuracy on the patient on which it is trained (or fine-tuned) can be expected. The main drawback would be the computational over-head due to the necessity to both generate a dataset and train a neural network for each new patient data, which may be unrealistic to apply in clinic. However, if the fine-tuning allows to reach a much better PVE reduction than the other methods with a higher level of confidence, it can be a suitable alternative. The other limitation is that this approach would rely on a patient ROI segmentation mask, while the PVCNet-methods did not require any segmentation at inference. However, only main ROIs masks would be needed, not lesions (artificially added in the training dataset), and could thus be obtained with an automatic segmentation software (such as TotalSegmentator), based on the CT image.

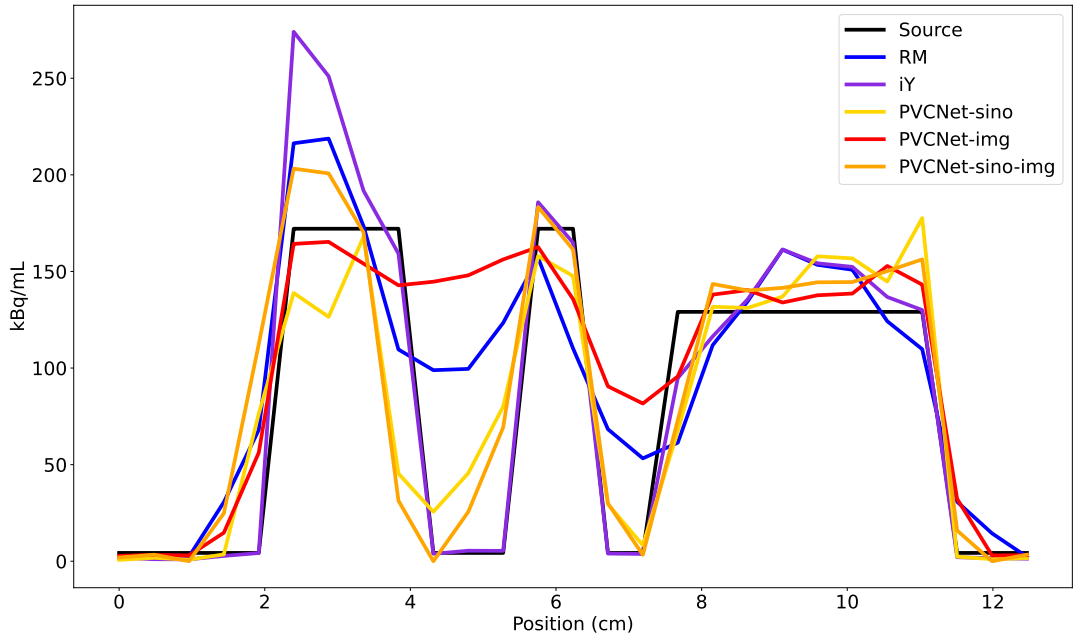
5.8 Conclusion

In this Chapter, we developed two additional deep learning based PVC models trained in the image domain with analytically generated data, either with or without

the output of PVCNet-sino (Chapter 4) as additional input information. Compared to sinogram-domain learning only, image domain learning resulted in better noise and artifact reduction on corrected SPECT images. We evaluated and compared all methods on a test dataset and a patient MC simulation. Overall, the dual domain training model (PVCNet-sino-img) allowed better ROI quantification, global error reduction and segmentation capability than the other explored methods. These findings suggest that mixing both domains for training is the most suitable approach for SPECT-PVC.



(A)



(B)

FIGURE 5.4 – Two profiles drawn on the true activity and reconstructed images for the patient MC simulation. (A) Profiles of lesion 4 (in red in Figure 5.3). (B) Profiles of the left kidney and spleen (in blue in Figure 5.3).

Conclusion

This final Chapter aims at recapitulating the main contributions presented in this thesis, their limitations as well as methodological and long-term perspectives for future works.

Contributions

QMC

Monte Carlo Simulations are known for their high computational cost and slowness to converge. This is particularly true in the context of nuclear imaging and dosimetric purposes. The slow $O\left(\frac{1}{\sqrt{N}}\right)$ convergence rate of MC simulations holds partly because of the uncorrelated nature of random numbers sampled. In Chapter 2 (Section 2.1), we investigated the potential interest of using correlated numbers generators, known to produce low-discrepancy sequences (the Sobol sequence), instead of conventional (uncorrelated) pseudo-random numbers generators. The use of a Low Discrepency (LD) sampler required to define *dimensions* adapted to particle tracking for medical physics to take into account possible multiple interactions and rejection loops. Results in vacuum showed encouraging results with up to 40% of error reduction, increasing with the number of primary particles sampled. In this very simple context, the number of QMC dimensions (5) was matching the total number of required dimensions, since no interaction could occur. Adding a waterbox instead of vacuum drastically increased the number of random numbers and dimensions needed, relegating the number of QMC dimensions to a minority. In this context and for dosimetry calculation in a more realistic setting, QMC simulations showed negligible improvement in terms of convergence rate, variance reduction and efficiency. Increasing the number of dimensions sampled with the quasi-random sampler does not appear as being a suitable solution since LD samplers tend to behave badly in high dimensions, showing irregular patterns and decreased convergence rate.

However, as outlined the discussion, further works should focus on highlighting which dimensions influence the most the output of a specific simulation in order to sample those dimensions with LD sequences.

Conditional GAN for SPECT simulation improvement

Also in Chapter 2 (Section 2.2), we exposed a method to replace particle tracking in a patient by a conditional GAN, allowing to drastically reduce the computational time of SPECT simulations. Although I was not the primary contributor to this work, I actively participated and proposed an improved implementation based on PyTorch. This implementation efficiently utilizes GPU memory, reducing copy from CPU

to GPU, resulting in a threefold improvement in computation time. This efficient implementation of the conGAN+nn-ARF combination, allowed to replace both sides of the simulation (in the patient / in the detector) by two successive neural networks. In total, this allowed to increase by almost a factor $\times 100$ the number of particles tracked per second.

However, one limitation that remains is the necessity to train a condGAN for each new patient, which includes the necessity to first generate a phase space for training. A possible improvement of the method could be to make the GAN also conditional with respect to the patient's CT.

Sinogram based PV mitigation

In Chapter 4, we proposed a deep learning PVC method for SPECT PVE correction. Since most DL-based PVC methods proposed in SPECT operate in the image domain, while the PVE formation process is inherently related to the projection domain because of collimator blurring, we investigated to possibility to train a DL model to correct PVE on projections, before reconstruction. For this purpose, we generated a large training dataset by forward-projecting 10,000 activity images, with and without collimator blurring. The training dataset was used to train a DL model $h_{\text{PVCNet-sino}}$ made of two successive 3D U-nets (one for denoising and one for PVC) in a supervised setting. To guide the optimization process, our proposed *PVCNet-sino* method was taking as input a sinogram denoted $P_{\text{rec10,fp}}$, obtained after forward-projection (without RM) of the previously reconstructed volume (with OSEM, RM, 10 iterations and 8 subsets). We showed in an ablation study that using this sinogram as an additional input helped to reduce the error and acted as a first estimation of the corrected sinogram. Overall, we showed that sinogram-based learning on a synthetic dataset was possible to reduce PV effect compared to RM and could even compete with established methods, such as the iterative Young method, without requiring any ROI segmentation mask at inference. However, reconstruction after correction on sinograms resulted in artifacts, and difficulty to correct small structures.

Image or sinogram domain training

In Chapter 5, the PVCNet-sino correction method was compared to two other DL-PVC models, both trained directly in the image domain: PVCNet-img, taking the attenuation map and the image reconstructed with RM as inputs and PVCNet-sino-img taking also the image reconstructed after correction with PVCNet-sino as additional input. Overall, these two image-based models achieved better denoising properties and error reduction than the image reconstructed with PVCNet-sino. However, compared to each other, the model using the output of the model previously trained on sinograms showed better spill-over correction and activity estimation in ROIs. This PVCNet-sino-img model, using both domain information also led to better segmentation capabilities. Still, recovery coefficients achieved by all the PVCNet methods on small or overlapping regions were not convincing enough. This concern is of importance, since PVEs are especially problematic for small or difficult to segment ROIs.

Overall limitations

Training Dataset

The quality of the training dataset has a strong impact on the quality of the expected network's output. The synthetic dataset we proposed was generated using only less than 20 patient's CT scans. We are conscious that this number is quite low and that our trained PVCNets could benefit from a larger dataset, with more variability. Moreover, the realism of the created sources is also debatable, since lesion shapes were much simpler than real life shapes. Finally, simulating the detection process by a simple forward-projection, even with attenuation, resolution modeling and Poisson noise, may also seem simplistic.

Evaluation dataset

A bottleneck of PVC algorithm evaluation consists in choosing an appropriate evaluation dataset. Even though we put a lot of effort to evaluate our proposed deep learning algorithms on several dataset types (large synthetic dataset, MC simulations, phantom experiments and real patient data), we were limited by the fact that reference PV-free images from clinical settings (true SPECT system and patient's activity distribution) do not exist.

Deep Learning "black box"

A frequent criticism of deep learning methods, specially for medical imaging applications, comes from their black-box aspect. A trained neural network will output a corrected image, but, at inference time, there is no mean to estimate how *confident* the model is in each voxels / region of the image. At present, with the proposed method, there is no guarantee on the generalizability of the proposed models to new and possibly very different patients anatomy or activity distributions.

Perspectives

Dual domain training

As mentioned in Chapter 5, the good results achieved by the dual-domain training approach suggests that a possible approach for future works could be to train both models $h_{\text{PVCNet-sino}}$ and $h_{\text{PVCNet-sino-img}}$ jointly, and not successively as it was the case in this Chapter. This way the reconstruction step would be included in the optimization process.

Single Patient

Fine-tuning on a dataset generated for a specific patient could be a potential way to increase the model's performance and trustability of the output. However, as

discussed already, this perspective may be too time consuming for a clinical routine implementation.

Data consistency

To ensure that the trained model is not hallucinating lesions, organ shapes or just misinterpreting the input data, guarantees in terms of *data-consistency* (DCC) preservation could be introduced. We already described and discussed in Section 5.7 a self-supervised fine-tuning approach to ensure the corrected image is still consistent with the measures. Another option is to use, during training in the sinogram domain (Section 4.3.1), an additional loss based on exponential data consistency conditions (eDCCs). eDCCs were proposed in [188, 25] to detect patient motion on SPECT projections and could be adapted as a training loss.

Unrolled methods

Unrolled methods can also be an alternative in order to combine the 1) model-based reconstruction framework provided by the MLEM algorithm which allows to model the physical process of image acquisition and 2) data-driven correction in an iterative scheme. To our knowledge, these methods were employed by [129] for image regularization but not specifically for PVC. The challenge induced by such methods is the necessity to track gradients across forward and backprojection steps of the OSEM algorithm, which can be computationally intensive.

Increasing the training dataset realism

Finally, the realism of the training dataset could be enhanced by using MC simulated data. However, such dataset would be either limited in size or computationnaly too expensive to obtain. One possibility would be to use the conGAN+nn-ARF combination proposed in Section 2.2.5, since the computational speed-up induced by this method is expected to be close to $\times 100$ compared to a full MC simulation with a comparable realism.

Long term perspectives

In this thesis we developed and tested some image correction techniques for SPECT image reconstruction. We believe that this area of research can benefit from recent advances in deep learning and compete with conventional model-based methods. However, in order for these methods to be applied in clinical routines, they still need to achieve stronger guarantees in terms of generalizability and explainability. For example, higher trustability can be achieved by adding physical constraints or data consistency conditions in the training process. When these conditions will be met, the benefits of deep learning methods, including image quality and accuracy, will considerably improve cancer treatments.

List of publications

— **Théo Kaprélian**, Bastien Doignies, David Coeurjolly, Victor Ostromoukhov, Nicolas Bonneel, David Sarrut. "Investigating Quasi-Monte Carlo (QMC) methods for Geant4/Gate simulations in medical physics". *IV Geant4 International User Conference at the physics-medicine-biology frontier*, Oct 2022, Napoli, Italy.

— **Théo Kaprélian**, Ane Etxebeste, David Sarrut. "Towards Deep-Learning Partial Volume Correction for SPECT". *17th International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*, Jul 2023, New York, United States.

— **Théo Kaprélian**, Ane Etxebeste, David Sarrut. "Partial Volume Correction on 177 Lu-SPECT sinogram with Deep Learning trained on synthetic data". *2024 IEEE Nuclear Science Symposium, Medical Imaging Conference, and Room-Temperature Semiconductor Detectors Symposium*, Oct 2024, Tampa, Florida, USA, United States.

— Albert Saporta, Ane Etxebeste, **Théo Kaprélian**, Jean Michel Létang et David Sarrut. "Modeling families of particle distributions with conditional GAN for Monte Carlo SPECT simulations". *Physics in Medicine & Biology 67.23*, 2022, p. 234001.

— David Sarrut, Ane Etxebeste, **Théo Kaprélian**, Albert Saporta et Jean Michel Létang. "Annihilation photon GAN source model for PET Monte Carlo simulation". *Physics in Medicine and Biology 68.13*, 2023, p. 135018.

Synthèse

Introduction

Au cours de l'année 2022, 20 millions de cancers ont été diagnostiqués dans le monde et presque 10 millions de personnes en sont décédées [75]. D'après l'OMS, le nombre de cas risque d'augmenter de 75% d'ici 2050. Le cancer désigne un ensemble de pathologies caractérisées par la multiplication incontrôlées de cellules formant des tumeurs qui peuvent s'étendre aux tissus voisins, voir, dans le cas de tumeurs métastatiques, dans d'autres endroits du corps que leur lieu de développement initial. Bien sûr, plusieurs traitements existent et permettent d'éliminer ces tumeurs cancéreuses et augmenter l'espérance de vie d'un patient. Les plus connus de ces traitements incluent la radiothérapie, la chirurgie ou l'immunothérapie. La Radiothérapie Interne Vectorisée (RIV) fait partie des traitements envisageables lorsque les tumeurs sont inopérables et métastatiques.

La RIV consiste à injecter par intraveineuse au patient un radiopharmaceutique composé d'un vecteur (sélectionné par sa capacité à se fixer aux récepteurs des cellules tumorales) et d'un radioisotope dont les rayonnements vont endommager les cellules tumorales. Un exemple de radioisotope utilisé pour la RIV est le ^{177}Lu, qui a pour particularité de ne pas seulement émettre des rayons β^- utiles pour la thérapie, mais également des rayons γ d'énergie moyenne dont la distribution peux être estimée par Tomographie par Emission Mono-Photonique (TEMP, ou tomoscintigraphie). Grâce à l'image TEMP, associée à un scanner CT, l'efficacité de la RIV peut être évaluée entre chaque injections. Pour cela, la dose absorbée (en Gy) dans chaque region, et en particulier dans les tumeurs et Organes à Risques (OAR) est calculée à partir de l'image TEMP. D'autre part, même si la plupart des traitements RIV au ^{177}Lu sont standardisés (c'est le cas du ^{177}Lu-PSMA et ^{177}Lu-DOTATATE), de nombreux travaux récents ont conclu qu'une personnalisation du traitement pouvait être bénéfique [47].

Que ce soit pour la recherche et le développement de nouveaux systèmes d'acquisition TEMP ou pour de la dosimétrie, les simulations Monte Carlo (MC) sont un outil important de la médecine nucléaire. Ces simulations consistent à simuler un grand nombre de particules de manière aléatoire, ainsi que leur probabilité d'interaction. Cependant, ces simulations sont très longues et coûteuses en temps de calcul. Une des raisons de cette lenteur vient de la manière dont les nombres aléatoires sont échantillonnés, c'est-à-dire avec des générateurs de nombre pseudo-aléatoires, qui produisent des séquences de nombres "indépendants" les uns des autres, ce qui limite la vitesse de convergence à $\frac{1}{\sqrt{N}}$.

La première contribution de cette thèse a été d'investiguer le potentiel des échantillonneurs *quasi*-aléatoires, dans le cadres des simulations MC pour la TEMP ou la dosimétrie. Ces échantillonneurs quasi-aléatoires sont conçus pour produire des

séquences de nombres dont l'uniformité est garantie (à l'inverse des générateurs pseudo-aléatoire) et ainsi permettre de réduire l'erreur d'intégration des méthodes Monte Carlo.

Dans le contexte de la RIV, l'aspect quantitatif des images TEMP joue un rôle primordial. Cependant, la résolution spatiale des images TEMP est toujours très limitée (plus de 10 mm) et bien moins bonne que celle d'autres types d'imagerie (TEP, CT, IRM). Cette mauvaise résolution spatiale des images TEMP est due à la réponse du collimateur et du détecteur (la PSF) des caméras gamma utilisées pour collecter les projections. Après reconstruction, l'effet de la PSF sur les images reconstruites se caractérise par les Effets de Volume Partiels (EVP), c'est-à-dire une sous-estimation de l'activité dans les volumes de taille inférieur à la résolution du système TEMP ($2\times$ FWHM, la Full-Width at Half Maximum).

La deuxième contribution de cette thèse a été de proposer une méthode de correction des EVPs sur les projections, avant reconstruction, grâce à un modèle d'apprentissage profond, entraîné sur des données synthétiques. Cette méthode a été évaluée sur différentes données (synthétiques, simulées, acquisition de fantômes et donnée de patient). D'autre part, cette approche a été comparée à une méthode d'entraînement dans le domaine image et une méthode mêlant les deux domaines.

Contexte

Contexte Clinique

Les thérapies RIV ^{177}Lu-DOTATATE et ^{177}Lu-PSMA consistent à injecter 7.4 GBq du radiopharmaceutique, 4 fois de suite avec entre 6 et 8 semaines d'écart entre chaque injections. L'intérêt des thérapies RIV au ^{177}Lu a été démontré par deux études: NETTER-1 [226] pour le ^{177}Lu-DOTATATE et VISION [207] pour le ^{177}Lu-PSMA. Ces deux études ont montré une augmentation de la durée de vie des patients, sans progression de la maladie et ont abouti à leur autorisation par les agences de santé.

Cependant, le problème principal lié à ces thérapies est que le vecteur ne se lie généralement pas uniquement aux cellules tumorales, mais également dans certains OAR, comme les reins, le foie, la rate ou les glandes salivaires et lacrymales. Pour évaluer l'efficacité du traitement et assurer la sécurité du patient, la dose (Équation (1.1)) est calculée dans chaque OAR et tumeurs, et ce, entre chaque injections. Dans le cadre des RIV personnalisées, l'objectif est donc de minimiser la dose absorbée par les OAR et maximiser la dose absorbée par les tumeurs. Pour cela, sachant que le ^{177}Lu émet des rayons gamma, dont la distribution peut donc être estimée grâce à la TEMP, et que la demi-vie du ^{177}Lu est de 6 jours, plusieurs acquisitions TEMP peuvent donc être réalisées entre chaque injections.

Le déroulement du traitement est représenté sur la Figure 1.1: entre chaque injection, entre une et trois acquisitions TEMP sont réalisée, puis les Régions d'intérêts (ROIs) sont segmentées pour calculer les Courbes Temps-Activité (TAC) pour enfin calculer la dose absorbée par chaque ROI. Dans ce cadre, la qualité des images TEMP est donc cruciale pour permettre au mieux l'évaluation du traitement et son optimisation.

Acquisition et reconstruction en TEMP

Une acquisition TEMP consiste, suite à l'injection du radiotraceur, à acquérir plusieurs (entre 60 et 120) projections 2D du patient, grâce à une gamma caméra qui tourne autour du patient. Une fois les projections acquises, elles sont utilisées par un algorithme de reconstruction tomographique pour reconstruire la distribution 3D de l'activité dans le patient. Le déroulement d'une acquisition TEMP est illustré sur la Figure 1.2.

Une gamma caméra typique est constituée de plusieurs éléments illustrés sur la Figure 1.3:

— Un collimateur, généralement en plomb, qui sert à ne sélectionner que les photons arrivant à la surface du détecteur avec un angle orthogonal à celui-ci. Le collimateur est indispensable pour la reconstruction de l'image afin de connaître la direction depuis laquelle les rayons gamma ont été émis.

— Le scintillateur, généralement en NaI(Tl), qui converti les rayons gamma en lumière visible.

— Des tubes photomultiplicateurs suivis d'un convertisseur analogique-numérique permettant de convertir le signal lumineux en faisceau d'électrons, puis en impulsion électrique et enfin en signal numérique.

Une fois que toutes les projections ont été acquises, la reconstruction tomographique permet de reconstruire la distribution de l'activité dans le patient. Les premiers algorithme proposés étaient basés sur l'opérateur adjoint de la transformée de Radon (Équation (1.5)). C'est le cas de la rétro-projection filtrée (Équation (1.8)). Cependant, cet algorithme a surtout été proposé pour la reconstruction d'images CT et, bien que rapide et simple, il est assez peu adapté à l'imagerie TEMP. A l'inverse, l'algorithme OSEM (Équation (1.25)) permet de prendre en compte l'aspect poissonien du bruit enregistré dans les projections: sachant le modèle de bruit et le projections mesurées, l'algorithme permet de maximiser la vraisemblance (Équation (1.16)) de l'image reconstruite, et ce de manière itérative.

D'autre part, l'algorithme OSEM permet de prendre en compte l'atténuation lors des projection/rétro-projections réalisées à chaque itération, et donc de la corriger.

La correction des rayonnements diffusés par interaction Compton en TEMP se fait généralement en acquérant plusieurs fenêtre d'énergies, proches de l'énergie du pic d'émission, pour ainsi le corriger avant la reconstruction (Équations (1.31) et (1.32)).

La correction de la résolution spatiale peut également se faire assez simplement en modélisant la PSF (préalablement estimée expérimentalement ou par simulation) lors des projection/rétro-projections (Équation (1.34)). La PSF est généralement modélisée par une gaussienne, dont la FWHM dépend linéairement de la distance entre la source et le détecteur (Équation (1.33)). Cette méthode, appelée RM pour *Resolution Modeling*, est très souvent utilisée car elle permet d'améliorer la résolution spatiale facilement mais n'est généralement pas suffisante pour corriger les EVPs. De plus, elle donne souvent lieu à une augmentation du bruit dans l'image lorsque le nombre d'itération augmente et à des artefacts de Gibbs.

Simulations Monte Carlo

L'objectif des simulations Monte Carlo (MC) est d'estimer une quantité inconnue en simulant des réalisations aléatoires des événements possibles qui y contribuent, en fonction de leur distribution probabilité qui est supposée connue. L'exemple typique consiste à évaluer l'intégrale d'une fonction (Équation (1.36)) par l'estimation donnée dans l'Équation (1.37) qui consiste à évaluer la fonction en un nombre N de points aléatoirement échantillonnés et de calculer leur moyenne. Cet estimateur de l'intégrale est non-biaisée et converge avec une vitesse en $O(\frac{1}{\sqrt{N}})$.

Une simulation MC peut toujours se ramener au calcul d'une intégrale comme l'Équation (1.36) mais dans le cadre des simulations en physique médicale, l'intégrande n'a pas d'expression explicite, étant donné la complexité des simulations. En effet, le nombre de type particules considéré peut être élevé (électrons, photons, protons...), de même pour le nombre d'interactions possibles (Compton, effet photo-électrique, production de pair, Bremsstrahlung...) et le milieu traversé peut être complexe (patient voxélisé, caméra gamma). Typiquement, une simulation MC consiste à:

1. Définir les paramètres de la simulation: géométrie, particules primaires, processus physiques considérés ainsi que la quantité d'intérêt (dose, projections...)
2. Echantilloner de manière aléatoire les particules primaires, crées par exemple par désintégration radioactive. Ces particules peuvent être des rayons gamma, des électrons, des protons...
3. Simuler aléatoirement leurs interactions avec la matière, jusqu'à leur absorption ou sortie du volume considéré.

Plusieurs types de codes de simulation MC pour le suivi des particules existent: les codes à caractère général comme Geant4 [1] qui fournissent un cadre général (définition d'un grand nombre de particules, de processus physiques, de géométries, de matériaux...) et qui peuvent donc être utilisés dans de nombreux domaines; et les codes à caractères spécifique, basés sur un code général comme Geant4, qui permettent de réaliser des simulations dans un domaine plus restreint. Par exemple, Gate [206] est un code spécifiquement conçu pour les simulations en tomographie par émission (TEP/TEMP), et les thérapies par radiations.

Bien que très utilisée, notamment pour leur réalisme et précisions, les simulations MC sont très lentes à converger. Pour réduire le temps de calcul, des Techniques de Réduction de Variance (TRV) peuvent être utilisées comme la division des particules qui consiste à de-multiplier en plusieurs copies identiques les particules qui atteignent une région d'intérêt, ou la roulette russe, qui consiste à faire l'inverse, c'est à dire "tuer" les particules qui atteignent des régions considérées moins importantes pour le résultat final. Les simulation quasi-aléatoires que nous avons proposées et qui seront décrites dans la prochaine section, s'inscrivent dans le cadre des techniques de réduction de variance.

Simulations Monte Carlo avancée pour la médecine nucléaire

Introduction

Comme indiqué précédemment, la lenteur de convergence des simulations MC est en partie due à la manière dont les nombre aléatoires sont échantillonnés, c'est à dire avec des générateur de nombre pseudo-aléatoires, qui tentent de mimer un "vrai" hasard en produisant des séquences de nombre qui sont décorrélés les uns des autres, c'est à dire un "bruit blanc". L'avantage de ces générateurs est que le résultats obtenu est assuré de converger mais, lorsque l'on regarde la distribution spatiale des nombres aléatoires obtenus (en 2d avec 1024 points sur la Figure 2.2), on se rend compte que l'on a aucune garantie sur homogénéité spatiale des points obtenus. A l'inverse, les générateurs de nombres *quasi*-aléatoires, comme Sobol ou Halton, s'assurent de l'uniformité des séquences produites (Figure 2.2).

La mesure de discrépance, fournie à l'équation (2.3), permet de mesurer l'uniformité d'un ensemble de points en dimension d, échantillonnés dans l'espace $[0, 1]^d$. Il peut être montré par l'inégalité de Koksma-Hlawka (Equation (2.4)) que l'erreur d'intégration de l'estimateur MC est majoré par la discrépance. Les suites à faible discrépance sont celles dont la discrépance converge en $O(\frac{(\ln N)^d}{N})$. C'est le cas de la Suite de Sobol, illustrée sur la Figure 2.2 et dont la construction est donnée par l'Équation (2.7). Cette convergence en $O(\frac{(\ln N)^d}{N})$ est bien meilleure que le $O(\frac{1}{\sqrt{N}})$ des méthodes MC classiques, à condition que d ne soit pas trop grand ou que N le soit suffisamment.

QMC dans Gate/Geant4

Dans ce contexte, notre contribution à consisté à utiliser la suite de Sobol dans les simulations Gate. Les suites quasi-aléatoires comme Sobol reposent sur des *dimensions* pour assurer l'uniformité de la séquence générée. Cependant, il n'y a pas de définition explicite de la dimension à laquelle correspond un nombre aléatoire lorsqu'il est généré par Gate ou Geant4. Nous avons donc proposé de définir un dimension par:

— La ligne de code dans laquelle le nombre aléatoire est appelé.

— Le nombre d'interactions réalisées par la particule suivie.

— L'itération de la boucle de réjection (si le nombre aléatoire est tiré à l'intérieur d'une boucle de réjection).

Sachant que le nombre de dimension total doit être connu avant le début de la simulation pour pouvoir utiliser la suite de Sobol, nous avons donc dû définir des nombres maximaux d'interactions et de boucle de réjection qui peuvent être échantillonnés avec la suite quasi-aléatoire. Toutes les dimensions supérieures à ces seuils seront échantillonnées avec un générateur pseudo-aléatoire. La Figure 2.4 résume notre définition des dimensions lorsque seule l'interaction Compton est considérée.

Résultats

Pour évaluer l'intérêt des méthodes QMC dans le cadre des simulations TEMP, nous avons réalisés plusieurs simulations, à chaque fois en comparant l'erreur obtenue avec la simulation MC classique avec celle obtenue avec la simulation utilisant la suite quasi-aléatoire, et ce pour un nombre variable de particules simulées N entre 10^6 et 2×10^7. La simulation MC de référence a été obtenue avec $N = 10^8$.

Dans la première simulation, des rayons gamma de 140.5 keV étaient émis depuis 6 sphères (Figure 2.6), placées dans le vide. Les particules sortant du volume ont été collectées et "rétro-projetées" de manière "idéale" avec la formule donnée par l'Équation (2.12). L'erreur a été calculée en comparant les erreurs quadratiques moyennes entre l'image obtenue par la simulation de référence et l'image obtenue avec/sans QMC avec N gamma primaires. Dans ce cadre, puisque la simulation se déroulait dans le vide, aucune interaction n'était possible entre l'émission et la sortie du volume et donc le nombre de dimension total était égal au nombre de dimension échantillonnées avec la suite de Sobol, c'est-à-dire 5 (3 pour la position initiale et 2 pour la direction). Les résultats obtenus dans cette simulations étaient très encourageants puisque, comme le montre la Figure 2.7, les erreurs obtenues avec la simulation QMC étaient entre 15% et 40% plus faibles qu'avec la simulation MC classique. De plus, ce gain semble augmenter de manière consistante avec N, ce qui indique que le gain est d'autant plus grand pour les simulations où une haute statistique est nécessaire, celles-là même qui sont actuellement très coûteuses.

Cependant, lorsque la même simulation est réalisée dans une boite d'eau et en restreignant les interactions possibles à seulement l'effet Compton avec $n_{\text{Compton}} = 1$ et $n_\gamma = n_{e^-} = 2$, les résultats obtenus ont été bien moins convaincants. En effet, dans ce contexte, la réduction de l'erreur s'est limitée à quelques pourcents tout au plus. Ce gain très minime peut être expliqué par le fait que le gain attendu diminue lorsque la dimension d de la suite de Sobol augmente. Cela est dû aussi au fait que, contrairement à l'expérience dans le vide, beaucoup de nombre aléatoires sont échantillonnés avec du bruit blanc dans ce cette simulation, ce qui réduit le gain introduit par le QMC.

D'autres simulations et comparaisons ont été réalisées dans le cadre d'une simulation pour le calcule de dose dans une volume contenant des matériaux de plusieurs densité (Figure 2.5). Cependant, dans ce cas aussi, la dimension du problème étant relativement grande, les résultats obtenus ont été très faibles. Il a donc été décidé de suspendre les travaux de projet puisque l'avantage introduit par les échantillonneurs à faibles discrépances dans un contexte réaliste, c'est à dire en grande dimension n'étaient pas suffisant. Cependant, nous pensons qu'une direction possible pour la suite serait de déterminer, pour un type de simulation donné, quels dimensions influencent le plus la quantité d'intérêt, pour n'échantillonner que ces dimensions avec des nombres quasi-aléatoires.

Effets de volume partiels (EVP) et méthodes de corrections (PVC)

Caractérisation des EVPs

Les Effets de Volume Partiels (EVP) en TEMP se manifestent par une sous-estimation de l'activité dans les volumes de taille inférieure à $2\times$ la FWHM et par une sur-estimation dans les volumes voisins (spill-out). De même, chaque volume peut être contaminé par l'activité des volumes environnants (spill-in). Ces EVP persistent malgré la correction d'atténuation (Section 1.2.4.1), des rayonnements diffusés (Équations (1.31) et (1.32)) et de la résolution spatiale (Équation (1.34)) lors de la correction. De plus, ils se manifestent d'autant plus dans les volumes de petite taille (Figure 3.2) et se quantifient, entre autre, par le Coefficient de Recouvrement décrit par l'Équation (3.1).

Les conséquences cliniques des EVP peuvent être de différentes nature. D'abord, lorsque l'image TEMP est utilisée pour diagnostiquer ou suivre l'évolution de tumeurs, les EVP altèrent à la fois l'activité moyenne et les contours du volume. Dans le cadre des thérapies RIV, les EVP rendent la segmentation des ROI plus délicate, ce qui par conséquent augmente l'incertitude de la dosimétrie.

Les EVPs sont dus à la résolution spatiale limitée des caméras gamma. La réponse de la caméra peut être divisée en deux effets principaux : la résolution intrinsèque et la réponse du collimateur. La résolution intrinsèque est proportionnelle à $\frac{1}{\sqrt{E_0}}$, où E_0 est l'énergie des gammas entrants dans le scintillateur. La résolution du collimateur peut être exprimée en fonction de la géométrie de celui-ci et de la distance entre la source et le détecteur (Équation (3.2)). Les paramètres de la réponse du détecteur ont été déterminés pour le système TEMP utilisé au Centre Léon Bérard, Lyon, France, qui est l'Intevo Bold commercialisée par Siemens, avec un collimateur d'énergie moyenne (MEGP) et une source de ^{177}Lu (cf Équation (3.11) et Figure 3.8).

Méthodes de corrections (PVC)

Plusieurs méthodes de correction des EVPs ont été proposées par le passé pour inverser l'effet de la PSF sur l'image reconstruite. Le problème inverse que ces méthodes tentent de résoudre est formulé dans l'Équation (3.13).

Le premier type de méthodes sont les méthodes de déconvolution. La méthode RM peut être considérée comme tel mais n'est pas suffisante. Pour limiter les artefacts que cette méthode crée, il est possible de sous-estimer la PSF lors de la reconstruction ou d'appliquer un filtre gaussien à l'image reconstruite, mais cela se fait au prix d'une résolution spatiale dégradée. Des méthodes basées sur la restauration de l'image par filtration ont également été proposées, comme le Filtre de Wiener, rappelé à l'Équation (3.15). Cependant, ces méthodes nécessitent de sélectionner un filtre approprié à chaque image et application [143] et mènent souvent à une augmentation du bruit de l'image [114]. Les méthodes de déconvolution itératives telles que Richardson-Lucy[184, 142] ont été appliqués en TEMP [231] mais nécessitent, pour ne pas que le bruit ne soit trop amplifié, d'appliquer une régularisation qui dégrade également la résolution spatiale.

Les méthodes basées sur une segmentation de l'image sont également nombreuses. La méthode basée sur les RC consiste à pré-calculer les coefficients de recouvrement en fonction de la taille des objets à corriger, puis de les appliquer à l'image reconstruite sur chaque objet segmenté. Bien que cette méthode soit assez fréquemment conseillée [50] et appliquée en clinique, elle se heurte à deux obstacles: le fait que les coefficients de recouvrements soit calculés expérimentalement sur un volume différent de celui de l'image à corriger et le fait que ces RC sont généralement uniquement paramétrés par la taille des objets. En effet, les EVP dépendent de nombreux paramètres comme la position de l'objet, du niveau de bruit, le ratio d'activité entre la ROI et le fond. Plus réalistes, les méthodes GTM (Équation (3.19)) et Labbé (Équation (3.21)) consistent à inverser la matrice contenant les activités moyennes dans chaque région, soit en prenant les contours exacts de la segmentation (méthode GTM), soit floutant les bords de chaque région par la PSF (méthode Labbé). Cependant, ces méthodes ne permettent d'avoir qu'une estimation région-par-région. La méthode MTC (Équation (3.22)) permet de corriger l'image voxel-par-voxel en corrigeant les régions de manière indépendante grâce à une première estimation de l'activité dans chaque région. La méthode de Yang itérative (Équation (3.26)) est une extension itérative de la méthode RBV (Équation (3.24)) qui converge en quelques itérations, permet de corriger assez bien le spill-in et le spill-out et de retrouver la vraie activité dans chaque région [194, 228].

Cependant, ces méthodes reposent sur la segmentation des ROIs qui est particulièrement délicate à obtenir, en particulier pour les lésions dont les contours sur l'image TEMP non-corrigée sont affectés par les EVP. Idéalement la segmentation de l'image ne devrait donc être faite qu'une fois l'image corrigée des EVPs. Pour la segmentation des ROIs sur le CT, la précision des contours peut être affectée par des erreurs de registration entre les deux images.

De manière générale, il y assez peu de consensus pour déterminer quelle méthode de correction des EVP à la fois la plus efficace pour les corriger et applicable en routine clinique. Certaines méthodes basées sur l'intelligence artificielle ont été proposées pour éviter d'avoir besoin d'une segmentation de l'image avant correction [255, 125].

Correction des EVP par apprentissage profond sur les projections

Création du jeu de donnée d'entraînement

Dans cette section, nous détaillons la méthode que nous proposons pour corriger les EVP par apprentissage profond sur les projections. Pour entraîner notre modèle, nous avons utilisé une large base de donnée synthétique dont nous allons détailler la créations dans le prochain paragraphe. Les raisons de ce choix viennent du fait qu'il n'est pas possible d'avoir une image de référence venant d'un jeu de donnée réel car les méthodes existantes de correction des EVP sont toujours imparfaites. D'autre part générer les données par simulation MC n'est pas option réaliste car nous serions soit limité par la taille du jeu de donnée produit, soit par les ressources en terme de temps de calcul nécessaire.

Pour créer un jeu de données synthétique mais réaliste, nous avons utilisé des scanners CT de 14 patient que nous avons segmentés de manière automatique avec le logiciel TotalSegmentator [248]. Nous avons ensuite généré 10,000 cartes d'activité aléatoires de la manière suivante: après avoir sélectionné de manière aléatoire un CT, l'activité dans le fond a été fixée à 1, puis, à chaque ROI a été assigné une ratio (cf le Tableau 4.1), puis entre 1 et 10 lésions de taille/position/ratio aléatoires ont été ajoutées. Enfin, la carte d'activité résultante a été multipliée par une fonction mimant une hétérogénéité réaliste (cf Equation (4.1)) et l'activité totale a été normalisée pour être comprise entre 200 MBq et 1 GBq en considérant le temps par projection égal à 15 secondes.

Pour chacune des carte d'activité générées, et la carte d'atténuation correspondante, nous avons généré des projections de différentes manières. D'abord, en projetant le volume avec le projecteur *Zeng*, sans modéliser la PSF, nous obtenons P_{noPVE}, le sinogramme sans PVE qui servira, pour l'entraînement, de référence. Ensuite, en projetant l'activité avec la PSF mesurée dans la section précédente, nous obtenons P_{PVE}. En appliquant un bruit de poisson à ce sinogramme , on obtient $P_{\text{PVE,noisy}}$. Tous les sinogrammes contiennent 120 projections de taille 128×128 avec une taille de pixel de 4.8×4.8 mm^2. La carte d'attenuation a également été projetée. Enfin, en reconstruisant I_{rec10} à partir des projections $P_{\text{PVE,noisy}}$ avec OSEM, RM, 10 itérations et 8 sous-ensembles, et en projetant cette image reconstruite, on obtient le sinogramme $P_{\text{rec10,fp}}$. Un exemple issu de l'ensemble de donnée généré de cette manière est donné sur la Figure 4.4.

Entraînement

Les Equations (4.7), (4.8) et la Figure 4.5 détaillent l'entraînement du réseau de neurones $h_{\text{PVCNet-sino}}$, composé de deux U-nets en 3d.

Résultats

Une étude par ablation sur un jeu de donnée de test contenant 1,000 données, générées de la même manière que le jeu de donnée d'entrainement mais avec 4 CT de patients differents, a permis de montrer l'intérêt de chaque spécificité de notre méthode, *i.e* les deux U-nets, leur architecture ainsi que l'utilisation de $P_{\text{rec10,fp}}$ en entrée. Pour ce dernier point, le simple fait de considérer ce sinogramme en entrée du réseau à permis d'obtenir une erreur moyenne absolue (NMAE) de 4.91×10^{-2}, contre 5.29×10^{-2} lorsque ce sinogramme n'était pas utilisé lors de l'entraînement.

Le modèle $h_{\text{PVCNet-sino}}$ a ensuite été évalué sur des données expérimentales obtenues sur le fantôme IEC contenant 6 spheres plongées dans une activité de fond avec un ratio de l'ordre de 11:1. Les images reconstruites et corrigées ainsi que les résultats en terme de coefficients de recouvrement sont présentés sur la Figure 4.14. On y voit que l'activité dans les sphères est bien mieux quantifiée qu'avec la méthode RM et que les RC sont assez proches de ceux obtenus avec la méthode de Yang itérative (iY), qui, pour rappel, utilise une segmentation exacte des sphères, contrairement à *PVCNet-sino*. En comparant les images corrigées, il est clair que PVCNet-sino permet d'améliorer la détection et le signal des sphères de taille 13/17/22/28/37 mm avec des RC de respectivement 0.56/0.73/0.90/1.04/1.08,

comparé à 0.24/0.35/0.52/0.66/0.77 avec RM et 0.69/0.79/1.01/1.07/1.11 avec iY. Cependant, la plus petite sphère (10 mm) disparaît complètement avec PVCNet-sino alors qu'elle est largement sur-estimée par iY.

Une simulation MC d'un patient a également été réalisée pour évaluer notre modèle et le comparer aux autres méthodes de correction des EVP. Les méthodes RM, iY et PVCNet-sino ont d'abord été comparées en terme de coefficients de recouvrement: la méthode PVCNet-sino permet d'obtenir la meilleure estimation dans toutes les ROI, hors lésions (fond, vésicule biliaire, reins, foie et rate). Cependant, sur les 4 lésions, 3 ont été mieux corrigées par iY, et souvent largement (RC de 0.97/1.06/1.15/1.10 pour iY, comparé à 0.58/0.66/0.80/0.90 pour PVCNet-sino). Sur les métriques globales comme l'erreur quadratiques moyenne et absolue, le VAA (défini Équation (4.14)), PVCNet-sino est systématiquement meilleure avec respectivement 0.90/0.26/21.24% pour les NRMSE/NMAE/VAA, à comparer avec les 1.26/0.42/7.55% de RM et 1.09/0.31/8.55% de iY. Visuellement, sur la Figure 4.17, bien que les contours des ROIs soient presque aussi bien corrigés avec PVCnet-sino qu'avec iY, notre méthode de correction et reconstruction souffre de quelques artefacts apparaissant après reconstruction (notamment sur les bords des reins et du foie).

Méthodes d'entraînement alternatives dans le domaine image

Méthodes post-reconstruction

Nous avons également comparé l'approche décrite dans la section précédente, qui n'opérait la correction que dans le domaine des projections, avec des méthodes d'entraînement et de correction sur les images reconstruites. Un premier réseau de neurones, PVCNet-img, a été entraîné en ne prenant en entrée que deux images I_{att} la carte d'atténuation, et I_{rec10}. La référence utilisée pour l'entraînement supervisée était la carte d'activité exacte générée dans la section précédente mais ré-échantillonnée à la résolution de la TEMP (c'est-à-dire $4.8 \times 4.8 \times 4.8$ mm^3). Un modèle, PVCNet-sino-img, a été entraîné d'une manière similaire mais avec une entrée additionnelle: l'image $I_{\text{PVCNet-sino}}$, obtenue après reconstruction de l'image avec les projections corrigées par le réseau entraîné à la section précédente, $h_{\text{PVCNet-sino}}$. Ce dernier modèle profite donc de l'entraînement à la fois sur les projections et sur les images reconstruites.

Résultats

Les différences et similitudes entre les trois méthodes comparées dans cette section sont rappelés dans le Tableau 5.1.

Sur le jeu de donnée de test, PVCnet-sino-img est le modèle qui a permis d'obtenir l'erreur quadratique moyenne la plus basse, des coefficients de recouvrements proches de 1, ainsi que une moyenne de coefficient Dice la plus élevée, comparée aux autre PVCNet (PVCNet-sino, PVCNet-img) et à la méthode de reconstruction avec RM. Ces résultats peuvent être consultés à la Figure 5.2

Sur la simulation MC du patient, c'est la méthode qui a permis de réduire le plus les erreurs NRMSE/NMAE/VAA (Tableau 5.3). De plus, les deux méthodes entraînées directement sur les images (PVCNet-img et PVCNet-sinot-img) ont fait montre de meilleure capacité à réduire le bruit de fond. Enfin, sur les RC et erreurs absolues locales sur les ROIs, les résultats indiquent une meilleure capacité de cette méthode à estimer l'activité dans la plupart des organes/ROI (fond, vésicule, reins et rate) mais la méthode iY surpasse toujours les autres sur les lésions. Pour rappel, aucune des méthodes PVCNet n'utilise la segmentation des ROIs pour corriger l'image lors de l'application du réseau, contrairement à iY. Visuellement, sur les images (Figure 5.3) et sur les profiles (Figure 5.4), lorsque l'on compare PVCNet-img avec PVCNet-sino-img, il est clair que le premier mène à une image plus flou, dont les contours des région d'intérêt (reins, lésion, foie) sont toujours légèrement affectés par des EVP, alors que bien que l'image obtenue avec PVCNet-sino-img ne soit pas parfaite, les EVPs y sont bien moins présents.

Les résultats de cette étude, bien que préliminaires, nous mènent à la conclusion que le fait d'utiliser deux réseaux (un agissant sur les projections et un autre agissant sur l'image) est l'approche qui mène au meilleur compromis entre la qualité visuelle et quantitative de l'image obtenue.

Conclusion

Dans cette thèse, nous avons montré la possibilité de réduire les EVPs sur les images TEMP grâce à l'apprentissage profond sur des données synthétiques. D'autre part nous avons comparé plusieurs approches et avons conclu que le fait d'utiliser à la fois des données dans le domaine des projections et des données directement dans le domaine image pour l'entraînement de modèles d'apprentissage profond, est l'approche permettant de mieux réduire les EVPs. Les EVPs étant particulièrement problématiques pour arriver à une personnalisation des traitements RIV, nous pensons qu'il sera prochainement possible de les corriger grâce à l'intelligence artificielle lors de la reconstruction de l'image TEMP, à condition d'atteindre des garanties de validité, généralisabilité et explicabilité assez fortes.

Bibliography

[1] S. Agostinelli et al. "Geant4—a simulation toolkit". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303. ISSN: 0168-9002.

[2] A. G. M. Ahmed and P. Wonka. "Screen-space blue-noise diffusion of monte carlo sampling error via hierarchical ordering of pixels". In: *ACM Transactions on Graphics* 39.6 (2020), pp. 1–15. ISSN: 0730-0301, 1557-7368.

[3] M. I. Akerele et al. "Comparison of Correction Techniques for the Spill in Effect in Emission Tomography". In: *IEEE transactions on radiation and plasma medical sciences* 4.4 (2020), pp. 422–432. ISSN: 2469-7311.

[4] E Amato, D Lizio, and S Baldari. "Absorbed fractions in ellipsoidal volumes for radionuclides employed in internal radiotherapy". In: *Physics in Medicine and Biology* 54.13 (2009), pp. 4171–4180. ISSN: 0031-9155, 1361-6560.

[5] H. Arabi, A. AkhavanAllaf, A. Sanaat, I. Shiri, and H. Zaidi. "The promise of artificial intelligence and deep learning in PET and SPECT imaging". In: *Physica Medica* 83 (2021), pp. 122–137. ISSN: 1120-1797.

[6] P. Arce et al. "Gamos: A framework to do Geant4 simulations in different physics fields with an user-friendly interface". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 735 (2014), pp. 304–313. ISSN: 01689002.

[7] K. Assié et al. "Monte Carlo simulation in PET and SPECT instrumentation using GATE". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 527.1 (2004), pp. 180–189. ISSN: 01689002.

[8] C. Bai, L. Shao, A. Da Silva, and Z. Zhao. "A generalized model for the conversion from CT numbers to linear attenuation coefficients". In: *IEEE Transactions on Nuclear Science* 50.5 (2003), pp. 1510–1515. ISSN: 1558-1578.

[9] V. Balaji, T.-A. Song, M. Malekzadeh, P. Heidari, and J. Dutta. "Artificial Intelligence for PET and SPECT Image Enhancement". In: *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine* 65.1 (2024), pp. 4–12. ISSN: 1535-5667.

[10] Bank, Data. "PENELOPE-2014: A code system for Monte Carlo simulation of electron and photon transport". In: (2015).

[11] P. E. Barbano and A. S. Fokas. "Multi-resolution inversion algorithm for the attenuated radon transform". In: *2011 IEEE International Workshop on Machine Learning for Signal Processing*. 2011 IEEE International Workshop on Machine Learning for Signal Processing. 2011, pp. 1–6.

[12] F. Beekman, E. Eijkman, M. Viergever, G. Borm, and E. Slijpen. "Object shape dependent PSF model for SPECT imaging". In: *IEEE Transactions on Nuclear Science* 40.1 (1993), pp. 31–39. ISSN: 1558-1578.

[13] F. J. Beekman, E. T. P. Slijpen, H. W. A. M. de Jong, and M. A. Viergever. "Estimation of the depth-dependent component of the point spread function of SPECT". In: *Medical Physics* 26.11 (1999), pp. 2311–2322. ISSN: 2473-4209.

[14] M. A. Belzunce, A. Mehranian, and A. J. Reader. "Enhancement of Partial Volume Correction in MR-Guided PET Image Reconstruction by Using MRI Voxel Sizes". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 3.3 (2019), pp. 315–326. ISSN: 2469-7303.

[15] V. Bettinardi, I. Castiglioni, E. De Bernardi, and M. C. Gilardi. "PET quantification: strategies for partial volume correction". In: *Clinical and Translational Imaging* 2.3 (2014), pp. 199–218. ISSN: 2281-7565.

[16] J. Biemond, R. Lagendijk, and R. Mersereau. "Iterative methods for image deblurring". In: *Proceedings of the IEEE* 78.5 (1990), pp. 856–883. ISSN: 00189219.

[17] C. M. Bishop. *Pattern recognition and machine learning*. Information science and statistics. New York: Springer, 2006. 738 pp. ISBN: 978-0-387-31073-2.

[18] J. F. Blinn. "Models of light reflection for computer synthesized pictures". In: *Proceedings of the 4th annual conference on Computer graphics and interactive techniques*. SIGGRAPH '77. New York, NY, USA: Association for Computing Machinery, 1977, pp. 192–198. ISBN: 978-1-4503-7355-5.

[19] L. Bodei et al. "Peptide receptor radionuclide therapy with 177Lu-DOTATATE: the IEO phase I-II study". In: *European Journal of Nuclear Medicine and Molecular Imaging* 38.12 (2011), pp. 2125–2135. ISSN: 1619-7089.

[20] G. Boening, P. Pretorius, and M. King. "Study of relative quantification of Tc99 m with partial volume effect and spillover correction for SPECT oncology imaging". In: *IEEE Transactions on Nuclear Science - IEEE TRANS NUCL SCI* 53 (2006), pp. 1205–1212.

[21] A. Bousse et al. "Point spread function optimization in SPECT". In: *IEEE Nuclear Science Symposuim Medical Imaging Conference*. IEEE Nuclear Science Symposuim Medical Imaging Conference. 2010, pp. 2061–2065.

[22] J. Bowsher et al. "Bayesian reconstruction and use of anatomical a priori information for emission tomography". In: *IEEE Transactions on Medical Imaging* 15.5 (1996), pp. 673–686. ISSN: 0278-0062, 1558-254X.

[23] T. J. Bradshaw et al. "Nuclear Medicine and Artificial Intelligence: Best Practices for Algorithm Development". In: *Journal of Nuclear Medicine* 63.4 (2022), pp. 500–510. ISSN: 0161-5505, 2159-662X.

[24] J. F. Briesmeister. "MCNPTM–A General Monte Carlo N–Particle Transport Code". In: *Version 4C, LA-13709-M, Los Alamos National Laboratory* 2 (2000).

[25] M. Bui, A. Robert, S. Rit, and A. Etxebeste. "Experimental evaluation of rigid patient motion detection and correction in SPECT with exponential data consistency conditions". In: *2024 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and Room Temperature Semiconductor Detector Conference (RTSD)*. 2024, pp. 1–1.

[26] M. Bui et al. "Validation of a model of the Symbia Intevo Bold SPECT scanner for Monte Carlo simulations". In: *Nuclear Science Symposium and Medical Imaging Conference (IEEE NSS MIC)*. Vancouver, Canada, 2023.

[27] I. Buvat and F. Orlhac. "The T.R.U.E. Checklist for Identifying Impactful Artificial Intelligence–Based Findings in Nuclear Medicine: Is It True? Is It Reproducible? Is It Useful? Is It Explainable?" In: *Journal of Nuclear Medicine* 62.6 (2021), pp. 752–754. ISSN: 0161-5505, 2159-662X.

[28] T. M. Buzug. *Computed Tomography*. Ed. by R. Kramme, K.-P. Hoffmann, and R. S. Pozos. Berlin, Heidelberg: Springer, 2011. ISBN: 978-3-540-74658-4.

[29] R. E. Caflisch. "Monte Carlo and quasi-Monte Carlo methods". In: *Acta numerica* 7 (1998), pp. 1–49.

[30] T. Cajgfinger, S. Rit, J. M. Létang, A. Halty, and D. Sarrut. "Fixed forced detection for fast SPECT Monte-Carlo simulation". In: *Physics in Medicine and Biology* 63.5 (2018).

[31] L. Carter, J. O. Ramos, and A. Kesner. "Personalized dosimetry of 177Lu-DOTATATE: a comparison of organ- and voxel-level approaches using open-access images". In: *Biomedical physics & engineering express* 7.5 (2021), 10.1088/2057–1976/ac1550. ISSN: 2057-1976.

[32] A. Celler et al. "Evaluation of dead-time corrections for post-radionuclide-therapy 177Lu quantitative imaging with low-energy high-resolution collimators". In: *Nuclear Medicine Communications* 35.1 (2014), p. 73. ISSN: 0143-3636.

[33] L.-T. Chang. "A Method for Attenuation Correction in Radionuclide Computed Tomography". In: *IEEE Transactions on Nuclear Science* 25.1 (1978), pp. 638–643. ISSN: 1558-1578.

[34] X. Chen. "Deep Learning Methods of Attenuation Correction, Reconstruction, Registration, and Denoising for Multi-Modal Cardiac SPECT - ProQuest". Doctor of Philosophy. Yale University, 2024.

[35] X. Chen et al. "DuDoSS: Deep-learning-based dual-domain sinogram synthesis from sparsely sampled projections of cardiac SPECT". In: *Medical Physics* 50.1 (2023), pp. 89–103. ISSN: 2473-4209.

[36] Z. Cheng, J. Wen, J. Zhang, and J. Yan. "Super-resolution reconstruction for parallel-beam SPECT based on deep learning and transfer learning: a preliminary simulation study". In: *Annals of Translational Medicine* 10.7 (2022), p. 396. ISSN: 2305-5839.

[37] S. R. Cherry, J. A. Sorenson, and M. E. Phelps. "Physics in Nuclear Medicine". In: *Physics in Nuclear Medicine (Fourth Edition)*. Ed. by S. R. Cherry, J. A. Sorenson, and M. E. Phelps. Philadelphia: W.B. Saunders, 2012. ISBN: 978-1-4160-5198-5.

[38] C. Chrysostomou, L. Koutsantonis, C. Lemesios, and C. N. Papanicolas. "SPECT Imaging Reconstruction Method Based on Deep Convolutional Neural Network". In: *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). 2019, pp. 1–4.

[39] S. Y. Chun. "The Use of Anatomical Information for Molecular Image Reconstruction Algorithms: Attenuation/Scatter Correction, Motion Compensation, and Noise Reduction". In: *Nuclear Medicine and Molecular Imaging* 50.1 (2016), pp. 13–23. ISSN: 1869-3474.

[40] S. Y. Chun, J. A. Fessler, and Y. K. Dewaraja. "Correction for collimator-detector response in SPECT using point spread function template". In: *IEEE transactions on medical imaging* 32.2 (2013), pp. 295–305. ISSN: 1558-254X.

[41] P. H. van Cittert. "Zum Einfluß der Spaltbreite auf die Intensitätsverteilung in Spektrallinien. II". In: *Zeitschrift für Physik* 69.5 (1931), pp. 298–308. ISSN: 0044-3328.

[42] C. O. da Costa-Luis and A. J. Reader. "Deep Learning for Suppression of Resolution-Recovery Artefacts in MLEM PET Image Reconstruction". In: *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). 2017, pp. 1–3.

[43] A. Da Silva, H. Tang, M. Wu, and B. Hasegawa. "Absolute quantitation of myocardial activity in phantoms". In: *IEEE Transactions on Nuclear Science* 46.3 (1999), pp. 659–666. ISSN: 1558-1578.

[44] A. Dash, M. R. A. Pillai, and F. F. Knapp. "Production of 177Lu for Targeted Radionuclide Therapy: Available Options". In: *Nuclear Medicine and Molecular Imaging* 49.2 (2015), pp. 85–107. ISSN: 1869-3474.

[45] A. R. De Pierro. "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography". In: *IEEE transactions on medical imaging* 14.1 (1995), pp. 132–137. ISSN: 0278-0062.

[46] D. Deidda et al. "Hybrid PET-MR list-mode kernelized expectation maximization reconstruction". In: *Inverse Problems* 35.4 (2019), p. 044001. ISSN: 0266-5611.

[47] M. Del Prete et al. "Personalized 177Lu-octreotate peptide receptor radionuclide therapy of neuroendocrine tumours: initial results from the P-PRRT trial". In: *European Journal of Nuclear Medicine and Molecular Imaging* 46.3 (2019), pp. 728–742. ISSN: 1619-7089.

[48] E. Demirci et al. "Normal distribution pattern and physiological variants of 68Ga-PSMA-11 PET/CT imaging". In: *Nuclear Medicine Communications* 37.11 (2016), p. 1169. ISSN: 0143-3636.

[49] Y. K. Dewaraja, K. F. Koral, and J. A. Fessler. "Regularized reconstruction in quantitative SPECT using CT side information from hybrid imaging". In: *Physics in medicine and biology* 55.9 (2010), pp. 2523–2539. ISSN: 0031-9155.

[50] Y. K. Dewaraja et al. "MIRD Pamphlet No. 23: Quantitative SPECT for Patient-Specific 3-Dimensional Dosimetry in Internal Radionuclide Therapy". In: *Journal of Nuclear Medicine* 53.8 (2012), pp. 1310–1325. ISSN: 0161-5505, 2159-662X.

[51] J. Dick, F. Y. Kuo, and I. H. Sloan. "High-dimensional integration: The quasi-Monte Carlo way". In: *Acta Numerica* 22 (2013), pp. 133–288. ISSN: 0962-4929, 1474-0508.

[52] B. Doignies. "Advanced Sampling for Medical Simulations Using Machine Learning Approaches". PhD thesis. Université Claude Bernard Lyon 1, 2024.

[53] A. Dosovitskiy et al. *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. 2021. arXiv: 2010.11929[cs].

[54] Y. Du, B. M. W. Tsui, and E. C. Frey. "Partial volume effect compensation for quantitative brain SPECT imaging". In: *IEEE transactions on medical imaging* 24.8 (2005), pp. 969–976. ISSN: 0278-0062.

[55] R. L. Eisner et al. "Use of cross-correlation function to detect patient motion during SPECT imaging". In: *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine* 28.1 (1987), pp. 97–101. ISSN: 0161-5505.

[56] K. Erlandsson, I. Buvat, P. H. Pretorius, B. A. Thomas, and B. F. Hutton. "A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology". In: *Physics in Medicine and Biology* 57.21 (2012), R119–R159. ISSN: 0031-9155.

[57] K. Erlandsson and B. Hutton. "A novel voxel-based partial volume correction method for single regions of interest". In: *Journal of Nuclear Medicine* 55 (supplement 1 2014), pp. 2123–2123. ISSN: 0161-5505, 2159-662X.

[58] K. Erlandsson and B. F. Hutton. "Partial volume correction in SPECT using anatomical information and iterative FBP". In: *Tsinghua Science and Technology* 15.1 (2010), pp. 50–55. ISSN: 1007-0214.

[59] K. Erlandsson, B. Thomas, J. Dickson, and B. F. Hutton. "Partial volume correction in SPECT reconstruction with OSEM". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 648 (2011), S85–S88. ISSN: 01689002.

[60] K. Erlandsson, A. Wong, R. Van Heertum, J. Mann, and R. Parsey. "An improved method for voxel-based partial volume correction in PET and SPECT". In: *NeuroImage* 31 (2006), T84. ISSN: 10538119.

[61] A. Etxebeste et al. "CCMod: a GATE module for Compton camera imaging simulation". In: *Physics in Medicine & Biology* 65.5 (2020), p. 055004. ISSN: 0031-9155.

[62] J. Fessler, N. Clinthorne, and W. Rogers. "Regularized emission image reconstruction using imperfect side information". In: *IEEE Transactions on Nuclear Science* 39.5 (1992), pp. 1464–1471. ISSN: 1558-1578.

[63] J. Fessler, Sonka, Milan, and Fitzpatrick, J Michael. "Statistical image reconstruction methods for transmission tomography". In: *Handbook of medical imaging*. Vol. 2. Bellingham, Wash, USA, 2000, pp. 1–70.

[64] N. Fields. *Collimators for Nuclear Medicine Gamma Camera*.

[65] D. Finocchiaro et al. "Partial volume effect of SPECT images in PRRT with 177Lu labelled somatostatin analogues: A practical solution". In: *Physica medica: PM: an international journal devoted to the applications of physics to medicine and biology: official journal of the Italian Association of Biomedical Physics (AIFB)* 57 (2019), pp. 153–159. ISSN: 1724-191X.

[66] A. Flaus et al. "PET image enhancement using artificial intelligence for better characterization of epilepsy lesions". In: *Frontiers in Medicine* 9 (2022), p. 1042706. ISSN: 2296-858X.

[67] C. Floyd, R. Jaszczak, S. Manglos, and R. Coleman. "Compensation for collimator divergence in SPECT using inverse Monte Carlo reconstruction". In: *IEEE Transactions on Nuclear Science* 35.1 (1988), pp. 784–787. ISSN: 1558-1578.

[68] G. Flux et al. "The Impact of PET and SPECT on Dosimetry for Targeted Radionuclide Therapy". In: *Zeitschrift für Medizinische Physik*. Schwerpunkt: Bildgebung in der Nuklearmedizin 16.1 (2006), pp. 47–59. ISSN: 0939-3889.

[69] M. Fontana, D. Dauvergne, J. M. Létang, J.-L. Ley, and Testa. "Compton camera study for high efficiency SPECT and benchmark with Anger system". In: *Physics in Medicine & Biology* 62.23 (2017), p. 8794. ISSN: 0031-9155.

[70] A. R. Formiconi. "Geometrical response of multihole collimators". In: *Physics in Medicine and Biology* 43.11 (1998), pp. 3359–3379. ISSN: 0031-9155, 1361-6560.

[71] E. C. Frey, B. M. W. Tsui, and G. T. Gullberg. "Improved estimation of the detector response function for converging beam collimators". In: *Physics in Medicine and Biology* 43.4 (1998), pp. 941–950. ISSN: 0031-9155, 1361-6560.

[72] A. Frezza et al. "Comprehensive SPECT/CT system characterization and calibration for 177Lu quantitative SPECT (QSPECT) with dead-time correction". In: *EJNMMI Physics* 7.1 (2020), p. 10. ISSN: 2197-7364.

[73] U. Garske-Román et al. "Prospective observational study of 177Lu-DOTA-octreotate therapy in 200 patients with advanced metastasized neuroendocrine tumours (NETs): feasibility and impact of a dosimetry-guided study protocol on outcome and toxicity". In: *European Journal of Nuclear Medicine and Molecular Imaging* 45.6 (2018), pp. 970–988. ISSN: 1619-7089.

[74] R. Gillen et al. "Towards accurate partial volume correction in 99mTc oncology SPECT: perturbation for case-specific resolution estimation". In: *EJNMMI Physics* 9.1 (2022), p. 59. ISSN: 2197-7364.

[75] *Global cancer burden growing, amidst mounting need for services*.

[76] P. Green. "Bayesian reconstructions from emission tomography data using a modified EM algorithm". In: *IEEE Transactions on Medical Imaging* 9.1 (1990), pp. 84–93. ISSN: 02780062.

[77] A. Grings, C. Jobic, T. Kuwert, and P. Ritt. "The magnitude of the partial volume effect in SPECT imaging of the kidneys: a phantom study". In: *EJNMMI Physics* 9.1 (2022), p. 18. ISSN: 2197-7364.

[78] G. T. Gullberg, R. H. Huesman, J. A. Malko, N. J. Pelc, and T. F. Budinger. "An attenuated projector-backprojector for iterative SPECT reconstruction". In: *Physics in Medicine & Biology* 30.8 (1985), p. 799. ISSN: 0031-9155.

[79] F. Gutbrod. "A fast random number generator for the Intel Paragon supercomputer". In: *Computer Physics Communications* 87.3 (1995), pp. 291–306. ISSN: 0010-4655.

[80] J. H. Halton. "On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals". In: *Numerische Mathematik* 2.1 (1960), pp. 84–90. ISSN: 0945-3245.

[81] J. M. Hammersley and D. C. Handscomb. *Monte Carlo Methods*. Dordrecht: Springer Netherlands, 1964. ISBN: 978-94-009-5821-0 978-94-009-5819-7.

[82] A. Hans Vija, A. Yahil, and E. Hawman. "Preprocessing of SPECT projection data: benefits and pitfalls". In: *IEEE Nuclear Science Symposium Conference Record, 2005*. IEEE Nuclear Science Symposium Conference Record, 2005. Vol. 4. 2005, pp. 2160–2164.

[83] K. He, X. Zhang, S. Ren, and J. Sun. "Deep Residual Learning for Image Recognition". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016, pp. 770–778. ISBN: 978-1-4673-8851-1.

[84] E. Henze et al. "Measurements of Regional Tissue and Blood-Pool Radiotracer Concentrations from Serial Tomographic Images of the Heart". In: *Journal of Nuclear Medicine* 24.11 (1983), pp. 987–996. ISSN: 0161-5505, 2159-662X.

[85] E. Hippeläinen, M. Tenhunen, H. Mäenpää, and A. Sohlberg. "Quantitative accuracy of 177Lu SPECT reconstruction using different compensation methods: phantom and patient studies". In: *EJNMMI Research* 6 (2016), p. 16. ISSN: 2191-219X.

[86] N. J. Hoetjes et al. "Partial volume correction strategies for quantitative FDG PET in oncology". In: *European Journal of Nuclear Medicine and Molecular Imaging* 37.9 (2010), pp. 1679–1687. ISSN: 1619-7089.

[87] E. J. Hoffman, S.-C. Huang, and M. E. Phelps. "Quantitation in Positron Emission Computed Tomography: 1. Effect of Object Size". In: *Journal of Computer Assisted Tomography* 3.3 (1979), pp. 299–308. ISSN: 0363-8715.

[88] X. Hong et al. "Enhancing the Image Quality via Transferred Deep Residual Learning of Coarse PET Sinograms". In: *IEEE Transactions on Medical Imaging* 37.10 (2018), pp. 2322–2332. ISSN: 1558-254X.

[89] M. Hotta et al. "PSMA PET Tumor–to–Salivary Gland Ratio to Predict Response to [177Lu]PSMA Radioligand Therapy: An International Multicenter Retrospective Study". In: *Journal of Nuclear Medicine* 64.7 (2023), pp. 1024–1029. ISSN: 0161-5505, 2159-662X.

[90] H. Hudson and R. Larkin. "Accelerated image reconstruction using ordered subsets of projection data". In: *IEEE Transactions on Medical Imaging* 13.4 (1994), pp. 601–609. ISSN: 1558-254X.

[91] D. M. V. Huizing, B. J. de Wit-van der Veen, M. Verheij, and M. P. M. Stokkel. "Dosimetry methods and clinical applications in peptide receptor radionuclide therapy for neuroendocrine tumours: a literature review". In: *EJNMMI Research* 8.1 (2018), p. 89. ISSN: 2191-219X.

[92] B. Hutton, A. Olsson, S. Som, K. Erlandsson, and M. Braun. "Reducing the influence of spatial resolution to improve quantitative accuracy in emission tomography: A comparison of potential strategies". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 569.2 (2006), pp. 462–466. ISSN: 01689002.

[93] B. F. Hutton, I. Buvat, and F. J. Beekman. "Review and current status of SPECT scatter correction". In: *Physics in Medicine & Biology* 56.14 (2011), R85. ISSN: 0031-9155.

[94] B. F. Hutton and Y. H. Lau. "Application of distance-dependent resolution compensation and post-reconstruction filtering for myocardial SPECT". In: *Physics in Medicine and Biology* 43.6 (1998), pp. 1679–1693. ISSN: 0031-9155, 1361-6560.

[95] L. Imbert et al. "Compared Performance of High-Sensitivity Cameras Dedicated to Myocardial Perfusion SPECT: A Comprehensive Analysis of Phantom and Human Images". In: *Journal of Nuclear Medicine* 53.12 (2012), pp. 1897–1903. ISSN: 0161-5505, 2159-662X.

[96] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein. "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation". In: *Nature Methods* 18.2 (2021), pp. 203–211. ISSN: 1548-7105.

[97] U. Jahn et al. "Peptide Receptor Radionuclide Therapy (PRRT) with 177Lu-DOTATATE; Differences in Tumor Dosimetry, Vascularity and Lesion Metrics in Pancreatic and Small Intestinal Neuroendocrine Neoplasms". In: *Cancers* 13.5 (2021), p. 962. ISSN: 2072-6694.

[98] F. James. "A review of pseudorandom number generators". In: *Computer Physics Communications* 60.3 (1990), pp. 329–344. ISSN: 0010-4655.

[99] S. Jan et al. "GATE: a simulation toolkit for PET and SPECT". In: *Physics in Medicine and Biology* 49.19 (2004), pp. 4543–4561. ISSN: 0031-9155.

[100] R. J. Jaszczak, K. L. Greer, C. E. Floyd, C. C. Harris, and R. E. Coleman. "Improved SPECT Quantification Using Compensation for Scattered Photons". In: *Journal of Nuclear Medicine* 25.8 (1984), pp. 893–900. ISSN: 0161-5505, 2159-662X.

[101] C. Jimenez-Mesa et al. "Applications of machine learning and deep learning in SPECT and PET imaging: General overview, challenges and future prospects". In: *Pharmacological Research* 197 (2023), p. 106984. ISSN: 1043-6618.

[102] S. Joe and F. Y. Kuo. "Constructing Sobol Sequences with Better Two-Dimensional Projections". In: *SIAM Journal on Scientific Computing* 30.5 (2008), pp. 2635–2654. ISSN: 1064-8275.

[103] J. Johnson, A. Alahi, and L. Fei-Fei. "Perceptual Losses for Real-Time Style Transfer and Super-Resolution". In: *Computer Vision – ECCV 2016*. Ed. by B. Leibe, J. Matas, N. Sebe, and M. Welling. Cham: Springer International Publishing, 2016, pp. 694–711. ISBN: 978-3-319-46475-6.

[104] T. Kangasmaa, A. Sohlberg, and J. T. Kuikka. "Reduction of Collimator Correction Artefacts with Bayesian Reconstruction in Spect". In: *International Journal of Molecular Imaging* 2011.1 (2011), p. 630813. ISSN: 2090-1720.

[105] T. Kaprelian et al. "Investigating Quasi-Monte Carlo (QMC) methods for Geant4/Gate simulations in medical physics". In: *IV Geant4 International User Conference at the physics-medicine-biology frontier*. Napoli, Italy, 2022.

[106] T. Kaprélian, A. Etxebeste, and D. Sarrut. "Partial Volume Correction on 177 Lu-SPECT sinogram with Deep Learning trained on synthetic data". In: 2024 IEEE Nuclear Science Symposium, Medical Imaging Conference, and Room-Temperature Semiconductor Detectors Symposium. Tampa, Florida, United States, 2024.

[107] T. Kaprélian, A. Etxebeste, and D. Sarrut. "Towards Deep-Learning Partial Volume Correction for SPECT". In: *17th International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*. Stony Brook, NY, United States, 2023.

[108] J. Khan et al. "Dosimetric implications of kidney anatomical volume changes in 177Lu-DOTATATE therapy". In: *EJNMMI Physics* 11.1 (2024), p. 71. ISSN: 2197-7364.

[109] R. Kikinis, S. D. Pieper, and K. G. Vosburgh. "3D Slicer: A Platform for Subject-Specific Image Analysis, Visualization, and Clinical Support". In: *Intraoperative Imaging and Image-Guided Therapy*. Ed. by F. A. Jolesz. New York, NY: Springer, 2014, pp. 277–289. ISBN: 978-1-4614-7657-3.

[110] H. Kim et al. "Physics-Guided Deep Scatter Estimation by Weak Supervision for Quantitative SPECT". In: *IEEE Transactions on Medical Imaging* (2023), pp. 1–1. ISSN: 1558-254X.

[111] M. A. King, R. B. Schwinger, P. W. Doherty, and B. C. Penney. "Two-dimensional filtering of SPECT images using the Metz and Wiener filters". In: *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine* 25.11 (1984), pp. 1234–1240. ISSN: 0161-5505.

[112] M. A. King, P. W. Doherty, R. B. Schwinger, and B. C. Penney. "A Wiener filter for nuclear medicine images". In: *Medical Physics* 10.6 (1983), pp. 876–880. ISSN: 2473-4209.

[113] D. P. Kingma and J. Ba. *Adam: A Method for Stochastic Optimization*. arXiv.org. 2014.

[114] V. Kohli, M. A. King, S. J. Glick, and T.-S. Pan. "Comparison of frequency-distance relationship and Gaussian-diffusion-based methods of compensation for distance-dependent spatial resolution in SPECT imaging". In: *Physics in Medicine and Biology* 43.4 (1998), pp. 1025–1037. ISSN: 0031-9155, 1361-6560.

[115] J. Kurth et al. "External radiation exposure, excretion, and effective half-life in 177Lu-PSMA-targeted therapies". In: *EJNMMI Research* 8.1 (2018), p. 32. ISSN: 2191-219X.

[116] C. Labbé, J. C. Froment, A. Kennedy, J. Ashburner, and L. Cinotti. "Positron Emission Tomography Metabolic Data Corrected for Cortical Atrophy Using Magnetic Resonance Imaging". In: *Alzheimer Disease & Associated Disorders* 10.3 (1996), p. 141. ISSN: 0893-0341.

[117] C. Labbé et al. "Absolute PET Quantification with Correction for Partial Volume Effects within Cerebral Structures". In: *Quantitative Functional Brain Imaging with Positron Emission Tomography*. Ed. by R. E. Carson, M. E. Daube-Witherspoon, and P. Herscovitch. San Diego: Academic Press, 1998, pp. 59–66. ISBN: 978-0-12-161340-2.

[118] K. Lange, M. Bahn, and R. Little. "A Theoretical Study of Some Maximum Likelihood Algorithms for Emission and Transmission Tomography". In: *IEEE Transactions on Medical Imaging* 6.2 (1987), pp. 106–114. ISSN: 1558-254X.

[119] K. Lange, D. R. Hunter, and I. Yang. "Optimization Transfer Using Surrogate Objective Functions". In: *Journal of Computational and Graphical Statistics* 9.1 (2000), pp. 1–20. ISSN: 1061-8600.

[120] P. L'Ecuyer. "Efficient and portable combined random number generators". In: *Communications of the ACM* 31.6 (1988), pp. 742–751. ISSN: 0001-0782.

[121] W. W. Lee and K-SPECT Group. "Clinical Applications of Technetium-99m Quantitative Single-Photon Emission Computed Tomography/Computed Tomography". In: *Nuclear Medicine and Molecular Imaging* 53.3 (2019), pp. 172–181. ISSN: 1869-3474.

[122] C. Lemieux. *Monte Carlo and quasi-Monte Carlo sampling*. Springer Series in Statistics. New York: Springer, 2009. ISBN: 978-0-387-78165-5.

[123] G. Leobacher. "A short introduction to quasi-Monte Carlo option pricing". In: *arXiv:1707.04293 [math, q-fin]* (2014), pp. 191–222. arXiv: 1707.04293.

[124] J. Leube, J. Gustafsson, M. Lassmann, M. Salas-Ramirez, and J. Tran-Gia. "A Deep-Learning-Based Partial-Volume Correction Method for Quantitative 177Lu SPECT/CT Imaging". In: *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine* 65.6 (2024), pp. 980–987. ISSN: 1535-5667.

[125] J. Leube, J. Gustafsson, M. Lassmann, M. Salas-Ramirez, and J. Tran-Gia. "Analysis of a deep learning-based method for generation of SPECT projections based on a large Monte Carlo simulated dataset". In: *EJNMMI Physics* 9.1 (2022), p. 47. ISSN: 2197-7364.

[126] J. Leube, J. Gustafsson, M. Lassmann, M. Salas-Ramirez, and J. Tran-Gia. "Deep learning based partial volume correction for 177Lu SPECT/CT imaging". In: *Journal of Nuclear Medicine* 64 (supplement 1 2023), P1010–P1010. ISSN: 0161-5505, 2159-662X.

[127] J. Leube et al. "Position dependence of recovery coefficients in 177Lu-SPECT/CT reconstructions – phantom simulations and measurements". In: *EJNMMI Physics* 11.1 (2024), p. 52. ISSN: 2197-7364.

[128] T. Leuliet. "Deep learning for tomographic reconstruction : Study and application to computed tomography and positron emission imaging". These de doctorat. Lyon, INSA, 2022.

[129] Z. Li, Y. K. Dewaraja, and J. A. Fessler. "Training End-to-End Unrolled Iterative Neural Networks for SPECT Image Reconstruction". In: *IEEE Transactions on Radiation and Plasma Medical Sciences* 7.4 (2023), pp. 410–420. ISSN: 2469-7311, 2469-7303. arXiv: 2301.09710[eess].

[130] Z. Li, J. A. Fessler, J. K. Mikell, S. J. Wilderman, and Y. K. Dewaraja. "DblurDoseNet: A deep residual learning network for voxel radionuclide dosimetry compensating for single-photon emission computerized tomography imaging resolution". In: *Medical Physics* 49.2 (2022), pp. 1216–1230. ISSN: 2473-4209.

[131] K. Liang, H. Yang, and Y. Xing. *Comparison of projection domain, image domain, and comprehensive deep learning for sparse-view X-ray CT image reconstruction*. arXiv:1804.04289. arXiv, 2019. arXiv: 1804.04289[physics].

[132] H. Lim, Y. K. Dewaraja, and J. A. Fessler. "SPECT Reconstruction With a Trained Regularizer Using CT-Side Information: Application to ^{177}Lu SPECT Imaging". In: *IEEE Transactions on Computational Imaging* 9 (2023), pp. 846–856. ISSN: 2333-9403.

[133] J.-S. Liow and S. C. Strother. "The convergence of object dependent resolution in maximum likelihood based tomographic image reconstruction". In: *Physics in Medicine & Biology* 38.1 (1993), p. 55. ISSN: 0031-9155.

[134] H. Liu et al. "Anatomical-based partial volume correction for low-dose dedicated cardiac SPECT/CT". In: *Physics in Medicine & Biology* 60.17 (2015), p. 6751. ISSN: 0031-9155.

[135] J. Liu et al. "Artificial Intelligence-Based Image Enhancement in PET Imaging: Noise Reduction and Resolution Enhancement". In: *PET Clinics* 16.4 (2021), pp. 553–576. ISSN: 15568598. arXiv: 2107.13595[eess].

[136] J. Liu, Y. Yang, M. N. Wernick, P. H. Pretorius, and M. A. King. "Deep learning with noise-to-noise training for denoising in SPECT myocardial perfusion imaging". In: *Medical Physics* 48.1 (2021), pp. 156–168. ISSN: 2473-4209.

[137] M. Ljungberg and S. E. Strand. "Attenuation and scatter correction in SPECT for sources in a nonhomogeneous object: a monte Carlo study". In: *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine* 32.6 (1991), pp. 1278–1284. ISSN: 0161-5505.

[138] M. Ljungberg. "Absolute Quantitation of SPECT Studies". In: *Seminars in Nuclear Medicine*. Instrumentation 48.4 (2018), pp. 348–358. ISSN: 0001-2998.

[139] M. Ljungberg and S.-E. Strand. "A Monte Carlo program for the simulation of scintillation camera characteristics". In: *Computer Methods and Programs in Biomedicine* 29.4 (1989), pp. 257–272. ISSN: 0169-2607.

[140] M. Ljungberg et al. "MIRD Pamphlet No. 26: Joint EANM/MIRD Guidelines for Quantitative 177Lu SPECT Applied for Dosimetry of Radiopharmaceutical Therapy". In: *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine* 57.1 (2016), pp. 151–162. ISSN: 1535-5667.

[141] I. Loshchilov and F. Hutter. *Decoupled Weight Decay Regularization*. 2019. arXiv: 1711.05101.

[142] L. B. Lucy. "An iterative technique for the rectification of observed distributions". In: *The Astronomical Journal* 79 (1974), p. 745. ISSN: 00046256.

[143] M. Lyra and A. Ploussi. "Filtering in SPECT Image Reconstruction". In: *International Journal of Biomedical Imaging* 2011 (2011), e693795. ISSN: 1687-4188.

[144] L. Lönnblad. "CLHEP—a project for designing a C++ class library for high energy physics". In: *Computer Physics Communications* 84.1 (1994), pp. 307–316. ISSN: 0010-4655.

[145] G. Marin et al. "A dosimetry procedure for organs-at-risk in 177Lu peptide receptor radionuclide therapy of patients with neuroendocrine tumours". In: *Physica Medica* 56 (2018), pp. 41–49. ISSN: 1120-1797.

[146] H. Marquis, K. P. Willowson, and D. L. Bailey. "Partial volume effect in SPECT & PET imaging and impact on radionuclide dosimetry estimates". In: *Asia Oceania Journal of Nuclear Medicine & Biology* 11.1 (2023), pp. 44–54. ISSN: 2322-5718.

[147] H. Marquis et al. "Theranostic SPECT reconstruction for improved resolution: application to radionuclide therapy dosimetry". In: *EJNMMI Physics* 8.1 (2021), p. 16. ISSN: 2197-7364.

[148] G. Marsaglia, A. Zaman, and W. Wan Tsang. "Toward a universal random number generator". In: *Statistics & Probability Letters* 9.1 (1990), pp. 35–39. ISSN: 0167-7152.

[149] K. Matsubara, M. Ibaraki, and T. Kinoshita. *DeepPVC: Prediction of A Partial Volume-Corrected Map for Brain Positron Emission Tomography Studies Via a Deep Convolutional Neural Network*. 2022.

[150] M. Matsumoto and T. Nishimura. "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator". In: *ACM Transactions on Modeling and Computer Simulation* 8.1 (1998), pp. 3–30. ISSN: 1049-3301.

[151] J. E. McNamara et al. "A flexible multicamera visual-tracking system for detecting and correcting motion-induced artifacts in cardiac SPECT slices". In: *Medical Physics* 36.5 (2009), pp. 1913–1923. ISSN: 2473-4209.

[152] T. Merlin et al. "CASToR: a generic data organization and processing code framework for multi-modal and multi-dimensional tomographic reconstruction". In: *Physics in Medicine & Biology* 63.18 (2018), p. 185005. ISSN: 1361-6560.

[153] N. Metropolis. "The beginning of the Monte Carlo method". In: *Los Alamos Science* 15.584 (1987), pp. 125–130.

[154] T. R. Miller and E. S. Rollins. "A practical method of image enhancement by interactive digital filtering". In: *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine* 26.9 (1985), pp. 1075–1080. ISSN: 0161-5505.

[155] G. S. P. Mok and Y. K. Dewaraja. "Recent advances in voxel-based targeted radionuclide therapy dosimetry". In: *Quantitative Imaging in Medicine and Surgery* 11.2 (2021), pp. 483–489. ISSN: 2223-4292.

[156] E. Mora-Ramirez et al. "Comparison of commercial dosimetric software platforms in patients treated with 177Lu-DOTATATE for peptide receptor radionuclide therapy". In: *Medical Physics* 47.9 (2020), pp. 4602–4615. ISSN: 2473-4209.

[157] Mueller, Jennifer L. and Siltanen, Samuli. "Chapter 2: Naïve reconstructions and inverse crimes". In: *Linear and Nonlinear Inverse Problems with Practical Applications*. Computational Science & Engineering. Society for Industrial and Applied Mathematics, 2012, pp. 7–34. ISBN: 978-1-61197-233-7.

[158] P. Mínguez Gabiña et al. "Activity recovery for differently shaped objects in quantitative SPECT". In: *Physics in Medicine & Biology* (2023). ISSN: 0031-9155.

[159] H. W. Müller-Gärtner et al. "Measurement of Radiotracer Concentration in Brain Gray Matter Using Positron Emission Tomography: MRI-Based Correction for Partial Volume Effects". In: *Journal of Cerebral Blood Flow & Metabolism* 12.4 (1992), pp. 571–583. ISSN: 0271-678X.

[160] H. Niederreiter. "Low-discrepancy point sets obtained by digital constructions over finite fields". In: *Czechoslovak Mathematical Journal* 42.1 (1992), pp. 143–166. ISSN: 0011-4642 (print).

[161] J. Nuyts. "Unconstrained image reconstruction with resolution modelling does not have a unique solution". In: *EJNMMI Physics* 1.1 (2014), p. 98. ISSN: 2197-7364.

[162] K. Ogawa, Y. Harata, T. Ichihara, A. Kubo, and S. Hashimoto. "A practical method for position-dependent Compton-scatter correction in single photon emission CT". In: *IEEE Transactions on Medical Imaging* 10.3 (1991), pp. 408–412. ISSN: 1558-254X.

[163] O. Oktay et al. *Attention U-Net: Learning Where to Look for the Pancreas*. 2018. arXiv: 1804.03999[cs].

[164] G. Ongie et al. "Deep Learning Techniques for Inverse Problems in Imaging". In: *IEEE Journal on Selected Areas in Information Theory* 1.1 (2020), pp. 39–56. ISSN: 2641-8770.

[165] A. B. Owen. "Scrambled net variance for integrals of smooth functions". In: *The Annals of Statistics* 25.4 (1997), pp. 1541–1562. ISSN: 0090-5364, 2168-8966.

[166] A. B. Owen. "Scrambling Sobol' and Niederreiter–Xing Points". In: *Journal of Complexity* 14.4 (1998), pp. 466–489. ISSN: 0885-064X.

[167] C. D. Pain, G. F. Egan, and Z. Chen. "Deep learning-based image reconstruction and post-processing methods in positron emission tomography for low-dose imaging and resolution enhancement". In: *European Journal of Nuclear Medicine and Molecular Imaging* 49.9 (2022), pp. 3098–3118. ISSN: 1619-7089.

[168] B. Pan et al. "Ultra high speed SPECT bone imaging enabled by a deep learning enhancement method: a proof of concept". In: *EJNMMI Physics* 9.1 (2022), p. 43. ISSN: 2197-7364.

[169] V. Panin, G. Zeng, and G. Gullberg. "Total variation regulated EM algorithm [SPECT reconstruction]". In: *IEEE Transactions on Nuclear Science* 46.6 (1999), pp. 2202–2210. ISSN: 1558-1578.

[170] B. M. Patrizia Boccacci. *Introduction to Inverse Problems in Imaging*. Boca Raton: CRC Press, 2020. 364 pp. ISBN: 978-0-367-80694-1.

[171] J. A. Patton and T. G. Turkington. "SPECT/CT Physical Principles and Attenuation Correction". In: *Journal of Nuclear Medicine Technology* 36.1 (2008), pp. 1–10. ISSN: 0091-4916.

[172] L. Paulin et al. "Cascaded Sobol' Sampling". In: *ACM Transactions on Graphics*. Proceedings of SIGGRAPH ASIA 2021 40.6 (2021), 274:1–274:13.

[173] L. Paulin et al. "MatBuilder: Mastering Sampling Uniformity Over Projections". In: *ACM Transactions on Graphics* 41.4 (2022), p. 84.

[174] J. Perl, J. Shin, J. Schümann, B. Faddegon, and H. Paganetti. "TOPAS: An innovative proton Monte Carlo platform for research and clinical applications". In: *Medical Physics* 39.11 (2012), pp. 6818–6837. ISSN: 2473-4209.

[175] S. M. B. Peters et al. "Towards standardization of absolute SPECT/CT quantification: a multi-center and multi-vendor phantom study". In: *EJNMMI Physics* 6.1 (2019), p. 29. ISSN: 2197-7364.

[176] L. Polson et al. *PyTomography: A Python Library for Quantitative Medical Image Reconstruction*. 2024. arXiv: 2309.01977.

[177] P. H. Pretorius et al. "Reducing the influence of the partial volume effect on SPECT activity quantitation with 3D modelling of spatial resolution in iterative reconstruction". In: *Physics in Medicine and Biology* 43.2 (1998), pp. 407–420. ISSN: 0031-9155, 1361-6560.

[178] P. H. Pretorius and M. A. King. "Diminishing the impact of the partial volume effect in cardiac SPECT perfusion imaging". In: *Medical Physics* 36.1 (2009), pp. 105–115. ISSN: 2473-4209.

[179] N. Protonotarios. "The Radon transform, its generalization and their applications in PET and SPECT medical imaging". PhD thesis. National Technical University of Athens School of Applied Mathematical and Physical Sciences Department of Mathematics, 2019.

[180] A. Rahmim, J. Qi, and V. Sossi. "Resolution modeling in PET imaging: Theory, practice, benefits, and pitfalls". In: *Medical Physics* 40.6 (2013), p. 064301. ISSN: 2473-4209.

[181] K. Ramonaheng, J. A. van Staden, and H. du Raan. "The effect of calibration factors and recovery coefficients on 177Lu SPECT activity quantification accuracy: a Monte Carlo study". In: *EJNMMI Physics* 8.1 (2021), p. 27. ISSN: 2197-7364.

[182] A. J. Reader. "Self-supervised and supervised deep learning for PET image reconstruction". In: *AIP Conference Proceedings* 3061.1 (2024), p. 030003. ISSN: 0094-243X.

[183] M. P. Reymann et al. "U-Net for SPECT Image Denoising". In: *2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). 2019, pp. 1–2.

[184] W. H. Richardson. "Bayesian-Based Iterative Method of Image Restoration". In: *Journal of the Optical Society of America* 62.1 (1972), p. 55. ISSN: 0030-3941.

[185] S. Rit et al. "The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)". In: *Journal of Physics: Conference Series* 489 (2014), p. 012079. ISSN: 1742-6596.

[186] P. Ritt, H. Vija, J. Hornegger, and T. Kuwert. "Absolute quantification in SPECT". In: *European Journal of Nuclear Medicine and Molecular Imaging* 38.1 (2011), pp. 69–77. ISSN: 1619-7089.

[187] A. Robert. "Correction du mouvement respiratoire en tomographie par émission mono-photonique". These de doctorat. Lyon, 2022.

[188] A. Robert, D. Sarrut, A. Etxebeste, J. M. Létang, and S. Rit. "Ability of exponential data consistency conditions to detect motion in SPECT despite other physical effects". In: *17th International Meeting on Fully 3D Image Reconstruction in Radiology and Nuclear Medicine*. Stony Brook, NY, United States, 2023, pp. 337–340.

[189] O. Ronneberger, P. Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Ed. by N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi. Cham: Springer International Publishing, 2015, pp. 234–241. ISBN: 978-3-319-24574-4.

[190] F. Rosenblatt. "The perceptron: A probabilistic model for information storage and organization in the brain". In: *Psychological Review* 65.6 (1958), pp. 386–408. ISSN: 1939-1471.

[191] O. G. Rousset, Y. Ma, and A. C. Evans. "Correction for Partial Volume Effects in PET: Principle and Validation". In: *Journal of Nuclear Medicine* 39.5 (1998), pp. 904–911. ISSN: 0161-5505, 2159-662X.

[192] L. I. Rudin, S. Osher, and E. Fatemi. "Nonlinear total variation based noise removal algorithms". In: *Physica D: Nonlinear Phenomena* 60.1 (1992), pp. 259–268. ISSN: 0167-2789.

[193] T. Rydén, M. Van Essen, I. Marin, J. Svensson, and P. Bernhardt. "Deep-Learning Generation of Synthetic Intermediate Projections Improves ^{177} Lu SPECT Images Reconstructed with Sparsely Acquired Projections". In: *Journal of Nuclear Medicine* 62.4 (2021), pp. 528–535. ISSN: 0161-5505, 2159-662X.

[194] J. Salvadori et al. "Anatomy-based correction of kidney PVE on 177Lu-SPECT images". In: *EJNMMI Physics* 11.1 (2024), p. 15. ISSN: 2197-7364.

[195] C. Sample et al. "Neural blind deconvolution for deblurring and supersampling PSMA PET". In: *Physics in Medicine & Biology* 69.8 (2024), p. 085025. ISSN: 0031-9155.

[196] J. C. Sanders, P. Ritt, T. Kuwert, A. H. Vija, and A. K. Maier. "Fully Automated Data-Driven Respiratory Signal Extraction From SPECT Images Using Laplacian Eigenmaps". In: *IEEE Transactions on Medical Imaging* 35.11 (2016), pp. 2425–2435. ISSN: 0278-0062, 1558-254X.

[197] A. Saporta, A. Etxebeste, T. Kaprelian, J. M. Létang, and D. Sarrut. "Modeling families of particle distributions with conditional GAN for Monte Carlo SPECT simulations". In: *Physics in Medicine & Biology* 67.23 (2022), p. 234001. ISSN: 0031-9155.

[198] H. Sari et al. "Estimation of an image derived input function with MR-defined carotid arteries in FDG-PET human studies using a novel partial volume correction method". In: *Journal of Cerebral Blood Flow & Metabolism* 37.4 (2017), pp. 1398–1409. ISSN: 0271-678X.

[199] D. Sarrut, N. Krah, J. N. Badel, and J. M. Létang. "Learning SPECT detector angular response function with neural network for accelerating Monte-Carlo simulations". In: *Physics in Medicine & Biology* 63.20 (2018), p. 205013. ISSN: 0031-9155.

[200] D. Sarrut, N. Krah, and J. M. Létang. "Generative adversarial networks (GAN) for compact beam source modelling in Monte Carlo simulations". In: *Physics in Medicine & Biology* 64.21 (2019), p. 215004. ISSN: 0031-9155.

[201] D. Sarrut, A. Etxebeste, T. Kaprelian, A. Saporta, and J. M. Létang. "Annihilation photon GAN source model for PET Monte Carlo simulation". In: *Physics in Medicine and Biology* 68.13 (2023), p. 135018.

[202] D. Sarrut, A. Etxebeste, N. Krah, and J. M. Létang. "Modeling complex particles phase space with GAN for Monte Carlo SPECT simulations: a proof of concept". In: *Phys.Med.Biol.* 66.5 (2021), p. 055014.

[203] D. Sarrut et al. "3D absorbed dose distribution estimated by Monte Carlo simulation in radionuclide therapy with a monoclonal antibody targeting synovial sarcoma". In: *EJNMMI Physics* 4.1 (2017), p. 6. ISSN: 2197-7364.

[204] D. Sarrut et al. "A review of the use and potential of the GATE Monte Carlo simulation code for radiation therapy and dosimetry applications". In: *Medical Physics* 41.6 (2014), p. 064301. ISSN: 2473-4209.

[205] D. Sarrut et al. "Advanced Monte Carlo simulations of emission tomography imaging systems with GATE". In: *Physics in Medicine & Biology* 66.10 (2021), 10TR03. ISSN: 0031-9155, 1361-6560.

[206] D. Sarrut et al. "The OpenGATE ecosystem for Monte Carlo simulation in medical physics". In: *Physics in Medicine & Biology* 67.18 (2022), p. 184001. ISSN: 0031-9155, 1361-6560.

[207] O. Sartor et al. "Lutetium-177–PSMA-617 for Metastatic Castration-Resistant Prostate Cancer". In: *New England Journal of Medicine* 385.12 (2021), pp. 1091–1103. ISSN: 0028-4793, 1533-4406.

[208] J. Seco and F. Verhaegen. *Monte Carlo Techniques in Radiation Therapy*. Boca Raton: CRC Press, 2013. 344 pp. ISBN: 978-1-4665-0792-0.

[209] W. Shao, M. G. Pomper, and Y. Du. "A Learned Reconstruction Network for SPECT Imaging". In: *IEEE transactions on radiation and plasma medical sciences* 5.1 (2021), pp. 26–34. ISSN: 2469-7311.

[210] W. Shao, S. P. Rowe, and Y. Du. "Artificial intelligence in single photon emission computed tomography (SPECT) imaging: a narrative review". In: *Annals of Translational Medicine* 9.9 (2021), p. 820. ISSN: 2305-5839.

[211] W. Shao, S. P. Rowe, and Y. Du. "SPECTnet: a deep learning neural network for SPECT image reconstruction". In: *Annals of Translational Medicine* 9.9 (2021), p. 819. ISSN: 2305-5839.

[212] S. Shcherbinin and A. Celler. "An enhancement of quantitative accuracy of the SPECT/CT activity distribution reconstructions: Physical phantom experiments". In: *Computerized Medical Imaging and Graphics* 34.5 (2010), pp. 346–353. ISSN: 0895-6111.

[213] L. A. Shepp and Y. Vardi. "Maximum Likelihood Reconstruction for Emission Tomography". In: *IEEE Transactions on Medical Imaging* 1.2 (1982), pp. 113–122. ISSN: 1558-254X.

[214] L. Shi, J. A. Onofrey, H. Liu, Y.-H. Liu, and C. Liu. "Deep learning-based attenuation map generation for myocardial perfusion SPECT". In: *European Journal of Nuclear Medicine and Molecular Imaging* 47.10 (2020), pp. 2383–2395. ISSN: 1619-7089.

[215] E. Shimron, J. I. Tamir, K. Wang, and M. Lustig. "Implicit data crimes: Machine learning bias arising from misuse of public data". In: *Proceedings of the National Academy of Sciences of the United States of America* 119.13 (2022), e2117203119. ISSN: 0027-8424.

[216] A. J. D. Silva et al. "Absolute Quantification of Regional Myocardial Uptake of 99mTc-Sestamibi with SPECT: Experimental Validation in a Porcine Model". In: *Journal of Nuclear Medicine* 42.5 (2001), pp. 772–779. ISSN: 0161-5505, 2159-662X.

[217] G. Singh et al. "Analysis of Sample Correlations for Monte Carlo Rendering". In: *Computer Graphics Forum* 38.2 (2019), pp. 473–491. ISSN: 1467-8659.

[218] K. Sjögreen Gleisner et al. "EANM dosimetry committee recommendations for dosimetry of 177Lu-labelled somatostatin-receptor- and PSMA-targeting ligands". In: *European Journal of Nuclear Medicine and Molecular Imaging* 49.6 (2022), pp. 1778–1809. ISSN: 1619-7070.

[219] I. H. Sloan and H. Woźniakowski. "When Are Quasi-Monte Carlo Algorithms Efficient for High Dimensional Integrals?" In: *Journal of Complexity* 14.1 (1998), pp. 1–33. ISSN: 0885-064X.

[220] D. L. Snyder and M. I. Miller. "The Use of Sieves to Stabilize Images Produced with the EM Algorithm for Emission Tomography". In: *IEEE Transactions on Nuclear Science* 32.5 (1985), pp. 3864–3872. ISSN: 1558-1578.

[221] D. L. Snyder, M. I. Miller, L. J. Thomas, and D. G. Politte. "Noise and Edge Artifacts in Maximum-Likelihood Reconstructions for Emission Tomography". In: *IEEE Transactions on Medical Imaging* 6.3 (1987), pp. 228–238. ISSN: 1558-254X.

[222] I. Sobol'. "On the distribution of points in a cube and the approximate evaluation of integrals". In: *USSR Computational Mathematics and Mathematical Physics* 7.4 (1967), pp. 86–112. ISSN: 00415553.

[223] X. Song, W. P. Segars, Y. Du, B. M. W. Tsui, and E. C. Frey. "Fast modelling of the collimator-detector response in Monte Carlo simulation of SPECT imaging using the angular response function". In: *Physics in Medicine and Biology* 50.8 (2005), pp. 1791–1804. ISSN: 0031-9155.

[224] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *The journal of machine learning research* 15.1 (2014), pp. 1929–1958.

[225] J. Stamos, W. Rogers, N. Clinthorne, and K. Koral. "Object-dependent performance comparison of two iterative reconstruction algorithms". In: *IEEE Transactions on Nuclear Science* 35.1 (1988), pp. 611–614. ISSN: 1558-1578.

[226] J. Strosberg et al. "Phase 3 Trial of 177Lu-Dotatate for Midgut Neuroendocrine Tumors". In: *New England Journal of Medicine* 376.2 (2017), pp. 125–135. ISSN: 0028-4793.

[227] K. Thielemans et al. "STIR: software for tomographic image reconstruction release 2". In: *Physics in Medicine and Biology* 57.4 (2012), pp. 867–883. ISSN: 1361-6560.

[228] B. A. Thomas et al. "PETPVC: a toolbox for performing partial volume correction techniques in positron emission tomography". In: *Physics in Medicine and Biology* 61.22 (2016), pp. 7975–7993. ISSN: 1361-6560.

[229] B. A. Thomas et al. "The importance of appropriate partial volume correction for PET quantification in Alzheimer's disease". In: *European Journal of Nuclear Medicine and Molecular Imaging* 38.6 (2011), pp. 1104–1119. ISSN: 1619-7089.

[230] J. Tohka and A. Reilhac. "A Monte Carlo study of deconvolution algorithms for partial volume correction in quantitative PET". In: *2007 IEEE Nuclear Science Symposium Conference Record*. 2006 IEEE Nuclear Science Symposium Conference Record. Vol. 6. 2006, pp. 3339–3345.

[231] J. Tohka and A. Reilhac. "Deconvolution-based partial volume correction in Raclopride-PET and Monte Carlo comparison to MR-based method". In: *NeuroImage* 39.4 (2008), pp. 1570–1584. ISSN: 1053-8119.

[232] L. D. Toso et al. "Improved 3D tumour definition and quantification of uptake in simulated lung tumours using deep learning". In: *Physics in Medicine & Biology* 67.9 (2022), p. 095013. ISSN: 0031-9155.

[233] J. Tran-Gia and M. Lassmann. "Optimizing Image Quantification for 177Lu SPECT/CT Based on a 3D Printed 2-Compartment Kidney Phantom". In: *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine* 59.4 (2018), pp. 616–624. ISSN: 1535-5667.

[234] B. M. W. Tsui and G. T. Gullberg. "The geometric transfer function for cone and fan beam collimators". In: *Physics in Medicine and Biology* 35.1 (1990), pp. 81–93. ISSN: 0031-9155, 1361-6560.

[235] B. Tsui, H. Hu, D. Gilland, and G. Gullberg. "Implementation of simultaneous attenuation and detector response correction in SPECT". In: *Nuclear Science, IEEE Transactions on* 35 (1988), pp. 778–783.

[236] K. Van Laere, M. Koole, I. Lemahieu, and R. Dierckx. "Image filtering in single-photon emission computed tomography: principles and applications". In: *Computerized Medical Imaging and Graphics* 25.2 (2001), pp. 127–133. ISSN: 0895-6111.

[237] B. Vandeghinste. "Iterative reconstruction in micro-SPECT/CT : regularized sparse-view CT and absolute in vivo multi-isotope micro-SPECT quantification". dissertation. Ghent University, 2014.

[238] Y. Vardi, L. A. Shepp, and L. Kaufman. "A Statistical Model for Positron Emission Tomography". In: *Journal of the American Statistical Association* 80.389 (1985), pp. 8–20. ISSN: 0162-1459.

[239] L. Verger et al. "Performance and perspectives of a CdZnTe-based gamma camera for medical imaging". In: *IEEE Transactions on Nuclear Science* 51.6 (2004), pp. 3111–3117. ISSN: 0018-9499.

[240] L. Vergnaud. "Dosimetry for 177Lu and 90Y radionuclide therapies through imaging and Monte Carlo simulations". PhD thesis. INSA de Lyon, 2023.

[241] L. Vergnaud, Y. K. Dewaraja, A.-L. Giraudet, J.-N. Badel, and D. Sarrut. "A review of 177Lu dosimetry workflows: how to reduce the imaging workloads?" In: *EJNMMI Physics* 11.1 (2024), p. 65. ISSN: 2197-7364.

[242] L. Vergnaud et al. "Dosimetric impact of 3D motion-compensated SPECT reconstruction for SIRT planning". In: *EJNMMI Physics* 10 (2023), p. 8. ISSN: 2197-7364.

[243] L. Vergnaud et al. "Patient-specific dosimetry adapted to variable number of SPECT/CT time-points per cycle for ^{177}Lu-DOTATATE therapy". In: *EJNMMI Physics* 9.1 (2022), p. 37. ISSN: 2197-7364.

[244] L. Vergnaud et al. "Performance study of a 360° CZT camera for monitoring 177Lu-PSMA treatment". In: *EJNMMI Physics* 10.1 (2023), p. 58. ISSN: 2197-7364.

[245] J. Violet et al. "Dosimetry of 177Lu-PSMA-617 in Metastatic Castration-Resistant Prostate Cancer: Correlations Between Pretherapeutic Imaging and Whole-Body Tumor Dosimetry with Treatment Outcomes". In: *Journal of Nuclear Medicine* 60.4 (2019), pp. 517–523. ISSN: 0161-5505, 2159-662X.

[246] G. Wang, J. C. Ye, and B. De Man. "Deep learning for tomographic image reconstruction". In: *Nature Machine Intelligence* 2.12 (2020), pp. 737–748. ISSN: 2522-5839.

[247] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. "Image Quality Assessment: From Error Visibility to Structural Similarity". In: *IEEE Transactions on Image Processing* 13.4 (2004), pp. 600–612. ISSN: 1057-7149.

[248] J. Wasserthal et al. "TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images". In: *Radiology: Artificial Intelligence* 5.5 (2023), e230024.

[249] J. F. Williamson. "Monte Carlo evaluation of kerma at a point for photon transport problems". In: *Medical Physics* 14.4 (1987), pp. 567–576. ISSN: 2473-4209.

[250] K. Willowson, D. L. Bailey, and C. Baldock. "Quantitative SPECT reconstruction using CT-derived corrections". In: *Physics in Medicine & Biology* 53.12 (2008), p. 3099. ISSN: 0031-9155.

[251] K. K. Wong et al. "Differences in tumor-to-normal organ SUV ratios measured with 68Ga-DOTATATE PET compared with 177Lu-DOTATATE SPECT in patients with neuroendocrine tumors". In: *Nuclear Medicine Communications* 43.8 (2022), p. 892. ISSN: 0143-3636.

[252] C. Wu et al. "Quantitative multi-pinhole small-animal SPECT: uniform versus non-uniform Chang attenuation correction". In: *Physics in Medicine & Biology* 56.18 (2011), N183. ISSN: 0031-9155.

[253] J. Wu and C. Liu. "Recent advances in cardiac SPECT instrumentation and imaging methods". In: *Physics in Medicine & Biology* 64.6 (2019), 06TR01. ISSN: 0031-9155.

[254] H. Xiang, H. Lim, J. A. Fessler, and Y. K. Dewaraja. "A deep neural network for fast and accurate scatter estimation in quantitative SPECT/CT under challenging scatter conditions". In: *European Journal of Nuclear Medicine and Molecular Imaging* 47.13 (2020), pp. 2956–2967. ISSN: 1619-7089.

[255] H. Xie et al. "Segmentation-free PVC for Cardiac SPECT using a Densely-connected Multi-dimensional Dynamic Network". In: *IEEE transactions on medical imaging* 42.5 (2023), pp. 1325–1336. ISSN: 0278-0062.

[256] B. Yang, L. Ying, and J. Tang. "Artificial Neural Network Enhanced Bayesian PET Image Reconstruction". In: *IEEE transactions on medical imaging* 37.6 (2018), pp. 1297–1309. ISSN: 0278-0062.

[257] J. Yang et al. "Investigation of partial volume correction methods for brain FDG PET studies". In: *IEEE Transactions on Nuclear Science* 43.6 (1996), pp. 3322–3327. ISSN: 1558-1578.

[258] Z. Yu et al. "Need for objective task-based evaluation of deep learning-based denoising methods: A study in the context of myocardial perfusion SPECT". In: *Medical Physics* 50.7 (2023), pp. 4122–4137. ISSN: 2473-4209.

[259] J. Zeintl, A. H. Vija, A. Yahil, J. Hornegger, and T. Kuwert. "Quantitative Accuracy of Clinical 99mTc SPECT/CT Using Ordered-Subset Expectation Maximization with 3-Dimensional Resolution Recovery, Attenuation, and Scatter Correction". In: *Journal of Nuclear Medicine* 51.6 (2010), pp. 921–928. ISSN: 0161-5505, 2159-662X.

[260] G. L. Zeng. "Gibbs Artifact Reduction by Nonnegativity Constraint". In: *Journal of nuclear medicine technology* 39.3 (2011), pp. 213–219. ISSN: 0091-4916.

[261] G. L. Zeng et al. "Iterative Reconstruction of Fluorine-18 SPECT Using Geometric Point Response Correction". In: *Journal of Nuclear Medicine* 39.1 (1998), pp. 124–130. ISSN: 0161-5505, 2159-662X.

[262] G. Zeng, C. Bai, and G. Gullberg. "A projector/backprojector with slice-to-slice blurring for efficient three-dimensional scatter modeling". In: *IEEE Transactions on Medical Imaging* 18.8 (1999), pp. 722–732. ISSN: 1558-254X.

[263] G. Zeng, G. Gullberg, B. Tsui, and J. Terry. "Three-dimensional iterative reconstruction algorithms with attenuation and geometric point response correction". In: *IEEE Transactions on Nuclear Science* 38.2 (1991), pp. 693–702. ISSN: 1558-1578.

[264] H. Zhu and J. Dick. "Discrepancy bounds for deterministic acceptance-rejection samplers". In: *Electronic Journal of Statistics* 8.1 (2014), pp. 678–707. ISSN: 1935-7524, 1935-7524.

FOLIO ADMINISTRATIF

THESE DE L'INSA LYON, MEMBRE DE L'UNIVERSITE DE LYON

NOM : KAPRELIAN
Prénoms : Théo Zadig François

DATE de SOUTENANCE : 21/02/2025

TITRE : Deep Learning for Partial Volume Correction in SPECT Imaging

NATURE : Doctorat

Numéro d'ordre : 2025ISAL0013

Ecole doctorale : Electronique, Electrotechnique, Automatique

Spécialité : Traitement du Signal et de l'Image

RESUME :
La tomographie par émission monophotonique (TEMP) est une modalité d'imagerie nucléaire notamment utilisée en oncologie pour la détection et le suivi des tumeurs. La quantification par TEMP est également devenue cruciale pour la dosimétrie dans le cadre des Radiothérapies Internes Vectorisées (RIV) au 177 Lu-PSMA ou 177 Lu-DOTATATE. Dans cette thèse, nous explorons deux sujets : 1) une méthode Quasi Monte Carlo (QMC) qui pourrait potentiellement accélérer les simulations 2) la correction des effets de volume partiel (EVPs) des images TEMP par apprentissage profond.
Les simulations Monte Carlo (MC) jouent un rôle essentiel en médecine nucléaire, notamment pour le calcul de dose, la conception de nouveaux systèmes TEMP et l'évaluation des algorithmes de reconstruction. Bien que très réalistes et précises les simulations MC sont coûteuses en calcul et lentes à converger.
La première contribution de cette thèse est l'investigation de l'usage des séquences quasi-aléatoires pour l'échantillonnage dans les simulations MC pour la médecine nucléaire. Les séquences quasi-aléatoires sont conçues pour remplir l'espace d'échantillonnage de manière plus homogène que les générateurs de nombres pseudo-aléatoires couramment utilisés. Une suite de Sobol a été utilisée dans Geant4/Gate pour échantillonner les positions, directions et l'énergie des gamma primaires, ainsi que leurs premières interactions Compton. Bien que les premiers résultats ont montré des réductions d'erreur prometteuses allant jusqu'à 40% pour le suivi des particules gamma dans le vide, d'autres expériences, notamment dans l'eau, ont montré que lorsque le nombre de dimensions requises augmentait, le gain induit par l'échantillonneur quasi-aléatoire se réduisait à seulement quelques pourcents.
Une autre limite de la dosimétrie basée sur l'image TEMP est que leur résolution spatiale reste limitée, en raison des EVP. Les EVPs sont une sous-estimation ou surestimation de l'intensité des voxels dans les images reconstruites. Généralement, les EVPs se manifestent par le flou des bords des petites régions ou par la contribution des régions environnantes à l'activité estimée d'un objet. Leur cause principale est la réponse du collimateur-détecteur, définie par sa fonction d'étalement du point (PSF), qui peut être corrigée en la modélisant dans les algorithmes de reconstruction itératifs. Cette méthode n'étant pas suffisante, la deuxième contribution de cette thèse est PVCNet-sino, une méthode de correction de EVPs basée sur l'apprentissage profond, entraînée sur un grand jeu de données de sinogrammes synthétiques pour corriger la PSF sur les projections, avant reconstruction. Les données d'entraînement ont été générées à l'aide de scanners de patients utilisés pour créer 10 000 sources d'activité aléatoires, projetées analytiquement 1) avec modélisation de la PSF et du bruit et 2) sans PSF. Un 3d-Unet a été entraîné pour inverser l'effet de la PSF sur ces sinogrammes. La méthode proposée ne repose pas sur une segmentation de l'image, à l'inverse de la plupart des méthodes de correction de EVPs. Après reconstruction, PVCNet-sino permet une meilleure réduction des EVPs par rapport à d'autres méthodes de correction et a été évaluée sur des données simulées, des expériences sur fantôme et des données de patients.
La correction avec PVCNet-sino a également été comparée à d'autres approches basées sur l'apprentissage profond : une dans le domaine image (PVCNet-img) et une approche opérant dans les deux domaines (PVCNet-sino-img). Les résultats indiquent que l'entraînement à la fois sur les projections et sur l'image mène à la meilleure correction, visuellement et quantitativement. Cependant, toutes les méthodes basées sur l'apprentissage profond proposées dans cette thèse ont montré des difficultés à corriger les petites lésions. A l'avenir, ce travail pourrait améliorer la quantification basée sur la TEMP dans le cadre des RIV et être une alternative aux méthodes basées sur la segmentation.

MOTS-CLÉS : TEMP, Monte Carlo, Apprentissage Profond, EVP, Quasi Monte Carlo
Laboratoire (s) de recherche : CREATIS
Directeur de thèse: David Sarrut
Président de jury : Irène Buvat

Composition du jury : Claude Comtat (Rapporteur), Andrew Reader (Rapporteur), Irène Buvat (Examinatrice), Johannès Tran-Gia (Examinateur).